


# Estimating genetic diversity, mating system and pollen dispersal to inform *ex situ* conservation of the tree *Genipa americana* L.

Marília Freitas de Vasconcelos Melo<sup>1</sup>, Alexandre Magno Sebbenn<sup>2</sup>,  
Bruno Cesar Rossini<sup>3\*</sup> , Ana Veruska Cruz da Silva Muniz<sup>4</sup>, Carlos Jose Rodrigues<sup>5</sup>,  
Celso Luis Marino<sup>3</sup> and Mario Luiz Teixeira de Moraes<sup>6</sup>

<sup>1</sup>Universidade Federal do Alagoas, Campus de Engenharias e Ciências Agrárias, CEP 57100-000, Rio Largo, AL, Brazil, <sup>2</sup>Instituto Florestal de São Paulo, C.P. 1322, CEP 01059-970, São Paulo, SP, Brazil, <sup>3</sup>Universidade Estadual Paulista, Instituto de Biotecnologia, UNESP, CEP 18607-440, Botucatu, SP, Brazil, <sup>4</sup>Embrapa Tabuleiros Costeiros, CEP 49025-040, Aracaju, SE, Brazil, <sup>5</sup>Cia Energética de São Paulo, Usina Porto Primavera, CEP 19274-000, Primavera, SP, Brazil and <sup>6</sup>Universidade Estadual Paulista, Faculdade de Engenharia, UNESP, CEP 15385-000, Ilha Solteira, SP, Brazil

Received 20 July 2020, Revised 4 January 2021; Accepted 26 January 2021 – First published online 1 March 2021

## Abstract

Using microsatellite loci, we assessed the mating system and genetic diversity of the dioecious tropical tree *Genipa americana* in a natural population (NP) and a progeny test (PT). For NP, we also estimated the paternity correlation within and among fruits and mean pollen dispersal distance. As expected for dioecious species, all offspring originated from outcrossing ( $t = 1$ ). Mating among relatives ( $1 - t_s$ ) and paternity correlation ( $r_p$ ) were variable among progenies ( $1 - t_s = 0.03$ – $0.19$ ;  $r_p = 0.04$ – $0.40$ ), but greater in NP than in PT. Fixation index ( $F$ ) was generally significant and lower in adults than in offspring, indicating selection against inbred individuals. Paternity correlation was higher within (0.40) than among (0.26) fruits, indicating a lower effective number of pollen donors ( $N_{ep}$ ) within (2.5) than among (3.8) fruits. Due to the higher  $r_p$  in NP, the effective size within progenies ( $N_e$ ) was lower (2.69) than PT (3.27). The pollen dispersal pattern was strongly leptokurtic, suggesting long-distance pollen dispersal (mean of 179 m). The results show that both populations can be used for seed collection in environmental reforestation programmes; however, considering that PT is structured in maternal progenies, NP is more suitable for seed collection due to the lower probability of mating among related trees.

**Keywords:** conservation genetics, gene flow, microsatellite loci, progeny test, tropical trees

## Introduction

In Brazil, forests have been indiscriminately exploited, which threatens the continuation of biomes (Francoso *et al.*, 2016; Silva *et al.*, 2020). Plans for conserving Brazil's natural ecosystems for future generations are urgently needed along with strategies for seed collection for environmental reforestation. One example is the Brazilian Atlantic Forest, where the conservation of the

ecosystem is a high priority due to its vast species diversity, significant levels of endemism and the extensive fragmentation of the biome (Dinerstein *et al.*, 1995; Colombo and Joly, 2010; Silva *et al.*, 2020).

*Ex situ* conservation is a complementary strategy to *in situ* conservation, both of which help to avoid the extinction of species, populations or individuals (Degen and Sebbenn, 2014; Hausman *et al.*, 2014). This can be done using experimental field designs such as provenance and progeny tests (PTs), which also enable the estimation of genetic variability and heritability of phenotypic traits (Sebbenn, 2006). Such provenance and PTs can be

\*Corresponding author. E-mail: [bruno.rossini@unesp.br](mailto:bruno.rossini@unesp.br)

transformed into seed orchards by selecting provenances, progenies, individuals within progenies and seeds that can be used as sources for environmental reforestation (Hodgkin *et al.*, 2003; Sebbenn *et al.*, 2007; Bakkali *et al.*, 2013). Forest reforestation based on seedlings that originate from a large number of seed trees (>25) can also be considered an *ex situ* conservation strategy (Sebbenn, 2006).

The levels of genetic diversity retained in *ex situ* conservation trials and reforestation areas depend on the reference effective population size ( $N_e$ ) of the population used to establish these areas (Sebbenn, 2006). The effective population size values for species with overlapping generations, such as trees, range from 70 to 150 (Nunney and Campbell, 1993; Caballero *et al.*, 2016), and values ranging from 1000 to 5000 (Nunney and Campbell, 1993; Lande, 1995) have been suggested for short- and long-term conservation purposes, respectively. The sampled  $N_e$  may change between seed germination (effective sample) and adult stages (realized sample) due to the effects of deterministic process, such as inbreeding and outbreeding depression, and stochastic process, including random mortality, disease and pest attacks, which can decrease the levels of inbreeding and change the levels of relatedness among progenies (Sebbenn, 2006; Feres *et al.*, 2012). Thus, understanding the mating pattern of the genetic material of source populations is critical to determine the number of seed trees needed to guarantee the reference effective population size indicated for *ex situ* conservation in trials and reforested areas (Sebbenn, 2006; Degen and Sebbenn, 2014).

Among the tree species that occur in the Atlantic Forest, the dioecious tree *Genipa americana* L. (Rubiaceae) is both ecologically and economically valuable: its wood is used in construction, and for furniture, tool handles and general carpentry; its bark is used for treating hides in tanneries as it is rich in tannins and the edible fruits also provide a bluish-green dye that is used for a variety of purposes. The species occurs in several forest formations throughout tropical America, from Mexico to Brazil (Carvalho, 1994), suggesting high-genetic diversity within and among populations. Population density is variable among sites (<1–25 trees/ha) and the trees have been used for reforestation in swampy and degraded areas (Durigan and Nogueira, 1990; Manoel *et al.*, 2017). It also commonly occurs in riparian forests. As adults, the tree can reach 25 m in height and 90 cm in diameter at breast height (DBH) (Carvalho, 1994). The dispersal of its fruits occurs by gravity, zoochory and hydrochory (Carvalho, 1994). The species is insect pollinated, particularly by bumblebees such as *Bombus morio* and *Epicharis rustica* var. *flava* (Crestana *et al.*, 2004). However, despite the ecological and economic importance of *G. americana*, there are few studies on its mating system and gene flow, and such studies are crucial for the effective conservation of the species (Manoel *et al.*, 2015a, 2017).

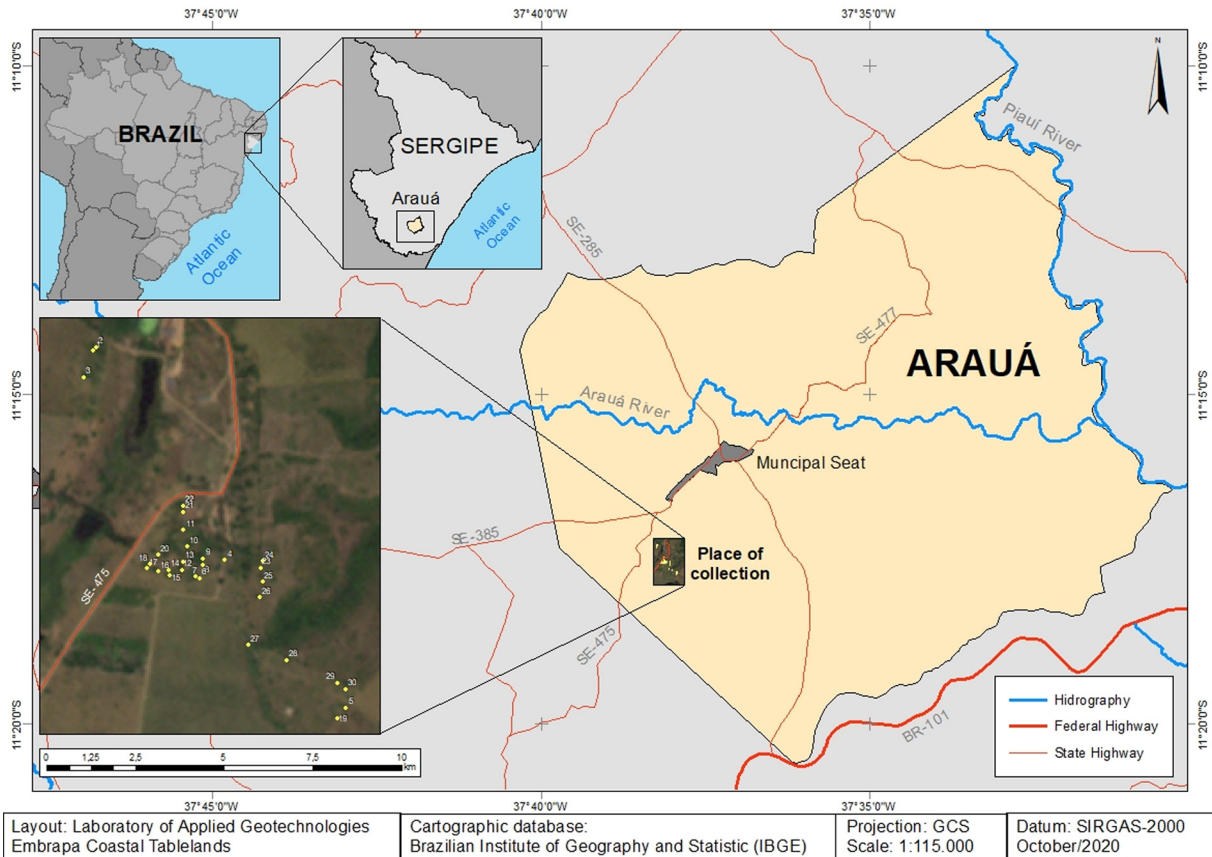
To investigate the differences between the effective and realized samples from natural populations (NPs) for *ex situ* conservation, we compared the mating system, genetic diversity, biparental inbreeding and effective size between two *G. americana* PTs, one with open-pollinated offspring at 6 months of age established in a nursery PT (effective mating system), and the other with open-pollinated trees at 14 years of age established in field PT (realized mating system). We specifically sought to answer the following: (i) Are there biparental inbred and full-sib individuals in the trials due to mating among relatives and correlated mating, respectively? (ii) Is the paternity correlation higher within than among fruits? (iii) What is the mean pollen dispersal distance? (iv) Is the effective population size different between effective and realized mating samples? (v) Is the effective population size of the trials sufficient for *ex situ* conservation of the populations?

## Materials and methods

### Study site and sampling

The study was carried out in two PTs originated from open-pollinated seeds sampled from two populations (distance >1900 km) in Brazil (Fig. 1). The first population is a forest fragment (21 ha) that is a remnant of a NP in the municipality of Arauá, Sergipe State, Brazil (11°15'30"S, 37°36'55"W, elevation of 73 m). The population occurs in a transition zone between the Atlantic Forest and Caatinga biomes. The climate is megathermal and sub-humid with an average annual temperature of 24.6°C and mean annual precipitation of 863.3 mm (Sergipe, 2000). In 2014, we sampled open-pollinated seeds from 30 seed trees, with 28 seeds per tree (four fruits, eight seeds per fruit) for a total of 960 seeds. Mother trees were georeferenced using GPS-III (Garmin, USA), measured for DBH, and leaves were sampled for DNA analysis. Identification of seeds, fruits and seed trees were maintained for hierarchical analysis of the mating system within and among fruits. Seeds were germinated in a nursery PT, established in a randomized complete block experimental design, with 30 open-pollinated progenies, four blocks and seven plants per plot, for a total of 840 individuals. We sampled leaves from all 488 surviving individuals (58.1%) for DNA analysis, for a total of 30 progenies, ranging from 4 to 20 individuals per progeny, from 2 to 4 fruits per tree and 1 to 8 seeds per fruit.

The second population is a PT established in 2000 as an Active Germplasm Bank of the Restoration and Ecosystem Conservation Division of CESP (Companhia Energética de São Paulo), in the municipality of Rosana, São Paulo State, Brazil (22°30'38.84"S, 52°57'23.20"W, elevation of 306 m). The PT was established to offset the impacts of the construction of the Sérgio Motta Hydroelectric Power Facility



**Fig. 1.** Geographic location of *G. americana* seed trees in the municipality of Arauá, Sergipe State, Brazil.

in the municipality of Porto Primavera, located in the Pontal do Paranapanema region. Seeds to establish the PT were collected from 30 seed trees located in the currently flooded reservoir of the dam, as such the population no longer exists. The experimental design to establish the PT was a randomized complete block, with 30 open-pollinated progenies, three blocks and eight plants per plot, using a  $3 \times 2$  m spacing, for a total of 720 individuals. The climate is megathermal and sub-humid with an average annual temperature of 21.9°C and mean annual precipitation of 1200 mm (Francisco, 1989). According to the Brazilian System of Soil Conservation, the soil of the site is an oxisol with a moderate, medium texture (Oliveira 1999), and the relief is relatively flat to undulating. For DNA analysis, we sampled leaves in 2014 from all 584 surviving trees in PT (81.1%), for a total sample of 30 progenies, ranging from 15 to 24 individuals per progeny.

### Microsatellite genotyping

DNA extraction was based on the method described by Doyle and Doyle (1990). We used seven species-specific primers developed for *G. americana* that present Mendelian segregation and are not genetically linked

(Manoel *et al.*, 2015b): Gam01, Gam06, Gam08, Gam11, Gam24, Gam28 and Gam41 (Manoel *et al.*, 2014). Genotyping was conducted on an ABI3130xl (Applied Biosystems, USA) using the protocol described by Schuelke (2000) with the addition of a M13 tail in the forward primer of each locus. Alleles were read in GeneMapper 5.0 software (Applied Biosystems, USA).

### Analysis of genetic diversity and structure

Genetic diversity for NP seed trees and offspring of both populations was estimated as a mean across loci for the following indices: total number of alleles per locus ( $K$ ); allelic richness ( $R$ ) and observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity. To investigate biparental inbreeding in NP seed trees, the fixation index ( $F$ ) was used and its statistical significance was tested using a Monte Carlo permutation of alleles among individuals, in FSTAT 2.9.3.2 software (Goudet, 1995). For offspring,  $F$  values estimated for open-pollinated progenies may be biased due to overestimates in the gene frequencies of maternal alleles, as each plant within the progeny received at least one maternal allele. Thus, for progenies we calculated the  $F$  values as described in Manoel *et al.* (2015b), which uses the pollen pool

frequency as the parental gene frequencies. The frequency of null alleles (*Null*) was estimated for adults and offspring of each population based on the inbreeding population model (IIM) in INEST 2.0 software (Chybicki and Burczyk, 2009). To test if these indices were significantly different between cohorts (seed trees and offspring), a Jackknife among loci test was used. As the trials were structured in progenies, with gene frequency likely affected by maternal alleles within progenies, we used the offspring's pollen allele pool frequencies to determine the genetic differentiation ( $G'_{st}$ ) between populations. This method may better represent the genetic diversity of the source populations by capturing a larger number of alleles. To do so, the method developed for microsatellite loci by Hedrick (2005) was used, and the pollen allele pool frequency was calculated for offspring of each population using MLTR 3.4 software (Ritland, 2002). From this, we estimated the expected heterozygosity for each locus ( $H_e = 1 - \sum_{i=1}^k p_i^2$ , where  $p_i$  is the frequency of the  $i$ -th allele and  $k$  is the number of alleles at the locus under consideration) and mean heterozygosity over loci within populations ( $\overline{H_e}$ ). We also estimated the mean genetic diversity within populations ( $H_s = (n_1 \overline{H_{e1}} + n_2 \overline{H_{e2}})/(n_1 + n_2)$ , where  $n_1$  and  $n_2$  are the sample size in NP and PT populations, respectively, and  $\overline{H_{e1}}$  and  $\overline{H_{e2}}$  are the expected heterozygosity in NP and PT populations, respectively) and for all populations ( $H_T = 1 - \sum_{i=1}^k \overline{p_i}^2$ , where  $\overline{p_i}$  is the mean frequency of the  $i$ -th allele between populations and  $k$  is the number of alleles in the locus in consideration). The indices  $H_s$  and  $H_T$  were used to estimate the genetic differentiation between populations, ( $G_{st} = 1 - (H_s/H_T)$ ) and to calculate the standardized genetic differentiation ( $G'_{st} = G_{st}(1 + H_s)/(1 - H_s)$ ) (Hedrick, 2005).

### Pollen dispersal

The distance of pollen dispersal was estimated only for NP, since the geographical coordinates of seed trees are known. As only seed trees and open-pollinated offspring were genotyped, mean pollen dispersal distance was estimated using the TWOGENER model, as implemented in POLDIST software (Robledo-Arnuncio *et al.*, 2007). This model requires information about the effective reproductive population density and because we did not have this value, we used the Kindist module as it does not depend on prior knowledge of the effective density of pollen donors. The Kindist model is based on the expected decrease in pairwise normalized correlated paternity ( $\psi_{(z)}$ ) in open-pollinated progenies on the standard distance ( $Z$ ) between pairwise progenies of the sampled seed trees (Robledo-Arnuncio *et al.*, 2007). We also estimated the global pollen pool differentiation among seed trees ( $\phi_{IT}$ ). To estimate the dispersion kernel curve, the paternity correlation must decrease with an increase in spatial

distance between seed trees. POLDIST gives the mean pollen dispersal distance and parameters of the dispersion curve as a function of an assumed dispersion. Here, we used the exponential dispersion model, which is the most suitable for species that may have long-distance dispersal (Smouse and Sork, 2004), such as *G. americana* which is pollinated by bees. The estimated parameters were the scale of pollen dispersion ( $a$ ), the shape of the dispersion ( $b$ ) and the mean and standard deviation of pollen dispersion.

### Mating system analysis

Analysis of the mating system was based on the mixed mating model and correlated mating model, implemented in MLTR 3.4 software (Ritland, 2002). Analyses were estimated at the population and individual progeny levels, using the expectation maximization (EM) method for numerical resolution. The estimated indices were: ovule and pollen gene frequencies; fixation index of seed trees ( $F_m$ ); multilocus outcrossing rate ( $t_m$ ); single-locus outcrossing rate ( $t_s$ ); mating among related individuals ( $1 - t_s$ ) and multilocus paternity correlation ( $r_p$ ). As in NP, our sample was hierarchical within and among fruits. Thus, the multilocus paternity correlation was also estimated within ( $r_{pw}$ ) and among ( $r_{pa}$ ) fruits, excluding fruits with only one offspring, for a total sample of 470. The 95% confidence interval of the indices was obtained by 1000 bootstrap resampling, using progenies as the resampling unit for population level analysis and individuals within progenies for the progeny level analysis. We also estimated the effective number of pollen donors as  $N_{ep} = 1/r_p$ , the average within-progeny coancestry coefficient ( $\Theta$ ) as  $\Theta = 0.125(1 + F_m)(1 + r_p)$  (Sousa *et al.*, 2005), and the variance effective size within progenies ( $N_e$ ) as:

$$N_e = \frac{0.5}{\Theta((n-1)/n) + (1+F)/2n},$$

where  $n$  is the sample size within progenies and  $F$  is the mean inbreeding coefficient within progenies (Sebbenn, 2006). The number of seed trees for seed collection was calculated as  $m = N_{e(tr)}/N_e$ , assuming that the objective was to retain a reference effective size ( $N_{e(tr)}$ ) of 150 (Sebbenn, 2006) in the total sampled progeny arrays. The total variance effective size ( $N_{e(T)}$ ) in the PTs was estimated assuming absence of relatedness between individuals of different progenies by  $N_{e(T)} = \sum_{i=1}^m N_e$ . To determine if the sample size within progeny ( $n$ ),  $H_o$ ,  $F_o$ ,  $N_{ep}$ ,  $\Theta$  and  $N_e$  were associated, we used the Spearman's rank correlation coefficient ( $\rho$ ).

## Results

### Genetic diversity

For the total sample from both populations, the number of alleles per locus ranged from 8 to 23, for a total of 96 alleles



**Table 1.** Genetic diversity and inbreeding of *G. americana* seed trees and offspring in a 6 months of age NP and in an 14 years of age PT

	NP: seed trees (SD)	NP: offspring (SD)	PT: offspring (SD)
Sample size: $n$	30	488	584
Total number of alleles: $K$	46	81	77
Allelic richness for 30 genotypes: $R$	6.5 (2.5)a	7.5 (3.4)b	7.4 (1.4)b
Observed heterozygosity: $H_o$	0.84 (0.11)a	0.73 (0.07)b	0.82 (0.17)a
Expected heterozygosity: $H_e$	0.70 (0.11)a	0.77 (0.13)b	0.79 (0.05)b
Fixation index: $F$	-0.19 (0.17)*a	0.05 (0.02)b	-0.04 (0.05)b
Range of null allele frequency: <i>Null</i>	0–0	0–0.06	0–0.11

SD is the standard deviation; for offspring the  $F$  was estimated using the  $H_e$  calculated from gene frequencies in the pollen pool; \* $P < 0.05$ . Different letters indicate significant differences ( $P < 0.05$ ) between samples, based on a Jackknife test.

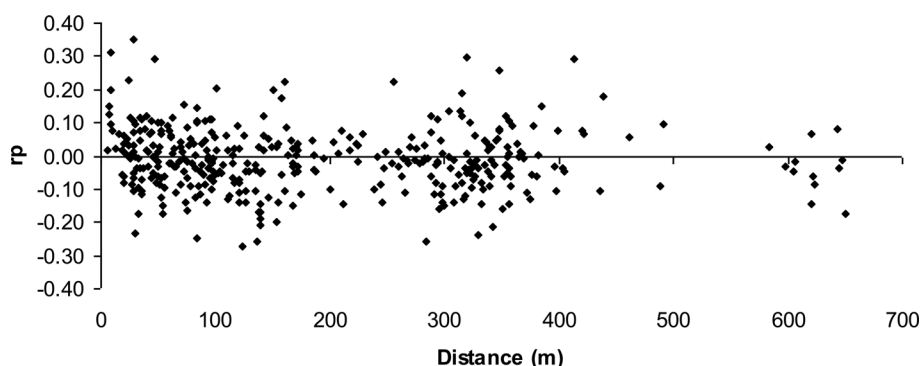
( $K$ ) and an average of 13.7. We detected 19 private alleles in NP and 15 in PT. The allelic richness ( $R$ ), expected heterozygosity ( $H_e$ ) and fixation index ( $F$ ) were significantly lower in NP seed trees ( $R=6.5$ ,  $H_e=0.70$ ,  $F=-0.19$ ) than in offspring of both populations ( $R=7.4-7.5$ ,  $H_e=0.77-0.79$ ,  $F=-0.04$  to 0.05) (Table 1). In contrast, the observed heterozygosity ( $H_o$ ) was significantly higher in NP seed trees ( $H_o=0.84$ ) and PT offspring ( $H_o=0.82$ ) than in NP offspring ( $H_o=0.73$ ). The fixation index ( $F$ ) was significant and lower than zero for the NP seed tree population, and not significantly different from zero for offspring of both populations (-0.04 to 0.05). These results suggest selection for heterozygous individuals among seed trees and an absence of biparental inbreeding in offspring. Null alleles (*Null*) were detected at low frequencies (maximum 0.11) in some loci among offspring, suggesting that null alleles did not significantly bias our estimates of mean  $F$  values. The genetic differentiation between PT and NP populations ( $G'_{st}$ ) was significantly greater than zero for the offspring pollen gene pool ( $0.52 \pm 0.08$ , mean  $\pm$  standard error), indicating that similar levels of genetic diversity (50%) was distributed within and between populations.

### Pollen dispersion

Spearman's rank correlation coefficient between pairwise correlated paternity and distance between seed trees in NP ( $\rho = -0.126$ ) was significantly negative ( $P=0.047$ ) and showed a downward trend with increasing distance between seed trees (Fig. 2 and online Supplementary Table S1). It was therefore possible to estimate the mean pollen dispersal distance in the population. Using the exponential power dispersion model, the scale of the pollen dispersion parameter ( $a$ ) was 89.5, the shape of the dispersion parameter ( $b$ ) was 0.415, and the mean pollen dispersal distance was  $179 \pm 18.9$  m ( $\pm$  standard deviation).

### Mating system

The outcrossing rate ( $t_m$ ) was not significantly different from unity (1) for the studied populations, indicating that all offspring originated from outcrossing, as expected for a dioecious species (Table 2). In NP, the rate of mating among relatives ( $1 - t_s = 0.15$ ) was significantly higher



**Fig. 2.** Distribution of pairwise normalized correlated paternity ( $r_p$ ) among *G. americana* seed trees as a function of the distance between seed trees from a NP in Arauá, Brazil.

**Table 2.** Mating system at the *G. americana* population level for a NP in Arauá and a PT in Rosana, Brazil

	NP (95% CI)	PT (95% CI)
Multilocus outcrossing rate: $t_m$	0.99 (0.98–1.0)	1.0 (1.0–1.0)
Mating among relatives: $1 - t_s$	0.15 (0.10–0.19)	0 (0–0.01)
Paternity correlation: $r_p$	0.34 (0.23–0.38)	0.08 (0.05–0.09)
Number of pollen donors: $N_{ep}$	2.9 (2.6–4.3)	13.0 (12.1–18.2)
Coancestry within family: $\Theta$	0.168 (0.15–0.17)	0.135 (0.13–0.135)
Effective size within family: $N_e$	2.69 (2.62–2.90)	3.27 (3.26–3.33)
Number of seed trees: $m$	56 (52–57)	46 (45–46)
Paternity correlation within fruits: $r_{pw}$	0.40 (0.32–0.47)	–
Paternity correlation among fruits: $r_{pa}$	0.26 (0.21–0.31)	–
Number of pollen donors within fruits: $N_{epw}$	2.5 (2.1–3.1)	–
Number of pollen donors among fruits: $N_{epa}$	3.8 (3.3–4.9)	–

95% CI is the 95% confidence interval.

than zero, indicating that some offspring present biparental inbreeding. The multilocus paternity correlation within and among fruits ( $r_p$ ) was significantly higher than zero in both populations, indicating that some offspring are full-sibs, particularly in the NP population. The effective number of pollen donors ( $N_{ep}$ ) was significantly higher in PT (13) than NP (2.9). Due to higher correlated mating in NP, the within progeny coancestry coefficient ( $\Theta$ ) was also higher in NP (0.168) than PT (0.135) and the effective size within progenies ( $N_e$ ) was lower in NP (2.69) than PT (3.27). Consequently, the number of seed trees ( $m$ ) required for seed collection was higher in NP (56) than PT (46). The total variance effective size ( $N_{eT}$ ) was higher in PT (98.1) than in NP (88.7). In NP, the paternity correlation within fruits ( $r_{pw} = 0.40$ ) was significantly higher than among fruits ( $r_{pa} = 0.26$ ), indicating that about 2.5 effective pollen donors fertilized the fruits ( $N_{epw}$ ) and 3.8 pollen donors fertilized each tree ( $N_{epa}$ ).

In general, the outcrossing rate ( $t_m$ ) estimated at the individual level was not significantly different from unity, ranging in NP from 0.93 to 0.99 and in PT from 0.96 to 0.99 (online Supplementary Tables S2 and S3). The Spearman’s rank correlation coefficient between the sample size ( $n$ ) of progenies and  $t_m$  was significant and higher than zero (0.678), indicating that lower  $t_m$  values are a product of progenies with small sample sizes. The lowest estimates (0.93) were detected in two NP progenies (16 and 28) with only four genotyped offspring, indicating that this result is an artefact of sample size. Furthermore, the other indices were also generally lower in these two progenies. All other estimated indices were variable among progenies ( $1 - t_s$ : NP = 0.03–0.20, PT = 0.04–0.10;  $r_p$ : NP = 0.06–0.40, PT = 0.04–0.11;  $N_{ep}$ : NP = 2.5–17.9, PT = 9.3–25.6;  $\Theta$ : NP = 0.132–0.235, PT = 0.130–0.139;  $N_e$ : NP = 1.73–3.29, PT = 3.12–3.40). In 22 progenies of NP

(73.3%) and 28 of PT (93.3%), the  $F_m$  was lower than  $F_o$ , suggesting selection against inbred individuals between seed/young trees and adult stages. In NP, for 28 progenies (93.3%) the  $r_{pw}$  was higher than  $r_{pa}$ , showing that the effective number of pollen donors was generally lower within than among fruits. The sample size within progenies ( $n$ ) was significantly negatively correlated with  $\Theta$  and positively correlated with  $N_e$  (Table 3). The  $N_{ep}$  was significantly positively correlated with  $H_o$  and  $N_e$  and significantly negatively correlated with  $F_o$  and  $\Theta$ . The  $H_o$  was significantly negatively correlated with  $F_o$  and  $N_e$ .

### Discussion

The indiscriminate exploitation of natural forest resources around the world has led to fragmentation, creating unbalanced ecosystems and the possible extinction of small populations. As such, understanding the mating system and pollen dispersal patterns of a species, along with its genetic diversity and population structure, become essential for their conservation. Our study compares offspring from

**Table 3.** Spearman’s rank correlation coefficient ( $\rho$ ) between pairwise indices of: sample size within family ( $n$ ); effective number of pollen donors ( $N_{ep}$ ); observed heterozygosity ( $H_o$ ); fixation index ( $F_o$ ); coancestry ( $\Theta$ ) and variance effective size ( $N_e$ )

	$N_{ep}$	$H_o$	$F_o$	$\Theta$	$N_e$
$n$	0.182	0.244	–0.215	–0.285*	0.581**
$N_{ep}$	–	0.335**	–0.246*	–0.922**	0.798**
$H_o$	–	–	–0.981**	0.215	–0.344**

\* $P < 0.05$ ; \*\* $P < 0.01$ .

two distinct *G. americana* PTs, one at 6 months of age (NP) and one at 14 years of age (PT), in order to understand the breeding dynamics and propose effective conservation strategies. However, it is important to highlight that our results for mating system in NP represent the effective mating pattern, as it is based on the analysis of germinated seeds, and in PT the results represent the realized mating pattern, as they are based on juveniles at 14 years of age. Thus, in both cases, our results may differ from the mating patterns that occurred during the respective reproductive events that gave origin to our samples, as some stochastic processes may have occurred, such as random mortality, seed predation and disease, as well as some deterministic processes, including selection against inbred individuals between the zygote stage and when sampling occurred.

### Genetic diversity and biparental inbreeding

The genetic differentiation between populations for the offspring pollen gene pool ( $G'_{st} = 0.52$ ) showed that the majority of genetic diversity is distributed between populations and not within populations. This result can be attributed to the distance between the populations (1900 km) and the observed pollen dispersal patterns of isolation by distance, which was reported previously for other *G. americana* populations (Rabbani *et al.*, 2012; Manoel *et al.*, 2017), indicating that genetic differentiation among populations increases with increased geographic distance.

The populations studied herein present high levels of genetic diversity in comparison with the results from a previous study using the same seven loci. Manoel *et al.* (2017) studying adults, juveniles and open-pollinated seeds from two reproductive events found a lower mean observed ( $H_o = 0.38\text{--}0.54$ ) and expected heterozygosity ( $H_e = 0.55\text{--}0.77$ ). Another study on three *G. americana* populations in the Amazon forest using six simple sequence repeat loci (Ruzza *et al.*, 2018), three of which were the same as those used herein, also reported lower  $H_o$  (0.13–0.21) and  $H_e$  (0.48–0.57) than our study. The differences between studies may be due to the sample size and natural history of the studied populations (bottleneck and founder effects). However, the high levels of heterozygosity detected for both populations assessed herein are significant for *in situ* conservation of the NP population and *ex situ* conservation of both populations, as a large number of new genotypic recombinations may occur through mating among individuals of NP and also among individuals of PT. Furthermore, this result suggests that both studied populations have the potential to be used in future breeding programmes and seed collection efforts aimed at the recovery of degraded areas.

### Mating system

The outcrossing rate ( $t_m$ ) estimated for NP and PT were similar to that expected for dioecious species ( $t_m = 1.0$ ), where there is no possibility of self-fertilization as all individuals originate from outcrossing. At the progeny level, the estimates of  $t_m$  were high (0.93–0.99), but lower than unity in some cases. Estimates of  $t_m$  lower than unity for the species (0.87 and 0.95) were also reported previously and were attributed to mating among related trees, since reproduction in dioecious species occurs only by outcrossing (Manoel *et al.*, 2015a). We observed that the  $t_m$  values at the progeny level were positively correlated with the sample size, where progenies with a small sample size presented the lowest values in NP, indicating that results for these progenies must be interpreted with caution. If we examine the multilocus genotypes of the progenies with the lowest estimates of crossbreeding, we can see that all descendants of all progenies had at least one different allele from the mother tree in at least one locus, indicating that they all originated from outcrossing. Furthermore, this suggests that the low estimates of  $t_m$  are an artefact of the numerical method of EM to solve the probability equations. To prove this hypothesis, we re-estimated the  $t_m$  using the Newton–Raphson numerical method and found that none of the estimates were significantly different from unity.

However, mating was not random due to some mating among relatives ( $1 - t_s$ ) and correlated mating ( $r_p$ ), especially in NP ( $1 - t_s = 0.15$ ,  $r_p = 0.34$ ). Mating among relatives is indicative of intrapopulation spatial genetic structure (SGS), which typically results from seed dispersal near to the seed trees (Hardy *et al.*, 2006; Tarazi *et al.*, 2013; Degen and Sebbenn, 2014). The presence of SGS in the population used to establish the PT is likely similar to our results for NP. Correlated mating ( $r_p$ ) can be attributed to the behaviour of pollen dispersers, in this case bees, visiting and fertilizing many flowers of the same female tree with pollen from a single or small number of previously visited male trees, as indicated by the effective number of pollen donors ( $N_{ep}$ : NP = 2.9; PT = 13). This pollinator behaviour also explains the lower levels of correlated mating among ( $r_{pa}$ ) than within ( $r_{pw}$ ) fruits and, consequently, the lower number of effective pollen donors within ( $N_{epw} = 2.5$ ) than among fruits ( $N_{epa} = 3.8$ ). Similarly, Manoel *et al.* (2015a) investigated annual variation in the mating system of a *G. americana* population and reported lower  $r_{pa}$  (0.364–0.714) than  $r_{pw}$  (0.553–0.842) and a lower  $N_{epw}$  (1.2–1.8) than  $N_{epa}$  (1.4–2.7). Furthermore, the same pattern of lower  $r_{pa}$  than  $r_{pw}$  has been identified for other multi-seed-fruit and insect- and animal-pollinated tree species (Fuchs *et al.*, 2003; Tamaki *et al.*, 2009; Silva *et al.*, 2011; Feres *et al.*, 2012; Wadt *et al.*, 2015; Baldoni *et al.*, 2017; Giustina *et al.*, 2018; Moraes *et al.*, 2018), indicating that this may be standard pollinator behaviour for such tree species.

Nevertheless, this fertilization pattern may also be the result of a limited number of flowering neighbour trees due to individual variation in flowering phenology and/or low population density as a result of natural species distribution, logging or forest fragmentation (Moraes *et al.*, 2018). Although the results indicate long-distance pollen dispersal in the NP population (mean of 179 m), the pattern was isolation by distance, which indicates more effective pollination events occurring at short distances and less chance of pollen dispersal over long distances. This pattern suggests that the observed mating among relatives ( $1 - t_s$ ) and correlated mating ( $r_p$ ) in NP is mainly the product of near-neighbour mating between related and unrelated male and female trees. *G. americana* is pollinated by insects (bees) capable of flying long distances (>5 km) and thus long-distance pollen dispersal (Steffan-Dewenter and Kuhn, 2003; Abou-Shaara, 2014). The result of short-distance pollen dispersal compared to potential long-distance flight by bees could also be related to foraging preferences of some species (De Vere *et al.*, 2017). Similarly, pollen dispersal following a pattern of isolation by distance has been found for other *G. americana* populations (mean of 166–313 m; Manoel *et al.*, 2017) and in studies on other tree species (Lander *et al.*, 2010; Braga and Collevatti, 2011; Fuchs and Hamrick, 2011; Ismail *et al.*, 2012; Manoel *et al.*, 2012; Baldauf *et al.*, 2014; Breed *et al.*, 2015; Reim *et al.*, 2015; Tambarussi *et al.*, 2015; Bezemer *et al.*, 2016; Baldoni *et al.*, 2017; Spoladore *et al.*, 2017; Moraes *et al.*, 2018; Souza *et al.*, 2018; Lander *et al.*, 2019; Potascheff *et al.*, 2019; Solís-Hernández and Fuchs, 2019). This is further supported by the presence of SGS and mating among related trees ( $1 - t_s$ ) in other *G. americana* populations (Sebbenn *et al.*, 1998; Manoel *et al.*, 2017) and can explain the mating among related trees detected within progenies of PT.

The effective number of pollen donors was four times lower in NP than in PT ( $N_{ep}$ : NP = 2.9; PT = 13). The difference between populations may be explained by the fact that NP represents the effective mating measured in offspring at 6 months of age and PT represents realized mating at 14 years of age. As previously mentioned, results may differ due to the occurrence of stochastic (random mortality, seed predation and disease) and deterministic processes (selection against inbred individuals between the offspring and adult stages, and nursery selection of more vigorous offspring to establish the PT population).

As noted above, the significant differences for  $1 - t_s$  and  $r_p$  between NP and PT can be attributed to the fact that our results for mating system in NP represent the effective mating pattern, measured in germinated seeds, while the result for PT represent the realized mating pattern, measured in 14-year-old trees. Thus, our results may differ from the actual mating patterns that occurred during the reproductive events due to stochastic and deterministic process.

However, the analysis of the PT population included the oldest and only surviving trees 14 years after establishing the PT; therefore, these trees were likely subject to stronger stochastic and deterministic processes than the NP population, particularly inbreeding depression which can cause higher mortality of inbred individuals originated from mating among related trees. The observed lower levels of biparental inbreeding in mothers ( $F_m$ ) than offspring ( $F_o$ ) is a strong indicator of selection against inbred individuals between seed/young trees and adult stages. Furthermore, the observed variation among seed trees within and among populations for mating among related trees and correlated mating can also be attributed to the mentioned stochastic and deterministic processes, as well as the site-specific individual variation in flowering phenology, limited number of flowering neighbour trees or low population density.

## Conclusions

PTs for *ex situ* conservation enable the maintenance of high levels of genetic diversity of a species, since they can preserve alleles from populations that no longer exist, despite sampling a small number of individuals from the original populations. The combined strategy of seed collection from both NP and PT populations is also valuable as it can support increases in genetic diversity levels. Commercial use of wood and fruits must be done with caution to ensure that a balanced sex ratio is preserved in the region. Our results show clear mating system differences between NP (effective) and PT (realized) offspring. Mating was not random in the populations, especially the effective mating (NP), due to the high rates of mating among related trees and correlated mating, resulting in a low effective size ( $N_e$ ). The high frequency of biparental inbred individuals (half-sibs and full-sibs) in the NP population is expected to decrease as the trees transition into adult stages due to inbreeding depression, which will increase the  $N_e$  and heterozygosity and decrease the genetic differences between the trials. Thus, the effective population size of both populations may reach the minimum level ( $N_e \geq 70$ ) for short-term *ex situ* conservation. As the species is dioecious and inbreeding can only occur through mating among relatives, the trial may be suitable for seed collection for environmental reforestation and breeding if submitted to selective logging, maintaining only male or female trees within each plot to avoid the occurrence of mating among related trees. Furthermore, to maximize the effective population size ( $N_e$ ) of the parental population for seed production, the same number of male ( $N_m$ ) and female ( $N_f$ ) trees of each progeny should be maintained. If a single male and a single female are selected per progeny, 30 males and 30 females will remain in the resulting seedling seed orchard after selective logging, which will result in an



effective population size [ $N_e = (4N_mN_f)/(N_m + N_f) = (4 \times 30 \times 30)/(30 + 30) = 60$ ] similar to that indicated for short-term *ex situ* conservation. However, it is important to note that the gene pool of the populations will be dispersed throughout the reforestation areas, thus creating new populations from the source material. For *ex situ* conservation and seed collection for reforestation of the NP, our results from the NP offspring population indicate that seeds must be collected from at least 56 trees to obtain samples with an  $N_e$  of 150.

### Supplementary material

The supplementary material for this article can be found at <https://doi.org/10.1017/S1479262121000022>.

### Acknowledgements

The authors are thankful for Sao Paulo Research Foundation (FAPESP; grant 13/19444-6 to MFVM). Research fellowships were provided by the National Council for Scientific and Technological Development (CNPq) for Alexandre M. Sebbenn and Mario L.T. Moraes. We also thank Dr Evelyn Nimmo for editing the English of the manuscript.

### References

- Abou-Shaara HF (2014) The foraging behaviour of honey bees, *Apis mellifera*: a review. *Veterinari Medicina* 59: 1–10.
- Baldauf C, Ciampi-Guillard M, Aguirra TJ, Corrêa CE, Santos FAM, Souza AP and Sebbenn AM (2014) Genetic diversity spatial genetic structure and realized seed and pollen dispersal of *Himatantibus drasticus* (Apocynaceae) in the Brazilian savanna. *Conservation Genetics* 15: 1073–1083.
- Baldoni AB, Wadt LHO, Campos T, Silva VS, Azevedo VCR, Mata LR, Botin AA, Mendes NO, Tardin FD, Tonini H, Hoogerheide ESS and Sebbenn AM (2017) Contemporary pollen and seed dispersal in natural populations of *Bertholletia excelsa* (Bonpl.). *Genetics and Molecular Research* 16: gmr16039756.
- Bakkali AE, Haouane H, Hadidou A, Oukabli A, Santoni S, Udupa SM, Van Damme P and Khadari B (2013) Genetic diversity of on-farm selected olive trees in Moroccan traditional olive orchards. *Plant Genetic Resources* 11: 97–105.
- Bezemer N, Krauss SL, Phillips RD, Roberts DG and Hopper SD (2016) Paternity analysis reveals wide pollen dispersal and high multiple paternity in a small isolated population of the bird-pollinated *Eucalyptus caesia* (Myrtaceae). *Heredity* 117: 460–470.
- Braga AC and Collevatti RG (2011) Temporal variation in pollen dispersal and breeding structure in a bee-pollinated Neotropical tree. *Heredity* 106: 911–919.
- Breed MF, Ottewill KM, Gardner MG, Marklund MHK, Dormont EE and Lowe AJ (2015) Mating patterns and pollinator mobility are critical traits in forest fragmentation genetics. *Heredity* 115: 108–114.
- Caballero A, Bravo I and Wang J (2016) Inbreeding load and purging: implications for the short-term survival and the conservation management of small populations. *Heredity* 118: 177–185.
- Carvalho PER (1994) Espécies florestais brasileiras: recomendações silviculturais, potencialidade e uso da madeira. Embrapa Florestas, Colombo.
- Chybicki IJ and Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100: 106–113.
- Colombo AF and Joly CA (2010) Brazilian Atlantic forest *lato Sensu*: the most ancient Brazilian forest, and a biodiversity hotspot, is highly threatened by climate change. *Brazilian Journal of Biology* 70: 697–708.
- Crestana MSM, Toledo Filho DV and Campos JB (2004) Florestas: sistemas e recuperação com essências nativas, produção de mudas e legislação Campinas: Coordenadoria de Assistência Técnica Integral, pp. 216.
- Degen B and Sebbenn AM (2014) Genetic and tropical forest. In: Pancel L and Köhl M (eds) *Tropical Forestry Handbook*. Berlin: Heidelberg: Springer-Verlag, pp. 1–30.
- De Vere N, Jones LE, Gilmore T, Moscrop J, Lowe A, Smith D, Hegarty MJ, Creer S and Ford CR (2017) Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Scientific Reports* 7: 42838.
- Dinerstein E, Olson DM, Graham DJ, Webster AL, Primm AS, Bookbinder MP and Ledec GA (1995) *Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Caribbean*. Washington: World Bank, pp. 176.
- Doyle JJ and Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.
- Durigan G and Nogueira JCB (1990) Recomposição de matas ciliares. *Série Registros: Instituto Florestal de São Paulo* 4: 1–14.
- Francisco CF (1989) Análise ambiental e consequência do desmatamento no município de Presidente Prudente no período de 1914 à 1986. Rio Claro. Dissertação (Mestrado em Geografia) Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista.
- Franco RD, Haidar RF and Machado RB (2016) Tree species of South America central savanna: endemism, marginal areas and the relationship with other biomes. *Acta Botanica Brasiliensis* 30: 78–86.
- Feres JM, Sebbenn AM, Guidugli MC, Mestriner MA, Moraes MLT and Alzate-Marin AL (2012) Mating system parameters at hierarchical levels of fruits, individuals and populations in the Brazilian insect-pollinated tropical tree, *Tabebuia roseo-alba* (Bignoniaceae). *Conservation Genetics* 13: 393–405.
- Fuchs EJ and Hamrick JL (2011) Mating system and pollen flow between remnant populations of the endangered tropical tree, *Guaiaicum sanctum* (Zygophyllaceae). *Conservation Genetics* 12: 175–185.
- Fuchs EJ, Lobo JA and 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*. *Conservation Biology* 17: 149–157.
- Giustina LD, Baldoni AB, Tonini H, Azevedo VCR, Neves LG, Tardin FD and Sebbenn AM (2018) Hierarchical outcrossing among and within fruits in *Bertholletia excelsa* Bonpl. (Lecythidaceae) open-pollinated seeds. *Genetics and Molecular Research* 17: gmr16039872.
- Goudet J (1995) FSTAT (version 2.9.3.2.): a computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.
- Hardy OJ, Maggia L, Bandou E, Breyne P, Henri Caron H, Chevallier M, Doliguez A, Dutech C, Kremer A, Latouche-Halle C, Troisoux V, Veron V and Degen B (2006) Fine-scale genetic structure and gene dispersal inferences

- in 10 Neotropical tree species. *Molecular Ecology* 15: 559–579.
- Hausman CE, Bertke MM, Jaeger JF and Rocha OJ (2014) Genetic structure of green ash (*Fraxinus pennsylvanica*): implications for the establishment of ex situ conservation protocols in light of the invasion of the emerald. *Plant Genetic Resources: Characterization and Utilization* 12: 286–297.
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* 59: 1633–1638.
- Hodgkin T, Rao VR, Cibrian-Jaramillo A and Gaiji S (2003) The use of ex situ conserved plant genetic resources. *Plant Genetic Resources* 1: 19–29.
- Ismail SA, Ghazoul J, Ravikanth G, Uma Shaanker R, Kushalappa CG and Kettle CJ (2012) Does long-distance pollen dispersal preclude inbreeding in tropical trees? Fragmentation genetics of *Dysoxylum malabaricum* in an agro-forest landscape. *Molecular Ecology* 21: 5484–5496.
- Lande R (1995) Mutation and conservation. *Conservation Biology* 9: 782–791.
- Lander TA, Boshier DH and Harris SA (2010) Fragmented but not isolated: contribution of single trees, small patches and long-distance pollen flow to genetic connectivity for *Gomortega keule*, an endangered Chilean tree. *Biological Conservation* 143: 2583–2590.
- Lander TA, Harris SA, Cremona PJ and Boshier DH (2019) Impact of habitat loss and fragmentation on reproduction, dispersal and species persistence for an endangered Chilean tree. *Conservation Genetics* 20: 973–985.
- Manoel RO, Alves PF, Dourado CL, Gaino APSC, Freitas MLM, Moraes MLT and Sebbenn AM (2012) Contemporary pollen flow, mating patterns and effective population size inferred from paternity analysis in a small fragmented population of the Neotropical tree *Copaifera langsdorffii* Desf. (Leguminosae-Caesalpinioideae). *Conservation Genetics* 13: 613–623.
- Manoel RO, Freitas MLM, Barreto MA, Moraes MLT, Souza AP and Sebbenn AM (2014) Development and characterization of 32 microsatellite loci in *Genipa americana* (Rubiaceae). *Applications in Plant Sciences* 2: 1300084.
- Manoel RO, Freitas MLM, Furlani Junior E, Alves PF, Moraes MLT and Sebbenn AM (2015a) Individual, fruit, and annual variation in correlated mating in a *Genipa americana* population. *Silvae Genetica* 69: 108–116.
- Manoel RO, Freitas MLM, Tambarussi EV, Cambuim J, Moraes MLT and Sebbenn AM (2015b) Study of Mendelian inheritance, genetic linkage and genotypic disequilibrium at six microsatellite loci of *Genipa americana* L. (Rubiaceae). *Genetics and Molecular Research* 14: 8161–8169.
- Manoel RO, Freitas MLM, Furlani Junior E, Alves PF, Moraes MLT and Sebbenn AM (2017) Low levels of pollen and seed flow in a riparian forest fragment of the dioecious tropical tree *Genipa americana* L. *Forestry Research and Engineering: International Journal* 1: 18–27.
- Moraes MA, Kubota TYK, Rossini BC, Marino CL, Freitas MLM, Moraes MLT, Silva AM, Cambuim J and Sebbenn AM (2018) Long-distance pollen and seed dispersal and inbreeding depression in *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) in the Brazilian savannah. *Ecology and Evolution* 8: 7800–7816.
- Nunney L and Campbell KA (1993) Assessing minimum viable population size: demography meets population genetics. *Tree* 8: 234–239.
- Oliveira JB. (1999) Mapa pedológico do Estado de São Paulo: legenda expandida. Instituto Agrônomo, Campinas.
- Potascheff CM, Oddou-Muratorio S, Klein EK, Figueira A, Bressan EA, Oliveira PE, Lander TA and Sebbenn AM (2019) Stepping stones or stone dead? Fecundity, pollen dispersal and mating patterns of roadside *Qualea grandiflora* Mart. trees. *Conservation Genetics* 20: 1355–1367.
- Rabbani ARC, Silva-Mann R and Ferreira RA (2012) Variabilidade genética de *Genipa americana* L. pertencente ao baixo curso do rio São Francisco. *Revista Arvore* 36: 401–409.
- Reim S, Proft A, Heinz S and Lochschmidt F (2015) Pollen movement in a *Malus sylvestris* population and conclusions for conservation measures. *Plant Genetic Resources: Characterisation and Utilisation* 15: 12–20.
- Ritland K (2002) Estimation of gene frequency and heterozygosity from pooled samples. *Molecular Ecology Notes* 78: 370–372.
- Robledo-Arnuncio JJ, Austerlitz F and Smouse PE (2007) POLDISP: a software package for indirect estimation of contemporary pollen dispersal. *Molecular Ecology* 7: 763–766.
- Ruzza DAC, Rossi AAB, Bispo RB, Tiago AV, Cochev JS, Rossi FS and Fernandes JM (2018) The genetic diversity and population structure of *Genipa americana* (Rubiaceae) in Northern Mato Grosso, Brazil. *Genetics and Molecular Research* 17: gm18017.
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.
- Sebbenn AM (2006) Sistema de reprodução em espécies arbóreas tropicais e suas implicações para a seleção de árvores matrizes para reflorestamentos ambientais. In: Higa AR and Silva LD (eds) *Pomares de Sementes de Espécies Nativas*. Curitiba: FUDEP, pp. 193–198.
- Sebbenn AM, Kageyama PY and Vencovsky R (1998) Variabilidade genética, sistema reprodutivo e estrutura genética especial em *Genipa americana* L. através de marcadores isoenzimáticos. *Scientia Forestalis* 53: 15–30.
- Sebbenn AM, Boshier D, Freitas MLM, Zanatto ACS, Sato AS, Etori LC and Moraes E (2007) Results of an international provenance trial of *Cordia alliodora* in São Paulo, Brazil at five and 23 years of age. *Silvae Genetica* 56: 110–117.
- Sergipe (2000) Secretaria de Estado do Planejamento e da Ciência e Tecnologia. Santana do São Francisco. In: Perfis Municipais Aracaju, pp. 75.
- Silva CRS, Albuquerque PSB, Ervedosa FR, Mota JWS, Figueira A and Sebbenn AM (2011) Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. *Heredity* 106: 973–985.
- Silva RFB, Millington JDA, Moran EF, Batistella M and Liu J (2020) Three decades of land-use and land cover changes in mountain regions of the Brazilian Atlantic Forest. *Landscape and Urban Planning* 204: 103948.
- Smouse PE and Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management* 197: 21–38.
- Solís-Hernández W and Fuchs EJ (2019) Effective gene flow patterns across a fragmented landscape in southern Costa Rica for *Symphonia globulifera* (Clusiaceae): a species with mobile seed and pollen dispersers. *Revista de Biología Tropical* 67: 95–111.
- Sousa VA, Sebbenn AM, Hattemer H and Ziehe M (2005) Correlated mating in populations of a dioecious Brazilian conifer, *Araucaria angustifolia* (Bert.) O. Ktze. *Forest Genetics* 12: 107–119.

- Souza FB, Kubota TYK, Tambarussi EV, Freitas MLM, Moraes MLT, Silva AM, Cambuim J and Sebbenn AM (2018) Historic pollen and seed dispersal in fragmented populations of the two largest trees of the Atlantic Forest. *Forestry Research and Engineering: International Journal* 2: 100–109.
- Spoladore J, Mansano VF, Lemes MR, Freitas LCD and Sebbenn AM (2017) Genetic conservation of small populations of the endemic tree *Swartzia glazioviana* (Taub.) Glaz. (Leguminosae) in the Atlantic Forest. *Conservation Genetics* 18: 1105–1117.
- Steffan-Dewenter I and Kuhn A (2003) Honeybee foraging in differentially structured landscapes. *Proceedings: Biological Sciences* 270: 569–575.
- Tamaki I, Setsuko S and Tomaru N (2009) Estimation of out-crossing rates at hierarchical levels of fruits, individuals, populations and species in *Magnolia stellata*. *Heredity* 102: 381–388.
- Tambarussi EV, Boshier D, Vencovsky R, Freitas MLM and Sebbenn AM (2015) Paternity analysis reveals significant isolation and near neighbour pollen dispersal in small *Cariniana legalis* Mart. Kuntze populations in the Brazilian Atlantic Forest. *Ecology and Evolution* 5: 5588–5600.
- Tarazi R, Sebbenn AM, Kageyama PY and Vencovsky R (2013) Edge effects enhance selfing and seed harvesting efforts in the insect-pollinated Neotropical tree *Copaifera langsdorffii* (Fabaceae). *Heredity* 110: 578–585.
- Wadt LHO, Baldoni AB, Silva VS, Campos T, Martins K, Azevedo VCR, Mata LR, Botin AA, Hoogerheide ESS, Tonini H and Sebbenn AM (2015) Mating system variation among populations, individuals and within and among fruits in *Bertolletia excelsa*. *Silvae Genetica* 64: 248–259.