

Genetic diversity and spatial structure within a natural stand of a tropical forest tree species, *Carapa procera* (Meliaceae), in French Guiana

AGNES DOLIGEZ^{†‡} & HELENE I. JOLY^{*†}

[†]U.R. Diversité et Amélioration Génétique, CIRAD-Forêt, Baillarguet, BP 5035, F34032 Montpellier Cedex, France
and [‡]Silvolab, BP 316, 97379 Kourou, French Guiana

Isozyme markers were used to characterize levels of diversity, genotypic structure and spatial genetic structure for the low-density tree species, *Carapa procera* (five adults per ha), within 300 ha of a continuous tropical rain forest stand. Both seed and adult stages were investigated, a high level of genetic diversity being found in both. Fixation indices showed excess homozygosity in seeds, and excess heterozygosity in adults, which might be caused by selection in favour of heterozygotes. Autocorrelation analysis of the spatial distribution of genotypes revealed no significant pattern in adults or in seeds before dispersal, and there was a high variability in correlogram shapes among alleles. This suggests that gene flow is extensive in *C. procera*, probably mainly through long-distance pollen dispersal, as seed dispersal is expected to be rather limited in this species (maximum distances of about 50 m). No clear-cut spatial pattern was observed in pollen allele frequencies, which supports the hypothesis of extensive pollen flow. This overall lack of structure is consistent with the data already available on the mating system of this predominantly outcrossing species.

Keywords: excess heterozygosity, gene flow, genetic structure, isozymes, spatial autocorrelation, tropical tree.

Introduction

Several theoretical studies have shown that nonrandom spatial distributions of genotypes can arise when differentiating factors, such as local genetic drift or selection favouring different genotypes in different locations, overcome homogenizing factors, such as gene flow through pollen and seed dispersal, local extinction and recolonization, or selection favouring the same genotypes in all locations (e.g. Sokal & Wartenberg, 1983; Epperson, 1990; Ohsawa *et al.*, 1993). Most of these theoretical studies were based on Wright's neighbourhood concept, in which each individual is assumed to mate at random with its surrounding neighbours. When present, spatial genetic structure at the population level will have important implications both for selection programmes and for conservation of genetic resources: it will lead to biased estimates of outcrossing rates in

natural populations (Ennos & Clegg, 1982); it will affect estimations of genetic variance as measured from open-pollinated progeny; it will influence sampling strategies in *ex situ* conservation programmes; in managed populations, its interaction with patterns of exploitation will affect the amount of genetic diversity. Moreover, most tree species display high levels of genetic variation, with the greatest part of this variation distributed within populations (Hamrick *et al.*, 1981; Moran, 1992). The aforementioned points justify special efforts to characterize and explain spatial genetic organization within tree populations.

Several authors have recently explored small-scale genetic structure in natural stands of temperate forest trees, using spatial autocorrelation analyses based on allozyme data. Some of these studies (Schnabel & Hamrick, 1990; Perry & Knowles, 1991; Schnabel *et al.*, 1991; Sork *et al.*, 1993, cited in Bacilieri *et al.*, 1994; Bacilieri *et al.*, 1994; Merzeau *et al.*, 1994; Berg & Hamrick, 1995; Shapcott, 1995)

*Correspondence. E-mail: joly.h@cirad.fr.

revealed more or less pronounced short-distance spatial clustering of genotypes in a variety of species. This structure, which is typical of isolation by distance, could generally be explained by restricted seed and/or pollen dispersal, sometimes with additional vegetative reproduction. Other work reported nearly random spatial structure, consistent with long-distance seed and pollen dispersal, except for a few loci that could be influenced by selection in some cases (Epperson & Allard, 1989; Knowles, 1991; Geburek & Tripp-Knowles, 1994; Leonardi *et al.*, 1996). Lastly, Boyle *et al.* (1990), Xie & Knowles (1991), Knowles *et al.* (1992) and Geburek (1993) found differences in spatial structure between stands of the same species, some of which could be interpreted as the result of different colonization histories.

Most of these studies of temperate species concern monospecific stands or species present at very high densities in forests comprising only a few species. In tropical forests, most tree species exhibit very low densities (often less than 1/ha; Hubbell & Foster, 1983). Thus, we could expect strong structure to develop in these species as a result of isolation by distance. On the other hand, outcrossing is predominant among these species, and many have the potential for very extensive seed dispersal through large mammals and birds. Hence, the resulting spatial genetic structure is not easy to predict.

Loveless (1992) stressed the importance and possible application of understanding genetic structure in the context of sustainable management of tropical forests. Ledig's (1992) report of local extirpation for some exploited forest tree species further supports the need to understand genetic structure in tropical forests. The present study is among the first to describe local spatial genetic structure for a tropical forest tree species. To our knowledge, only two other such works have been published so far. Boshier *et al.* (1995a) showed that in a high density, almost monospecific, stand of *Cordia alliodora*, a neotropical tree, near-neighbours were more highly related genetically than were more distant trees. Hamrick *et al.* (1993) showed that spatial family structure was present in small- and intermediate-diameter classes for three tropical tree species in Panama. For two of them (*Platypodium elegans* and *Alseis blackiana*), however, this structure disappeared in the largest diameter classes, probably because of high mortality rates; for the third one (*Swartzia simplex*), the structure was still present in adults.

The objective of the present work was to describe genetic diversity for the low-density tree species,

Carapa procera, within a forest stand of about 300 ha (adult density in this stand, 5/ha). This implies: (1) quantifying the amount of diversity within this stand; (2) describing genotypic structure; and (3) describing spatial genetic structure in adults and obtaining preliminary information on the extent of pollen flow. Preliminary knowledge of the amount of genetic diversity is useful in interpreting the lack or presence of structure, because if diversity is low and/or there are many heterozygotes, spatial structuring of genotypes will be more difficult to detect.

The biological and ecological characteristics of this species are as follows: a divided geographical distribution over a wide range (Surinam, French Guiana, state of Amazonas in Brazil, West and Central Africa; Pennington, 1981); a random spatial distribution of trees (CIRAD-Forêt, unpublished data); seed dispersal by large caviomorph rodents, and thus probably up to about 25 m (as for *Vouacoupa americana*, a species with the same dispersal agents; Forget, 1990) or even up to twice that distance as a result of secondary dispersal; much variation in reproductive phenology among individuals and among years (Forget, 1996); monoecious sexual system (Pennington, 1981); pollination probably mediated by small insects (deduced from the size and shape of flowers). Moreover, a study of the *C. procera* mating system in the same stand showed that it is mainly allogamous (multilocus estimate of the outcrossing rate: 0.78, SD 0.05), with outcrosses showing some deviations from random mating, probably resulting from a small amount of biparental inbreeding; finally, correlation of paternity within open-pollinated sibships was low (Doligez & Joly, 1997).

Materials and methods

This study was based on enzyme polymorphisms found in *Carapa procera* seedlings grown from seed. All seeds were collected within the plot limits of the Paracou study site (40 km west of Kourou, French Guiana), which consists of 16 natural rain forest plots dispersed throughout an area of about 300 ha. Within each plot, all trees of diameter ≥ 10 cm were identified, mapped and regularly censused by CIRAD-Forêt. These plots are each 6.25 ha, except one which is 25 ha (Fig. 1). A more detailed description of the site is given in Schmitt & Bariteau (1990) and Doligez & Joly (1997). Seeds were collected within plots from all trees that produced fruit from early April to mid-May 1994; there were thus uncollected *C. procera* trees between plots.

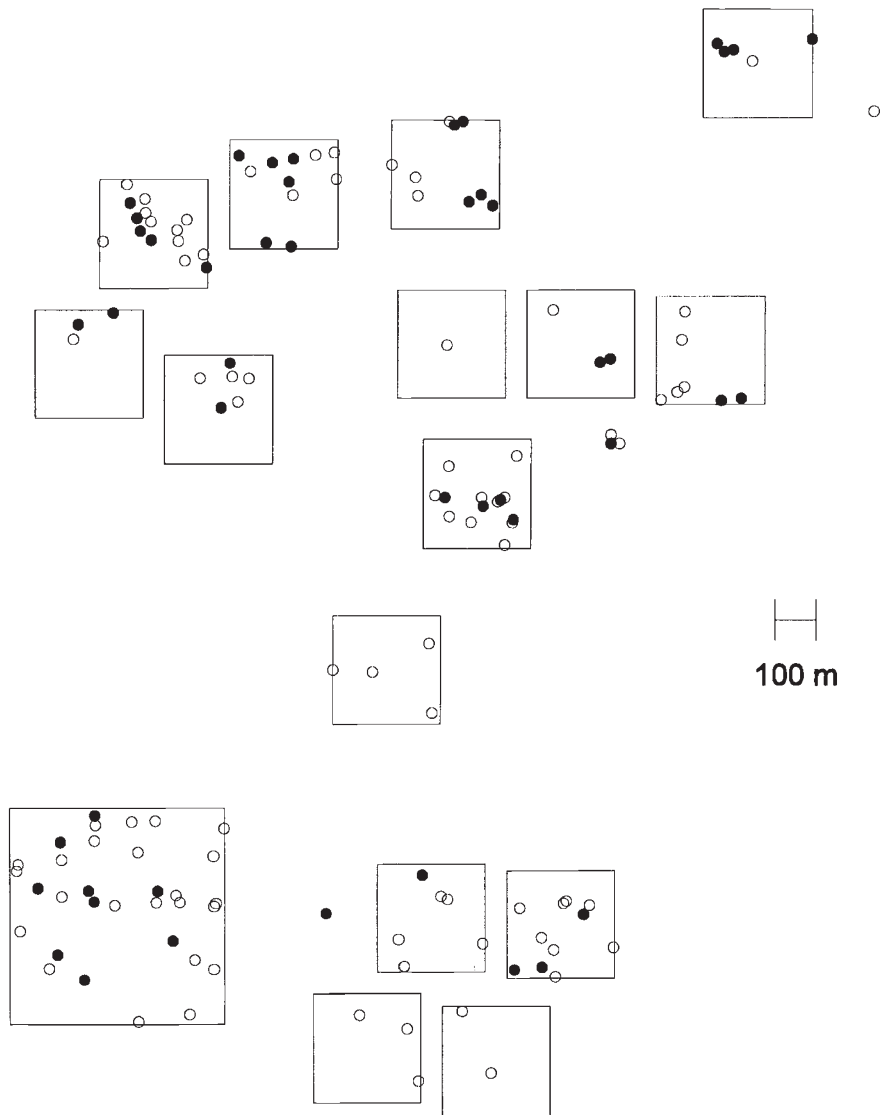


Fig. 1 Map of collected trees: (●) trees for which at least six seedlings were analysed; (○) trees for which only one seedling was analysed. Squares indicate limits of the plots in the Paracou study site.

For the present study, we used data from two sets of seeds. The first set was that used for the analysis of mating system parameters; it consisted of open-pollinated progeny arrays of 6–15 seeds each from 47 mother trees (Doligez & Joly, 1997). For the second set, one seed was collected from each of 90 additional trees (Fig. 1). Over both sets, a total of 573 seeds were germinated and seedlings were analysed for nine enzyme systems (AAP, alanine aminopeptidase, EC 3.4.11.1; AAT, aspartate aminotransferase, EC 2.6.1.1; ENP, endopeptidase, EC 3.4.-.-; GDH, glutamate dehydrogenase, EC 1.4.1.2; IDH, isocitrate dehydrogenase, EC 1.1.1.42; MDH, malate dehydrogenase, EC 1.1.1.37; ME, malic enzyme, EC 1.1.1.40; PGM, phosphoglucomutase, EC 5.4.2.2; PRX, peroxidase, EC 1.11.1.7), using either starch or polyacrylamide gel electrophoresis.

Details of the electrophoresis procedure and genetic interpretations of zymograms are given in Doligez & Joly (1997); 15 loci were resolved. On gels stained for AAP, two banding zones could also be revealed using *L*-leucine- β -naphthylamide as the substrate; the corresponding loci were therefore named *Lap-1* and *Lap-2*. For the 47 trees of the first set, maternal genotypes were inferred from progeny data by the method of Brown & Allard (1970), as implemented by Ritland (1990) in his multilocus outcrossing estimation program, *MLT*. Unfortunately, practical constraints made it impossible for us to characterize adult genotypes directly from leaf material collected from the tree. However, this should have only negligible consequences on further analysis of the maternal population because 10 offspring per progeny array are enough to infer the exact maternal geno-

type with a probability close to one in a predominantly outcrossing species such as *C. procera* (Brown & Allard, 1970).

For diversity analysis, we studied two sets of individuals, which we will refer to as 'populations' for convenience: the adult population (sample size $n = 47$, inferred genotypes) and the seed population (sample size $n = 137$, one seed chosen at random from each of the 47 progeny arrays plus the additional seeds corresponding to 90 different trees).

The following diversity and genetic structure parameters were estimated using the program BIOSYS-1 (Swofford & Selander, 1981): percentage polymorphic loci (P_{100} and P_{95} , with 100 per cent and 95 per cent criteria, respectively), mean number of alleles per locus (A), mean effective number of alleles per locus ($A_e = 1/\sum p_i^2$, where p_i is the allelic frequency at the locus studied), mean observed frequency of heterozygotes (H_o), mean Nei's diversity index (i.e. the heterozygosity expected under random mating, $H_e = 1 - \sum p_i^2$) and fixation indices for polymorphic loci ($F = 1 - H_o/H_e$). The mean fixation index was computed as $1 - \bar{H}_o/\bar{H}_e$, where \bar{H}_o and \bar{H}_e are averages over polymorphic loci. Average values of A , A_e , H_o and H_e were computed for all loci and for polymorphic loci only. A χ^2 -test of homogeneity of genotypic frequencies between seeds and adults was carried out using exact probabilities.

The primary concern in this study was description of the spatial genetic structure within the adult population, but the spatial genetic structure within the population of seeds before dispersal was also studied, because study of the structure at this stage allowed larger sample sizes and thus provided additional information on the spatial structure in adults, even although modified by pollination events and genetic drift. Thus, a spatial autocorrelation analysis was carried out on both the adult and seed populations, using Moran's I statistic, which is defined as (Upton & Fingleton, 1985):

$$I = \frac{n \sum_{ij} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{ij} w_{ij} \sum_i (x_i - \bar{x})^2},$$

where n is the number of individuals in the population, w_{ij} is the geographical distance between individuals i and j , x_i and x_j are the genetic scores of individuals i and j , and \bar{x} is the mean genetic score of the population; the genetic score of a given individual is defined as the number of copies (0, 1 or 2) of the allele studied that this individual possesses. We ascribed to each seed the coordinates of its mother

tree, which means that what we investigated at the seed stage was spatial structure before dispersal, which is principally indicative of pollen flow. Alleles with a frequency of less than 0.1 or more than 0.9 were not used in the analysis.

We computed Moran's coefficient for different distance classes by setting w_{ij} to zero when the distance between individuals i and j did not belong to the distance class under consideration. For the adult population, we used the following 20 distance classes: (1) to (18), 0–1800 m by 100 m; (19), 1800–2000 m; (20), 2000–3000 m. For the seed population, 20 different distance classes were used: (1) to (10), 0–500 m by 50 m; (11) to (15), 500–1000 m by 100 m; (16) to (19), 1000–2000 m by 250 m; (20), 2000–3000 m. Distance classes were chosen so that they contained at least 30 pairs for each allele (except the first class for adults, which contained 28 pairs), as a smaller sample size is not adequate (D. Wartenberg, cited in Waser & Mitchell, 1990). For this reason, the smallest distance class could not be less than 100 m wide for adults with our data. Owing to the low density of *C. procera* adult trees, the use of narrower distance classes would require much greater sample sizes; this would mean the mapping of additional individuals outside the plots and genetic study of their progeny, which was not possible here. However, Epperson's (1995) theoretical simulation study of spatial genetic structure under isolation by distance showed that areas of homozygote patches are much larger than neighbourhood areas. Thus, if spatial distribution of genotypes is significantly nonrandom, structure is expected to be more coarse-grained than seed and pollen dispersal distances. Hence, it should be detectable even when distance classes larger than gene dispersal distances are used.

The plot of autocorrelation coefficients against distance classes is called a correlogram. One correlogram was constructed for each of eight alleles from five different loci. For diallelic loci, a single correlogram was drawn, as both alleles yield exactly the same information. To test for significance at each point of the correlogram, we simulated 1000 new samples by randomization (i.e. by randomly reassigning the n genotypes to the n individual locations) and compared Moran's I -values for these samples to the observed values in our study populations. Sidák's technique was used to assess the overall significance of each correlogram (Oden, 1984): the correlogram was considered significant at the 5 per cent bilateral level if the probability, under the null hypothesis, of obtaining a more extreme correlogram (i.e. a correlogram that shows more departure from randomness

than the one observed) was greater than 0.975 or less than 0.025. This probability is $1 - (1 - m)^k$, where m is the minimum individual probability value in a k -class correlogram.

Finally, to obtain more information on pollination patterns, which should help explain any observed spatial genetic structure, correlograms were constructed for pollen allele frequencies of individual trees. These frequencies were estimated as follows: in each progeny, the number of alleles assumed to be of maternal origin (either through ovules or through selfing pollen) was removed from the total zygotic pool of alleles; 50:50 segregation of alleles was assumed for heterozygous mothers. Allele frequencies were then computed among those alleles that were left. As progeny sizes were limited in our study, this analysis should only be considered as preliminary.

Results

Genetic variation and genotypic structure

Among the 15 loci that could be resolved, 10 had at least two alleles at both seed and adult stages. Diversity indices (Table 1) were not different between stages, except H_o , which was higher for adults. For seeds, all fixation indices were positive or near zero, whereas for adults, all fixation indices were negative, with absolute values as great as 0.31 (Table 2). However, only three loci for seeds and one locus for adults showed significant departures from Hardy–Weinberg genotypic proportions. The fixation index was significantly lower in adults than in seeds (χ^2 -test, $P < 0.05$) for four loci out of 10.

Mean fixation indices were -0.21 and 0.18 for adults and seeds, respectively.

Spatial genetic structure

Correlograms obtained for adult and seed stages are shown in Figs 2 and 3, respectively. Patterns of variation of genetic autocorrelation with distance classes were very different among alleles.

For adults, seven correlograms were constructed instead of eight, as only two different alleles were present at locus *Lap-2*. Seventeen Moran's I statistics out of the 140 computed (i.e. 12.1 per cent) were significant at the $\alpha = 5$ per cent bilateral level. Of these coefficients, seven were for *Pgm* and five for *Aap*. However, when visually inspecting maps of maternal genotypes for these loci, we were not able to detect any clear spatial pattern.

For the seed population, only six autocorrelation coefficients out of 160 (i.e. 3.8 per cent) were significant at $\alpha = 5$ per cent. This is less than the frequency of spurious significance (5 per cent) expected when such a series of tests is carried out, and thus we did not consider it meaningful. This means that the spatial distribution of seed genotypes cannot be distinguished from a random one.

Values of autocorrelation coefficients ranged from -0.40 to 0.31 and from -0.29 to 0.18 for adults and seeds, respectively. Šidák's tests of overall correlogram significance showed that only one correlogram out of 15 (*Aap* for adults) significantly departed from randomness at the 5 per cent level.

With the same 20 distance classes that were used for adults, only 7.5 per cent of the 160 autocorrela-

Table 1 Diversity parameters for both seed and adult stages of *Carapa procera*

Parameter		Seed population	Adult population
Percentage polymorphic loci, P	95% criterion	53%	47%
	100% criterion	67%	67%
Mean number of alleles per locus, A	Polymorphic loci*	2.2	2.1
	All loci	1.8	1.7
Mean effective number of alleles, A_e (SD)	Polymorphic loci*	1.47 (0.49)	1.41 (0.40)
	All loci	1.32 (0.46)	1.28 (0.38)
Observed heterozygosity, H_o (SD)	Polymorphic loci*	0.21 (0.16)	0.30 (0.24)
	All loci	0.14 (0.17)	0.20 (0.24)
Expected heterozygosity, H_e (SD)†	Polymorphic loci*	0.26 (0.21)	0.25 (0.19)
	All loci	0.17 (0.21)	0.17 (0.19)

*100% criterion.

†Nei's unbiased estimate (Swofford & Selander, 1981).

Table 2 Allelic frequencies and Wright's fixation indices *F* for both stages of *Carapa procera*

Locus	Allele	Seeds			Adults			χ^2 -test of homogeneity of genotypic frequencies between seeds and adults (exact probabilities)
		Allelic frequency	<i>F</i>	χ^2	Allelic frequency	<i>F</i>	χ^2	
<i>Aap</i>	<i>a</i>	<u>0.65</u>	0.25	*	<u>0.75</u>	-0.23	NS	*
	<i>b</i>	0.35			0.25			
<i>Aat</i>	<i>a</i>	0.05	-0.05	NS	0.05	-0.06	NS	NS
	<i>b</i>	0.95			0.95			
<i>Enp</i>	<i>a</i>	0.94	0.07	NS	0.92	-0.09	NS	NS
	<i>b</i>	0.06			0.08			
<i>Idh</i>	<i>a</i>	<u>0.27</u>	0.40	**	0.14	-0.16	NS	*
	<i>b</i>	<u>0.73</u>			0.86			
<i>Lap-1</i>	<i>a</i>	0.96	-0.04	NS	0.96	-0.04	NS	NS
	<i>b</i>	0.04			0.04			
<i>Lap-2</i>	<i>b</i>	<u>0.47</u>	0.04	NS	<u>0.49</u>	-0.28	NS	NS
	<i>c</i>	<u>0.52</u>			<u>0.51</u>			
	<i>d</i>	0.01			0.00			
<i>Mdh</i>	<i>a</i>	<u>0.16</u>	0.16	NS	<u>0.23</u>	-0.31	*	**
	<i>b</i>	<u>0.84</u>			<u>0.77</u>			
<i>Me</i>	<i>a</i>	0.99	-0.01	NS	0.99	-0.01	NS	NS
	<i>b</i>	0.01			0.01			
<i>Pgm</i>	<i>a</i>	<u>0.28</u>	0.24	**	<u>0.23</u>	-0.28	NS	*
	<i>b</i>	<u>0.55</u>			<u>0.63</u>			
	<i>c</i>	<u>0.17</u>			<u>0.14</u>			
<i>Prx</i>	<i>a</i>	0.95	-0.05	NS	0.95	-0.06	NS	NS
	<i>b</i>	0.05			0.05			

Underlined frequencies correspond to the alleles that were used in autocorrelation analysis. Significance of chi-square statistics was tested using exact probabilities; NS, $P > 0.05$; * $P < 0.05$, ** $P < 0.01$.

tion coefficients for individual pollen allele frequencies were significantly different from expected under the randomization assumption, and these were not systematically associated with any particular distance (data not shown). On the whole, there is no indication that near-neighbours receive pollen pools more similar than individuals further apart, as would be expected if pollen flow were spatially restricted; rather, individual maternal trees seem to be sired by enough fathers for pollen pools to largely overlap.

Discussion

Levels of enzymatic polymorphism within the population

In a recent review of isozyme variation in tropical woody species, Loveless (1992) found an average percentage polymorphic loci P_{100} of 39.0 (SD 24.3),

and an average gene diversity H_e of 0.13 (SD 0.09) within populations of native woody taxa (including shrubs). The level of intrapopulation enzymatic polymorphism we found for *Carapa procera*, as measured by H_e averaged over all loci, was in this range, but slightly higher than the mean of diversity values for this category of species, whereas P_{100} suggests much more diversity for *C. procera* than for other tropical woody species. Thus, *C. procera* seems to be one of the most variable species studied to date, with values of within-population diversity not far from those reported for temperate gymnosperms ($P_{100} = 67.7$ and H_e averaged over all loci = 0.21; Hamrick *et al.*, 1981).

Hall *et al.* (1994) estimated diversity levels for nine populations of the other species of the genus *Carapa*, *C. guianensis*. Compared with *C. procera*, they found a lower proportion of polymorphic loci,

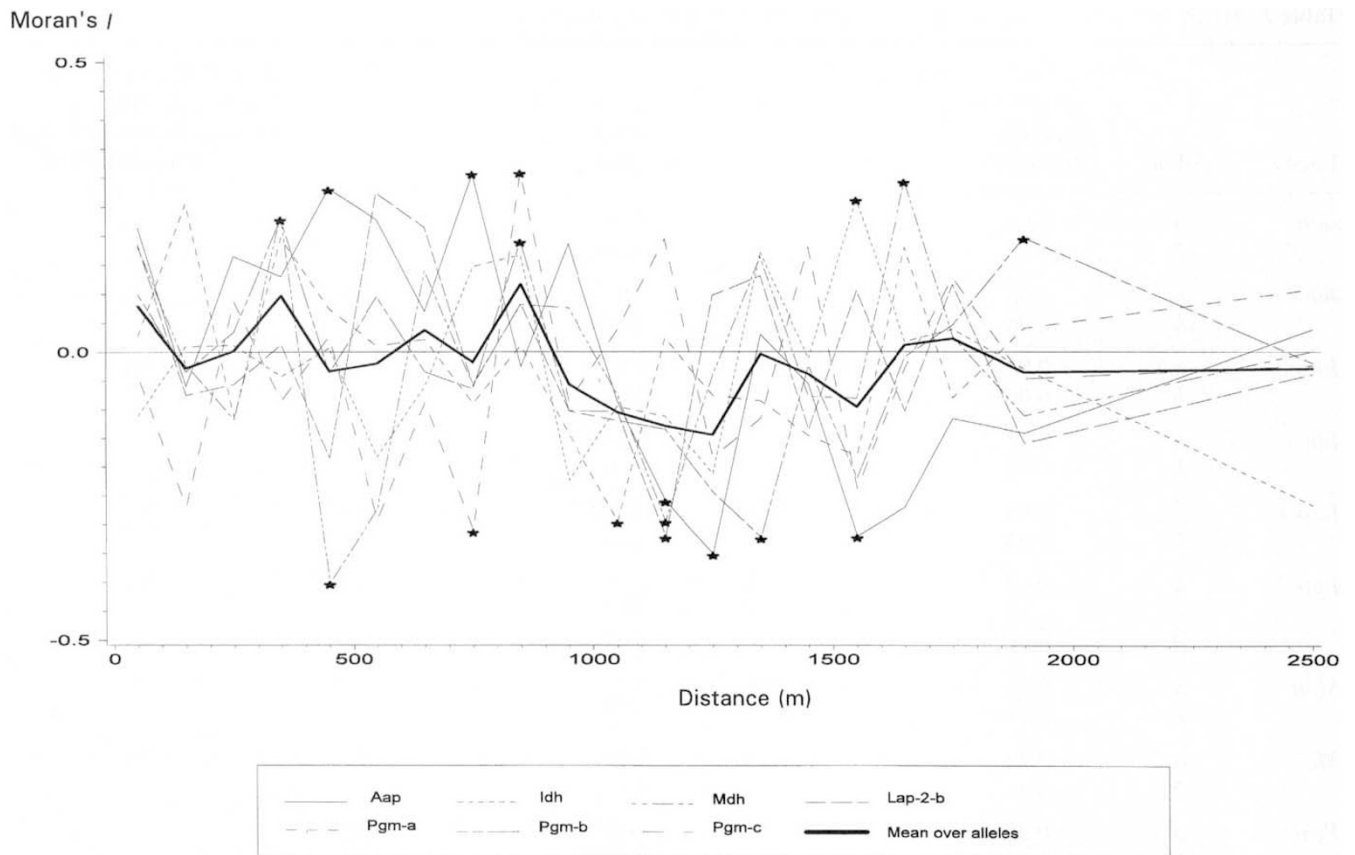


Fig. 2 Correlograms for seven alleles in the adult population, and mean correlogram over alleles. Stars indicate autocorrelation coefficients that are significantly different from their expected value at the 5 per cent bilateral level.

but higher diversity among polymorphic loci, resulting in lower intrapopulation diversity when all loci were included. As far as is known (Pennington, 1981; Hall *et al.*, 1994), the main ecological difference between the two *Carapa* species seems to be tree density; *C. guianensis* may locally have more than 15 mature trees per hectare (Hall *et al.*, 1994), whereas the density of adults of *C. procera* at the Paracou study site was 5/ha. Thus, comparison of intrapopulation diversity between these species does not seem to support Hamrick & Murawski's (1991) assumption that tropical tree species with the highest densities of reproductive individuals should have the highest diversity levels. However, more loci can be considered as showing low substrate specificity in our study than in the study of *C. guianensis*. Furthermore, only one population was studied here and there is no indication of how representative the genetic diversity within this population is for the species as a whole. Thus, the comparison of diversity levels between the two *Carapa* species should be interpreted with caution.

Genotypic structure

Among recent studies of genotypic structure in temperate tree species, one can find many reports of heterozygote excess at the adult stage in natural populations (references in Mitton & Jeffers, 1989; Bush & Smouse, 1992). Evidence for increasing heterozygosity with age (e.g. between seed embryos and mature trees) has also been found in some of these or other studies (references in Bush & Smouse, 1992). In tropical forest tree species, a growing body of recent allozyme-based studies seems to yield similar results (see references in Doligez & Joly, 1997).

To date, the only theoretical models that can be invoked to explain negative F -values in adults, such as those that were found for *C. procera* and the other species mentioned above, are models of overdominance with close linkage between marker and selected loci, and selection of sufficient strength to overcome the homozygote excess resulting from inbreeding at the zygotic stage (Charlesworth, 1991).

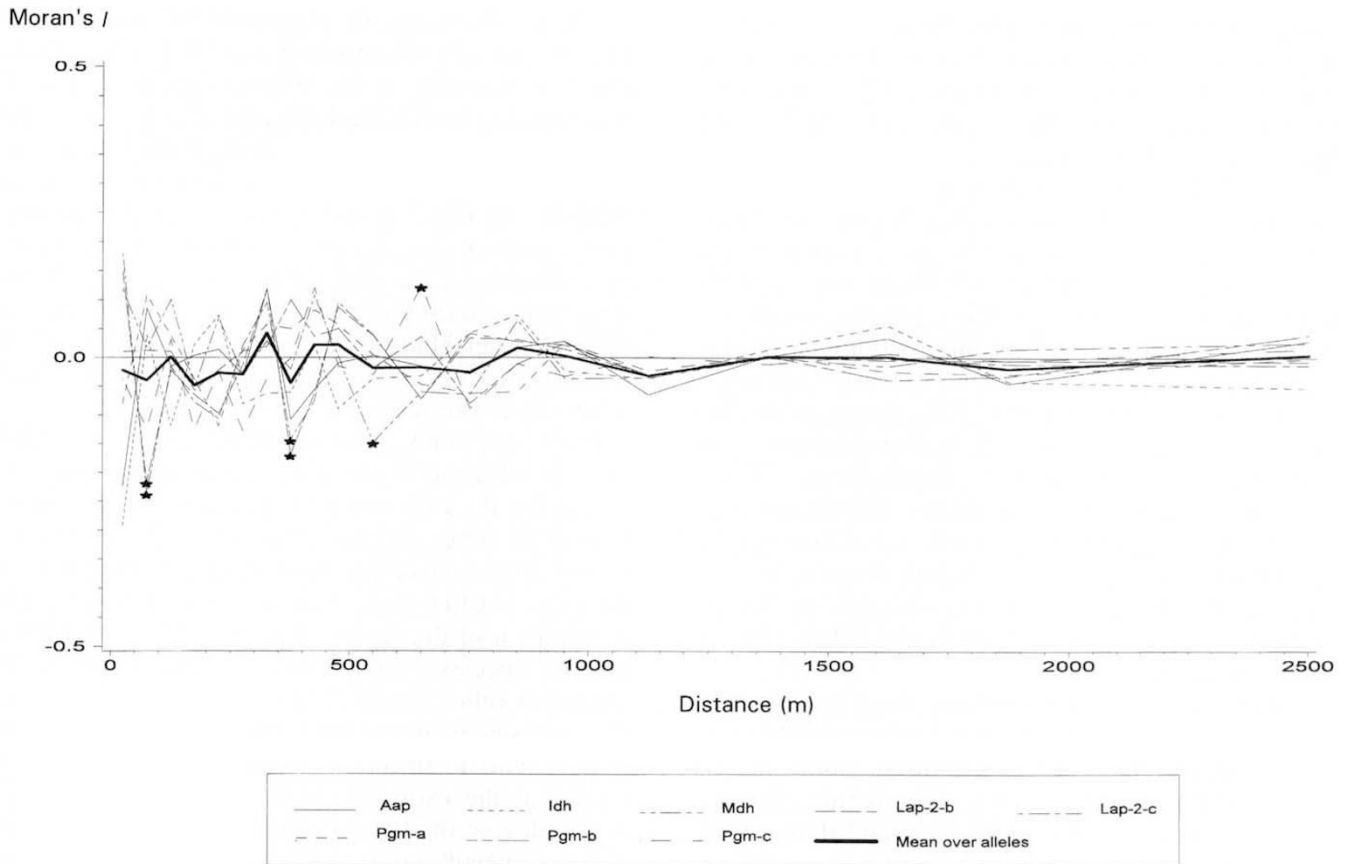


Fig. 3 Correlograms for eight alleles in the seed population, and mean correlogram over alleles. Stars indicate autocorrelation coefficients that are significantly different from their expected value at the 5 per cent bilateral level.

However, further work would be necessary to test directly for this hypothesis.

Spatial genetic structure within the population

No systematic spatial genetic pattern indicative of isolation by distance could be detected for adults of *C. procera* within the Paracou study site. Note that the sampling strategy used here (i.e. sampling within noncontiguous plots) should not affect the detection of spatial structure owing to isolation by distance. Even if it led to missing entire homozygote patches for some alleles, this could not be so for all alleles. In our study, no autocorrelation coefficient was significant in the first three distance classes, and there was no systematic decrease in autocorrelation either. This leads to the conclusion that there is no strong spatial genetic structure caused by isolation by distance in *C. procera* adults within this stand.

Seed dispersal occurs over relatively short distances, compared with intertree distances in this low-density species. The lack of structure observed

in adults thus suggests that gene flow through pollen dispersal was extensive enough in past generations to prevent local differentiation arising as a result of genetic drift. Alternatively, repeated local events of extinction and recolonization because of stochastic variations in local density of *C. procera* could have generated sufficient gene flow to overcome the local differentiating effect of genetic drift (Slatkin, 1985). Demographic data that could help document the latter hypothesis are as yet too scarce. As for the former hypothesis, it is supported by the three following independent observations. First, study of the *C. procera* mating system within the same stand revealed predominant outcrossing (Doligez & Joly, 1997), thus allowing for a great amount of gene flow through pollen (even though this is not a sufficient condition for the occurrence of long-distance pollen dispersal). Secondly, the low correlation of paternity observed in the same study suggested that any given fruiting tree had been sired by about 6–7 different fathers, and thus pollen donors could not be only the very nearest neighbours. Thirdly, the present

study gave preliminary results about autocorrelation of pollen allele frequencies, showing that pollen flow was extensive enough for pollen pools received by neighbours not to be significantly more similar than those received by distant trees.

It was also found that there was probably a small amount of biparental inbreeding among outcrosses (Doligez & Joly, 1997). As the present study yields no evidence that related individuals are spatially closer than unrelated ones, these positive assortative matings might rather reflect the existence of phenological grouping of genotypes. Related individuals might have more overlapping flowering periods than unrelated ones, and thus a higher probability of mating, as in *Cordia alliodora* (Boshier *et al.*, 1995b); however, in contrast to this species, individuals of *C. procera* showing flowering synchrony would not be neighbours. Unfortunately, no data about reproductive phenology of individual trees for this species are available yet, so further work is needed to test for this hypothesis.

Alternatively, it is possible that weak spatial structure was present but could not be detected because of the small adult sample size used. Study of one offspring per tree at the seed stage before dispersal allowed a larger sample to be used and thus a more powerful autocorrelation analysis to be carried out. At this stage, no clear-cut spatial pattern could be detected either. This confirms the overall lack of spatial structure observed in adults, because, assuming that the population is at equilibrium, any spatial structure that would result from isolation by distance in *C. procera* adults should also be detectable in seeds before dispersal, or at any stage after dispersal. Pollen dispersal tends to blur the structure existing in adults but, on the other hand, an important part of the gene sampling responsible for genetic drift takes place during fertilization. Dispersal of *C. procera* seeds by rodents burying only one seed per cache (Forget, 1996) may contribute to the erosion of spatial structure; and after seed dispersal, further gene sampling occurs through mortality of seedlings and juveniles up to the adult stage.

Hamrick *et al.* (1993) found no spatial pattern in adults of *Platypodium elegans* and *Alseis blackiana*, two neotropical tree species, although strong family structure was present in immature individuals. Both species are wind-dispersed and probably pollinated by small bees, but they have very different adult densities. *Carapa procera* has an intermediate density of mature individuals. All three species exhibit sufficient gene flow to prevent development of local genetic structure, whatever their density, and even when their seeds and pollen are dispersed

to similar distances. In two other tropical species, *Swartzia simplex* (Hamrick *et al.*, 1993) and *Cordia alliodora* (Boshier *et al.*, 1995a), significant spatial structure was found in adults, even though they were at high densities within the stands studied and gene flow was not highly restricted. Thus, the few data available to date, including results of the present study, suggest that density is not one of the major determinants of the amount of spatial genetic structure within stands of tropical trees. Variations in effective neighbourhood sizes among different reproduction periods might have much more effect on structure in some cases.

From this study, we can conclude that the high amount of genetic diversity found for *C. procera* within the Paracou study site (compared with populations of other tropical tree species) exhibits no strong spatial structure as a result of isolation by distance, but that there is an excess of heterozygotes in adults; also that pollen flow is probably extensive in this species. Comparison of our results with studies of other tropical forest tree species will allow the assessment of the exact role of density in determining genetic structure. Finally, it is necessary to point out the overall lack of theoretical knowledge of the degree of structure that can be expected with different neighbourhood sizes for species at various densities, in which individuals are irregularly spaced. Theoretical studies giving the minimum sample sizes necessary to detect different levels of spatial structure arising from different degrees of isolation by distance would greatly facilitate interpretation of results such as ours.

Acknowledgements

The authors thank H. Caron and E. Bandou for welcoming A.D. to the Silvolab electrophoresis laboratory of the Institut National de la Recherche Agronomique in Kourou, S. Carcelle for her help with the isozyme analysis, L. Houde for advice on the use of autocorrelation statistics, and M.-H. Chevallier, V. Lebot and two anonymous referees for helpful comments on the manuscript. CIRAD-Forêt gave permission to collect seeds on the Paracou plots and supplied assistance for field work. This work was supported by a PhD fellowship from the Ministère de la Recherche et de l'Espace to A.D. and funds from CIRAD-Forêt.

References

- BACILIERI, R., LABBE, T. AND KREMER, A. 1994. Intraspecific genetic structure in a mixed population of *Quercus*

- petraea* (Matt.) Leibl and *Q. robur* L. *Heredity*, **73**, 130–141.
- BERG, E. E. AND HAMRICK, J. L. 1995. Fine-scale genetic structure of a turkey oak forest. *Evolution*, **49**, 110–120.
- BOSHIER, D. H., CHASE, M. R. AND BAWA, K. S. 1995a. Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighborhood, and population substructure. *Am. J. Bot.*, **82**, 484–490.
- BOSHIER, D. H., CHASE, M. R. AND BAWA, K. S. 1995b. Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 2. Mating system. *Am. J. Bot.*, **82**, 476–483.
- BOYLE, T., LIENGSIRI, C. AND PIEWLUANG, C. 1990. Genetic structure of black spruce on two contrasting sites. *Heredity*, **65**, 393–400.
- BROWN, A. H. D. AND ALLARD, R. W. 1970. Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. *Genetics*, **66**, 133–145.
- BUSH, R. M. AND SMOUSE, P. E. 1992. Evidence for the adaptive significance of allozymes in forest trees. *New Forests*, **6**, 179–196.
- CHARLESWORTH, D. 1991. The apparent selection on neutral marker loci in partially inbreeding populations. *Genet. Res.*, **57**, 159–175.
- DOLIGEZ, A. AND JOLY, H. I. 1997. Mating system of *Carapa procera* (Meliaceae) in the French Guiana tropical forest. *Am. J. Bot.*, **84**, (in press).
- ENNOS, R. A. AND CLEGG, M. T. 1982. Effect of population substructuring on estimates of outcrossing rate in plant populations. *Heredity*, **48**, 283–292.
- EPPELSON, B. K. 1990. Spatial autocorrelation of genotypes under directional selection. *Genetics*, **124**, 757–771.
- EPPELSON, B. K. 1995. Fine-scale spatial structure: correlations for individual genotypes differ from those for local gene frequencies. *Evolution*, **49**, 1022–1026.
- EPPELSON, B. K. AND ALLARD, R. W. 1989. Spatial autocorrelation analysis of the distribution of genotypes within populations of lodgepole pine. *Genetics*, **121**, 369–377.
- FORGET, P.-M. 1990. Seed-dispersal of *Vouacapoua americana* (Caesalpiniaceae) by caviomorph rodents in French Guiana. *J. Trop. Ecol.*, **6**, 459–468.
- FORGET, P.-M. 1996. Removal of seeds of *Carapa procera* (Meliaceae) by rodents and their fate in rainforest in French Guiana. *J. Trop. Ecol.*, **12**, 751–761.
- GEBUREK, T. 1993. Are genes randomly distributed over space in mature populations of sugar maple (*Acer saccharum* Marsh.)? *Ann. Bot.*, **71**, 217–222.
- GEBUREK, T. AND TRIPP-KNOWLES, P. 1994. Genetic architecture in bur oak, *Quercus macrocarpa* (Fagaceae), inferred by means of spatial autocorrelation analysis. *Pl. Syst. Evol.*, **189**, 63–74.
- HALL, P., ORRELL, L. C. AND BAWA, K. S. 1994. Genetic diversity and mating system in a tropical tree, *Carapa guianensis* (Meliaceae). *Am. J. Bot.*, **81**, 1104–1111.
- HAMRICK, J. L. AND MURAWSKI, D. A. 1991. Levels of allozyme diversity in populations of uncommon Neotropical tree species. *J. Trop. Ecol.*, **7**, 395–399.
- HAMRICK, J. L., MITTON, J. B. AND LINHART, Y. B. 1981. Levels of genetic variation in trees: influence of life history characteristics. In: Conkle, M. T. (ed.) *Isozymes of North American Forest Trees and Forest Insects*. General Technical Report PSW-48, pp. 35–41. USDA Forest Service, Pacific Southwest Forest and Range Experimental Station, Berkeley, CA.
- HAMRICK, J. L., MURAWSKI, D. A. AND NASON, J. D. 1993. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio*, **107/108**, 281–297.
- HUBBELL, S. P. AND FOSTER, R. B. 1983. Diversity of canopy trees in a neotropical forest and implications for conservation. In: Sutton, S. L., Whitmore, T. C. and Chadwick, A. C. (eds) *Tropical Rain Forest: Ecology and Management*, pp. 25–41. Blackwell Scientific Publications, Oxford.
- KNOWLES, P. 1991. Spatial genetic structure within two natural stands of black spruce (*Picea mariana* (Mill.) B.S.P.). *Silvae Genet.*, **40**, 13–19.
- KNOWLES, P., PERRY, D. L. AND FOSTER, H. A. 1992. Spatial genetic structure in two tamarack (*Larix laricina* (du Roi) K. Koch) populations with differing establishment histories. *Evolution*, **46**, 572–576.
- LEDIG, F. T. 1992. Human impacts on genetic diversity in forest ecosystems. *Oikos*, **63**, 87–108.
- LEONARDI, S., RADDI, S. AND BORGHETTI, M. 1996. Spatial autocorrelation of allozyme traits in a Norway spruce (*Picea abies*) population. *Can. J. Forest Res.*, **26**, 63–71.
- LOVELESS, M. D. 1992. Isozyme variation in tropical trees: patterns of genetic organization. *New Forests*, **6**, 67–94.
- MERZEAU, D., COMPS, B., THIEBAUT, B., CUGUEN, J. AND LETOUZEY, J. 1994. Genetic structure of natural stands of *Fagus sylvatica* L. (beech). *Heredity*, **72**, 269–277.
- MITTON, J. B. AND JEFFERS, R. M. 1989. The genetic consequences of mass selection for growth rate in Engelmann Spruce. *Silvae Genet.*, **38**, 6–12.
- MORAN, G. F. 1992. Patterns of genetic diversity in Australian tree species. *New Forests*, **6**, 49–66.
- ODEN, N. L. 1984. Assessing the significance of a spatial correlogram. *Geogr. Analyt.*, **16**, 1–16.
- OHSAWA, R., FURUYA, N. AND UKAI, Y. 1993. Effect of spatially restricted pollen flow on spatial genetic structure of an animal-pollinated allogamous plant population. *Heredity*, **71**, 64–73.
- PENNINGTON, T. D. 1981. *Flora Neotropica*. Meliaceae. Hafner Publishing Company, New York.
- PERRY, D. J. AND KNOWLES, P. 1991. Spatial genetic structure within three sugar maple (*Acer saccharum* Marsh.) stands. *Heredity*, **66**, 137–142.
- RITLAND, K. 1990. A series of FORTRAN computer programs for estimating plant mating systems. *J. Hered.*, **81**, 235–237.
- SCHMITT, L. AND BARITEAU, M. 1990. Gestion de l'écosystème forestier guyanais. Etude de la croissance et de la régénération naturelle. Dispositif de Paracou. *Bois et Forêts des Tropiques*, **220**, 3–23.
- SCHNABEL, A. AND HAMRICK, J. L. 1990. Organization of

- genetic diversity within and among populations of *Gleditsia triacanthos* (Leguminosae). *Am. J. Bot.*, **77**, 1060–1069.
- SCHNABEL, A., LAUSHMAN, R. H. AND HAMRICK, J. L. 1991. Comparative genetic structure of two co-occurring tree species, *Maclura pomifera* (Moraceae) and *Gleditsia triacanthos* (Leguminosae). *Heredity*, **67**, 357–364.
- SHAPCOTT, A. 1995. The spatial genetic structure in natural populations of the Australian temperate rainforest tree *Atherosperma moschatum* (Labill.) (Monimiaceae). *Heredity*, **74**, 28–38.
- SLATKIN, M. 1985. Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.
- SOKAL, R. R. AND WARTENBERG, D. E. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics*, **105**, 219–237.
- SWOFFORD, D. L. AND SELANDER, R. B. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *J. Hered.*, **72**, 281–283.
- UPTON, G. J. G. AND FINGLETON, B. 1985. *Spatial Data Analysis by Example*. Vol. I. *Point Pattern and Quantitative Data*. John Wiley & Sons, Chichester.
- WASER, N. M. AND MITCHELL, R. J. 1990. Nectar standing crops in *Delphinium nelsonii* flowers: spatial autocorrelation among plants? *Ecology*, **71**, 116–123.
- XIE, C. Y. AND KNOWLES, P. 1991. Spatial genetic substructure within natural populations of jack pine (*Pinus banksiana*). *Can. J. Bot.*, **69**, 547–551.