www.nature.com/hdy

ORIGINAL ARTICLE

Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf

AM Sebbenn[1], ACM Carvalho[2], MLM Freitas[1], SMB Moraes[2], APSC Gaino[2], JM da Silva[2], C Jolivet[3] and MLT Moraes[2]

[1]*Instituto Florestal de São Paulo, Seção de Melhoramento e Conservação Genética Florestal, São Paulo, São Paulo, Brazil;* [2]*Faculdade de Engenharia de Ilha Solteira/UNESP, Departamento de Fitotecnia, Ilha Solteira, São Paulo, Brazil and* [3]*Johann Heinrich von Thünen Institut, Institut für Forstgenetik, Grosshansdorf, Germany*

Over the past century, the Brazilian Atlantic forest has been reduced to small, isolated fragments of forest. Reproductive isolation theories predict a loss of genetic diversity and increases in inbreeding and spatial genetic structure (SGS) in such populations. We analysed eight microsatellite loci to investigate the pollen and seed dispersal patterns, genetic diversity, inbreeding and SGS of the tropical tree *Copaifera langsdorffii* in a small (4.8 ha), isolated population. All 112 adult trees and 128 seedlings found in the stand were sampled, mapped and genotyped. Seedlings had significantly lower levels of genetic diversity ($A = 16.5 \pm 0.45$, mean $\pm$ 95% s.e.; $H_e = 0.838 \pm 0.006$) than did adult trees ($A = 23.2 \pm 0.81$; $H_e = 0.893 \pm 0.030$). Parentage analysis did not indicate any seed immigration ($m_{seeds} = 0$) and the pollen immigration rate was very low ($m_{pollen} = 0.047$). The average distance of realized pollen dispersal within the stand was 94 m, with 81% of the pollen travelling < 150 m. A significant negative correlation was found between the frequency and distance of pollen dispersal ($r = -0.79$, $P < 0.01$), indicating that short-distance pollinations were more frequent. A significant SGS for both adults (~ 50 m) and seedlings (~ 20 m) was also found, indicating that most of the seeds were dispersed over short distances. The results suggested that the spatial isolation of populations by habitat fragmentation can restrict seed and pollen gene flow, increase SGS and affect the genetic diversity of future generations.
Heredity (2011) **106**, 134–145; doi:10.1038/hdy.2010.33; published online 7 April 2010

Introduction

Over the past century, the Brazilian Atlantic forest has experienced significant destruction and fragmentation. The once large and continuous forest has been split into small, well-separated forest fragments (Ribeiro *et al.*, 2009), which have become the source of future tree populations. Approximately 80% of the remaining forest fragments are smaller than 50 ha and are separated from other fragments by an average of 1440 m (Ribeiro *et al.*, 2009). Forest fragmentation decreases the size of the reproductive population, isolates populations and reduces the area available for species regeneration, especially in species that naturally occur at low density (< 5 trees ha^{-1}). This can affect processes such as pollen and seed dispersal as well as mating systems. Consequences include increases in genetic differentiation among populations and intra-population spatial genetic structure (SGS), as well as reductions in genetic diversity

and effective population size (Young *et al.*, 1996; Lowe *et al.*, 2005; Aguilar *et al.*, 2008). The reproductive isolation of small populations is expected to increase the SGS; mating among relatives and inbreeding also increases if the pollen is dispersed at high frequency over short distances. However, although some studies of contemporary gene flow in fragmented landscapes have found that habitat fragmentation can promote long-distance pollen gene flow (Aldrich and Hamrick, 1998; White *et al.*, 2002), others have observed the opposite, finding that isolation of forest stands by fragmentation reduces pollen and seed immigration (Schuster and Mitton, 2000; Robledo-Arnuncio and Gil, 2005; Bittencourt and Sebbenn, 2007). Thus, more studies are necessary to understand the effects of forest fragmentation on the gene flow and mating systems of tree species.

Pollen and seed dispersal (gene flow) are the principal determinants of genetic structure and diversity in tree species (Dick *et al.*, 2008). Estimates of both pollen and seed gene flow are required to predict the effects of forest fragmentation on the genetic structure of a tree species, and special attention should be given to established seedlings and saplings (Aguilar *et al.*, 2008). Contemporary pollen gene flow in plant species has been intensively studied using microsatellite markers and paternity analysis approaches in open-pollinated seeds sampled

Correspondence: Dr AM Sebbenn, Instituto Florestal de São Paulo, Seção de Melhoramento e Conservação Genética Florestal, CP 1322, 01059-970, São Paulo, São Paulo, Brazil.
E-mail: alexandresebbenn@yahoo.com.br

from individual seed-plants (White *et al.*, 2002; Robledo-Arnuncio and Gil, 2005; Bittencourt and Sebbenn, 2007; Bacles and Ennos, 2008). These studies describe effective pollen dispersal (Burczyk *et al.*, 2004), that is, the movement of pollen between the pollen donor and the mother plant before seed establishment. In contrast, gene movement has been studied on a small scale in established seedlings (Meagher and Thompson, 1987; Aldrich and Hamrick, 1998; Asuka *et al.*, 2005; Bacles *et al.*, 2006; Nakanishi *et al.*, 2008), which represent the realized pollen and seed gene flow (Meagher and Thompson, 1987; Burczyk *et al.*, 2004). The use of open-pollinated seeds to study gene flow allows the estimation of the rate of pollen immigration, the pattern and distance of pollen dispersal, differences in male and female fertility and their correlation with tree size. In contrast, the use of established seedlings permits the study of not only the aforementioned factors, but also the rate of seed immigration, distance of seed dispersal and pattern of seed dispersal. Thus, contemporary gene flow can be studied in more detail using established seedlings to provide important insights into the dynamics of natural populations (Burczyk *et al.*, 2004). Furthermore, the use of established seedlings to study contemporary gene flow allows comparisons between the dispersal distances (and patterns) of seeds and pollen.

In many tree species, most of the seeds are dispersed close to the mother tree (Asuka *et al.*, 2005; Bacles *et al.*, 2006; Bittencourt and Sebbenn, 2007; Geng *et al.*, 2008; Nakanishi *et al.*, 2008). Thus, seedlings tend to grow around their mother tree in mixtures of different maternal sib families, such as half-sib (Asuka *et al.*, 2005), full-sib, self-sib and self-half-sibs. In addition, seedlings from different mother trees fathered by the same paternal tree will be paternal half-sibs, thus highlighting the family structure of the grouping (Asuka *et al.*, 2005). The family structure depends on both seed and pollen dispersal. If pollen is dispersed more widely than seeds, the SGS will be influenced more by maternal alleles and determined more by seed dispersal than by pollen dispersal. Meanwhile, if seeds are dispersed more widely than pollen, SGS will be influenced more by paternal alleles and determined more by pollen dispersal distances. If neither seeds nor pollen are dispersed widely, SGS will be influenced by both maternal and paternal alleles and determined by both seed and pollen dispersal. This last case is expected to produce the strongest SGS because of the aggregation of parental alleles in patches around trees of both parents. However, SGS can also be affected by anthropogenic processes such as habitat fragmentation. The isolation of small populations can decrease both seed and pollen dispersal, increasing the levels of relatedness within the remaining stands and, consequently, the level of SGS.

Pollen and seed dispersal can be studied simultaneously using direct methods based on parentage analysis (Meagher, 1986; Marshall *et al.*, 1998) and indirect methods using spatial autocorrelation analysis (Vekemans and Hardy, 2004; Hardy *et al.*, 2006; Nakanishi *et al.*, 2008). Likelihood-based parentage analysis, initially introduced in plants by Meagher (1986), is the most commonly used approach overall. However, SGS analysis, which is based on the estimation of pairwise coancestry coefficients (kinship coefficients) among plants within different distance classes, is the most

commonly used indirect method. Both methods provide information regarding the scale of seed dispersal, but parentage analysis has the advantage of estimating the current rates of seed and pollen immigration inside the sampled population. Parentage analysis can also estimate the pollen and seed dispersal distance for each mating event (Smouse and Sork, 2004; Hardy *et al.*, 2006) in dioecious species if the sex of the plants has been determined (Meagher and Thompson, 1987) or in hermaphroditic and monoecious species if suppositions regarding seed and pollen dispersal distance have been established (Bacles *et al.*, 2006; Nakanishi *et al.*, 2008). Parentage analysis may identify (i) no parent, (ii) at least one parent or (iii) at least one parent-pair for a given seedling (Bacles *et al.*, 2006). Case (i) may represent seed immigration. Meanwhile, case (iii) may represent pollen immigration; if the closest parent was assumed to be the maternal parent, it may also indicate a conservative distance of seed dispersal (Bacles *et al.*, 2006). Using the same assumptions as for case (iii), case (ii) may also represent a conservative distance of seed dispersal (Bacles *et al.*, 2006). In contrast, SGS measures only the scale on which pairwise relatedness is higher than expected for a random distribution, and thus reveals historical patterns of pollen and seed dispersal rather than contemporary ones (Vekemans and Hardy, 2004; Hardy *et al.*, 2006).

Copaifera langsdorffii Desf. is a particularly important tree species in Brazil because of the economic value of its wood and the oil extracted from its trunk. This species is found in Brazil between the Amazon and Santa Catarina state, generally at high population densities (>5 individuals ha^{-1}, Carvalho, 2003). However, most of the species' natural habitats in Brazil (the Atlantic forest) have been destroyed; only 11.4–16% of the original area remains (Ribeiro *et al.*, 2009), and now the species is found only in small and isolated forest fragments. *C. langsdorffii* is a hermaphroditic tropical tree that is pollinated by insects (predominantly bees, *Apis mellifera* and *Trigona* spp.; Carvalho, 2003) and reproduces in a mixed mating system (Oliveira *et al.*, 2002). Trees can begin flowering at approximately 5 years of age, but the flowering and fruiting events are not annual. Years with high fructification are followed by years with low fructification. The fruits are ovoid, surrounded by an abundant coloured aril and carry a single seed. Dispersal probably occurs by zoochory (birds: *Ramphastos toco*, *Cyanocorax cristatellus* and *Turdus rufiventris*; mammals: *Brachyteles arachnoides* and *Cebus apella nigritus*; Carvalho, 2003). Trees can reach 25–40 m in height and 100 cm in diameter at breast height (Carvalho, 2003). However, the annual average growth in height (0.67 m year^{-1}) and diameter at breast height (0.68 cm year^{-1}; Carvalho, 2003) is low.

This work aimed to investigate the patterns of seed and pollen dispersal and the SGS in a small, isolated and fragmented population of *C. langsdorffii*. We assessed these patterns by parentage and autocorrelation analyses using nuclear microsatellite markers. This approach allowed us to test the following two hypotheses: (i) given that this *C. langsdorffii* population is surrounded by residential areas and sugar cane plantations and is isolated from the next conspecific neighbour by 1.2 km, we expected the population to be totally isolated and to experience no seed and pollen immigration; and (ii) if the

population is reproductively isolated and the seeds are dispersed to nearby seed-trees, we expected to find fine-scale SGS, lower genetic diversity, a higher inbreeding coefficient, a higher coancestry coefficient and a lower effective population size in seedlings than in adult trees. The following questions were investigated: (i) Are there differences in genetic diversity between adults and seedlings in the population? (ii) What is the outcrossing rate at the seedling stage? (iii) What is the rate of seed and pollen immigration in the stand? (iv) What are the patterns, including spread distance, of seed and pollen dispersal within the study area? (v) Are there differences in SGS between adults and seedlings in the population? (vi) Are there differences in coancestry and effective population size between adults and seedlings in the population?

Materials and methods

Study site and sampling
This study was conducted in São José do Rio Preto Park (20° 46′ 44. 14′ S, 49° 21′ 17.70′′ W), São José do Rio Preto municipality, São Paulo state, southeast Brazil (Figure 1). The climate in this area is tropical, with dry winters and humid summers. The altitude is 489 m, and the average monthly rainfall ranges from 16 to 241 mm. The annual average temperature is 25.3 °C, averaging 18.9 °C in the coldest month (July) to 27.8 °C in the warmest month (January). The vegetation in this region is characterized as seasonal semi-deciduous Atlantic forest. São José do Rio Preto Park is a small public park of 4.8 ha with a high diversity of tree species (approximately 70), including *Aspidosperma cuspa*, *C. langsdorffii*, *Cedrela fissilis*, *Cordia trichotoma*, *Didymopanax morototonii*, *Schinus terebinthifolius* and *Tabebuia ochracea*, among many others (Stranghetii *et al.*, unpublished). This stand is a remnant of a once continuous semi-deciduous forest. The fragmentation of this area for agricultural purposes occurred 60–80 years ago. The *C. langsdorffii* population in the park is now geographically isolated; the nearest conspecific tree is 1.2 km away. The park is enclosed by a 3 m high wire fence and is surrounded by the town and sugarcane plantations (Figure 1). The stand has also been exploited by selective logging in the past. The stand contains 112 adult *C. langsdorffii* trees (23.3 trees ha^{-1}) with diameter at breast height's ranging from 15 to 93 cm (40.2 cm average) and ages probably ranging from 22 to 140 years. Many of these trees existed pre-fragmentation. In this area, it is also possible to observe *C. langsdorffii* seedlings, with heights ranging from 7 to 116 cm (21.4 cm average), suggesting that these individuals are <2 years of age and probably originated from different reproductive events. These plants are generally found near reproductive trees (Figure 1).

A census was performed in the stand to conduct a genetic study of seed and pollen dispersal. All 112 adult trees and 128 seedlings were sampled and mapped using a GPS III (Garmin, Olathe, KS, USA) (Figure 1). We also measured the diameter at breast height of the adult trees and the height of the seedlings. Leaf tissues and cambium from adult trees were sampled for DNA analysis; only leaf tissues were sampled from seedlings. The distances between adult trees ranged from 0.43 to 257.4 m (average 92.5 m, median 85.7 m), and the
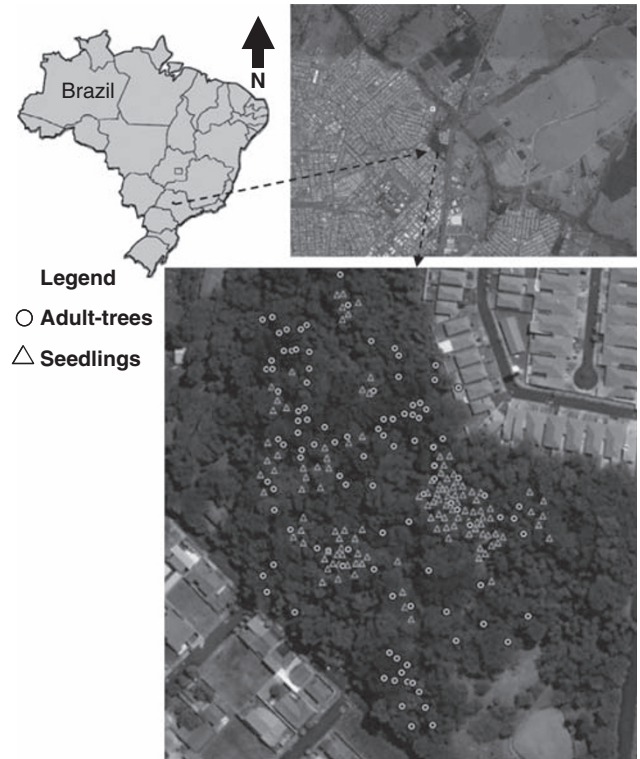
Figure 1 Locations of *C. langsdorffii* adults and seedlings in the analysed forest fragment.

distances between seedlings ranged from 0.34 to 232.2 m (average 67.6 m, median 65.9 m).

DNA extraction and microsatellite analysis
DNA analysis was based on eight nuclear simple sequence repeat markers developed for *C. langsdorffii* by Ciampi *et al.* (2000) (CL01, CL02, CL06, CL20, CL27, CL32, CL34 and CL37). DNA extraction was carried out using the protocol published by Doyle and Doyle (1990). After extraction, DNA was quantified in 1% agarose gels with a lambda DNA standard. After DNA quantification, microsatellite regions were PCR amplified using conditions described by Ciampi *et al.* (2000). The amplification reactions were conducted in a thermocycler (Eppendorf AG, Hamburg, Germany) programmed for the following conditions: 5 min at 95 °C, followed by 29 cycles composed of 1 min denaturation at 94 °C, 1 min annealing at 50–68 °C (depending on the annealing temperature of the primer) and 1 min elongation at 72 °C, followed by a final elongation step of 7 min at 72 °C. The amplified fragments were denatured and separated in a gel composed of 5% polyacrylamide, 8 M urea (mp = 60.06) and 10 × TBE (Tris-Borate 90 mM, EDTA 1 mM, pH 8.0) and visualized with silver nitrate (20%). Fragments of different sizes were scored as different alleles.

Genetic diversity and fixation index analysis
The genetic diversity of adults and seedlings was estimated by the number of alleles (*A*), observed heterozygosity (H_o) and expected heterozygosity at Hardy–Weinberg equilibrium (H_e) for each locus and across all loci. The level of inbreeding among adults and seedlings was estimated using the fixation index (*F*). In

seedlings, the intra-individual fixation index was calculated using reference allele frequencies in adult trees with SPAGEDI version 1.2 (Hardy and Vekemans, 2002). Linkage disequilibrium between all pairs of loci in each sample (adults and seedlings) was calculated using the log-likelihood ratio G-statistic (Goudet, 1995). The significance of the H_o values and linkage disequilibrium were calculated using permutations (10 000) and sequential Bonferroni correction for multiple comparisons (95%, $\alpha = 0.05$). Except for intra-individual fixation indexes in seedlings, all analyses were run using the FSTAT programme, version 2.9.3.2 (Goudet, 1995). To compare the average values of A, H_o and H_e between adult trees and seedlings, the 95% confidence interval (CI) of the standard error (s.e. $\pm 1.96 \times$ s.e.) of these parameters was calculated using a jackknife procedure across all loci.

Parentage analysis

The theoretical power to exclude the first parent was calculated using the CERVUS 3.0 programme (Marshall *et al*., 1998; Kalinowski *et al*., 2007). Parentage analysis was carried out by maximum-likelihood maternity and paternity assignment (Meagher, 1986) based on the multilocus genotypes of the 128 seedlings and all 112 adult reproductive trees of the population. These analyses were performed using CERVUS 3.0. The most likely parents and parent pairs were determined by the Δ statistic (Marshall *et al*., 1998) using the reference allele frequencies calculated in the adult population, as suggested by Meagher and Thompson (1987). To determine the putative mother and father of the seedling, we considered all 112 reproductive trees to be parent candidates. Significance for Δ was determined through maternity and paternity tests simulated by the software (critical Δ) using confidence levels of 80 and 95%, a genotyping error ratio of 0.01 and 1 00 000 repetitions. The calculation of critical Δ values was based on the assumption that 90% of the candidates sampled were located within the plot. If a mother or father candidate or a parent pair had a Δ value higher than the critical Δ value calculated by simulations, it was considered the true parent or the true parent pair. If the same individual was found to be the maternal and paternal parent, this seedling was considered self-progeny. Thus, the estimate of the realized outcrossing rate (t) was calculated as the number of outcrossed ($n_{outcrossed}$) seedlings divided by the total number of assigned ($n_{tatal\_assigned}$) seedlings: $t = n_{outcrossed}/n_{tatal\_assigned}$. If a single parent was identified, it was assumed to be the maternal parent. If two parents were identified inside the population, the closer parent was considered the seed parent (Dow and Ashley, 1998; Bacles *et al*., 2006; Nakanishi *et al*., 2008). These assumptions were used to determine the putative seed-tree and pollen tree because seeds of *C. langsdorffii* are dispersed by frugivore birds and monkeys and pollen is dispersed primarily by bees (Pedroni, 1993). Pedroni (1993) reported that monkeys are the most frequent seed disperser for this species, eating the aril and dispersing the seed close to the seed-tree. This may favour shorter seed dispersal than pollen seed dispersal in the species. Thus, seedlings should probably be nearer to mother trees than to father trees, and the estimates of seed dispersal will represent minimum seed dispersal distance. The cryptic gene flow, or probability of assign-

ing a candidate mother or father inside the population when the true father is outside of the population, was calculated as described by Dow and Ashley (1996). The seed and pollen immigration rate (m) was calculated as the proportion of seedlings that had no parents ($n_{immigrant(seed)}$) or had only one parent ($n_{immigrant(pollen)}$) inside the population, relative to the total number of sampled seedlings (n_{total}): $m = n_{immigran}/n_{total}$ (Burczyk *et al*., 1996). As all sampled individuals have known spatial positions, the effective seed dispersal distance was calculated based on the position of the seedlings relative to their putative mothers, and the distance of pollen dispersal was based on the position of the putative mothers relative to the putative fathers. The coancestry coefficient between seedlings and putative parents was estimated to confirm the parentage assignments. The expected coancestry coefficient value for parent-sib pairs is 0.25. These analyses were conducted with SPAGEDI (Hardy and Vekemans, 2002), using the J Nason estimator as described in Loiselle *et al*. (1995). As coancestry estimates are affected by reference population allele frequencies (Hardy and Vekemans, 2002), we used the reference allele frequencies calculated in the adult population.

SGS analysis

SGS was studied using the estimated average coancestry coefficient (θ_{xy}) between pairs of adult trees and pairs of seedlings. The θ_{xy} coefficients were calculated using SPAGEDI. For analysis of the seedlings, we used the same method applied to the parentage analysis and for analysis of adult trees; we assumed that the reference population has the same allele frequencies found in the present adult generation. To visualize the SGS, θ_{xy} values were averaged over a set of distance classes and then plotted against the distances. We used 16 m distance intervals for seedlings and adults, with a maximum distance of 160 m (the maximum distance that gave a good representation of pairs of individuals for each sample) in each analysis. To test whether the SGS deviated significantly from a random structure, the 95% CI was calculated for each observed value and each distance class from 10 000 permutations of individuals among locations. To compare the SGS of adults and seedlings, the statistic S_p (Vekemans and Hardy, 2004) was calculated as $-b_k/(1-\theta_1)$, where θ_1 is the average pairwise coancestry coefficient calculated between all individuals within the first distance class (0–10 m) and b_k is the slope of the regression of coancestry coefficient on the logarithm of spatial distance (0–160 m). To test for SGS, the spatial position of the individuals was permuted (10 000 times) to obtain the frequency distribution of b_k under the null hypothesis that θ_1 and $\ln(d_{xy})$ were uncorrelated. We also estimated the coancestry coefficients between the seedlings and their putative fathers and mothers assigned by parentage analysis (see above) and between seedlings determined to be full-sibs, half-sibs and non-sibs. Following Hardy *et al*. (2004) and Nakanishi *et al*. (2008), we also conducted a separate analysis of the SGS among gametes (haplotypes) derived from seed and pollen parental genotypes for seedlings in which at least one parent was assigned (Hardy *et al*., 2004). Then, maternal and paternal haplotypes were converted into diploid homozygous genotypes to

calculate the coancestry coefficients for 20 distance classes at 5-m intervals (0–100 m). In cases wherein both seedlings and maternal genotypes were heterozygous for the same alleles (ambiguous paternal contribution), the maternal and paternal haplotypes were converted within the corresponding heterozygous genotypes (Hardy *et al*., 2004; Nakanishi *et al*., 2008). The SGS was then analysed using the method applied to the parentage analysis of seedlings.

Estimating historical gene dispersal from SGS

Historical gene dispersal for adults and seedlings was estimated from SGS assuming that the observed SGS represents an equilibrium isolation-by-distance pattern (Hardy *et al*., 2006). The historical gene dispersal in terms of neighbourhood size (Nb) was estimated as $= -(1-\theta_1)/b_k$ (Vekemans and Hardy, 2004), where b_k is the regression slope within a distance class of $\sigma_g < d_{ij} < 20\sigma_g$. This estimation of Nb is dependent on the effective density D_e (Hardy *et al*., 2006). Thus, D_e was estimated as $D_{e=} D(N_e/N)$, where the effective density is the ratio of the effective population size to the census populations size (Vekemans and Hardy, 2004). After other studies in plants (Hardy *et al*., 2006), we used $D/10$ and $D/2$ as minimum and maximum estimates of D_e. With a fixed D_e, the lower and upper bounds for the 95% CI of Nb were estimated as $Nb_{(lower)} = (\theta_1-1)/(b_k-2SE_b)$ and $Nb_{(upper)} = (\theta_1-1)/(b_k-2SE_b)$, respectively, where SE_b is the s.e. of b_k, calculated by jackknifing data over each loci (Hardy *et al*., 2006). The 95% CI of σ_g was estimated as $\sigma_g = \sqrt{Nb/4\pi D_e}$ using the lower and upper bounds for Nb (Hardy *et al*., 2006). When $b_k < SE_b$, the upper bound was reported as infinite, ∞ (Hardy *et al*., 2006).

Estimates of group coancestry and effective population size

Group coancestry (Lindgren and Mullin, 1998) was estimated separately for adults and seedlings with the average coancestry coefficient between all pairs of individuals (Θ), using the J Nason estimator (Loiselle *et al*., 1995) in the SPAGEDI programme,

$$\Theta = \frac{0.5n(1+F_p) + \sum_{i=1}^{n}\sum_{i\neq j}^{n} \theta_{xy}}{n^2}$$

(Lindgren and Mullin, 1998), where n is the number of sampled individuals, F_p is the inbreeding coefficient for the parental population and θ_{xy} is the pairwise coancestry coefficient between individuals x and y. The effective population size was calculated from Θ as $N_e = 0.5/\Theta$ (Cockerham, 1969). Finally, the relationship between the effective population size and the population size (n) N_e/N was calculated.

Results

Genetic diversity and fixation index in adult trees and seedlings

Using the log-likelihood ratio G-statistic and a sequential Bonferroni correction (95%, $\alpha = 0.05$), no significant linkage disequilibrium was detected in either the adult or the seedling cohorts, indicating that these loci are adequate for use in parentage analysis. All analysed genotypes (adults and seedlings) were unique, indicating an absence of vegetative propagation.

A very high level of polymorphism was detected in the studied population (Table 1). The 112 adult trees had 186 alleles among eight loci, while the 128 seedlings showed 132 alleles. The adult trees had 54 exclusive alleles, yet no exclusive alleles were found in the seedlings. This indicates that not all adult trees successfully participated in reproduction and that there was limited pollen and seed immigration in the forest fragment. According to the 95% CI calculated by the jackknife method, the average number of alleles ($A = 23.2 \pm 0.81$, mean $\pm CI_{95\%}$), observed heterozygosity ($H_o = 0.757 \pm 0.013$) and expected heterozygosity ($H_e = 0.893 \pm 0.030$) in adult trees all differed significantly from the values calculated in the seedlings ($A = 16.5 \pm 0.45$; $H_o = 0.788 \pm 0.008$; $H_e = 0.838 \pm 0.006$). In adults, four (CL02, CL06, CL27 and CL37) of the eight loci and the average fixation index over loci ($F = 0.152$) showed significant deviation from Hardy–Weinberg equilibrium after Bonferroni correction (95%, $\leqslant 0.05$) (Table 1). In the seedlings, five loci (CL01, CL06, CL32, CL34, and CL37) and the average fixation index over loci ($F = 0.124$) showed significant deviation from Hardy–Weinberg equilibrium.

Table 1 Genetic diversity and fixation indexes of adult trees and seedlings in a small, isolated population of *Copaifera langsdorffii*

Locus	Adults (N = 112)					Seedlings (N = 128)			
	A	H_o	H_e	F	$P_{1^0 parent}$	A	H_o	H_e	F_1^1
CL01	33	0.909	0.912	0.003	0.308	17	0.850	0.885	0.144*
CL02	29	0.723	0.909	0.205*	0.311	21	0.815	0.881	0.085
CL06	13	0.477	0.829	0.425*	0.509	7	0.637	0.716	0.257*
CL20	14	0.818	0.855	0.043	0.459	14	0.885	0.865	-0.011
CL27	20	0.757	0.910	0.168*	0.315	16	0.832	0.830	0.039
CL32	34	0.860	0.931	0.076	0.254	20	0.763	0.805	0.149*
CL34	24	0.832	0.892	0.067	0.362	20	0.770	0.876	0.133*
CL37	19	0.682	0.906	0.247*	0.330	17	0.754	0.847	0.177*
Mean	23.2	0.757	0.893	0.152*	—	16.5	0.788	0.838	0.124*
CI$_{95\%}$	± 0.81	± 0.013	± 0.030	—	—	± 0.45	± 0.008	± 0.006	—
Total	186	—	—	—	0.9998	132	—	—	—

*$P < 0.05$. A is the number of alleles, H_o is the observed heterozygosity, H_e is the expected heterozygosity, F is the fixation index, and $P_{1^0 Parent}$ is the power to exclude the first parent. CI$_{95\%}$: 95% standard error calculated by the jackknife method. The parameters A, H_o and H_e are significantly different between adults and seedlings according to the 95% confidence interval. 1: intra-individual fixation index calculated with reference allele frequencies obtained for the adult trees using the SPAGEDI program.

Parentage analysis

The hypothetical total parentage exclusion probability over the eight loci of the first parent was very high ($P_{1^\circ Parent} = 0.9998$; Table 1). Consequently, the probability of cryptic pollen and seed gene flow was low, 0.022 ($1-0.9998^{112}$). This indicates that the observed levels of cryptic pollen and seed flow should not have biased our pollen and seed gene flow estimates.

Mother trees within the stand were assigned to all 128 seedlings (Table 2) with a 95% confidence level ($m_{seeds} = 0$). Pollen donors within the stand were assigned to 122 of the 128 seedlings. For 119 of these 122 seedlings, pollen donors were assigned with a 95% confidence level; three of the donors were assigned with 80% confidence. The remaining six seedlings ($m_{pollen} = 0.047$) most likely represent pollen that originated from trees outside the studied forest fragment. The average pairwise coancestry coefficient (and s.e.) was 0.255 ± 0.018 between seedlings and assigned mother trees and 0.252 ± 0.015 between seedlings and assigned father trees. The 95% confidence s.e. indicates that all of these coancestry values were significantly different from zero. The 128 seedlings were apparently mothered by 45 (40%) of the 112 adult trees in the stand. The 122 seedlings for which the father was found inside the stand were apparently fathered by 45 (40%) of the 112 adult trees. One single seedling was likely a product of selfing ($s = 0.078$). In total, 64 (57%) of 112 reproductive trees had at least one sibling.

The pollen dispersal distance ranged from 4.9 to 229 m, with an average of 94 m. (Table 2, Figure 2a). Approximately 50% of the assigned pollen travelled <86 m and 81% travelled <150 m (Figure 3). There was a significant negative correlation between the number of seedlings ($r = -0.79$, df $= 8$, $P < 0.01$) fertilized by pollen donors and the distance between paternal and maternal trees, suggesting isolation by distance.

Our estimates of minimum seed dispersal distance ranged from 1 to 170 m, with an average of 61 m (Table 2; Figure 3b). Approximately 50% of the seedlings were found within 52 m of the mother tree, and 82% were within 100 m (Figure 2b). A significant and strong negative correlation was found between number of seedlings within a distance class ($r = -0.85$, df $= 9$, $P < 0.01$) and distance between mother tree and seedling, indicating short-distance seed dispersal.

SGS in adults and seedlings

Significant SGS was found within 50 m of adult trees and within 20 m of seedlings (Figures 3a and b). In adults, θ_{xy} values were significantly higher than the upper limit of the 95% CI in the 0–50 m distance class. These decreased to negative values at over 60 m and were significantly lower than the lower limit of the 95% CI between 70 and 160 m. In the seedlings, all coancestry coefficient values were positive and higher than those detected in adult trees of similar distance classes. Coancestry coefficients were significantly higher than the upper limit of the 95% CI within 20 m and significantly lower than the lower limit of the 95% CI between 85 and 130 m.

The average θ_{xy} value for all pairs of adult trees within the first distance class (0–10 m) was lower ($\theta_{xy} = 0.071$) than that observed in seedlings for the same spatial scale; the average values for all pairs of seedlings within the two first classes (0–10 m) was 0.120. The regression slope b_k of the pairwise coancestry coefficient on a logarithm of spatial distance scale (0–160 m) was significantly negative for both adults and seedlings (Table 3), confirming isolation by distance. The intensity of SGS, measured by S_p, was similar in adults ($S_p = 0.0259$) and seedlings ($S_p = 0.0246$).

Estimation of historic gene dispersal

Our estimates of historical gene dispersal distance, based on restricted b_k (0–160 m), produced a $\sigma_g = 44$ m and a Nb of 32 for seedlings (assuming $D_e = D/2$). For seedlings, assuming $D_e = D/10$, the estimates of σ_g did not converge. For the adults, the estimates of σ_g did not converge for the tested effective densities, even if we used the observed density as the effective density (results not shown).

SGS of maternal and paternal alleles

We analysed 128 seedlings for maternal alleles and 122 seedlings for paternal alleles to study the spatial patterns of seed and pollen dispersal. The patterns of SGS detected for both maternal and paternal alleles were very similar (Figure 4). The θ_{xy} values estimated for both maternal and paternal alleles were positive in all studied distance classes. However, significant positive θ_{xy} coefficients were found between 0 and 10 m for both maternal and paternal alleles. In the two distance classes within this range, the estimated θ_{xy} values were close to those expected for half-sibs.

## Coancestry coefficients and effective population size in adults and seedlings

The average coancestry coefficient (Θ) among all pairs of adult trees was 0.008. Assuming random mating, this indicates that the expected rate of biparental inbreeding is very low ($<1\%$). The estimated effective size of the

Table 2 Realised pollen and seed dispersal and average coancestry between pollen donors and seed parents with respective assigned seedlings in a small, isolated population of *Copaifera langsdorffii*

Sample	n	s (%)	Dispersal		Dispersal distance (m)			Coancestry	
			Within	Outside	Mean	Median	Min/max	Mean	Min/max
Pollen	128	0.78 (1)[a]	95% (122)	5% (6)	94 ± 55	86	4.9/229	0.252 ± 0.015	0.001/0.527
Seeds	128	—	100% (128)	0% (0)	61 ± 41	52	1/170	0.255 ± 0.018	−0.110/0.527

Min/max are the minimum and maximum dispersal distances. *n* is the sample size and *s* is the selfing rate. Mean coancestry $\pm$ s.d.
[a]Estimated individual inbreeding coefficient of 0.322 (expected = 0.5).
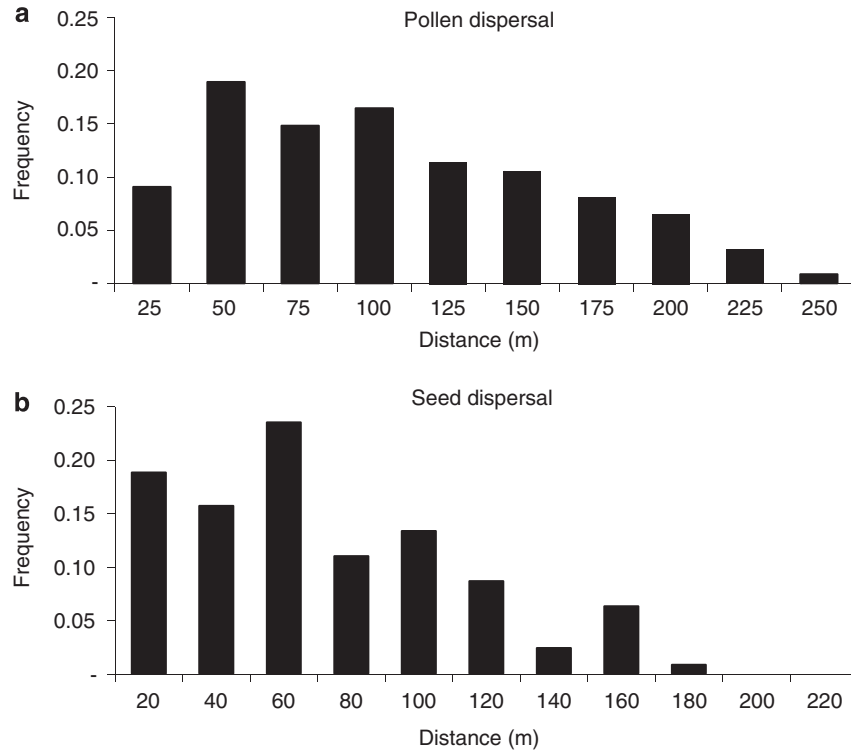
Figure 2 Frequency distributions of (**a**) realized (black bars) pollen dispersal distances and (**b**) seed dispersal distances.
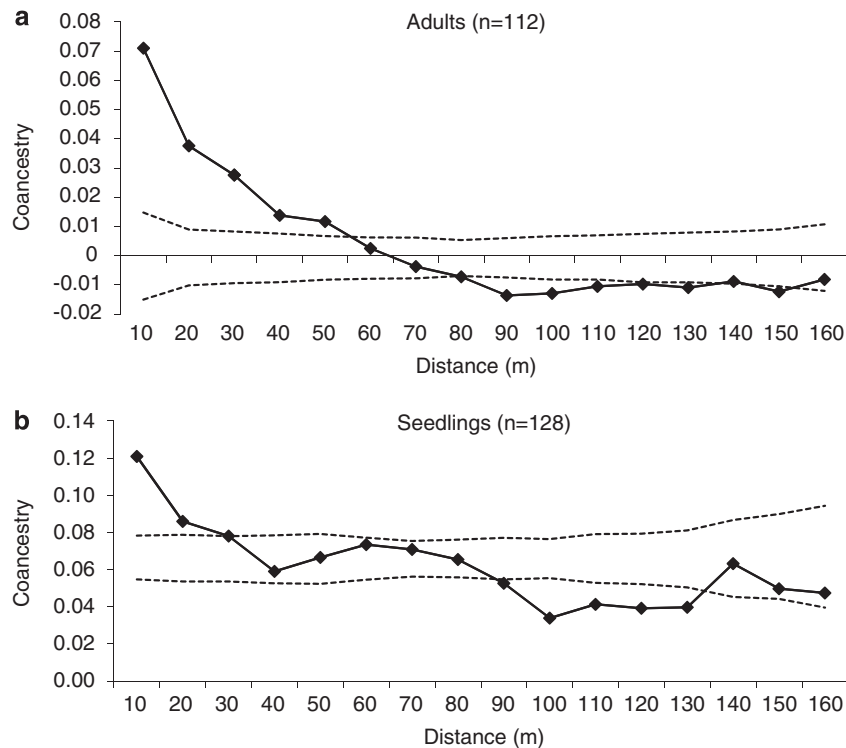
Figure 3 Correlograms of average coancestry coefficients (θ_{xy}) of *C. langsdorffii* adults (**a**) and seedlings (**b**) for 16 distance classes with intervals of 10 m for adults and 5 m for seedlings. The solid line represents the average θ_{xy} value. The dashed lines represent the 95% (two-tailed) CI of the average θ_{xy} distribution calculated from 10 000 permutations of spatial distance among pairs of adults and seedlings.

reproductive population indicates that the 112 adult trees correspond to 64 unrelated and non-inbred individuals ($N_e/N = 0.57$). Similarly, the average Θ among all pairs of seedlings was 0.011. The estimated N_e indicates that the 128 seedlings correspond to 44 unrelated and non-inbred individuals ($N_e/N = 0.35$).

Table 3 Estimates of spatial genetic structure parameters for *Copaifera langsdorffii* adults and seedlings

Sample	θ_1	b_k (0–160)	S_p (s.e.)	Nb (s.e.)	σ_g s.e. (m) $(D_e = D/2)$	σ_g s.e. (m) $(D_e = D/10)$
Adults	0.071***	−0.0241***	0.0259 (0.0054)	NC	NC	NC
Seedlings	0.120***	−0.0217***	0.0246 (0.0070)	32 (13–∞)	44 (1.46–∞)	NC

Abbreviation: NC, not converged.

*** $P < 0.01$. θ_1 is the coancestry coefficient in the first distance class; b_k is the of the regression slope of the coancestry coefficient value on the logarithm of spatial distance between individuals separated by 0 to 160 m (the maximum distance that gave a good representation of pairs of individuals for each sample); S_p is the statistic that measures the extent of spatial genetic structure in the first distance class (0–160 m); Nb is the neighbourhood size; σ_g is the historical gene dispersal distance for two estimates of effective density ($D_e = D/2$ and $D_e = D/10$).
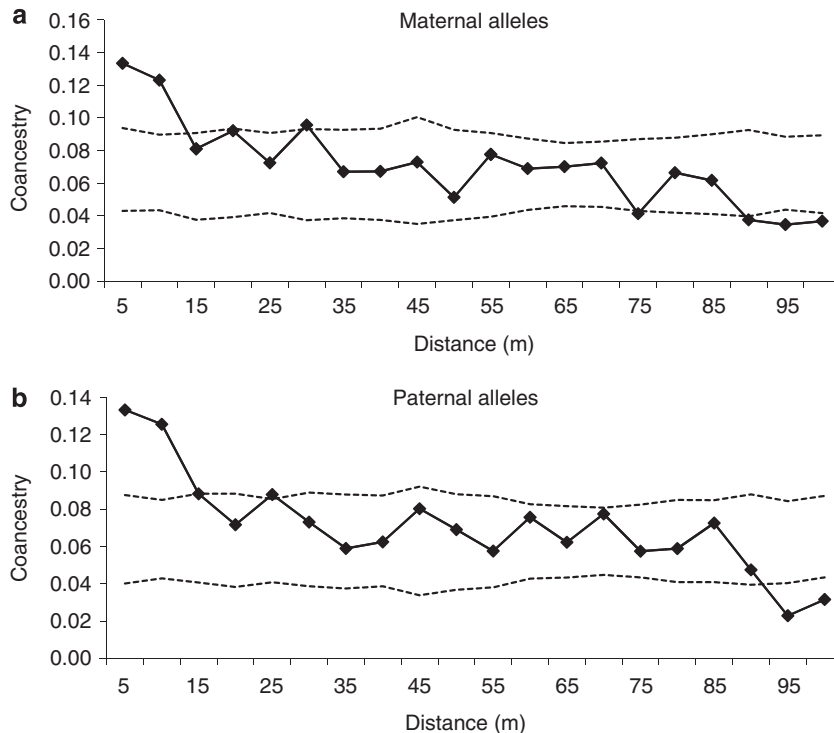
Figure 4 Correlograms of average coancestry coefficients (θ_{xy}) of (**a**) maternal alleles and (**b**) paternal alleles of *C. langsdorffii* for 10 distance classes with intervals of 5 m. The solid line represents the average θ_{xy} value. The dashed lines represent the 95% (two-tailed) CI of the average θ_{xy} distribution calculated from 10 000 permutations of spatial distance among pairs of adults and seedlings.

Discussion

Genetic diversity in adult trees and seedlings
Forest fragmentation can cause genetic drift because of non-random mating and thus it can reduce the genetic diversity of populations (Young *et al.*, 1996; Jump and Peñuelas, 2006). Although the seedling sample (128) was bigger than the adult tree sample (112), the adults had approximately 30% more alleles and significantly higher expected heterozygosity than did the seedlings. This loss of genetic diversity in seedlings is possibly because of deviations from random mating through, for instance, asynchronism in flowering phenology, resulting in genetic drift. That is, for non-random mating because of differences in male or female fertility, individual variation in flowering phenology can reduce the number of trees involved in mating events, thus reducing the number of trees that contribute their genes to the next generation. As a consequence, not all alleles present in the parental population will be present in regeneration stage plants. In this population, individual variation in flowering phenology was observed, and not all adult trees produced flowers and seeds every year (Stranghetii *et al.*, unpublished). We believe that this is the main cause of genetic erosion in the seedlings. Other potential factors include the intensity of predation on the seedlings and/or high inbreeding depression, resulting in lower seed viability. Inbreeding could be attributed to selfing and/or mating among relatives, considering that many near neighbour adult trees are related, according to SGS analysis.

Inbreeding in adult trees and seedlings
A significant excess of homozygosity was detected in adult trees and seedlings. Excess homozygosity suggests inbreeding, the presence of null alleles in the analysed loci or the occurrence of the Wahlund effect. As *C. langsdorffii* is a hermaphroditic tree, the observed excess homozygosity could be explained by selfing and mating among relatives. As the inbreeding in one

generation is equal to the coancestry coefficient among the parents, the observed levels of inbreeding in adults ($F = 0.152$) and seedlings ($F = 0.124$) can be explained by mating among half-sibs ($\theta_{xy} = 0.125$).

Null alleles increase the estimate of the fixation index because heterozygous individuals are misclassified as homozygous because of the lack of amplification of the null allele in PCR analysis. Thus, the positive and significant fixation index may be partly artifactual because of the presence of null alleles. The last explanation for the high fixation index could be subpopulation differentiation resulting in the Wahlund effect. As the fixation index was estimated from the whole population, the apparent inbreeding may reflect an excess of homozygotes because of subdivision of the population resulting from SGS (Bittencourt and Sebbenn, 2007). These three factors may have contributed to the very high fixation index observed in both adult trees and seedlings.

Outcrossing rate

As the high fixation indices observed in both adults and seedlings may have been caused by selfing, we analysed the mating system to test the likelihood of that scenario. Our results suggest a very high realized outcrossing rate among the studied seedlings ($t = 0.9992$). This value is higher than has been previously estimated for open-pollinated seeds ($t_m = 0.917$, Oliveira *et al.*, 2002). Both results confirm that most populations of the species reproduce predominantly by outcrossing. However, the higher outcrossing rate detected herein is probably partly because of the fact that this was estimated at the seedling stage (realized outcrossing rate), whereas Oliveira *et al.* (2002) analysed open-pollinated seeds (effective outcrossing rate). Studies of mating systems in tropical tree species have suggested that the outcrossing rate can change between fertilization and seedling establishment. For example, Hufford and Hamrick (2003), using the same cohort in the tropical tree *Platypodium elegans,* found a lower outcrossing rate in mature seeds (0.82) than in seedlings (0.91). The decrease in the selfing rate ($s = 1\text{-}t$) caused by early selection was attributed to inbreeding depression between the mature seed stage and the seedling stage. Similar results were reported for the tropical tree *Neobalanocarpus heimii* (Naito *et al.*, 2005). Thus, the higher outcrossing rate we observed at the seedling stage compared with the rate observed by Oliveira *et al.* (2002) for mature seeds very likely resulted from inbreeding depression. This is consistent with the excess homozygosity found in the population, as discussed above.

Seed and pollen immigration

No seed immigration ($m_{seeds} = 0$) was observed in the study population. We expected this result because the forest fragment is physically isolated by a 3 m wire fence and is geographically isolated from other conspecifics by at least 1.2 km by the town and sugar-cane plantations. *C. langsdorffii* seeds are dispersed mainly by monkeys and birds. The study area is home to two species of monkeys, *Alouatta guariba* and *Callithrix penicillata*, which are both probable dispersers of *C. langsdorffii* seeds. However, the physical isolation of the forest fragment probably limits their movement

between stands. Monkeys also eat seed arils and regurgitate the seeds in or near the canopy of the tree from which they collect the fruits, thus dispersing seeds near the seed-tree (Pedroni, 1993). As birds are able to disperse seeds over relatively long distances, the geographical isolation of the stand was not expected to limit seed dispersal by birds to the same extent as seed dispersal by monkeys. Many bird species known to be dispersers of *C. langsdorffii* seeds were observed in the stand, including *Turdus rufiventris* and *Ramphastos toco*. Thus, a possible explanation for the absence of seed immigration is the absence of *C. langsdorffii* seed resources near the study stand. Natural forests in this region are strongly fragmented, and today only small forest fragments, which are isolated by sugar cane plantations, remain. The absence of reproductive *C. langsdorffii* individuals and populations close to the study stand could be the main explanation for the absence of seed immigration in the current population.

The results also showed a very low pollen immigration rate ($m_{pollen} = 4.7\%$). This rate of pollen immigration was lower than the rates found for populations of tree species located in more isolated stands: minimum 31.6% in *Swietenia humilis*, isolated by more than 1 km (White *et al.*, 2002), minimum 33.3% in *Eucalyptus wandoo*, isolated by 180–1080 m (Byrne *et al.*, 2008), and 39.4% in *Sorbus terminalis* isolated by 0.4–6 km (Hoebee *et al.*, 2007). The behaviour of pollinators partially determines the distance over which pollen can be dispersed (Dick *et al.*, 2008). *C. langsdorffii* is predominantly pollinated by bees (*Apis mellifera* and *Trigona* spp., Carvalho, 2003). Bees are capable of long-distance pollen dispersal and can transfer pollen between trees over distances up to several kilometres (Ghazoul, 2005). Studies of tropical bee-pollinated trees have also reported long-distance pollen dispersal (>1 km; White *et al.*, 2002; Dick *et al.*, 2003; Byrne *et al.*, 2008). Native and exotic bees are common insects in the study region, and both *Apis mellifera* and *Trigona* spp. were observed in the stand. Thus, the low pollen immigration rate is another strong indicator of the low frequency of other reproductive individuals of the species in the region.

Seed dispersal distance

The results showed a high frequency of short-distance seed dispersal (average of 61 m), with 82% of seeds dispersed within 100 m of the seed-tree (Figure 2b). Furthermore, there was a significant strong negative correlation between the number of seedlings and the distance between a mother tree and its seedlings, as well as a significantly positive SGS. Our assumption that the closest parent detected by parentage analysis was the seed parent may have affected our seed dispersal distance estimate. In fact, this assumption can result in an underestimation of the distance of seed dispersal (Hardesty *et al.*, 2006). Thus, the observed distance of seed dispersal must be interpreted with caution because it represents a minimum estimate. However, a high frequency of short-distance seed dispersal has been reported in many studies of gene flow in trees based on paternity analysis and genetic markers: 37% within 100 m in *Fraxinus excelsior* (Bacles *et al.*, 2006); 50% within 51 m in *Prunus mahaleb* (Jordano *et al.*, 2007); and 47% within 60 m in *Araucaria angustifolia* (Bittencourt and

Sebbenn, 2007). In contrast, Hardesty *et al*. (2006) found long-distance seed dispersal for the tropical dioecious tree *Simarouba amara*, with 74% of the seedlings established more than 100 m from the mother tree (mean of 197.7 m). The main animal seed dispersers of *C. langsdorffii* are monkeys and birds, which both swallow arils and eventually regurgitate the seeds in the canopy of the seed-tree or on a neighbouring tree (Pedroni, 1993; Carvalho, 2003). This could explain the short seed dispersal distances. Thus, the behaviour of the seed dispersers of this species probably explains the short-distance seed dispersal that we observed. The small size of the stand and its use as a public park may also influence the distance of seed dispersal by limiting the available seedling establishment areas to those nearest the seed-trees.

In spite of the assumption that the closest parent of the seedlings was the seed parent and the more distant parent was the pollen parent, our estimate of historic gene dispersal distance measured in seedlings (σ_g) was shorter (average of 44 m) than the direct estimates of seed (average of 61 m) and pollen (average 94 m) dispersal distance. This result confirms the limited gene dispersal distance in the present population of *C. langsdorffii* and suggests a neighbourhood area of 33 individuals.

Pollen dispersal distance

The pattern of realized pollen dispersal detected in the study population of *C. langsdorffii* was that of nearest neighbour pollen dispersal, with an average pollen dispersal distance of 94 m and 50% of pollen dispersed within 86 m (Figure 2a). This pattern of pollen dispersal has been observed in many other animal-pollinated tropical tree species (Hoebee *et al*., 2007; Lourmas *et al*., 2007; Lacerda *et al*., 2008; Naito *et al*., 2008) and may be explained by the pollinator behaviour (bees) and flowering density. Some bee species are expected to forage between near neighbours (Dick *et al*., 2008), which can increase the frequency of short-distance pollen dispersal. However, the distance to the next near neighbour will depend on the individual flowering phenology. In our study, not all trees present flowers each year, introducing a stochastic component in the distance and pattern of pollen dispersal that may change between different reproductive events. We will study these processes in more detail in the future using open-pollinated tree families.

SGS in adults and seedlings

A strong positive SGS was detected in both adult trees and seedlings, suggesting isolation by distance (Table 3, Figures 3a and b). The extent of SGS measured by S_p (Table 3) was similar between adults and seedlings, although seedlings showed a higher coancestry coefficient than adults within short-distance classes (0–10 m). This phenomenon likely occurs because seedlings reflect recent seed and pollen dispersal events, while adults reflect historic seed and pollen dispersal events. However, increased mortality of juveniles because of local competition could provide another explanation for the lower levels of coancestry in adults compared with seedlings. Parentage analysis of seedlings also showed a high frequency of short seed and pollen dispersal

distances. This explains the observed pattern of SGS in the seedlings and perhaps also in adults.

The results showed relatively strong and similar levels of SGS for both maternal and paternal alleles (Figures 3a and b). In the first two distance classes (0–5 and 5–10 m), the coancestry coefficient was estimated to be equal to that expected between half-sibs ($\theta_{xy} = 0.125$). In a similar study of a population of *Quercus salicina* in Japan, Nakanishi *et al*. (2008) found large differences in the SGS of maternal and paternal alleles. Maternal alleles showed stronger SGS than did paternal alleles, and the coancestry coefficient in the first distance class was 0.275 for maternal alleles and only 0.046 for paternal alleles. According to the investigators, these differences were probably because of different pollen and seed dispersal distances, with pollen dispersal greater on average than seed dispersal. Thus, limited seed dispersal was the main cause of SGS in that study. In contrast, in this study, as seeds were apparently dispersed near to seed-trees and pollen was also frequently dispersed over short distances, the observed pattern was similar for both maternal and paternal alleles. This suggests that seeds and pollen contributed equally to the observed SGS at the seedling stage, although inbreeding depression and the hypothesis that the closest tree to a seedling was its mother may have affected these results.

Effective population size

Forest fragmentation has exposed tree populations to the deleterious effects of reduced gene flow (Bacles *et al*., 2006). Reduced gene flow in small populations is expected to increase the coancestry and inbreeding within populations. The average adult population coancestry coefficient, or group coancestry ($\Theta = 0.007798$), suggests that when random mating occurs, a very low level of biparental inbreeding is expected ($<1\%$). This is lower than the level of inbreeding detected in the seedlings ($F_p = 0.124$) and indicates a deviation from random mating, in contrast to the observed random pollen dispersal. On the basis of the group coancestry, we estimated that the 112 adults and 128 seedlings represent an effective population size (N_e) of only 64 and 44, respectively non-relatives and non-inbred individuals. This low N_e reflects the high proportion of related and inbred individuals within the population. If the level of isolation observed herein continues for many generations in this population, it could become extinct through the deleterious effects of genetic drift (increased coancestry and inbreeding, loss of genetic diversity and reduced effective population size). The introduction of different genotypes from different populations in the same region (to avoid the effects of outbreeding depression) could reduce the level of group coancestry and delay the deleterious effects of genetic drift. However, the geographic scale at which inbreeding and outbreeding depression occur in this species is still poorly characterized.

In summary, our results suggest that the spatial isolation of populations by habitat fragmentation may reduce genetic diversity and effective population size, restrict pollen and seed gene flow and increase the SGS of new generations. However, it is important to note that our study was not repeated in other populations to confirm the observed low levels of gene flow. This is

significant because the gene flow as well as the mating system can change from one reproductive event to another, and the pattern observed here could be very different from that observed in other populations or in other reproductive events of the same population. Thus, it is important to repeat this study in other populations and/or reproductive events to confirm the observed results.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

References

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Loco J (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* **17**: 5177–5188.

Aldrich PR, Hamrick JL (1998). Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science* **281**: 103–105.

Asuka Y, Tomaru N, Munehara Y, Tani N, Tsumaru Y, Yamamoto S (2005). Half-sib family structure of *Fagus crenata* samplings in an old-growth beech-dwarf bamboo forest. *Mol Ecol* **14**: 2565–2575.

Bacles CFE, Ennos RA (2008). Paternity analysis of pollen-mediated gene flow for *Fraxinus excelsior* L in a chronically fragmented landscape. *Heredity* **101**: 368–380.

Bacles CFE, Lowe A, Ennos R (2006). Effective seed dispersal across a fragmented landscape. *Science* **311**: 628.

Bittencourt JM, Sebbenn AM (2007). Patterns of pollen and seed dispersal in a small fragmented population of a wind pollinated *Araucaria angustifolia* in southern Brazil. *Heredity* **99**: 580–591.

Burczyk J, Adams WT, Shimizu JY (1996). Mating patterns and pollen dispersal in a natural knobcone pine (*Pinus attenuata* Lemmon) stand. *Heredity* **77**: 251–260.

Burczyk J, DiFazio SP, Adams WT (2004). Gene flow in forest trees: how far do genes really travel? *Forest Genet* **11**: 1–14.

Byrne M, Elliott CP, Yates CJ, Coates DJ (2008). Maintenance of high pollen dispersal in *Eucalyptus wandoo*, a dominant tree of the fragmented agricultural region in Western Australia. *Conserv Genet* **9**: 97–105.

Carvalho PER (2003). *Espécies arbóreas brasileiras*. Embrapa Informação Tecnológica: Brasília.

Ciampi AY, Brondani R, Grattapaglia D (2000). Desenvolvimento de Marcadores Microssatélites para *Copaifera langsdorffii* Desf. (*Copaíba*)—Leguminosae-Caesalpinioideae e Otimização de Sistemas Fluorescentes de Genotipagem Multiloco. Boletim de Pesquisa: Embrapa, **16**, pp 1–40.

Cockerham CC (1969). Variance of gene frequencies. *Evolution* **23**: 72–84.

Dick CW, Etchelecu G, Austerlitz F (2003). Pollen dispersal of Neotropical trees (*Dinizia excelsa*: Fabaceae) by native insects and Africa honeybees in pristine and fragmented Amazonian rainforest. *Mol Ecol* **12**: 753–764.

Dick CW, Hardy OJ, Jones FA, Petit RJ (2008). Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropic Plant Biol* **1**: 20–33.

Dow BD, Ashley MV (1996). Microsatellite analysis of seed dispersal and parentage of sampling in bur oak, *Quercus macrocarpa*. *Mol Ecol* **5**: 615–627.

Dow BD, Ashley MV (1998). High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *J Hered* **89**: 62–70.

Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.

Geng Q, Lian C, Goto S, Tao J, Kimura M, Islam MS *et al.* (2008). Mating system, pollen and propagule dispersal, and spatial genetic structure in a high-density population of the mangrove tree *Kendelia candel*. *Mol Ecol* **17**: 4724–4739.

Ghazoul J (2005). Pollen and seed dispersal among dispersed plants. *Biol Rev* **80**: 413–443.

Goudet J (1995). Fstat (version 2.9.3.2.): a computer program to calculate *F*-statistics. *J Heredity* **86**: 485–486.

Hardesty BD, Hubbell SP, Bermingham E (2006). Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. *Ecol Lett* **9**: 516–525.

Hardy O, 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.

Hardy OJ, González-Martínez S, Colas B, Fréville H, Mignot A, Olivieri I (2004). Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. *Genetics* **168**: 1601–1614.

Hardy OJ, Maggia L, Bandou E, Breyne P, Caron J, Chevallier MH *et al.* (2006). Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Mol Ecol* **15**: 559–571.

Hoebee SE, Arnold U, Düggelin C, Gugerli F, Brodbeck S, Rotach P *et al.* (2007). Mating patterns and contemporary gene flow by pollen in a large continuous and a small isolated population of the scattered forest tree *Sorbus torminalis*. *Heredity* **99**: 47–55.

Hufford KM, Hamrick JL (2003). Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). *Evolution* **57**: 518–526.

Jordano P, García C, Godoy JA, Garcia-Castaño JL (2007). Differential contribution of frugivores to complex seed dispersal patterns. *Proc Natl Acad Sci USA* **104**: 3278–3282.

Jump AS, Peñuelas J (2006). Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proc Natl Acad Sci USA* **103**: 8096–8100.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* **16**: 1099–1106.

Lacerda EBL, Sebbenn AM, Kanashiro M (2008). Long-pollen movement and deviation of random mating in a low-density continuous population of *Hymenaea courbaril* in the Brazilian Amazon. *Biotropica* **40**: 462–470.

Lindgren D, Mullin TJ (1998). Relatedness and status number in seed orchard crops. *Can J For Res* **28**: 276–283.

Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420–1425.

Lourmas M, Kjellberg F, Dessard H, Joly HI, 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.

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.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* **7**: 639–655.

Meagher TR (1986). Analysis of paternity within a natural population of *Chamaelirium luteum*.1. Identification of most-likely male parents. *Am Nat* **128**: 199–215.

Meagher TR, Thompson E (1987). Analysis of parentage for naturally established seedlings of *Chamaelirim luteum* (Liliaceae). *Ecology* **68**: 803–812.

Naito Y, Konuma A, Iwata YS, Seiwa K, Okuda T, Lee SL *et al.* (2005). Selfing and inbreeding depression in seeds and seedlings of *Neobalanocarpus heimii* (Dipterocarpaceae). *J Plant Res* **118**: 423–430.

Naito Y, Kanzaki M, Iwata H, Obayashi K, Lee SL, Muhammad N *et al.* (2008). Density-dependence selfing and its effects on seed performance in a tropical canopy tree species, *Shorea acuminata* (Dipterocarpaceae). *For Ecol Manage* **256**: 375–383.

Nakanishi A, Tomaru N, Yoshimaru H, Manabe T, Yamamoto S (2008). Effects of seeds- and pollen-mediated gene dispersal on genetic structure among *Quercus salicina* samplings. *Heredity* **102**: 182–189.

Oliveira AF, Carvalho D, Rosado SCS (2002). Taxa de cruzamento e sistema reprodutivo de uma população natural de *Copaifera langsdorffii* Desf. na região de Lavras (MG) por meio de isoenzimas. *Rev Bras Bot* **25**: 331–338.

Pedroni F (1993). *Dispersão e recrutamento de Copaifera langsdorffii Desf. (Caesalpiniaceae) na reserva municipal de Santa Genebra*. MSC Dissertation in Biology Science, UNICAMP: Campinas, SP, Brazil.

Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FJ, Hirato MM (2009). The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implictions for conservation. *Biol Conser* **142**: 1141–1153.

Robledo-Arnuncio JJ, Gil L (2005). Patterns of pollen dispersal in a small population of *Pinus sylvestris* L revealed by total-exclusion paternity analysis. *Heredity* **94**: 13–22.

Schuster WSF, Mitton JB (2000). Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity* **84**: 348–361.

Smouse PE, Sork VL (2004). Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecol Manage* **197**: 21–38.

Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analysis in plant populations. *Mol Ecol* **13**: 921–935.

White GM, Boshier DH, Powell W (2002). Increased pollen flow counteracts fragmentation in tropical dry Forest: an example from *Swietenia humilis* Zuccarini. *Proc Natl Acad Sci USA* **99**: 2038–2042.

Young AG, Boyle T, Brown ADH (1996). The population genetic consequences of habitat fragmentation for plants. *Tren Ecol Evol* **11**: 413–418.