

1 **Inbreeding depression from selfing and mating between relatives in the Neotropical tree *Cariniana legalis* Mart.**

2 **Kuntze**

3 Evandro Vagner Tambarussi^{1,5}, David Boshier², Roland Vencovsky³, Miguel Luiz Menezes Freitas⁴ and Alexandre

4 Magno Sebbenn⁴

5
6 ¹ Universidade Estadual do Centro-Oeste, PR 153, Km 7, Irati, PR, 84500-000, Brazil1,

7 ² University of Oxford, Department of Plant Sciences, South Parks Road, Oxford, OX1 3RB, UK,

8 ³ Escola Superior de Agricultura “Luiz de Queiroz” Universidade de São Paulo, Av. Pádua Dias, 11, Caixa Postal 9,
9 Piracicaba, SP, 13418-900, Brazil,

10 ⁴ Instituto Florestal de São Paulo, CP 1322, São Paulo, SP, 01059-970, Brasil.

11 ⁵ Corresponding author: tambarussi@gmail.com (Telephone: +55 42 3421 3087 and fax: + 55 42 3421 3067)

12

13 **Keywords:** conservation genetics; inbreeding; mating among relatives; microsatellite loci; tropical tree species.

14

15

16 **Abstract**

17 Selfing or mating between related individuals in self-compatible hermaphroditic tree species may lead to inbreeding
18 depression (ID) due to homozygosis in recessive, identical by descent alleles. In general, studies of ID in tree species
19 have been based on comparisons of selfed individuals (produced by controlled pollination) with outcrossed individuals
20 for quantitative traits in progeny tests. However, this approach requires a long time to quantify the extent of ID. Thus,
21 we used an approach based on genetic markers to estimate coancestry coefficients between assigned parents from
22 paternity analysis in two populations of the Neotropical tree *Cariniana legalis*. Using this method, we were able to
23 determine which seedlings in a nursery trial originated from; i) outcrossing between un-related trees, ii) mating between
24 related trees and iii) selfing. We detected a low selfing rate (< 10%), but a substantial quantity of seedlings from mating
25 between related parents (minimum of 35.7%). In general, the outcrossed seedlings from unrelated parents exhibited
26 significantly greater genetic diversity than those resulting from selfing and mating among relatives. The extent of ID
27 varied among traits and populations. Outcrossed seedlings originating from unrelated trees generally showed greater
28 survival than seedlings originating from selfing and related parents. Inbreeding depression was greater in the selfed
29 seedlings than in those from mating among related parents. The results are discussed in terms of implications for genetic
30 conservation, breeding and environmental restoration using the species.

31

32 Introduction

33 Tree populations across much of the world have suffered fragmentation or intensive logging, with associated decreases
34 in population size and physical isolation. A consequence may be increased rates of selfing and mating between related
35 trees, with a higher incidence of inbred individuals in succeeding generations (Lowe et al. 2005; Chaves et al. 2011;
36 Manoel et al. 2012; Ismail et al. 2012; Rymer et al. 2015; Tambarussi et al. 2015). Inbreeding in plants may lead to
37 inbreeding depression (ID), where homozygosity increases in identical by descent (IBD) alleles produce individuals
38 with reduced mean phenotypic values for reproductive, physiological efficiency and adaptive traits, compared to those
39 originating from outcrossing among unrelated individuals (Charlesworth and Charlesworth 1987; Ritland 1996;
40 Falconer and Mackay 1997; Byers and Waller 1999; Winn et al. 2011). As ID may decrease survival, growth and
41 reproductive fertility, the possibility and extent of ID in tree populations is of direct practical importance for genetic
42 conservation, breeding and environmental restoration, in particular for small isolated (Lowe et al. 2005). Use of inbred
43 seeds should be avoided in *ex situ* conservation due to the decrease in effective population size, while higher mortality
44 and reduced growth increase the costs of conservation and reforestation.

45
46 In natural populations of insect-pollinated tree species, inbreeding can result from pollinator visitation of flowers on the
47 same tree (selfing) or on neighbouring trees (mating among relatives) that show intra-population spatial genetic
48 structure (SGS) generally resulting from limited seed dispersal (e.g. Bittencourt and Sebbenn 2007; Dick et al. 2008).
49 Selfing produces at least 50% inbreeding ($F_s = 0.5(1 + F_m)$), where F_m is the inbreeding coefficient of the mother,
50 while mating between two related individuals produces inbreeding (F_r) of similar strength as the coancestry (θ)
51 among the related parents, $F_r = \theta$ (Lindgren et al. 1997). Thus, inbreeding due to mating among related trees is
52 expected to be less deleterious than that due to selfing, because mating between full-sibs ($\theta = 0.25$) or an offspring and
53 one parent ($\theta = 0.25$) will produce a maximum of 25% inbreeding.

54
55 Inbreeding depression can be expressed at different life stages of an organism, while also varying across populations,
56 individuals and traits (Griffin and Cotterill 1988; Wu et al. 1998; Koelewijn et al. 1999). In the first case, the genes
57 responsible for deleterious effects at different stages of life are not necessarily the same. Variation in ID among
58 individual plants and populations can be expected as they differ in the numbers and frequency of deleterious alleles they
59 carry, with mutations differentially accumulated across populations and progeny (Koelewijn et al. 1999). Many studies
60 of coniferous and deciduous tree species comparing selfed with outcrossed progeny in provenance and progeny tests,
61 have revealed ID for flowering, seed production and germination, survival, and growth traits, as well as increases in
62 phenotypic variation among progeny arrays (e.g. Griffin and Cotterill 1988; Sorensen 1997; Wu et al. 1998; Koelewijn
63 et al. 1999; Chaves et al. 2011).

65 In the present study, we used a different approach to estimate ID in a large tree of the Atlantic rain forest, i.e.,
 66 *Cariniana legalis* L (Lecythidaceae), using seven codominant nuclear microsatellite loci, paternity analysis and
 67 estimates of pairwise coancestry between seed trees and pollen donors assigned by paternity analysis. Paternity analysis
 68 permits the detection of selfed seeds and the father of outcrossed seeds where the seed tree is known. Outcrossed seeds
 69 are not inbred if they result from mating between unrelated parents and inbred if from related parents. The degree of
 70 relatedness among seed trees and assigned pollen donors (seed parents) of sampled seeds can be inferred from pairwise
 71 coancestry coefficient estimates (θ). Seed parents can be non-relatives ($\theta \leq 0$) or relatives ($\theta > 0$). Thus, the
 72 combination of paternity analysis and estimates of the pairwise coancestry coefficient between seed parents permits
 73 estimation of the rate of mating among relatives. If genotyped open-pollinated seeds used in the paternity analysis are
 74 established in a progeny test, quantitative and qualitative traits can be measured. Thus, each plant can be grouped as
 75 originating from selfing, mating among related individuals or outcrossing among non-relatives, and the means of traits
 76 can be estimated for each of these three groups. Consequently, ID from selfing and mating among relatives can be
 77 estimated for a phenotypic trait, as the relative performance of inbred individuals in relation to the performance of
 78 outcrossed individuals from unrelated parents. Using this approach, we investigated the extent of ID due to selfing and
 79 mating among relatives, for survival, growth in height (H) and root-collar diameter (RCD) in 8 month old seedlings
 80 originating from seed collected in three fragmented *C. legalis* populations and established in a nursery provenance and
 81 progeny test. We addressed the following questions: i) Which of the study traits was most strongly affected by ID? ii)
 82 Are there differences in the levels of ID between populations? iii) What are the differences in ID among seedlings that
 83 originated from selfing compared to mating among relatives? iv) Are there differences in the levels of genetic diversity
 84 between seedlings produced by selfing, mating among relatives and outcrossing among non-relatives? v) Based on
 85 random samples of seedlings from these populations, what is the expected level of ID in the offspring generation?

86

87 **Material and Methods**

88 **The studied species**

89 *Cariniana legalis* Mart. O. Kuntze (Lecythidaceae) is an endemic, upper canopy tree of the Atlantic forest that occurs
 90 between latitudes 8 and 23° S, generally at low density (< 1 tree/ha), reaching 60 m in height and 4 m in diameter at
 91 breast height (Carvalho, 2003). The species is semi-heliophytic and evergreen, with hermaphroditic flowers pollinated
 92 by bees of the genera *Melipona* and *Trigona* (Prance and Mori 1979), and seed dispersed by gravity and wind (Carvalho
 93 2003). The wood is light and used in civil construction and furniture. The species is considered endangered (Sebbenn et
 94 al. 2000), and a better understanding of ID will provide useful information for effective *in* and *ex situ* conservation
 95 measures.

96

97 **Study sites and sampling**

98 The study was performed using open-pollinated seeds collected from three remnant *C. legalis* populations. The first site
99 was the Ibicatu State Forest (22°46' S and 47°43' W, mean altitude of 542 m) located in Piracicaba, São Paulo State,
100 Brazil, which has a total area of 72 ha. The second site was the Mata da Figueira (MGI), a small fragment (7.2 ha) of
101 riparian forest on a semi-deciduous plateau within the Ecological Station Mogi-Guaçu, Farm "Campininha" of the
102 Forestry Institute of São Paulo (22°16' S and 47°11' W, mean altitude of 600 m). Approximately 2.9 km from MGI and
103 within the same Ecological Station are four *C. legalis* trees (MGII). For our analysis, we pooled MGI and MGII, and
104 refer to the combined population as MG. Ibicatu is isolated from other *C. legalis* populations by at least 2.5 km, while
105 MGI and MGII are 3 km apart and approximately 75 km from Ibicatu. We mapped all trees using a GPS III (Garmin,
106 USA) and collected stem bark tissue for microsatellite analysis of all individuals at Ibicatu (65 trees) and MG (26 trees).
107 More than 50 open-pollinated fruit were collected from different parts of the canopy of 15 randomly selected seed trees
108 in Ibicatu, five in MG1 and two in MGII. Sampling was limited at the MG sites, owing to a lack of fruit production in
109 some trees. After harvesting, fruits were packaged in plastic bags separately by seed tree and left in the shade for about
110 15 days to facilitate DNA seed extraction.

111

112 **Nursery provenance and progeny test**

113 In August of 2011, a nursery provenance and progeny trial of *C. legalis* was established at the University of São Paulo
114 (ESALQ/USP), Piracicaba, São Paulo State, Brazil, using a sub-set of the collected seeds, owing to space limitations.
115 The trial was established in a randomized complete block design at 10 x 10 cm spacing with 22 families (15 and 7
116 families from Ibicatu and MG, respectively), five replicates (blocks) and six trees per plot, totalling 30 plants per family
117 (seed tree), and 450 (n_{IB}) plants from Ibicatu and 210 (n_{MG}) from MG (total in the trial: $n_T = 660$ plants). The total
118 number of germinated seeds (n_g) in the trial was 553 of 660 (83.8%), being 378 (84.0%) from Ibicatu and 175 (83.3%)
119 from MG. Seedlings were evaluated individually for total height (H), root-collar diameter (RCD), and survival (SUR)
120 eight months after germination. Total height and RCD were measured using an IDF Digital Caliper IP67 (IDF
121 Metrology, Pontoglio, Italy) with a resolution of 0.01 cm. The SUR trait presented a binomial distribution as individuals
122 may be living (1) or dead (0), with percentage survival rate calculated as $= (n_{survival} / n_g)100\%$, where $n_{survival}$ is the
123 number of surviving seedlings. As we could not determine if seeds did not germinate due to inbreeding from selfing or
124 mating among relatives, or from stochastic causes, such as disease and predation, we estimated the realized level of
125 seedling survival from outcrossing among unrelated individuals, related parents and from selfing, based on the
126 genotypes of germinated seedlings from each population (Table 1).

127

128 **DNA extraction and microsatellite analysis**

129 For all adult trees, DNA was extracted from 100 mg of stem bark tissue per individual using AnalytikJena DNA
130 isolation kits. Seeds were germinated in vermiculite until the cotyledons emerged, and DNA was then extracted from
131 the first true leaf pairs using the method of Doyle and Doyle (1990). Germinated seeds (seedlings) were genotyped for
132 seven microsatellite loci as follows: Ibicatu 600 seeds (from eight fruits of each seed tree, five seeds per fruit, totalling
133 40 seeds per tree), MGI 250 seeds (50 seeds from ten fruits of each seed tree; five seeds per fruit) and MGII 200 seeds
134 (100 seeds from 20 fruits of each seed tree, five seeds per fruit). We used seven specific microsatellite loci developed
135 for *C. legalis* (Cle01, Cle04, Cle05, Cle08, Cle09, Cle10, and Cle12), with details on DNA amplification and
136 genotyping for these loci described in Tambarussi et al. (2013a). These microsatellite loci were selected because they
137 exhibited Mendelian inheritance, with high levels of polymorphism and were unlinked (Tambarussi et al. 2013b). The
138 amplification products were separated on a Fragment Analyzer™ Automated CE System (Advanced Analytical
139 Technologies, Inc. [AATI], Ames, Iowa, USA) using dsDNA Reagent Kit, 35-500 bp. The raw data were analyzed
140 using the software PROSize version 2.0 (AATI). Samples with questionable or missing alleles (peaks), or for which
141 there were mismatches between mothers and progeny, were genotyped a second time.

142

143 **Paternity analysis**

144 Paternity analyses were performed using a categorical likelihood method implemented in the Cervus 3.0 program
145 (Marshall et al. 1998; Kalinowski et al. 2007). The power of our seven loci to correctly detect putative pollen donors in
146 the populations was estimated by the non-exclusion probability of a second parent when the first parent is known (P_2),
147 using Cervus. Paternity analyses for the Ibicatu and MG populations were conducted with the genotypes of adult trees,
148 seed trees and the seeds from each population. To determine putative pollen donors, all adult trees of each population
149 were considered as candidate pollen donors. In the paternity analyses, the most likely pollen donor was determined
150 using the reference allele frequencies from the two adult populations, as suggested by Meagher and Thompson (1987).
151 The paternity of each seed was determined based on the Δ statistic (Marshall et al. 1998), which is defined as the
152 difference between the "LOD score" of the most likely candidate father and the "LOD score" of the second-most likely
153 candidate. Significance was determined with paternity tests simulated by the Cervus program, with the Δ statistic
154 determined to a confidence level of 80%, as suggested by Marshall et al. (1998), using 10,000 repetitions. As we did not
155 detect any mismatches between mother and sibs, suggesting no (or very few) genotyping errors, the Cervus default was
156 used for the ratio of genotyping errors (0.01). The proportion of pollen donors sampled was held at 70%, due to the high
157 degree of isolation of the Ibicatu (at least 2.5 km) and MG (3 km) populations and the fact that in insect pollinated trees,
158 pollen dispersal generally occurs among near-neighbors, following a pattern of isolation by distance (Dick et al. 2008;
159 Degen and Sebbenn 2014), decreasing the probability of pollen immigration into stands. The minimum number of loci
160 examined to determine the paternity of a seed was fixed at six; thus, one single mismatch among the trio of seed, tree-
161 seed and putative pollen donor was accepted. If a seed had no potential pollen donor within the population, that seed

was considered to have a pollen donor located outside the population (pollen immigration). We estimated the selfing rate (s), outcrossing rate (t), outcrossing rate between unrelated parents (t_u) and from related parents (t_r) based on genotypes of the germinated seedlings (n_g). The selfing rate (s) was estimated as the proportion of seeds determined to have the same seed tree as the pollen donor (n_s) in relation to n_g by $s = n_s / n_g$ and the outcrossing rate was calculated as $t = 1 - s$. To determine seedlings originating from mating among related parents we estimated the relatedness of seed trees and fathers (assigned by paternity analysis) using the coancestry coefficient (θ) calculated by the method described in Loiselle et al. (1995), and implemented in the Spagedi 1.3 program (Hardy and Vekemans 2002). Following Ismail et al. (2014), if the coancestry between parents (θ) was ≥ 0.1 , we assumed the seedling was inbred due to mating between related parents with $t_r = n_r / n_g$, while mating between unrelated trees was calculated by $t_u = n_u / n_g$, where n_u and n_r are the number of seedlings that originated from unrelated parents and related parents, respectively. As all sampled adult trees in the populations were genotyped and their spatial position known (x and y coordinates), the seedlings assigned to a pollen donor were used to determine the mean pollen dispersal distance (D), and the standard deviation calculated according to the Euclidean distance between the two points. The 95% confidence interval (95% CI) was estimated as $= 2(SD/\sqrt{n})$, where SD is the standard deviation.

Analyses of genetic diversity and inbreeding

Genetic diversity and inbreeding were characterized by combining individuals within the populations into three classes: i) outcrossing between unrelated individuals (t_u), ii) mating among related individuals (t_r), and iii) selfing (s). The estimated indices were the total number of alleles (k), allelic richness (R), and observed heterozygosity (H_o). Inbreeding was estimated by the fixation index (F), and significance tested by permuting the alleles among individuals. These analyses were performed using the FSTAT program (Goudet 1995). To investigate if the indices R , H_o and F were significantly different between seedlings originating from t_u , t_r , and s an unpaired t -test (Sokal and Rohlf 1995) was used based on single-locus data.

Inbreeding depression

Differences in survival and mean growth traits between classes (t_u , t_r , and s) were tested using ANOVA and Tukey procedures in the SAS program (SAS, 1999). Mean inbreeding depression (ID) was estimated for SUR, H and RCD at eight months of age from seedlings in the progeny test, in terms of mating among relatives (ID_r), selfing (ID_s) and the total for mating among relatives and selfing (ID_T) according to the following expressions:

191 $ID_r = 100\% [1 - (\bar{x}_r / \bar{x}_u)],$

192 $ID_s = 100\% [1 - (\bar{x}_s / \bar{x}_u)],$ and

193 $ID_T = 100\% [1 - (\bar{x}_i / \bar{x}_u)],$

194 where \bar{x}_u , \bar{x}_r , \bar{x}_s and \bar{x}_i are the means of traits for individuals that were from mating between unrelated trees,
 195 between relatives, those that were selfed, and all inbred individuals (i.e. selfed + mating among relatives), as determined
 196 by paternity analyses and the coancestry coefficient between parents.

197

198 **Results**

199 **Paternity analysis and mating system**

200 The non-exclusion probability of a second parent (P_2) across seven microsatellite loci in adults was very low in Ibicatu
 201 (0.00215) and MG (0.00045), indicating that this set of loci have good power to detect the pollen donors of seedlings in
 202 the populations. The paternity analysis assigned one putative pollen parent for all germinated seedlings (Table 1) and
 203 detected one pollen immigrant in MG1, assigned from a MGII tree, indicative of 2% pollen immigration. Of 378
 204 germinated seedlings from Ibicatu, 20 were assigned to the same seed tree, indicating a realized selfing rate (s) of
 205 5.3% (Table 1). In MG, of 175 germinated seedlings, 16 were assigned to the same seed tree, indicating a selfing rate of
 206 9.1%. Thus, the realized outcrossing rate was 94.7% and 90.9% in Ibicatu and MG, respectively. Assuming that
 207 seedling parents with coancestry coefficient ≥ 0.1 are related (95% CI), the mean pairwise coancestry coefficient ($\bar{\theta}$)
 208 among seed trees and assigned unrelated pollen donors was significantly lower [Ibicatu= mean 0.08 (95% CI
 209 0.07/0.09); MG= -0.02 (-0.03/-0.01)] than that of the seed trees and assigned related pollen donors that resulted in
 210 mating among relatives [Ibicatu= mean 0.22 (0.20/0.23); MG= 0.12 (0.11/0.14)], which suggest that mating among
 211 relatives occurred between full-sib parents in Ibicatu ($\theta = 0.25$) and half-sib parents in MG ($\theta = 0.125$). The
 212 outcrossing rate among unrelated individuals (t_u) was 59.0% and 46.9% in Ibicatu and MG, respectively and the
 213 mating among relatives rate (t_r) was 35.7% and 44.0% in Ibicatu and MG, respectively. Pollen dispersal distances
 214 (D) were significantly longer for mating resulting in outcrossed seedlings (t_u) than for those produced by mating
 215 between relatives (t_r). These results also show that the survival of seedlings originating from t_u was higher than from
 216 t_r and s .

217

218 **Genetic diversity and inbreeding**

219 Outcrossed seedlings (t_u) of the Ibicatu population exhibited significantly greater mean allelic richness per locus (R)
 220 and observed heterozygosity (H_o) than seedlings from mating among relatives (t_r) and selfing (s , Table 1).

221 Additionally, seedlings that originated from mating among relatives exhibited significantly greater R and H_o than
222 selfed seedlings. In the MG population, outcrossed seedlings exhibited significantly greater R than seedlings
223 originating from selfing and significantly greater H_o than seedlings from mating among relatives and selfing.
224 Seedlings that originated from mating among relatives showed significantly greater H_o than selfed seedlings. In both
225 populations, the fixation index (F) was lower for outcrossed seedlings than for seedlings from mating between related
226 individuals and selfed seedlings. The F values were also lower for outcrossed seedlings from related individuals than
227 selfed seedlings. Thus, the genetic diversity indices were always higher and inbreeding lower in seedlings originating
228 from mating among unrelated parents than for seedlings from mating between related individuals or selfing.

229

230 **Survival and growth rate at 8 months**

231 Total survival at 8 months after planting of germinated seedlings was high (479 of 553 = 86.6%), being lower in Ibicatu
232 (81.7%) than in MG (97.1%) population (Table 2). The ANOVA detected significant differences (F test) among classes
233 (t_u , t_r and s) for all traits in Ibicatu. Survival of outcrossed seedlings from Ibicatu (93.7%) was significantly higher
234 than seedlings originating from mating among related individuals (65.2%) and selfing (60.0%). Outcrossed seedlings
235 from Ibicatu were significantly taller (16.5 cm) than those from mating between related individuals (14.4 cm).

236

237 **Inbreeding depression**

238 We observed greater ID in the selfed seedlings in terms of SUR in both populations and H in MG population than in the
239 seedlings produced by mating among relatives, but greater ID in the seedlings produced by mating among relatives than
240 produce from selfing for H in Ibicatu and RCD in both populations (Table 3). Comparisons of the populations revealed
241 that the total ID in terms of SUR and H was greater in Ibicatu than MG. However, the MG population exhibited the
242 greatest ID in terms of RCD.

243

244 **Discussion**

245 To our knowledge this is the second study of ID in a tree species based on results from paternity analyses and estimates
246 of coancestry between seed trees and assigned pollen donors. Ismail et al. (2014) used this method to investigate ID in
247 the tree *Viteria indica*. In general, ID studies are based on hand pollination, selfing trees or performing matings between
248 individuals with different levels of relatedness and comparing the performance for heritable traits between inbred and
249 non-inbred individuals established in provenance and progeny tests. By contrast, we reconstructed the pedigrees of
250 open-pollinated seeds of *C. legalis* collected in three forest fragments and used a nursery provenance/progeny test to
251 determine which seeds were produced by outcrossing among unrelated trees, selfing and mating between relatives. The

approach was efficient in showing that ID in *C. legalis* varied between populations at an early stage of development (8 months of age) in terms of SUR and H.

However, we note that our estimates of mating among relatives may be biased because pairwise relatedness estimates using genetic markers are not straight forward, as some individuals may be related but heterozygous for different alleles at a locus (A_1A_2 and A_3A_4). Assuming that the frequency of each of four alleles (A_1 , A_2 , A_3 and A_4) in parental (F0) population is 0.25, using Loiselle et al. (1995) method for estimating the coancestry coefficient for pairwise genotypes A_1A_2 and A_3A_4 (F1) with no IBD alleles will result in a coancestry coefficient of -0.33. Consequently, the pairwise estimate will be biased, in comparison to the expected coancestry between full-sibs (0.25). However, this represents an extreme case for one locus, where related parents (F1) present no IBD alleles in a locus. In other loci, both parents may present the same alleles, producing estimates of pairwise relatedness close to the true relatedness. To discriminate unrelated individuals from half- and full-sibs, requires about 20 polymorphic loci (Blouin et al. 1996).

Realized mating system

We detected the pollen parent for all seedlings established in the provenance/progeny test. However, as we assigned seedling pollen donors using an 80% confidence level in the paternity analysis, our results may include some false positive pollen donor assignments (Ismail et al. 2014). At the same confidence level, not all seedlings were assigned in fathers in a pollen flow study based on all seeds sampled from these populations, with 8 and 4% of pollen flow detected into the Ibicatu and MG populations, respectively (Tambarussi et al. 2015).

We detected higher outcrossing between unrelated individuals in both populations (minimum t_u of 47%) relative to the crossing rate between relatives (minimum t_r of 36%) and selfing (minimum s of 5%), which gave total outcrossing rates ($t_t = t_u + t_r$) of 95% in Ibicatu and 91% in MG. Thus, the majority of seedlings were produced by outcrossing. These results confirm those of a previous study in three populations of the same species, which revealed that reproduction occurs predominantly by outcrossing, but a portion of mating occurs among related individuals due to intra population spatial genetic structure, with a small amount of selfing (Sebbenn et al. 2000; Tambarussi et al. 2015). However, as our mating system results were estimated using seedlings in a nursery provenance and progeny test, the results represent the realized mating system and probably not that which occurred during fertilization due to the occurrence of ID and the resultant mortality of some inbred individuals between fertilization and the nursery stage. Decreases in outcrossing rates due to ID have been reported for the tropical trees *Platypodium elegans* (Hufford and Hamrick 2003) and *Neobalanocarpus heimii* (Naito et al. 2005). Our study detected greater inbreeding due to mating between relatives than from selfing (Table 1). Levels of inbreeding depression varied between seedlings originating from outcrossing among unrelated individuals (t_r) and selfing (s), among traits and among populations.

285

286 **Which study trait was most strongly affected by ID?**

287 Our results revealed differences between traits in ID levels due to mating among relatives (ID_r) and selfing (ID_s)
288 (Table 3). Generally, SUR showed the greatest ID (ranging from 0.2 to 36%), and RCD the lowest (ranging from -3.2 to
289 5.5%), with the ID for SUR influenced by the trait's binomial distribution (presence or absence) rather than a
290 continuous distribution. Differences in levels of ID between heritable quantitative traits (H and RCD) may be due to
291 dominance and over-dominance effects, with the highest dominance or over-dominance deviations showing the greatest
292 ID (Falconer and Mackay 1997). Variation in ID between traits has also been reported in other tree studies. For
293 example, height, diameter at breast height (dbh) and volume of outcrossed *Eucalyptus regnans* progeny at 45 months of
294 age were on average 11, 18 and 37% greater, respectively, than values for selfed progeny (Griffin and Cotterill 1988).

295

296 **Are there differences between populations in ID?**

297 Inbreeding depression was highest for SUR in the Ibicatu progeny resulting from selfing and mating among relatives,
298 and for H in the Ibicatu progeny from mating among relatives. The MG population exhibited the greatest ID due to
299 selfing in terms of H and in RCD due to mating between relatives. The greater ID in Ibicatu was probably due to a
300 higher frequency of deleterious alleles or genetic load. For example, Karkkainen et al. (1996) studied geographic
301 variation in ID of *Pinus sylvestris* and observed variation between regions in the mean rate of seed abortion and the
302 number of lethal embryos among progeny derived from selfing, suggesting that different populations may have different
303 genetic load levels.

304

305 **What is the difference in ID between individuals originating from selfing and mating among relatives?**

306 The effects of ID due to selfing and mating among relatives were variable across the studied traits with differences
307 between H and RCD compared to SUR partly due to the first two being continuous variables, whereas SUR presents a
308 binomial distribution. For SUR, ID was greater among selfed seedlings (ID_s) than in seedlings originating from mating
309 between related individuals (ID_r). By contrast, for H and RCD, selfed individuals exhibited lower ID than those
310 resulting from mating between relatives, with the exception of H in the MG population, which showed the opposite
311 pattern. Selfing produced the minimum of 50% homozygosity for IBD alleles in heterozygous loci of the seed trees, and
312 mating between relatives produced inbreeding equal to the coancestry among the parents. Thus, selfing is expected to
313 produce greater ID than mating between related individuals, as was observed for SUR in both populations and H in the
314 MG population. The contrasting results for H in the Ibicatu population and RCD in both populations can be linked to
315 the greater ID observed for SUR. As SUR was lower among the selfed seedlings than those resulting from mating
316 between relatives, it is probable that only selfed individuals with low genetic loads, and consequently low ID due to

317 selfing, survived to be measured for H and RCD eight months after germination. Thus, ID for these traits was higher
318 among seedlings resulting from mating between relatives, possibly because these seedlings exhibited greater survival
319 than those originating from selfing, due to the low combination likelihood of homozygous IBD alleles. Furthermore,
320 some embryo mortality may occur before germination, and our estimate represents a rate of realized selfing and mating
321 among relatives, not the rate at fertilization. However, our results do represent an early effect of ID. Inbreeding
322 depression may act at various stages of seedling development, and it is possible that ID that results from both selfing
323 and mating among relatives will continue to act on these traits. For example, Sorensen and Miles (1982) found that ID
324 for survival in *Pseudotsuga menziesii*, *Pinus ponderosa* and *Abies procera*, was greatest within the first two years of
325 planting (range of 3 to 16%), whereas these values were 0.4 to 3% ten years after planting.

326
327 **Are there differences in genetic diversity between seeds produced by selfing, mating between relatives and**
328 **outcrossing?**

329 Our results revealed clear associations of mating system indices (t_u , t_r and s), survival, allelic richness (R),
330 observed heterozygosity (H_o) and fixation index (F). Generally, in both populations, the greater survival of
331 seedlings resulting from outcrossing among unrelated individuals gave greater R and H_o and lower F compared to
332 seedlings from mating among related trees, and from selfing. As expected, we also detected that seedlings from mating
333 between relatives exhibited significantly higher R and H_o and lower F than those from selfing (Table 1). As
334 already mentioned, selfing and mating between relatives result in increases in homozygosity in IBD deleterious alleles,
335 and decreases in genetic diversity relative to outcrossing between unrelated trees. Furthermore, selfing increases IBD to
336 a level greater than that resulting from mating between relatives, as only maternal IBD alleles are present in selfed
337 individuals.

338
339 **What are the expected levels of ID in offspring generations based on random seedling samples from these**
340 **populations?**

341 The probabilities of observing ID in seedlings from Ibicatu (ID_T) were higher than in the MG seedlings for SUR, with
342 the reverse pattern for RCD. Rates of selfing and mating between relatives, inbreeding (estimated from the fixation
343 index), and mean homozygosity for seedlings originating from selfing were lower in Ibicatu than in MG, leading to the
344 expectation that ID would be greater in the MG population. As this expectation was not met, a lower genetic load is the
345 probable explanation due to different frequencies of deleterious alleles for these traits in each population.. As such it is
346 likely that other fragmented populations of *C. legalis* will also show ID, but may well vary in ID levels for different
347 variables.

348 This study shows the effectiveness of combining paternity analysis and relatedness estimates with a progeny trial for
349 studying ID, as part of a broader study on gene flow and mating system, and without the complication and cost of
350 conducting controlled pollinations. The approach is likely to be most effective in tree species without a self-
351 incompatibility mechanism, where human disturbance (e.g. fragmentation, logging) produces sufficiently large
352 increases in inbreeding (selfing and mating among related individuals) to produce adequate numbers of inbred seedlings
353 to compare in field trials. Our results show the importance of studying ID over a number of populations and traits, while
354 it is also necessary to study a population over many years to understand the temporal dynamics of ID.

355

356 Implications

357 The detected levels of ID from selfing and mating between related individuals in *C. legalis* populations have practical
358 implications for genetic conservation, breeding and environmental restoration in terms of their impacts on survival and
359 growth. The collection of open-pollinated seed from trees in fragmented populations is common practice and it is
360 impossible to stop inbreeding in open-pollinated seeds as the species is self-compatible and intra-population spatial
361 genetic structure also leads to some bi-parental inbreeding (Tambarussi et al. 2015). However, the following measures
362 can be applied to minimize the use and establishment of inbred seedlings and ensure genetic diversity in *ex situ* genetic
363 conservation, breeding and environmental restoration. 1) *Keep seedlings identified by mother tree in the nursery* – it
364 will become evident which families show higher mortality and thus higher levels of inbreeding. This will also allow
365 equal numbers of seedlings from each seed tree to be used in planting, ensuring both genetic balance and diversity. 2) *In*
366 *the nursery, eliminate seedlings with indications of inbreeding* i.e. those with diseases, chlorosis, stunted growth or
367 other abnormalities. This will ensure that survival and growth in the field will not be greatly constrained by inbreeding
368 impacts. 3) *Collect from a large number of mother trees across a range of fragments* - this will ensure that sufficient
369 genetic diversity is maintained even while eliminating inbred material. These strategies decrease the likelihood of
370 inbred seedlings becoming established in *ex situ* gene banks, breeding programmes and at restoration sites, thus
371 ensuring that the fitness of established seedling populations is higher than that of the seed population and increasing the
372 evolutionary adaptive potential of such programmes.

373

374 Acknowledgements

375 We thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - 2010/10704-7) and Conselho
376 Nacional de Desenvolvimento Científico e Tecnológico (CNPq - 470491/2010-8) for funding this project. Evandro V.
377 Tambarussi was supported by a FAPESP PhD scholarship (grant 2010/12354-3). Alexandre M. Sebbenn, Miguel L.M.
378 Freitas and Roland Vencovsky are supported by CNPq research fellowships. We thank Tiago Gabassi for help in the
379 laboratory and Wladimir Correa and Dirceu de Souza for their help in collecting the plant samples. We also thank
380 NPGLE for corrections to the English.

381

382 **References**

- 383 Bittencourt JM, Sebbenn AM (2007) Patterns of pollen and seed dispersal in a small fragmented population of a wind
384 pollinated *Araucaria angustifolia* in southern Brazil. *Heredity* 99: 580–591.
- 385 Blouin MS, Parsons M, Lacaille V, Lotz S (1996) Use of microsatellite loci to classify individuals by relatedness. *Mol*
386 *Ecol* 5: 393–401.
- 387 Byers DL, Waller DM (1999) Do plant populations purge their genetic load? Effects of population size and mating
388 history on inbreeding depression. *Annual Review of Ecology and Systematics* 30: 479–513.
- 389 Carvalho PER (2003) Espécies arbóreas brasileiras. Brasília: Embrapa Informação Tecnológica.
- 390 Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. *Ann Rev Ecol Syst*
391 18: 237–268.
- 392 Chaves LJ, Vencovsky R, Silva RSM, Telles MPC, Zucchi MI, Coelho ASG (2011) Estimating inbreeding depression
393 in natural plant populations using quantitative and molecular data. *Conserv Genet* 12: 569–576.
- 394 Degen B, Sebbenn AM (2014) Genetics and tropical forests. In: Pancel L, Köhl M (eds) *Tropical Forestry Handbook*,
395 2nd edn. Springer, Berlin, pp 885–920.
- 396 Dick CW, Hardy OJ, Jones FA, Petit RJ (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain
397 forest trees. *Tropical Plant Biol* 1: 20–33.
- 398 Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.
- 399 Falconer DS, Mackay TFC (1997) *Introduction to Quantitative Genetics*. Longman Group Ltd, England.
- 400 Goudet J (1995) Fstat (Version 2.9.3.2.): A computer program to calculate F-statistics. *J Hered* 86: 485–486.
- 401 Griffin AR, Cotterill PP (1988) Genetic variation in growth of outcrossed, selfed and open-pollinated progenies of
402 *Eucalyptus regnans* and some implications for breeding strategy. *Silvae Genet* 37: 124–131.
- 403 Hardy O, Vekemans X (2002) SPAGeDI: a versatile computer program to analyze spatial genetic structure at the
404 individual or population levels. *Mol Ecol Notes* 2: 618–620.
- 405 Hufford KM, Hamrick JL (2003) Viability selection at three early life stages of the tropical tree, *Platypodium elegans*
406 (Fabaceae, Papilionoideae). *Evolution* 57: 518–526.
- 407 Ismail SA, Ghazoul J, Ravikanth G, Uma Shaanker R, Kushalappa CG, Kettle CJ (2012) Does long-distance pollen
408 dispersal preclude inbreeding in tropical trees? Fragmentation genetics of *Dysoxylum malabaricum* in an agro-
409 forest landscape. *Mol Ecol* 21: 5484–5496.
- 410 Ismail SA, Ghazoul J, Ravikanth G, Uma Shaanker R, Kushalappa CG, Kettle CJ (2014) Fragmentation genetics of
411 *Vateria indica*: implications for management of forest genetic resources of an endemic dipterocarp. *Conserv*
412 *Genet* 15: 533–545.

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

415 Kärkkäinen K, Savolainen O (1996) The degree of early inbreeding depression determines the selfing rate at the seed
 416 stage: Model and results from *Pinus sylvestris* (Scots pine). *Heredity* 71: 160–166.

417 Koelewijn HP, Koski V, Savolainen O (1999) Magnitude and timing of inbreeding depression in Scots Pine (*Pinus*
 418 *sylvestris* L.). *Evolution* 53: 758–768.

419 Lindgren D, Luigi DG, Jefferson PA (1997) Status number for measuring genetic diversity. *Forest Genet* 4: 69–76.

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

422 Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005) Genetic resource impacts of habitat loss and degradation;
 423 reconciling empirical evidence and predicted theory for neotropical trees. *Heredity* 95: 255–273.

424 Manoel RO, Alves PF, Dourado CL, Gaino APSC, Freitas MLM, Moraes MLT, Sebbenn AM (2012) Contemporary
 425 pollen flow, mating patterns and effective population size inferred from paternity analysis in a small fragmented
 426 population of the Neotropical tree *Copaifera langsdorffii* Desf. (Leguminosae-Caesalpinioideae). *Conserv*
 427 *Genet* 13: 613–623.

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

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

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

434 Prance GT, Mori SA (1979) Lecythidaceae – Part 1. *Flora Neotropica Monograph*. 21: 1–272.

435 Ritland K (1996) Inferring the genetic basis of inbreeding depression in plants. *Genome* 39: 1–8.

436 Rymer PD, Sandiford M, Harris SA, Billingham MR, Boshier DH (2015) Remnant *Pachira quinata* pasture trees have
 437 greater opportunities to self and suffer reduced reproductive success due to inbreeding depression. *Heredity* 115:
 438 115–124.

439 Sebbenn AM, Kageyama PY, Siqueira ACMF, Zanatto ACS (2000) Sistema de cruzamento em populações de
 440 *Cariniana legalis* Mar. O. Ktze.: implicações para a conservação e o melhoramento genético. *Sci For* 58: 24–40.

441 SAS (1999) Institute Inc. SAS procedures guide. Version 8 (TSMO). SAS, Cary.

442 Sokal RR, Rohlf FJ (1995) Biometry: principles and practices of statistics in biological research. Third Edition. W.H.
 443 Freeman and Company, New York.

444 Sorensen FC, Miles R (1982) Inbreeding depression in height, height growth and survival of Douglas-Fr, Ponderosa
 445 Pine, and Noble Fir to 10 years of age. *For Sci* 28: 283–292.

446 Tambarussi EV, Sebbenn AM, Moreno MA, Ferraz EM, Kageyama PY, Vencovsky R (2013a) Microsatellite
447 markers for *Cariniana legalis* (Lecythidaceae) and their transferability to *Cariniana estrellensis*. Appl Plant Sci
448 1: 1200493.

449 Tambarussi EV, Vencovsky R, Freitas MLM, Sebbenn AM (2013b) Mendelian inheritance, genetic linkage and
450 genotypic disequilibrium at nine microsatellite loci of *Cariniana legalis* (Mart.) O. Kuntze. Genet Mol Res 12:
451 5442–5457.

452 Tambarussi EV, Boshier DH, Vencovsky R, Freitas MLM, Di-Dio O-J, Ferraz EM, Moreno MA, Sebbenn AM (2015)
453 Paternity analysis reveals significant isolation and near neighbour pollen dispersal in small *Cariniana legalis*
454 Mart. Kuntze populations in the Brazilian Atlantic Forest. Ecology and Evolution 5: 5588–5600.

455 Winn AA, Elle E, Kalisz S, Cheptou PO, Eckert CG, Goodwillie C, Johnston MO, Moeller DA, Ree RH, Sargent RD,
456 Vallejo-Marín M (2011) Analysis of inbreeding depression in mixed-mating plants provides evidence for
457 selective interference and stable mixed mating. Evolution 65: 3339–3359.

458 Wu HX, Matheson AC, Spencer D (1998) Inbreeding in *Pinus radiata*. 1. The effect of inbreeding on growth, survival
459 and variance. Theor Appl Genet 97: 1256–1268.

460

461

462

463 **Table 1** Results of the mating system, mean coancestry coefficient, pollen dispersal distance between parents, genetic
464 diversity and inbreeding of the Ibicatu and Mogi-Guaçu (MG) populations of *Cariniana legalis*. t_u , t_r and s are the
465 rate of outcrossing among unrelated individuals, outcrossing among related individuals, and selfing, respectively; n_g
466 and n_a are the number of germinated and assigned seedlings for pollen donors, respectively; 95% CI is the 95%
467 confidence interval; min-max is the minimum and maximum values. The different letters indicate significant differences
468 at the 5% level by the unpaired t-test; * $P < 0.05$.

	t_u	t_r	s
Ibicatu: $n_g = 378$			
Assigned: n_a [n_a / n_g 100%]	223 [59.0%]	135 [35.7%]	20 [5.3%]
Coancestry between parents: $\bar{\theta}$ (95% CI)	0.08 (0.07/0.09)*	0.22 (0.20/0.23)*	-
Distance between parents: D (m) (min/max)	378 (322/434)	268 (176/360)	-
Total number of alleles: k	99	80	50
Allelic richness for five individuals: R	6.1A	5.3B	4.3C
Observed heterozygosity: H_o	0.85A	0.76B	0.52C
Fixation index: F	0.02A	0.06A*	0.26B*
Mogi-Guaçu: $n_g = 175$			
Assigned: n_a [n_a / n_g 100%]	82 [46.9%]	77 [44.0%]	16 [9.1%]
Coancestry between parents: $\bar{\theta}$ (95% CI)	-0.02 (-0.03/-0.01)*	0.12 (0.12/0.14)*	-
Distance between parents: D (m) (min/max)	144 (1/288)	28 (12/44)	-
Total number of alleles: k	83	86	38
Allelic richness for five individuals: R	5.8A	5.8A	4.4B
Observed heterozygosity: H_o	0.91A	0.73B	0.34C
Fixation index: F	-0.08A*	0.14B*	0.58C*

469

470

471

472 **Table 2** Survival (SUR), total height (H) and root-collar diameter (RCD) at 8 months of the Ibicatu and Mogi-Guaçu
473 (MG) populations of *Cariniana legalis*. t_u , t_r and s are the rate of outcrossing among unrelated individuals,
474 outcrossing among related individuals, and selfing, respectively; n_a and $n_{survival}$ are the number of germinated seedling
475 and survival seedlings at 8 months, respectively. The different letters indicate significant differences at the 5% level by
476 Tukey test. P_{value} is the probability of F value of the ANOVA be significant among trait for groups (t_u , t_r and
477 s).

	n_g	$n_{survival}$	SUR (%)	H (cm)	RCD (cm)
Ibicatu					
Assigned for t_u	223	209	93.7A	16.49A	3.17A
Assigned for t_r	135	88	65.2B	14.40B	3.09A
Assigned for s	20	12	60.0B	16.46A	3.28A
Mean			81.7	15.93	3.15
P_{value}			0.001	0.004	0.001
Assigned for $t_r + s$	155	100	65.0	14.76	3.11
Total	378	309			
MG					
Assigned for t_u	82	80	97.6A	10.76A	2.30A
Assigned for t_r	77	75	97.4A	10.66A	2.18A
Assigned for s	16	15	93.7A	10.41A	2.38A
Mean			97.1	10.69	2.25
P_{value}			0.438	0.886	0.697
Assigned for $t_r + s$	93	90	96.8	10.62	2.21
Total	175	170			

478
479

480
 481 **Table 3** Inbreeding depression (ID) in percent (%) due to mating between relatives (ID_r), selfing (ID_s) and total
 482 (ID_T) for survival (SUR), height (H), and root-collar diameter (RCD) in the Ibicatu and Mogi-Guaçu (MG) populations
 483 of *Cariniana legalis*.

	Ibicatu			MG		
	SUR	H	RCD	SUR	H	RCD
ID due to mating between relatives: ID_r	30.4%	12.8%	3.9%	0.2%	1.0%	5.5%
ID due to selfing: ID_s	36.0%	0.0%	-3.2%	3.9%	3.3%	-3.2%
Total ID: ID_T	31.2%	11.5%	3.1%	0.8%	1.3%	4.1%

484
 485
 486