

## ORIGINAL ARTICLE

# Edge effects enhance selfing and seed harvesting efforts in the insect-pollinated Neotropical tree *Copaifera langsdorffii* (Fabaceae)

R Tarazi<sup>1</sup>, AM Sebbenn<sup>2</sup>, PY Kageyama<sup>3</sup> and R Vencovsky<sup>1</sup>

Edge effects may affect the mating system of tropical tree species and reduce the genetic diversity and variance effective size of collected seeds at the boundaries of forest fragments because of a reduction in the density of reproductive trees, neighbour size and changes in the behaviour of pollinators. Here, edge effects on the genetic diversity, mating system and pollen pool of the insect-pollinated Neotropical tree *Copaifera langsdorffii* were investigated using eight microsatellite loci. Open-pollinated seeds were collected from 17 seed trees within continuous savannah woodland (SW) and were compared with seeds from 11 seed trees at the edge of the savannah remnant. Seeds collected from the SW had significantly higher heterozygosity levels ( $H_0 = 0.780$ ;  $H_e = 0.831$ ) than seeds from the edge ( $H_0 = 0.702$ ;  $H_e = 0.800$ ). The multilocus outcrossing rate was significantly higher in the SW ( $t_m = 0.859$ ) than in the edge ( $t_m = 0.759$ ). Pollen pool differentiation was significant, however, it did not differ between the SW ( $\Phi'_t = 0.105$ ) and the edge ( $\Phi'_t = 0.135$ ). The variance effective size within the progenies was significantly higher in the SW ( $N_e = 2.65$ ) than at the edge ( $N_e = 2.30$ ). The number of seed trees to retain the reference variance effective size of 500 was 189 at the SW and 217 at the edge. Therefore, it is preferable that seed harvesting for conservation and environmental restoration strategies be conducted in the SW, where genetic diversity and variance effective size within progenies are higher.

*Heredity* (2013) **110**, 578–585; doi:10.1038/hdy.2013.8; published online 13 March 2013

**Keywords:** fragmentation; savannah; mating system; TwoGener analysis; microsatellite loci; effective size

## INTRODUCTION

The physical and biotic alterations at the edge (boundaries) of fragmented habitats, called edge effects, have serious impacts on ecosystem functionality (Laurence *et al.*, 2007). In a fragmented forest, there is a reduction in the density of reproductive trees and neighbour size, as well as changes in the behaviour of pollinators, all of which affect the mating system of reproductive tree species (Fuchs *et al.*, 2003). The knowledge of how and whether edge effects influence the mating system of tree populations and, consequently, the coancestry and variance effective size within open-pollinated progenies is crucial for the development of seed harvesting, conservation, breeding and environmental restoration strategies (Moraes and Sebbenn, 2011). This is especially important in Brazil because of the intense destruction and fragmentation of its rain forests (Ribeiro *et al.*, 2009) and savannahs (Klink and Machado, 2005). In this context, one important and practical question remains unclear: can we collect seeds at the edge of the remnants for conservation, breeding and environmental restoration purposes or must we collect seeds at the centre of the remnants to maximise the variance effective size of the harvested seed lot? As the genetic diversity, relatedness and variance effective size within open-pollinated progenies is determined by mating patterns (Murawski and Hamrick, 1991), our hypothesis is that habitat boundaries may not be the most suitable place to conduct seed harvesting, although these places provide easy access and low operational costs.

In hermaphroditic tree species, the level of outcrossing depends on the following factors: the reproductive population density or the flowering tree density; the foraging patterns of pollinators within and among conspecific individuals; mechanical, physiological and genetic barriers that prevent selfing; and the selective abortion of selfed fruits and seeds (Murawski and Hamrick, 1991; Goodwillie *et al.*, 2005). Ecological variables, such as spatial isolation (Fuchs *et al.*, 2003; Sork and Smouse, 2006), flowering phenology, plant density (Murawski and Hamrick, 1991) and changes in pollinator activity (Hirao *et al.*, 2006), have been identified as factors that may affect the mating patterns in tree species. Moreover, temporal isolation (differences in flowering phenology) and spatial isolation (low density of flowering trees) have been proposed as causes of high rates of self-pollination in some species of trees (Murawski and Hamrick, 1991; Moraes and Sebbenn, 2011). A reduction in the density of reproductive trees is expected to increase the degree of spatial isolation, which increases selfing (Fuchs *et al.*, 2003; Robledo-Arnuncio *et al.*, 2004; Aguilar *et al.*, 2008; Moraes and Sebbenn, 2011). Despite the increase in selfing that occurs with a decrease in population density, the reported results for correlated mating have conflicted. Some studies have indicated that an increase in correlated mating occurs with a decrease in population density (Fuchs *et al.*, 2003; Moraes and Sebbenn, 2011), while the opposite was observed in another study (Robledo-Arnuncio *et al.*, 2004). Selfing increases inbreeding in the offspring generation,

<sup>1</sup>Departamento de Genética, Universidade de São Paulo, CP 9, Piracicaba, SP, Brazil; <sup>2</sup>Instituto Florestal de São Paulo, CP 339, Piracicaba, SP, Brazil and <sup>3</sup>Departamento de Ciências Florestais, Universidade de São Paulo, CP 9, Piracicaba, SP, Brazil  
Correspondence: Dr R Tarazi, Departamento de Genética, Universidade de São Paulo, Avenida Pádua Dias, 11 CP 9, Piracicaba, SP 13418-900, Brazil.  
E-mail: betotarazi@yahoo.com.br

Received 2 March 2012; revised 5 November 2012; accepted 8 January 2013; published online 13 March 2013

which produces inbreeding depression. Selfing, mating among relatives and correlated mating increase the frequency of identical-by-descent alleles and decrease the variance effective size within open-pollinated progenies (Murawski and Hamrick, 1991).

Genetic markers, such as microsatellite loci, have been used successfully to quantify the mating system in tree species (Ashley, 2010). In maternal open-pollinated seeds, the mating system can be measured as the proportion of seeds that result from outcrossing or selfing events, and the outcrossing fraction can be split into the proportions of seeds that originated from the same pollen donor (full-sibs) and those from different pollen donors (half-sibs; Goodwillie *et al.*, 2005). These mating system events can be quantified using the mixed mating system (Ritland and Jain, 1981) and correlated mating models (Ritland, 1989) that have been implemented in the classical MLTR programme (Ritland, 2002). Complementarily, pollen dispersal studies within continuous and dense populations of plants that do not require sampling potential pollen donors can be investigated using *TWOGENER* (Austerlitz and Smouse 2001; Smouse *et al.*, 2001) and *KINDIST* (Robledo-Arnuncio *et al.*, 2007) analyses. These two methods require mother-offspring diploid codominant genotypic data. The *TWOGENER* analysis estimates the pollen pool structure using the multi-locus male gamete genotypes of single seeds and subjects them to an analysis of molecular variance, yielding the intra-class correlation coefficient of pollen genotypes ( $\Phi_{it}$ ), which is a measure of the differentiation within the pollen pool (Austerlitz and Smouse, 2001; Smouse *et al.*, 2001). This method also allows conclusions to be made about the pollen dispersal distance ( $\delta$ ), the effective pollination neighbourhood area (Smouse *et al.*, 2001), the effective number of pollen donors ( $N_{ep}$ ), the coancestry coefficient ( $\Theta$ ), the effective size within progenies and the number of seed trees required to collect seeds (Bittencourt and Sebbenn, 2008). As an upgrade to the *TWOGENER* analysis, *KINDIST* is independent of the effective density of the pollen donors and provides a description of the relationship between among-sibship normalised correlated paternity estimates ( $\Psi$ ), distance ( $z$ ) and pollen dispersal distance ( $\delta$ ) (Robledo-Arnuncio *et al.*, 2007).

Here, we investigated the mating system and pollen pool differentiation among seed trees in the insect-pollinated Neotropical tree *C. langsdorffii* within a continuous savannah woodland (SW) and its edge using a mixed mating system (Ritland and Jain, 1981) and correlated mating models (Ritland, 1989) and *TWOGENER* (Smouse *et al.*, 2001) and *KINDIST* (Robledo-Arnuncio *et al.*, 2007) analyses. As the mating system of a tree species may be affected by the population density of reproductive plants (Murawski and Hamrick, 1991; Naito *et al.*, 2008; Moraes and Sebbenn, 2011), the lower density in the edge than in the centre of a forest fragment may result in increased rates of selfing, correlated matings, pollen pool differentiation among seed trees and inbreeding and decrease in genetic diversity and variance effective size in open-pollinated progenies. Thus, we specifically addressed three questions about edge effects: (i) is the average number of alleles, heterozygosity and outcrossing rate higher in progenies of seed trees located within continuous SW than its edge? (ii) Is the pollen pool differentiation lower in the continuous SW than in the edge of the remnant? (iii) Is the variance effective size within progenies higher in seed trees located within the continuous SW or at its edge?

## MATERIALS AND METHODS

### *Copaifera langsdorffii* Desf. (Fabaceae)

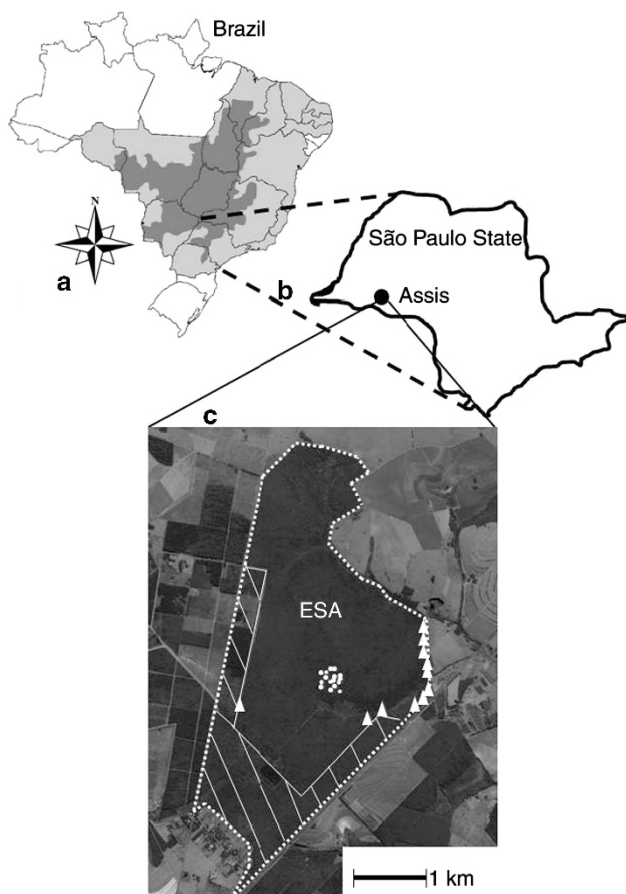
*Copaifera langsdorffii* is a hermaphroditic tree with wide distribution in the Brazilian savannah and Atlantic rain forest (Carvalho, 2003). This species produces seeds mainly by outcrossing, but selfing has been detected even in the

seedlings of natural regeneration (Sebbenn *et al.*, 2011). Manual crosses revealed that geitonogamy was the only observed selfing phenomenon, indicating that this species is highly dependent on its pollinators (Freitas and Oliveira, 2002). Its bisexual flowers exhibiting herkogamy are pollinated mainly by small native bees, such as *Scaptotrigona*, *Trigona* and the naturalised bee *Apis mellifera*. A high quantity of flowers is produced in several weeks, and the blossoming of the tree varies from annual to supra-annual (Freitas and Oliveira, 2002). In natural populations, the reproductive stage begins between 20 and 30 years because of the tree's medium to slow growth rate; however, some individuals in experimental plantations with a 6 cm diameter at breast height may flower (Carvalho, 2003). *C. langsdorffii* seeds are mainly dispersed by gravity and birds. In addition to its ecological role, *C. langsdorffii* also has economic importance because of its high quality wood and oil with pharmaceutical properties that can be extracted directly from its trunk. However, the species is listed as being at risk of extinction in some Brazilian states as a result of the intense destruction and fragmentation of the Atlantic rain forest and savannah environments (reviewed in Carvalho, 2003).

### Study site and sampling

The Brazilian savannah, known as 'cerrado', covers an area of 2 000 000 km<sup>2</sup> from 3°N to 24°S and represents approximately 23% of the country's land surface (Klink and Machado, 2005). The cerrado consists of a mosaic of physiognomies that vary along a gradient from open grasslands to dense woodland formations controlled by interactions among climatic, edaphic and disturbance factors, such as fire and herbivory. The typical savannah 'cerrado *sensu stricto*' is characterised by open grasslands, shrubs and up to 50% of discontinuous tree cover. On the other hand, SW 'cerradão', has a forest physiognomy, where the woody vegetation forms a continuous stratum, typically with >90% coverage, and the average height of the trees varies between 8 and 15 m (Pinheiro and Durigan, 2009). The Brazilian southern savannahs originally covered an area of 50 000 km<sup>2</sup> (2.5% of the Brazilian savannah) in the State of São Paulo and parts of Minas Gerais and Paraná. The southern savannah is distinguished from the core savannah domain because it occurs under lower temperatures and has a shorter dry season (Durigan, 2006). Consequently, the southern savannahs form a distinct phytogeographical region, which is divided into two areas: one covered by SW, which represents 70% of the surveyed savannah sites in the State of São Paulo, and another covered by typical savannah. Remarkably, the third most frequently recorded woody species in the surveyed southern savannahs sites in the State of São Paulo is *C. langsdorffii* (Durigan, 2006).

This study was conducted at the Ecological Station of Assis (ESA), which is a reserve of the State of São Paulo, Brazil (22°33'20'' to 22°37'41''S and 50°24'4'' to 50°21'27''W). The reserve covers 1760 ha (Figure 1): 1312.18 ha is occupied by native vegetation, and 1115.68 ha is composed of SW. The ESA, one of the largest and among the few protected southern savannah remnants in São Paulo, is isolated in an agricultural landscape. The ESA boundaries are composed of open SW juxtaposed with shrubby savannah, open savannah and an agricultural landscape (Pinheiro and Durigan, 2009) that contains very few, far-between (>1 km) and small (<10 ha) isolated savannah fragments. The absolute density (trees ha<sup>-1</sup>) of *C. langsdorffii* individuals with a diameter at breast height >5 cm throughout the ESA in the SW was 296, making *C. langsdorffii* the most dominant woody species (Pinheiro, 2008). Within the ESA, we focused our study on the edge of the remnant and a pre-existing 10.24 ha permanent plot of SW that is at least 500 m from the edge of the remnant. In the plot, a total of 4009 *C. langsdorffii* individuals with a diameter at breast height that was >5 cm were found, measured and mapped by the BIOTA-FAPESP Permanent Plot Consortium in 2006; *C. langsdorffii* was the plot's dominant species. To address our questions, the following sampling scheme in the ESA was conceived (Figure 1). We sampled seeds from all of the *C. langsdorffii* trees with a diameter at breast height ≥25 cm that flowered and fructified within the same period of 2007 that were in the plot ( $N=17$ ) or that were up to 10 m from the edge of the ESA ( $N=11$ ). The reproductive tree density was 1.66 trees ha<sup>-1</sup> in the plot and 0.66 trees ha<sup>-1</sup> at the edge. The open-pollinated offspring (seeds) of the 28 seed trees were directly collected from the crown of the trees. Twenty random offspring per seed tree (here referred as a progeny) were used in our analysis, totalling 560 offspring or 28 progenies.



**Figure 1** (a) Geographical distribution of *C. langsdorffii* (light grey) and the Brazilian savannah (dark grey) in Brazil. (b) Location of the city of Assis in São Paulo State. (c) Satellite image of the ESA and its surrounding agricultural landscape. Boundaries of the ESA are represented by continuous white dots, and areas inside the ESA that are not characterised as SW are represented by diagonal white lines. Distribution of the 17 seed trees inside the 10.24 ha plot in the continuous SW (○) and the 11 seed trees throughout the edge of the ESA (Δ).

### Microsatellite genotyping

DNA was extracted from the leaflets of the offspring (seeds germinated in a greenhouse) and the leaves of the seed trees using the CTAB method (Doyle and Doyle, 1990). After extraction, DNA was quantified on 1% agarose gels with a lambda DNA standard and diluted to a concentration of  $2.5 \text{ ng } \mu\text{l}^{-1}$ . The eight microsatellite loci used in this study and the PCR conditions were reported by Ciampi *et al.* (2000). The amplifications were performed using an MJ Research PT-100 thermal cycler (MJ Research, Matertown, MA, USA) adjusted to the following conditions: 5 min at  $94^\circ\text{C}$ ; 30 cycles of 1 min at  $95^\circ\text{C}$ , 1 min at the specific temperature for each primer, and 1 min at  $72^\circ\text{C}$ ; and a final elongation step of 7 min at  $72^\circ\text{C}$ . The amplification products were separated on 5% (w/v) polyacrylamide gels and stained with silver nitrate. As microsatellite genotyping errors can lead to biased results, we double-checked all gels. We used MICRO-CHECKER 2.2.3 (Oosterhout *et al.*, 2004) to test possible stuttering, large allelic dropout and null alleles in the adult individuals, using a 95% confidence interval based on 1000 Monte-Carlo permutations. No evidence for scoring errors because of stuttering or for large allelic dropout was found. Null alleles may be present at loci CL02, CL32 and CL34, as is suggested by the general excess of homozygotes for most allele size classes ( $P < 0.05$ ). However, loci CL02 and CL34 had a mistyping of zero, and locus CL32 had only one mistyping, indicating that the homozygote excess may be due to selfing in the population and not necessarily to null alleles. Another method that may demonstrate that all loci have no null alleles is the check for

genotyping errors and null alleles that we performed using MLTR 3.4 (Ritland, 2002) to detect seed tree–offspring mismatches. We obtained zero seed tree–offspring mismatches and an absence of null alleles in the offspring for all loci. In a previous study, Tarazi *et al.* (2010) demonstrated that the SSR marker of *C. langsdorffii* exhibited Mendelian inheritance and an absence of linkage in the studied offspring population.

### Hypothesis tests

As the seed trees sample sizes were different between the sites, to compare the genetic diversity, mating system and *TWOGENER* parameters, and to verify their significance, we used a 95% bootstrap percentile confidence interval based on 1000 replications ( $CI_{95\%}$ ). The lower and upper confidence intervals were determined by taking the 2.5 and 97.5 percentiles after ranking the bootstrap values. In all comparisons, we used progenies as replication units. In addition, we tested if sample size influenced our estimates by bootstrapping 1000 times 11 progenies (number of progenies of the edge) from the SW and re-estimated all of the parameters. We also tested if the isolated tree in the western edge of ESA could affect the parameters cited above. This analysis was conducted by comparing the overall 1000 bootstrapped estimates of 11 progenies of the edge with estimates without the isolated tree.

### Genetic diversity and fixation index analyses

The genetic diversity of the offspring inside the SW and at the edge was estimated by the number of alleles per locus ( $k$ ), allelic richness ( $R$ ), using rarefaction method (El Moussadik and Petit, 1996), total number ( $k_{\text{total}}$ ) and average number of alleles per locus as well as the observed ( $H_o$ ) and expected heterozygosity at Hardy–Weinberg equilibrium ( $H_e$ ). The level of inbreeding was estimated using the fixation index ( $F$ ). All parameters were calculated using the FSTAT programme (Goudet, 1995).

### Mating system analyses

The mating system was analysed for site (SW and edge) and for progeny level based on the mixed mating model and a correlated mating model using the MLTR 3.4 programme (Ritland, 2002). Analyses at the population level were carried out using the Newton–Raphson numerical method and at the progeny level using the expectation-maximisation numerical method. The estimated parameters were the multilocus outcrossing rate ( $t_m$ ), the single-locus outcrossing rate ( $t_s$ ), the outcrossing rate among relative trees ( $t_m - t_s$ ), the multilocus paternity correlation ( $r_{pm}$ ), the single-locus paternity correlation ( $r_{ps}$ ) and the average effective number of pollen donors ( $N_{ep} = 1/r_{pm}$ ). We obtained a 95% bootstrap percentile confidence interval based on 1000 replications. The bootstrap analysis at the population level was carried out using progenies as replication units and at the progeny level using offspring within progenies as replication units.

### TwoGener and KinDist analyses

To test the hypothesis that there is higher genetic differentiation in the pollen pool of seed trees from the edge than in the SW, we used *TWOGENER* (Smouse *et al.*, 2001) and *KINDIST* (Robledo-Arnuncio *et al.*, 2007) analyses. For these analyses, we used the progeny arrays of the 17 seed trees in the SW, which totalled 340 seeds, and the progeny of the 11 seed trees of the edge of the ESA, which totalled 220 seeds, as well as all of the genotypes and the spatial position of the sampled seed trees. The global differentiation of the pollen pool ( $\Phi_{ft}$ ) among seed trees, the relationship between among-sibship normalised correlated paternity estimates ( $\Psi$ ), distance ( $z$ ) and pollen dispersal distance ( $\delta$ ) was estimated using the Poldisp 1.0 programme (Robledo-Arnuncio *et al.*, 2007). The  $\Phi_{ft}$  value can be affected by inbreeding and selfing (Austerlitz and Smouse, 2001; Burczyk and Koralewski, 2005). However, our estimation of inbreeding in the seed trees of both sites was zero ( $F_{CSW} = 0.002$  and  $F_{\text{edge}} = 0.000$ ), and consequently, the estimate of  $\Phi_{ft}$  was not influenced by the inbreeding of the parental generation:  $\Phi_{ft} = \Phi_{ft}/(1 - F_p)$  (Austerlitz and Smouse, 2001). Therefore,  $\Phi_{ft}$  was only adjusted for selfing ( $s$ ) as follows:  $\Phi_{ft}'' = (2\Phi_{ft}' - s^2)/2(1 - s)^2$  (Burczyk and Koralewski, 2005). From  $\Phi_{ft}''$ , we also estimated the average effective number of pollen donors ( $N_{ep} = 1/2\Phi_{ft}''$ , Smouse *et al.*, 2001). If a significant negative relationship between among-sibship normalised correlated paternity estimates ( $\Psi$ ) and distance ( $z$ ) were to

be detected in the sites, pairwise  $\Psi(z)$  estimates between all seed trees in the population should be calculated to derive the pollen dispersal distance ( $\delta$ ) and the estimated effective density of reproductive trees ( $d_e$ ) in the normal and exponential dispersal function (Robledo-Arnuncio et al., 2007).

**Estimation of the variance effective size and the number of seed trees for seed harvesting**

The minimum number of seed trees ( $m$ ) required for harvesting seeds for conservation purposes was calculated:  $m = N_{e(\text{reference})}/N_e$ , where  $N_{e(\text{reference})}$  is the reference variance effective size of the seeds to be harvested (500) and  $N_e$  is the average variance effective size within progenies (Bittencourt and Sebbenn, 2008). The  $N_e$  was estimated as follows:  $N_e = 0.5 / \{\Theta[(n-1)/n] + (1 + F_e)/2n\}$  (Cockerham, 1969), where  $F_e$  is the inbreeding coefficient of the seeds,  $\Theta$  is the average coancestry within progenies and  $n$  is the sample size (here assumed to be 20). The average coancestry coefficient within progenies was estimated from the mating system parameter by the equation  $\Theta = 0.125(1 + F_p)[4s + (t_m^2 + st_m r_{ps})(1 + r_{pm})]$ . This estimator corresponds to half of the relatedness coefficient within progenies ( $\Theta = r/2$ ) as derived by Ritland, 1989. The other parameters were previously described. The average coancestry coefficient was also estimated through the global differentiation of the pollen pool ( $\Phi_{it}'$ ):  $\Theta = 0.125(1 + F_p)[4s + t_m^2(1 + 2\Phi_{it}')]$  (Bittencourt and Sebbenn, 2008).

**Analyses of seed germination**

To evaluate whether there were differences in germination between the seeds collected from the plot in the SW and those from the edge, we performed a germination experiment. Seeds from seed trees were germinated separately in germination boxes using a completely randomised design with three repetitions and 25 seeds per germination box. To test whether there were significant differences in the absolute germination values between the plot in SW and the edge, we used the F test of variance analysis within the SAS programme (SAS Institute Inc., 1999). To test whether there was a significant association between the average progeny germination (absolute values) and selfing, the fixation index, the observed heterozygosity or the expected heterozygosity, we estimated the Spearman rank coefficient of correlation also using the SAS programme (SAS Institute Inc., 1999).

**RESULTS**

**Genetic diversity and inbreeding**

The total ( $k_{\text{total}}$ ) number of alleles per locus and the observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosities were significantly higher in seeds collected from the SW than in those from the edge (Table 1).

Even after sampling 11 progenies from SW, the observed ( $H_o = 0.780$ ;  $CI_{95\%} = 0.775-0.802$ ) and expected ( $H_e = 0.843$ ;  $CI_{95\%} = 0.823-0.860$ ) heterozygosities were significantly higher in seeds collected from the SW than in the edge.

The removal of the isolated tree in the western edge of ESA from the analysis did not affected the observed ( $H_o = 0.713$ ;  $CI_{95\%} = 0.689-0.736$ ) and expected ( $H_e = 0.796$ ;  $CI_{95\%} = 0.756-0.822$ ) heterozygosities. In the seeds collected from the SW, we found a total of 136 different alleles in the eight loci, with an average of 17.0 alleles among loci. In seeds from the edge, we found 131 alleles, with an average among loci of 16.4 alleles. We found 14 private alleles (frequency ranging from 0.002 to 0.042) in the seeds collected from the SW and 10 (frequency ranging from 0.002 to 0.065) from the edge. Although the fixation index ( $F$ ) was significantly positive in the SW and at the edge (Table 1), which suggested inbreeding, it did not differ between these two sites. No difference was found even after sampling 11 progenies from the SW ( $F = 0.063$ ;  $CI_{95\%} = 0.021-0.099$ ) and removing the isolated tree of the edge from the analysis ( $F = 0.104$ ;  $CI_{95\%} = 0.026-0.166$ ).

**Mating system**

The estimated outcrossing rate was significantly lower than unity (1.0) at both sites (Table 2a) and was significantly higher in the SW ( $t_m = 0.859$ ) than its edge ( $t_m = 0.759$ ). Furthermore, in the multi-locus outcrossing rate after sampling 11 progenies from SW ( $t_m = 0.861$ ;  $CI_{95\%} = 0.810-0.922$ ) and removing the isolated tree of the edge from the analysis ( $t_m = 0.749$ ;  $CI_{95\%} = 0.645-0.805$ ), no significant differences were found from the previous estimates based on 17 and 11 progenies, respectively. The outcrossing rate among related trees ( $t_m - t_e$ ) was significantly different from zero in the SW and at the edge but was similar between the sites (Table 2a). The multilocus ( $r_{pm}$ ) and single-locus ( $r_{ps}$ ) paternity correlation was also significantly different from zero, but did not differ between the sites (Table 2a). The effective number of pollen donors ( $N_{ep}$ ) was similar at both sites, indicating low reproductive dominance (Table 2a).

The progeny multilocus outcrossing rate ( $t_m$ ) was variable among seed trees, ranging from 0.47 to 0.99 in the SW and from 0.40 to 1.00 in the edge (Table 3). The multilocus outcrossing rate in the edge was  $<0.75$  in 7 of the 11 open-pollinated progenies (64%),

**Table 1 Genetic diversity parameters and fixation index in the *C. langsdorffii* offspring from the continuous SW and the edge of the ESA**

| Locus              | SW offspring (n = 340) |      |       |       |        | Edge offspring (n = 220) |      |       |       |        |
|--------------------|------------------------|------|-------|-------|--------|--------------------------|------|-------|-------|--------|
|                    | k                      | R    | $H_o$ | $H_e$ | F      | k                        | R    | $H_o$ | $H_e$ | F      |
| CL01               | 12                     | 12.0 | 0.770 | 0.825 | 0.067  | 12                       | 11.8 | 0.735 | 0.839 | 0.124  |
| CL02               | 18                     | 17.1 | 0.667 | 0.877 | 0.240  | 16                       | 15.7 | 0.564 | 0.707 | 0.202  |
| CL06               | 17                     | 16.8 | 0.920 | 0.890 | -0.035 | 17                       | 16.7 | 0.832 | 0.878 | 0.054  |
| CL20               | 22                     | 21.9 | 0.838 | 0.905 | 0.075  | 22                       | 21.7 | 0.843 | 0.855 | 0.014  |
| CL27               | 13                     | 12.5 | 0.847 | 0.889 | 0.047  | 12                       | 12.0 | 0.704 | 0.850 | 0.171  |
| CL32               | 21                     | 18.7 | 0.808 | 0.820 | 0.014  | 16                       | 15.9 | 0.777 | 0.801 | 0.031  |
| CL34               | 17                     | 15.8 | 0.599 | 0.795 | 0.246  | 15                       | 15.0 | 0.580 | 0.745 | 0.222  |
| CL39               | 16                     | 15.8 | 0.871 | 0.875 | 0.004  | 21                       | 20.9 | 0.725 | 0.885 | 0.181  |
| Mean               | 16.2                   | 16.1 | 0.780 | 0.849 | 0.081* | 15.0                     | 14.8 | 0.702 | 0.800 | 0.122* |
| $CI_{95\%}$ lower  | 15.5                   | 15.4 | 0.776 | 0.831 | 0.040  | 13.8                     | 13.6 | 0.688 | 0.761 | 0.024  |
| $CI_{95\%}$ upper  | 16.8                   | 16.7 | 0.798 | 0.861 | 0.098  | 16.0                     | 15.8 | 0.743 | 0.820 | 0.161  |
| $k_{\text{total}}$ | 136                    | —    | —     | —     | —      | 131                      | —    | —     | —     | —      |

Abbreviations: ESA, Ecological Station of Assis; SW, savannah woodland.

\* $P < 0.05$ ; mean was obtained from 1000 bootstrap values among progenies;  $CI_{95\%}$  lower and  $CI_{95\%}$  upper are the 95% bootstrap confidence interval after sampling progenies 1000 times;  $k$  is the number of alleles per locus;  $R$  is the allelic richness based on the minimum sample size of 220 individuals;  $k_{\text{total}}$  is the total number of alleles;  $H_o$  is the observed heterozygosity;  $H_e$  is the expected heterozygosity at Hardy-Weinberg equilibrium;  $F$  is the fixation index.

**Table 2** Mating system parameters estimated from the mixed and correlated mating model (a) and *TWOGENER* analysis (b) from continuous SW and the edge of the ESA in *C. langsdorffii*

| Parameters   | SW                  | Edge                 |
|--|---------------------|----------------------|
| <i>(a) Mixed and correlated mating models</i>                  |                     |                      |
| Multilocus outcrossing rate: $t_m$                             | 0.859 (0.829–0.888) | 0.759 (0.718–0.805)  |
| Single-locus outcrossing rate: $t_s$                           | 0.778 (0.752–0.805) | 0.675 (0.636–0.720)  |
| Outcrossing rate among relatives: $t_m - t_s$                  | 0.081 (0.067–0.094) | 0.083 (0.065–0.102)  |
| Multilocus paternity correlation: $r_{pm}$                     | 0.059 (0.032–0.076) | 0.056 (0.024–0.080)  |
| Single-locus paternity correlation: $r_{ps}$                   | 0.116 (0.090–0.129) | 0.094 (0.065–0.101)  |
| Effective number of pollen donors: $N_{ep}$                    | 17 (13–31)          | 18 (13–42)           |
| Coancestry within progenies: $\Theta$                          | 0.170 (0.162–0.179) | 0.199 (0.185–0.212)  |
| Variance effective size: $N_e$                                 | 2.647 (2.537–2.760) | 2.305 (2.183–2.454)  |
| Number of seed trees for $N_{e(\text{reference})} = 500$       | 189 (181–197)       | 217 (203–229)        |
| <i>(b) TWOGENER analysis</i>                                   |                     |                      |
| Pollen pool differentiation: $\Phi_{ft}$                       | 0.088 (0.068–0.108) | 0.107 (0.061–0.142)  |
| Pollen pool differentiation: $\Phi'_{ft}$                      | 0.105 (0.078–0.129) | 0.135 (0.041–0.190)  |
| Multilocus paternity correlation: $r_p = 2\Phi'_{ft}$          | 0.210 (0.156–0.258) | 0.270 (0.082–0.380)  |
| Effective number of pollen donors: $N_{ep}$                    | 4.726 (3.876–6.437) | 3.695 (2.635–12.137) |
| Coancestry within progenies: $\Theta$                          | 0.182 (0.177–0.187) | 0.212 (0.198–0.220)  |
| Variance effective size: $N_e$                                 | 2.495 (2.444–2.556) | 2.180 (2.110–2.309)  |
| Number of seed trees for $N_{e(\text{reference})} = 500$ : $m$ | 200 (195–204)       | 229 (217–237)        |

Abbreviation: ESA, Ecological station of Assis; SW, savannah woodland.

Values inside ( ) are the 95% bootstrap percentile confidence interval based on 1000 replications. Replication units are progenies for the mixed and correlated mating models and for *TWOGENER* analysis.**Table 3** Estimates of mating system parameter for each progeny of the continuous SW and the edge of the ESA

| Progeny              | $t_m$ (s.d.) | $t_m - t_s$ (s.d.) | $r_{pm}$ (s.d.) | $N_{ep}$ | $\Theta$ | $N_e$ |
|----------------------|--------------|--------------------|-----------------|----------|----------|-------|
| SW 1                 | 0.90 (0.06)  | 0.12 (0.03)        | 0.05 (0.02)     | 20.8     | 0.156    | 2.95  |
| SW 2                 | 0.95 (0.05)  | 0.10 (0.02)        | 0.05 (0.02)     | 20.4     | 0.144    | 3.16  |
| SW 3                 | 0.80 (0.09)  | 0.04 (0.04)        | 0.10 (0.03)     | 9.6      | 0.187    | 2.51  |
| SW 4                 | 0.99 (0.00)  | 0.10 (0.02)        | 0.05 (0.03)     | 19.6     | 0.133    | 3.39  |
| SW 5                 | 0.95 (0.04)  | 0.10 (0.02)        | 0.03 (0.01)     | 37.0     | 0.141    | 3.22  |
| SW 6                 | 0.95 (0.05)  | 0.03 (0.03)        | 0.07 (0.03)     | 13.9     | 0.146    | 3.12  |
| SW 7                 | 0.90 (0.06)  | 0.02 (0.03)        | 0.09 (0.05)     | 11.1     | 0.160    | 2.88  |
| SW 8                 | 0.47 (0.10)  | -0.06 (0.04)       | 0.11 (0.04)     | 9.0      | 0.295    | 1.65  |
| SW 9                 | 0.85 (0.11)  | 0.06 (0.08)        | 0.10 (0.05)     | 10.0     | 0.174    | 2.67  |
| SW 10                | 0.90 (0.06)  | 0.06 (0.03)        | 0.06 (0.03)     | 16.9     | 0.157    | 2.93  |
| SW 11                | 0.95 (0.05)  | 0.03 (0.03)        | 0.06 (0.03)     | 16.9     | 0.145    | 3.14  |
| SW 12                | 0.76 (0.09)  | -0.06 (0.04)       | 0.07 (0.03)     | 14.5     | 0.198    | 2.38  |
| SW 13                | 0.90 (0.06)  | 0.10 (0.03)        | 0.02 (0.00)     | 45.5     | 0.153    | 3.00  |
| SW 14                | 0.71 (0.10)  | 0.03 (0.04)        | 0.08 (0.03)     | 14.5     | 0.212    | 2.24  |
| SW 15                | 0.90 (0.06)  | 0.05 (0.03)        | 0.05 (0.02)     | 21.3     | 0.156    | 2.95  |
| SW 16                | 0.95 (0.05)  | 0.06 (0.03)        | 0.04 (0.01)     | 25.0     | 0.143    | 3.18  |
| SW 17                | 0.80 (0.09)  | 0.03 (0.03)        | 0.11 (0.04)     | 9.1      | 0.187    | 2.51  |
| Edge 1               | 0.75 (0.10)  | 0.07 (0.03)        | 0.04 (0.03)     | 26.3     | 0.198    | 2.38  |
| Edge 2               | 0.75 (0.14)  | 0.01 (0.10)        | 0.02 (0.02)     | 40.0     | 0.197    | 2.39  |
| Edge 3               | 0.90 (0.07)  | 0.05 (0.04)        | 0.02 (0.01)     | 62.5     | 0.153    | 3.00  |
| Edge 4               | 0.75 (0.09)  | -0.04 (0.03)       | 0.02 (0.00)     | 62.5     | 0.196    | 2.40  |
| Edge 5               | 0.90 (0.06)  | 0.10 (0.05)        | 0.05 (0.03)     | 19.2     | 0.157    | 2.93  |
| Edge 6               | 1.00 (0.00)  | 0.03 (0.01)        | 0.02 (0.01)     | 41.7     | 0.128    | 3.50  |
| Edge 7               | 0.75 (0.09)  | 0.08 (0.04)        | 0.11 (0.04)     | 8.7      | 0.203    | 2.33  |
| Edge 8               | 0.55 (0.12)  | 0.06 (0.05)        | 0.09 (0.04)     | 10.5     | 0.265    | 1.82  |
| Edge 9               | 0.40 (0.13)  | 0.03 (0.09)        | 0.08 (0.02)     | 12.0     | 0.320    | 1.53  |
| Edge 10 <sup>a</sup> | 0.85 (0.07)  | 0.15 (0.04)        | 0.18 (0.08)     | 5.6      | 0.182    | 2.57  |
| Edge 11              | 0.75 (0.10)  | -0.01 (0.05)       | 0.07 (0.04)     | 14.3     | 0.200    | 2.36  |

Abbreviations: ESA, Ecological Station of Assis;  $N_{e_i}$ , average variance effective size within progeny;  $N_{ep}$ , effective number of pollen donors;  $r_{pm}$ , multilocus paternity correlation; SW, savannah woodland;  $t_m$ , multilocus outcrossing rate;  $t_m - t_s$ , outcrossing rate among relatives;  $\Theta$ , average coancestry coefficient within progenies.<sup>a</sup>Indicates the isolated tree in the western edge of ESA.

in contrast to 2 of the 17 progenies (12%) sampled in the SW. Individual differences in  $t_m - t_s$  were also variable among the seed trees, ranging from -0.06 to 0.12 in the SW and from -0.04 to 0.15 in the edge (Table 3). Multilocus paternity correlation ( $r_{pm}$ ) ranged from 0.02 to 0.11 in the SW and from 0.02 to 0.18 in the edge. Consequently, the effective number of pollen donors ( $N_{ep}$ ) were also variable among the seed trees, ranging from 9.0 to 45.5 in the SW and from 5.6 to 62.5 in the edge (Table 3).

#### Pollen pool differentiation and pollen dispersal

The global differentiation of the pollen pool ( $\Phi_{ft}$ ) was significantly different from zero among seed trees in the SW ( $\Phi_{ft} = 0.088$ ) and at the edge ( $\hat{\Phi}_{ft} = 0.107$ ), however, it did not differ between these two sites (Table 2b). Selfing significantly influenced the estimated value of  $\hat{\Phi}_{ft}$ . Excluding the effect of selfing,  $\Phi'_{ft}$  was 19% higher in the SW and 26% higher at the edge compared with  $\Phi_{ft}$  (Table 2b). Moreover, for pollen pool differentiation when we sampled 11 progenies from SW ( $\Phi_{ft} = 0.088$ ;  $CI_{95\%} = 0.064-0.115$  and  $\Phi'_{ft} = 0.106$ ;  $CI_{95\%} = 0.070-0.132$ ) or removed the isolated tree of the edge from the analysis ( $\Phi_{ft} = 0.104$ ;  $CI_{95\%} = 0.056-0.142$  and  $\Phi'_{ft} = 0.128$ ;  $CI_{95\%} = 0.058-0.163$ ), no differences were found from the previous estimates based on 17 and 11 progenies, respectively. The multilocus paternity correlation ( $r_p = 2\Phi'_{ft}$ ) and  $N_{ep}$  estimates (Table 2b) were significantly different than the values estimated using the MLTR programme (Table 2a). No significant decrease in  $\Psi$  was detected with the increase in the distance between seed trees in the SW ( $r = -0.137$ ;  $P = 0.11$ ) and at the edge ( $r = -0.059$ ;  $P = 0.67$ ). Thus, the *KINDIST* estimation of the pollen dispersal distance becomes unreliable and, consequently,  $\delta$  and  $d_c$  were not estimated.

#### Coancestry, variance effective size and the number of seed trees for seed harvesting

The average coancestry coefficient ( $\Theta$ ) within progenies estimated from mating system parameters and from *TWOGENER* analysis was

significantly lower in the SW ( $\Theta = 0.170$ ;  $\Theta = 0.182$ , respectively) than at the edge ( $\Theta = 0.199$ ;  $\Theta = 0.212$ , respectively; Table 2). The estimated parameters  $\Theta$ ,  $N_e$  and  $m$  did not differ between MLTR and TWOGENER procedures (Table 2). In progeny level, the estimated coancestry coefficient ( $\Theta$ ) was variable, ranging from 0.133 to 0.295 in the SW and from 0.128 to 0.320 in the edge (Table 3).

To both estimates from mating system parameters and TWOGENER analysis, the average variance effective size within a progeny was significantly lower than expected for a panmictic population ( $N_e = 4$ ) and was significantly lower at the edge than in the SW (Table 2). These results suggest that seed collected from the edge have generally higher relatedness and lower variance effective size than seeds collected from the SW. Among progenies, the estimate of the variance effective size ranged from 1.65 to 3.39 in the SW and from 1.53 to 3.50 in the edge (Table 3). Therefore, the minimum number of seed trees required to harvest seeds was significantly higher at the edge ( $\hat{m} = 217$  for mating system parameters and  $\hat{m} = 229$  for TWOGENER analysis) than in the SW ( $\hat{m} = 189$  for mating system parameters and  $\hat{m} = 200$  for TWOGENER analysis).

### Seed germination

The average seed germination value was significantly higher ( $P < 0.024$ ) in seeds from the SW (23.4; s.e. = 0.2) than in those from the edge (22.1; s.e. = 0.6). However, at the progeny level, we do not find any significant association between germination and selfing, observed heterozygosity, expected heterozygosity and fixation index, using a linear Spearman rank correlation.

## DISCUSSION

### Genetic diversity, inbreeding and mating system

The results of our study supported our initial hypothesis that seeds collected from the centre of a remnant have higher heterozygosity levels ( $H_o$  and  $H_e$ ) than seeds collected at the edge (Table 1). This is most likely a consequence of the low population density of reproductive trees at the edge of the remnant, which results in the selfing rate higher than 0.25 in 64% of the open-pollinated progenies at the edge, in contrast to 12% of progenies from in the SW. This provides direct evidence of low genetic recombination in seeds generated at the edge. Selfing is expected at a higher rate in harsh environments (Goodwillie *et al.*, 2005; Lowe *et al.*, 2005) and in low-density populations (Murawski and Hamrick, 1991). Significant levels of inbreeding were detected at both sites, the results suggest that selfing decreased the heterozygosity levels and increase inbreeding in seeds from the edge. Thus, the higher levels of selfing may explain the low observed heterozygosity at the edge. However, in general, selfing in tropical tree species produces inbreeding depression, causing the mortality of selfed seeds and individuals as was observed in *Eucalyptus argutifolia* (Kennington and James, 1997), *Eucalyptus marginata* (Millar *et al.*, 2000), *Pseudobombax munguba* (Gribel and Gibbs, 2002), *Platypodium elegans* (Hufford and Hamrick, 2003) and *Neobalanocarpus heimii* (Naito *et al.*, 2005). We observed significant higher germination in the SW than the edge, suggesting that inbreeding depression was higher in seeds from the edge, where selfing was highest. The lack of association between germination and selfing, fixation index, observed heterozygosity or expected heterozygosity may be caused by the inbreeding depression. However, inbreeding depression occurs in many ontogenic stages and, consequently, the levels of heterozygosity and the observed inbreeding in *C. langsdorffii* for both sites may substantially change between the seed and adult stages. If inbreeding

depression were to affect selfed seeds, the observed levels of inbreeding in the SW and the edge would decrease significantly until the adult stage. Although data regarding inbreeding depression in the *C. langsdorffii* are not available, this phenomenon most likely also occurs in this tropical tree species. In *C. langsdorffii*, as in many tropical tree species, outcrossing is important for promoting the maintenance of genetic diversity and the adaptive potential of the species. A high adaptive potential counteracts the effects of environmental stochasticity, especially in harsh environments, such as remnant edges or newly colonised areas.

### Pollen pool differentiation

Although, the average global differentiation of the pollen pool ( $\Phi'_{ft}$ ) seemed to be higher in edge than in the SW, no significant difference was found between sites. The possible reasons for the lack of difference in  $\Phi'_{ft}$  values between both sites are related to sample design and edge effects: (1) the spatial distance among seed trees in the edge was larger than in the SW (Figure 1); (2) in the edge, there is a lower reproductive density of trees and a probable scarcity of pollinators. The large distance among the sampled seed trees in the edge, associated to the low reproductive density of trees and the scarcity of pollen donors probably increased pollen pool differentiation in the edge in comparison with the SW, because of the low overlapping pollen pool among seed trees in the edge, but increased also the  $CI_{95\%}$  due the higher individual variation in outcrossing rate (Table 3). In contrast, the spatial distance among the seed trees of the SW was lower, probably increasing overlapping pollen pool among seed trees and decreasing average global pollen pool differentiation in comparison with the edge. Thus, it is not possible to attribute the higher genetic differentiation in pollen pool only to the edge effects because there is a mixture of sampling effect with edge effects. To separate these effects, a further study must be conducted, based on sampling seed trees in the SW more distant than has been sample in the present study.

The paternity correlation estimate ( $r_{pm}$ ) and the effective number of pollen donors ( $N_{ep}$ ) were significantly different between correlated matings model (Ritland, 1989) and TWOGENER analysis (Smouse *et al.*, 2001). The correlated mating model estimated a low paternity correlation and consequently, high number of pollen donors in both the SW and the edge, and no differences were found between the two sites (Table 2a). In contrast, the TWOGENER analysis estimated a significantly higher paternity correlation and a lower number of pollen donors than the correlated mating model for both the SW and the edge (Table 2b). The difference in the results between the two approaches is associated to the differences in the methods of estimation. In the TWOGENER analysis, both parameters  $r_p$  and  $N_{ep}$  were estimated indirectly from the pollen pool differentiation ( $\Phi'_{ft}$ ), a parameter strongly affected by the sample design. For example, the distances between seed trees in SW were lower than seed trees located in the edge, which resulted in a higher average global pollen pool differentiation in the edge (Table 2) and consequently, resulted in the significant higher  $r_p$  and lower  $N_{ep}$  in the edge than in SW. In contrast, in the correlated mating model, the parameter  $r_p$  is estimated from the expression,  $r_p = [2/(1+F)]f$ , where  $F$  is the inbreeding in the parental parents and  $f$  is the correlation of paternal gametes (Ritland, 1989). Thus, joint estimation of inbreeding levels in the parental and offspring generation are necessary to estimate the paternity correlation (Ritland, 1989), but this estimate is independent of distance among seed trees. In summary, the correlating mating model estimates  $r_p$  from the

inbreeding in the parental parents and the correlation of paternal gametes and does not directly depend on the distance of seed trees as is the case of TWOGENER analysis. This can explain the observed difference between the two methods in the parameters  $r_p$  and  $N_{ep}$ . Thus, as our sample design, based on seed trees located in very different distance among seed trees between the SW and edge does not allow a robust comparison between these two sites. For our specific case, correlated mating model is more robust to compare the estimated parameters.

Although the estimated paternity correlation and the effective number of pollen donors were significantly different between the methods, in general no significant differences were observed for the coancestry coefficient ( $\Theta$ ), the variance effective size within progenies ( $N_e$ ), or the necessary number of seed trees from which seeds should be collected ( $m$ ). This absence of differences is caused by the fact that the parameters  $N_e$  and  $m$  were estimated from the parameter  $\Theta$ , and this last parameter is mainly affected by the levels of selfing (or outcrossing). As we used the same values for selfing (and outcrossing), which were calculated from a mixed mating model to estimate the parameter  $\Theta$ , for both models, no differences were detected in the estimate of the three parameters  $\Theta$ ,  $N_e$  and  $m$ .

#### Implications for conservation genetics and evolution

Selfing, mating among relatives and correlated matings increase the frequency of identical-by-descent alleles within open-pollinated progenies. In another word, these three cited process increase inbreeding and the relatedness and decrease the heterozygosities within the progenies, as observed comparing the SW and the edge (Table 1). This has strong implications for harvesting seeds for conservation, breeding programs and environmental restoration, as the objective is to achieve maximal genetic diversity and variance effective size in the seed samples. Therefore, seed harvesting must preferably be conducted where selfing, mating among relatives and correlated matings are less frequent. Thus, according to the estimates of genetic diversity, selfing, coancestry within progenies and average variance effective size ( $N_e$ ), seed harvesting must be performed within the SW. At this site, seeds must be harvested from at least 189 seed trees so that the sampled progeny arrays reach the required  $N_e$  of 500. However, it is important to note that some seed trees should be selected at the edge of the remnant to compound the total sample, as these seeds contain private alleles that originated from other demes.

#### DATA ARCHIVING

Data deposited in the Dryad repository: doi:10.5061/dryad.73g3d.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for funding this project (grant 07/06648-1), the Instituto Florestal do Estado de São Paulo (IF) for the research and field activities licence in the ESA (licence 44.184/2006), the Laboratório de Reprodução e Genética de Espécies Arbóreas (University of São Paulo) for laboratory help and Luciano Angelo de Souza Bernardes (Santa Cruz State University, NCBGIB) for bioinformatics help. Roberto Tarazi was supported by a FAPESP PhD scholarship (grant 2006/04490-9), and Alexandre M Sebbenn, Paulo Y Kageyama and Roland Vencovsky are supported by a National Counsel of Technological and Scientific Development (CNPq)

research fellowship. We also thank three anonymous reviewers for their important and very constructive suggestions for the genetic analyses in the manuscript.

- Aguilar R, Quesada M, Ashworth L, Herreras-Diego Y, Lobo J (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* **17**: 5177–5188.
- Ashley MV (2010). Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. *Critical Rev Plant Sci* **29**: 148–161.
- Austerlitz F, Smouse PE (2001). Two-generation analysis of pollen flow across a landscape. II. Relation between  $\Phi_{ft}$ , pollen dispersal and interfemale distance. *Genetics* **157**: 851–857.
- Bittencourt JVM, Sebbenn AM (2008). Pollen movement and spatial genetic structure in a continuous forest of wind-pollinated *Araucaria angustifolia*, inferred from paternity and TwoGener analysis. *Conserv Genet* **9**: 855–868.
- Burczyk J, Koralewski TE (2005). Parentage versus two-generation analyses for estimating pollen-mediated gene flow in plant populations. *Mol Ecol* **14**: 2525–2537.
- Carvalho PER (2003). *Espécies Florestais Brasileiras: recomendações silviculturais, potencialidades e uso da madeira*. EMBRAPA-CNPq: Colombo.
- Ciampi AY, Brondani RPV, Grattapaglia D (2000). Desenvolvimento de marcadores microsatélites para *Copaifera langsdorffii* Desf. (copaiba) Leguminosae—Caesalpinioideae e otimização de sistemas fluorescentes de genotipagem multilocus. *Bolet Pesq Embrapa Recursos Genéticos e Biotecnologia* **16**: 1–40.
- Cockerham CC (1969). Variance of gene frequency. *Evolution* **23**: 72–84.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.
- Durigan G (2006). Observations on the southern cerrados and their relationship with the core area. In Pennington RT, Lewis GP, Ratter JA (eds). *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography and Conservation*. Taylor & Francis: London, pp 67–77.
- El Moussadiq A, Petit RJ (1996). High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theor Appl Genet* **92**: 832–839.
- Freitas CV, Oliveira PE (2002). Reproductive biology of *Copaifera langsdorffii* Desf. (Leguminosae, Caesalpinioideae). *Rev Bras Bot* **25**: 311–321.
- Fuchs EJ, Lobo JA, Quesada M (2003). Effects of forest fragmentation and flowering phenology on the reproductive success and mating patterns of the tropical dry forest tree *Pachira quinata*. *Conserv Biol* **17**: 149–157.
- Goodwillie C, Kalisz S, Eckert CG (2005). The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annu Rev Ecol Syst* **36**: 47–79.
- Goudet J (1995). FSTAT. (Version 2.9.3.2.): a computer program to calculate F-statistics. *J Heredity* **86**: 485–486.
- Gribel R, Gibbs PE (2002). High outbreeding as a consequence of selfed ovule mortality and single vector bat pollination in the Amazonian tree *Pseudobombax munguba* (Bombacaceae). *Int J Plant Sci* **163**: 1035–1043.
- Hufford KM, Hamrick JL (2003). Variability selection at three early life stages of the tropical tree *Platypodium elegans* (Fabaceae, Papilionoideae). *Evolution* **57**: 518–526.
- Hirao AS, Kameyama Y, Ohara M, Isagi Y, Kudo G (2006). Seasonal changes in pollinator activity influence pollen dispersal and seed production of the alpine shrub *Rhododendron aureum* (Ericaceae). *Mol Ecol* **15**: 1165–1173.
- Kennington WJ, James SH (1997). The effect of small population size on the mating system of a rare mallee, *Eucalyptus argutifolia* (Myrtaceae). *Heredity* **78**: 252–260.
- Klink CA, Machado R (2005). Conservation of the Brazilian Cerrado. *Conserv Biol* **19**: 707–713.
- Laurance WF, Nascimento HEM, Laurance SG, Andrade A, Ewers RM, Harms KE *et al.* (2007). Habitat fragmentation, variable edge effects, and the landscape-divergence hypothesis. *PLoS One* **10**: e1017.
- 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.
- Millar MA, Byrne M, Coates DJ, Stukely MJC, McCom JA (2000). Mating system studies in jarrah, *Eucalyptus marginata* (Myrtaceae). *Aust J Bot* **48**: 475–479.
- Moares MLT, Sebbenn AM (2011). Pollen dispersal between isolated trees in the Brazilian savannah: a case study of the neotropical tree *Hymenaea stigonocarpa*. *Biotropica* **43**: 192–199.
- Murawski DA, Hamrick JL (1991). The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity* **67**: 167–174.
- Naito Y, Konuma A, Iwata H, Suyama Y, Seiwa K, Okudo T *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-dependent selfing and its effects on seed performance in a tropical canopy tree species, *Shorea acuminata* (Dipterocarpaceae). *Forest Ecol Manage* **256**: 375–383.
- Oosterhout CV, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Eco Notes* **4**: 535–538.

- Pinheiro ES (2008). *Ecological and remote sensing analyses applied to estimate the cerrado phytomass in the Assis Ecological Station, São Paulo State, Brazil. PhD Thesis 192p.* University of São Paulo: Brazil.
- Pinheiro ES, Durigan G (2009). Spatial and temporal dynamics (1962–2006) of Cerrado vegetation types in a protected area, southeastern Brazil. *Rev Bras Bot* **32**: 441–454.
- Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FJ, Hirota MM (2009). The Brazilian Atlantic forest: how much is left, and how is the remaining forest distributed? Implications for conservation. *Biol Conserv* **142**: 1141–1153.
- Ritland K (1989). Correlated mating in the partial selfer *Mimulus guttatus*. *Evolution* **43**: 848–859.
- Ritland K (2002). Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity* **88**: 221–228.
- Ritland K, Jain S (1981). A model for the estimation of outcrossing rate and gene frequency using independent loci. *Heredity* **47**: 35–52.
- Robledo-Arnuncio JJ, Alía R, Gil L (2004). Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. *Mol Ecol* **13**: 2567–2577.
- Robledo-Arnuncio JJ, Austerlitz F, Smouse PE (2007). POLDISP: a software package for indirect estimation of contemporary pollen dispersal. *Mol Ecol Notes* **7**: 763–766.
- SAS Institute Inc. (1999). *SAS Procedures Guide. Version 8 (TSMO)*. SAS Institute Inc. Cary, NC, 27513USA.
- Sebbenn AM, Carvalho ACM, Freitas MLM, Moraes SMB, APSC Gaino, Silva JM *et al.* (2011). 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. *Heredity* **106**: 134–145.
- Smouse PE, Dyer RJ, Westfall RD, Sork VL (2001). Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. *Evolution* **55**: 260–271.
- Sork VL, Smouse PE (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecol* **21**: 821–836.
- Tarazi R, Sebbenn AM, Mollinari M, Vencovsky R (2010). Mendelian inheritance, linkage and linkage disequilibrium in microsatellite loci of *Copaifera langsdorffii* Desf. *Conserv Genet Resources* **2**: 201–204.