# Genomic diversity is similar between Atlantic Forest restorations and natural remnants for the native tree *Casearia sylvestris* Sw.

João Paulo Gomes Viana[1], Marcos Vinícius Bohrer Monteiro Siqueira[2], Fabiano Lucas Araujo[3], Carolina Grando[1], Patricia Sanae Sujii[1], Ellida de Aguiar Silvestre[1], Mariana Novello[1], José Baldin Pinheiro[4], Marcelo Mattos Cavallari[5], Pedro H. S. Brancalion[6], Ricardo Ribeiro Rodrigues[7], Anete Pereira de Souza[8], Julian Catchen[9], Maria I. Zucchi[10]*
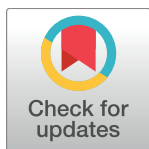
1 Graduate Program in Genetics and Molecular Biology, University of Campinas, Campinas–SP, Brazil, **2** Universidade do Sagrado Coração, Bauru–SP, Brazil, **3** Graduate Program in Tropical and Subtropical Agriculture, Agronomic Institute of Campinas, Campinas–SP, Brazil, **4** Department of Genetics, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba-SP, Brazil, **5** Embrapa Pecuária Sudeste, Embrapa, São Carlos–SP, Brazil, **6** Department of Forest Sciences, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba-SP, Brazil, **7** Department of Biology, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba-SP, Brazil, **8** Department of Plant Biology, University of Campinas, Campinas–SP, Brazil, **9** Department of Animal Biology, University of Illinois at Urbana—Champaign, Champaign–IL, United States of America, **10** Agência Paulista de Tecnologia dos Agronegócios, Piracicaba–SP, Brazil

* mizucchi@apta.sp.gov.br

Abstract

The primary focus of tropical forest restoration has been the recovery of forest structure and tree taxonomic diversity, with limited attention given to genetic conservation. Populations reintroduced through restoration plantings may have low genetic diversity and be genetically structured due to founder effects and genetic drift, which limit the potential of restoration to recover ecologically resilient plant communities. Here, we studied the genetic diversity, genetic structure and differentiation using single nucleotide polymorphisms (SNP) markers between restored and natural populations of the native tree *Casearia sylvestris* in the Atlantic Forest of Brazil. We sampled leaves from approximately 24 adult individuals in each of the study sites: two restoration plantations (27 and 62 years old) and two forest remnants. We prepared and sequenced a genotyping-by-sequencing library, SNP markers were identified *de novo* using Stacks pipeline, and genetic parameters and structure analyses were then estimated for populations. The sequencing step was successful for 80 sampled individuals. Neutral genetic diversity was similar among restored and natural populations (A_R = 1.72 ± 0.005; H_O = 0.135 ± 0.005; H_E = 0.167 ± 0.005; F_{IS} = 0.16 ± 0.022), which were not genetically structured by population subdivision. In spite of this absence of genetic structure by population we found genetic structure within populations but even so there is not spatial genetic structure in any population studied. Less than 1% of the neutral alleles were exclusive to a population. In general, contrary to our expectations, restoration plantations were then effective for conserving tree genetic diversity in human-modified tropical landscapes. Furthermore, we demonstrate that genotyping-by-sequencing can be a useful tool in restoration genetics.

Introduction

Increasing environmental degradation has led to the promotion of ecological restoration worldwide, particularly for forest ecosystems [1–3]. Forest restoration methods are selected based upon the degree of ecosystem degradation [4,5], the availability of native forest fragments to act as seed sources in the landscape [6,7] the goals of restoration interventions [4], and funding constraints [8,9]. For all restoration approaches, forest succession must proceed in time, which is the primary ecological process controlling the recovery of forest ecosystems [10]. However, several factors drive tropical forest succession. Although some of these factors are unpredictable and difficult to manipulate at the site level, such as dispersal by fauna, others can be controlled at some level during restoration planning, implementation, or adaptive management; these are the factors that receive special attention from practitioners. Genetic diversity of reintroduced populations is a key factor for the long-term persistence of native species in forest restoration sites that can be manipulated by practitioners, based on the selection of natural populations and seed-trees for producing planting stocks [9,11,12].

Genetic diversity is defined as the variety of alleles and genotypes in a population, which drives morphological, physiological and behavioral differences in individuals and populations [13]. It is important for the long-term viability of populations because of the influence on adaptive potential, inbreeding, and genetic drift [14–16]. Moreover, the risk of extinction of natural populations increases at reduced genetic diversity [16]. One of the causes of increased vulnerability is the limitation of evolutionary potential to respond to environmental changes [17], a critical issue in the Anthropocene [18]. Tree species, in particular, are highly vulnerable to reduced genetic diversity in founder populations, because of the lack of mobility and susceptibility to significant environmental changes within a lifetime [19]. In spite of the numerous conceptual arguments for considering genetic diversity in biodiversity conservation, a large gap occurs between theory and practice.

Forest restoration projects, for example, began long before the recognition of the importance of genetic diversity for their long-term sustainability [2]. The primary ecological goals have been the recovery of forest structure and woody species diversity, with less focus on genetic diversity. More recently, many studies have addressed the importance of genetic diversity for restoration success [11,12,20,21]. However, few studies have measured genetic diversity in restored areas, which limits our understanding of the genetics of restoration.

Forest restoration plantings are often implemented using seedlings from few mother trees, and the level of genetic diversity and kinship among planted seedlings has received little attention [22,23]. Populations reintroduced through restoration plantings may have low genetic diversity and be genetically structured due to founder effects and genetic drift, which limit the potential of restoration to recover ecologically resilient plant communities. These limitations are a concern because inbreeding can lead to reduced fertility and individual survival in some species. However, natural remnants located near forest restoration areas can conserve genetic diversity [24] and supply a diverse set of alleles for the population undergoing restoration, acting as stepping stones that promote genetic connectivity among forest fragments across human-modified landscapes [25,26]. Few studies have assessed genetic diversity following forest restoration interventions, particularly in older restoration projects, implemented before the recognition of the importance of genetic diversity for the long-term viability of plant populations. Consequently, the expectation is that populations established through restoration plantations will have low genetic diversity and be structured due to founder effects, as observed for regenerating palm populations in second-growth tropical forests [27].

The advent of short-read sequencing, which can assess large-scale sets of single nucleotide polymorphisms (SNPs), has eased the study of population genetics in a genomic scale using

populations of non-model species [28], such as tropical trees. Additionally, SNPs offer the potential for large-scale scanning of the genome by providing a large number of markers that improve the accuracy of population genetic estimates [29]. Among the methodologies used for SNP discovery, Elshire et al. (2011) [30] proposed a method that uses restriction enzymes to reduce the complexity of the genome. This genotyping-by-sequencing (GBS) approach is a low cost, highly specific, and reproducible method to detect SNPs and can be applied in non-model species. Although each SNP represents a point measurement, we can combine these SNPs into a continuous distribution across the genome to conduct scans for genetic variation, which can be applied to native species [31], such as those used in forest restoration projects, and advance the knowledge on restoration genetics.

In this context, we assessed the genetic diversity and genetic structure of populations of the native tree *C. sylvestris* in restoration plantations and natural remnants of the Brazilian Atlantic Forest to verify if there are differences in genetic diversity level between these type of forest fragments.

Materials and methods

Study species

Casearia sylvestris Sw. (Saliaceae; *Salix* sp. and *Populus* sp. are tree species of this family widely used in population genetic studies in a genomic scale) is a pioneer tree species highly abundant in secondary forests [32–35], distributed from Mexico to Argentina [36]. The mating system has not been studied using molecular markers approach yet, but the flowers are hermaphroditic, and the species is self-compatible [37,38]. Similar to other species of this genus, *C. sylvestris* is usually pollinated by various species of flies, whereas seed dispersal occurs by small birds over short distances [39].

Study areas and sampling

We performed the study in the Brazilian Atlantic Forest, a global hot spot for biodiversity conservation that is estimated to contain from 1% to 8% of the world's terrestrial species [40] and more than 8000 endemic species. However, less than 8% of the original area remains within continuous forests [41] and only 11–16% when secondary forests and small fragments (<100 ha) are considered [42]. We selected two natural remnants (Campinas and Tietê) and two forest restoration plantations (Cosmópolis and Iracemápolis) in São Paulo State, southeastern Brazil, for this study (S1 Appendix). The natural remnant in Campinas (Mata de Santa Genebra, hereafter SG, 22°49.36' S, 47°6.60' W) is 252-ha and is embedded within one of the largest urban centers in Brazil, in a landscape dominated by pastures, sugarcane plantations, and small-scale forestry. Tietê (hereafter TT, 23°0.11' S, 47°43.66' W) is a 70-ha forest fragment surrounded by extensive pastures, similar to most natural remnants of the Brazilian Atlantic Forest. Cosmópolis (hereafter CO, 22°40.30' S, 47°12.25' W) is a 25-ha forest restoration plantation implemented between 1955 and 1960 in the riparian buffer of the Jaguari River in a landscape matrix dominated by sugarcane plantations, with few natural fragments remaining. The Iracemápolis (hereafter IR, 22°34.60 S, 47°30.18' W) study site is a 50-ha forest restoration plantation established between 1988 and 1990 at the border of a water reservoir that supplies drinking water to the city of Iracemápolis [9]. All four forests had a high abundance of *C. sylvestris* individuals, particularly near the borders and in the understory. The CO population was established with nursery-grown seedlings produced with seeds harvested in an old landscaping planting in the "Luiz de Queiroz" College of Agriculture of the University of São Paulo in Piracicaba–SP, whereas the IR population was established with nursery-grown seedlings bought from a forest nursery in Campinas–SP and another in São Paulo—SP. Leaves from

approximately 24 adult individuals, separated by approximately 12 m, were collected from each of the four populations. Using GPS equipment, a minimum distance of 12 m was established between individuals, with a margin of error of a maximum of 4 m (S1 Dataset). In the forest restorations we tried to select individuals of the same generation, but as *C. sylvestris* is a rapid growing plant species, both planted (i.e., adult individuals growing in the original planting lines) and regenerating individuals (i.e., adult individuals growing in between the original planting lines) may have been sampled.

SNP discovery and data analysis

Library preparation and sequencing. DNA was extracted from leaf tissue samples of individuals from each of the four populations following the method described by Doyle and Doyle (1987) [43]. To increase the chances of obtaining the same proportion of reads per individual, the samples with high quality DNA were diluted to the same concentration based on electrophoresis in agarose 1% gel. Preliminary tests with different restriction enzymes were performed. PstI digestion showed the highest concentration of fragments with the greatest distribution in size. The library was then prepared according to the method described by Elshire et al. (2011) [30], using barcodes with at least three nucleotides of difference to identify each individual. Finally, single-end sequencing of the 95 samples was performed in a single lane of an Illumina HiSeq 2000 (Illumina, San Diego, CA) producing sets of 100bp reads.

Data processing. Sequences of each sample were demultiplexed and cleaned using the *process_radtags* program from the *Stacks* package v. 1.44 [44]. A quality score filter (-q, *process_radtags*) was applied across each sample to retain only high-quality sequencing reads. Reads with any non-called base were discarded (-c, *process_radtags*). Barcodes were recovered when they had a maximum of one mismatch (-r, *process_radtags*). Because a reference genome is not available for *C. sylvestris*, the loci were built *de novo*. Sequences from each sample were aligned using a minimum depth of 3 and allowing up to 2 mismatches with the *ustacks* program (-m 3, -M 2, *ustacks*). Subsequently, a catalog of loci was built from subsets of samples, allowing 2 mismatches (-n 2, *cstacks*) between loci from distinct samples. All samples were matched against the catalog, using *sstacks*, to determine the set of loci in each. We then implemented a correction step with *rxstacks*, discarding loci when the mean log likelihood of the locus in the metapopulation was less than -10; confounded loci were also removed. Using the *populations* program, a filter was applied across loci and individuals to retain loci with a maximum of 20% missing data (S2 Dataset). To separate the outlier loci in the set of SNPs, we used an F_{ST}-based approach to assess population pairwise comparisons using the infinite alleles model and a false discovery rate (FDR) of 0.05, as implemented in the *Lositan* software [45]. Subsequently, neutral candidate loci were separated from outlier loci using 99% confidence intervals boundaries. All below estimates were calculated using the set of neutral candidate loci (hereafter neutral loci).

Genetic diversity. To compare the genetic diversity in forest restoration populations and natural remnants, we used the following population parameters: number of different alleles, number of private alleles, observed heterozygosity, expected heterozygosity and the inbreeding coefficient (F_{IS}). We tested for significant differences in heterozygosities and F_{IS} using confidence intervals calculated based on 1000 bootstrap resamples. Estimates of population genetic parameters were calculated using the *PopGenKit* [46] and *hierfstat* [47] packages. Venn diagrams were generated using the *VennDiagram* package [48]. All packages were developed for use in the *R* software [49].

Genetic differentiation and structure. To understand the partitioning of genetic diversity, we calculated overall and pairwise F-statistics using the *diveRsity* package [50]. The

significance of these estimates was determined using 95% confidence intervals based on 1000 bootstrap resamples across individuals. A Discriminant Analysis of Principal Components (DAPC) was conducted to explore the genetic structure in these populations using *adegenet* [51,52].

Spatial genetic structure. The spatial genetic structure (SGS) analysis was performed based on the kinship coefficient proposed by Loiselle et al. (1995) [53]. This kinship coefficient is robust and unbiased by the frequency of rare alleles. The number and distance intervals of classes were defined to suit each population. To summarize the spatial genetic structure, we used the *Sp-statistic* proposed by Vekemans and Hardy (2004) [54], which is calculated as $S_p = -\beta_{log} / 1 - F_1$; where β_{log} is the regression slope of the kinship coefficient on the log of geographic distance, and F_1 is the mean of the kinship coefficient between individuals estimated for the first distance class. For each distance class, the 95% confidence interval was obtained based on 1000 permutations. For all analyses of SGS within the populations, we used the *SPA-GeDi* program developed by Hardy and Vekemans (2002) [55], whereas the spatial autocorrelation graphs were plotted with the *R* software [49].

Results

SNP discovery and data processing

Considering the high number of reads obtained (317,802,888 reads), the sequencing was successful. After cleaning, demultiplexing and filtering of missing data, 861 high-quality biallelic SNPs were identified, which were used to characterize the genetic diversity of *C. sylvestris* individuals distributed across the four populations. For these results, the mean sequencing depth per individual ranged from 1.25 to 43.40, whereas mean depth per locus ranged from 4.70 to 83.22 (SD = 8.58). After all filtering steps it was possible to retain 80 individuals for the subsequent analyses. Using the F_{ST} approach in pairwise population comparisons to detect outlier loci, we identified 662 neutral candidate loci and 199 outlier candidate loci.

Genetic diversity

We found no significant differences in observed and expected heterozygosities and inbreeding coefficient between the natural remnants and forest restoration populations (Table 1).

Of the 1324 different alleles from neutral loci detected in the four populations, all populations shared approximately 79%, 11% corresponded to alleles shared by at least three populations, 9% were in at least two populations, and less than 1% were exclusive to a population (Fig 1). These results are consistent with those presented in Table 1, indicating no genetic differences among the forest restoration and natural remnant populations.

Table 1. Genetic diversity parameters for *Casearia sylvestris* populations in natural remnants and restoration plantations in the Atlantic Forest of southeastern Brazil.

Pop.	N	A	P	H_O	H_E	F_{IS}
CO	22	1217	2	0.13 (0.13–0.14)	0.17 (0.16–0.17)	0.19 (0.15–0.23)
IR	18	1213	1	0.14 (0.13–0.16)	0.17 (0.15–0.17)	0.15 (0.10–0.19)
SG	21	1223	4	0.14 (0.13–0.15)	0.17 (0.15–0.17)	0.14 (0.10–0.18)
TT	19	1216	2	0.13 (0.12–0.15)	0.16 (0.15–0.17)	0.17 (0.12–0.22)

CO–Forest restoration in Cosmópolis, SP; IR–Forest restoration in Iracemápolis, SP; SG–Natural remnant in Campinas, SP; TT—Natural remnant in Tietê, SP. n–Sample size; A–Number of different alleles; P–Number of private alleles; H_O–Observed heterozygosity; H_E–Expected heterozygosity; F_{IS}–Within population fixation index.
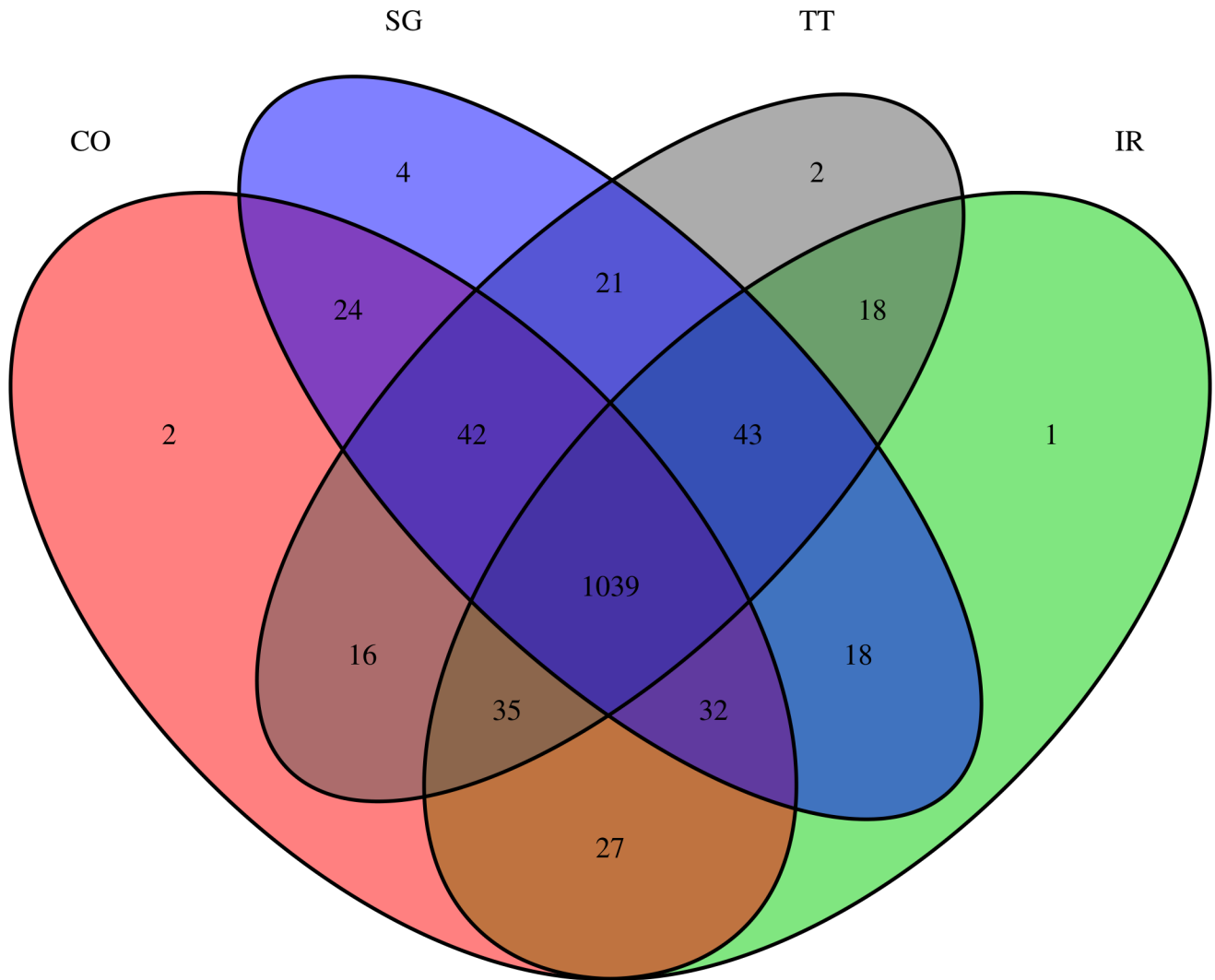
https://doi.org/10.1371/journal.pone.0192165.t001

Fig 1. Venn diagram showing the number of shared alleles among groups of *Casearia sylvestis* populations from restoration plantations and natural remnants in the Atlantic Forest of southeastern Brazil. Seventy-nine percent of all alleles were found in all populations. CO–Samples from forest restoration in Cosmópolis, SP; IR–Samples from forest restoration in Iracemápolis, SP; SG–Samples from natural forest in Campinas, SP; TT–Samples from natural forest in Tietê, SP. Different colors represent each of the four studied populations.

Genetic differentiation and structure

The populations were not genetically differentiated. Moreover, all pairwise F_{ST} estimates (Table 2, upper triangular) were not significantly different from zero ($P > 0.05$). Based on

Table 2. Pairwise Nei's F_{ST} value (upper triangular. F_{ST} value were not significantly different from zero based on 95% confidence intervals calculated from 1000 bootstrap resamplings.

	CO	IR	SG	TT
CO	-	-0.0006	0.0011	0.0004
IR	-	-	0.0020	-0.0004
SG	-	-	-	0.0008
TT	-	-	-	-

CO–Forest restoration in Cosmópolis, SP; IR–Forest restoration in Iracemápolis, SP; SG–Natural remnant in Campinas, SP; TT—Natural remnant in Tietê, SP.
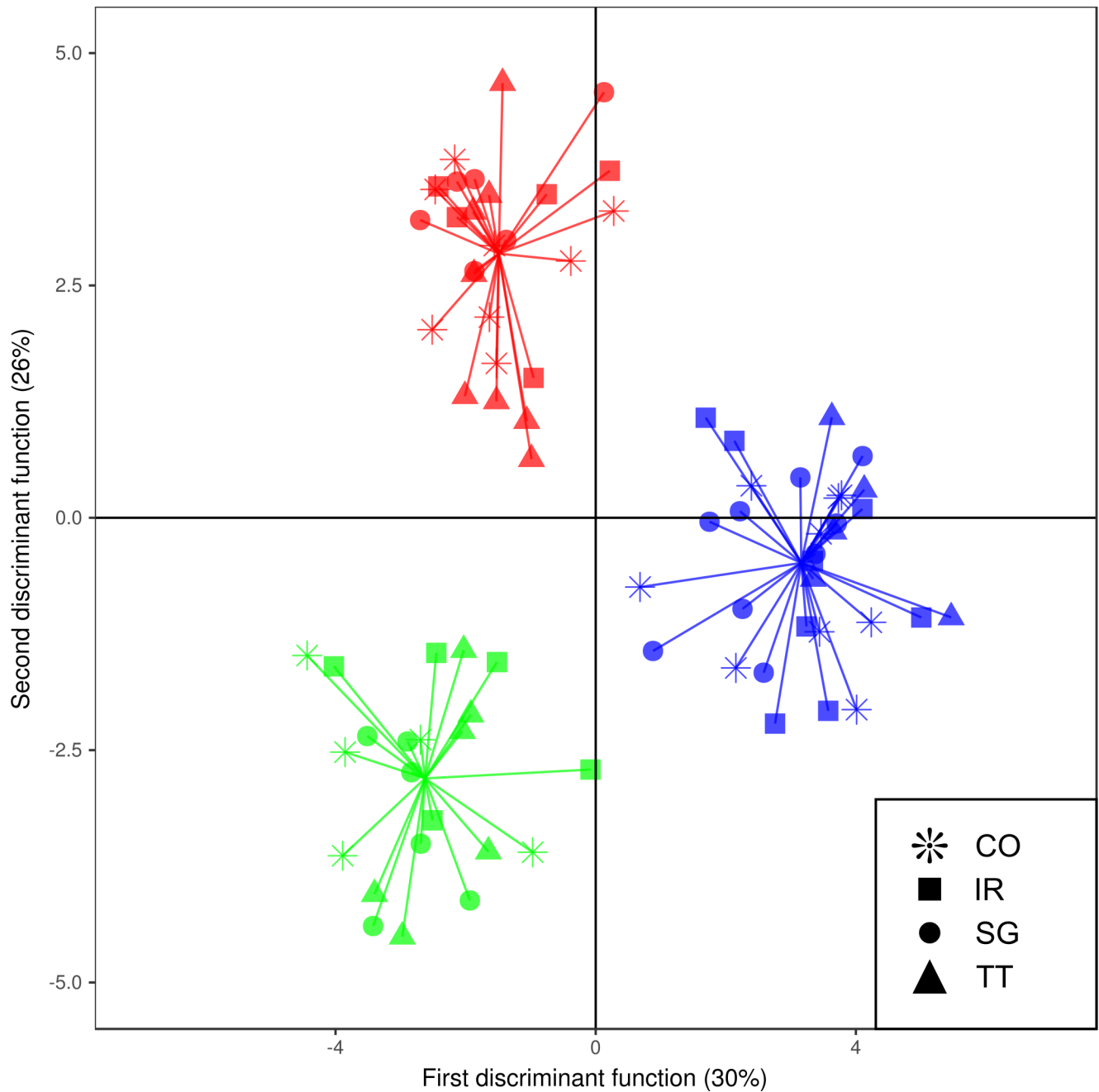
Fig 2. Scatterplot from DAPC showing the two discriminant functions for *Casearia sylvestris* populations from restoration plantations and natural remnants in the Atlantic Forest of southeastern Brazil. Different symbols represent populations from four studied forest fragments. Different colors represent tree different clusters found using BIC of DAPC. These clusters were built with samples from all pre-defined populations, indicating that genetic structure was not based on population subdivision. CO–Samples from forest restoration in Cosmópolis, SP; IR–Samples from forest restoration in Iracemápolis, SP; SG–Samples from natural forest in Campinas, SP; TT–Samples from natural forest in Tietê, SP.

these estimates, which consider the set of neutral loci, no population genetic structure was detected. Additionally, DAPC did not show any structure by pre-defined population subdivision (Fig 2).
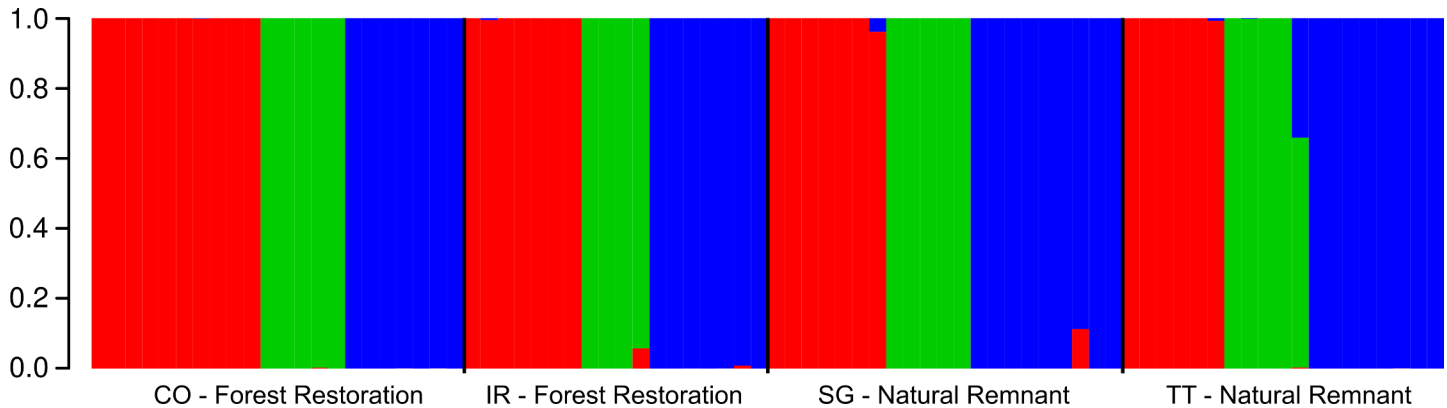
Fig 3. Bar plot of membership probability from Discriminant Analysis of Principal Components for *Casearia sylvestis* populations from restoration plantations and natural remnants in the Atlantic Forest of southeastern Brazil. Different colors represent tree different clusters found using BIC of DAPC. Each clusters is composed of samples from all pre-defined populations; thus, the genetic structure was not based on population subdivision. CO–Samples from forest restoration in Cosmópolis, SP; IR–Samples from forest restoration in Iracemápolis, SP; SG–Samples from natural forest in Campinas, SP; TT–Samples from natural forest in Tietê, SP.

Each of the groups generated based on Bayesian information criterion (BIC) was formed by samples from all populations (Fig 3), reinforcing the lack of genetic structure resulting from population subdivision.

We found no significant positive spatial autocorrelation at any distance interval (S2 Appendix). Moreover, no evidence of spatial genetic structure was detected in the forest fragments (Table 3). The regression slope of the kinship coefficient on the log of distance was not significantly different from zero (P > 0.05). Therefore, individuals located close to one another were not necessarily more genetically related than individuals separated by greater distances.

Discussion

The value of research involving population genetics in a genomic scale using non-model species without previous knowledge of the genome size and complexity has been previously demonstrated in conservation genomics. With the advent of inexpensive sequencing and the advancement of techniques that reduce genome complexity, such as GBS, the study of genetic diversity and structure in natural populations can be based on hundreds of SNPs [56,57]. Here, we demonstrate the usefulness of this approach for restoration genetics.

We found no differences in the estimates of genetic diversity for *C. sylvestris* populations in natural remnants and forest restorations. The similar values between forest restorations and natural remnants are promising results for Brazilian Atlantic Forest conservation, because they indicate the effectiveness of forest restoration in maintaining the genetic diversity of the

Table 3. Estimates of fine-scale spatial genetic structure for all *C. sylvestris* populations.

Pop.	β_{log}	P-value (β_{log})	F_1	Sp
CO	0.0038	0.1968	0.0646	-0.0041
IR	0.0007	0.8641	0.0144	-0.0007
SG	-0.0016	0.5165	-0.0065	0.0015
TT	-0.0025	0.4026	-0.0035	0.0025

β_{log}–Regression slope of kinship coefficient on log of spatial distance; F_1 –Kinship coefficient for the first distance class; *Sp*—Vekemans and Hardy's (2004) estimator of fine-scale SGS intensity. CO–Forest restoration in Cosmópolis, SP; IR–Forest restoration in Iracemápolis, SP; SG–Natural remnant in Campinas, SP; TT—Natural remnant in Tietê, SP.

species, and consequently, the potential to self-perpetuate in the restored communities. Heterozygosity and allelic richness are good indicators of individual fitness in a population [15,58] and represent an initial estimate of the adaptive potential of a population to respond to environmental changes [59]. In previous studies developed by Cavallari et al. (2010) [60] on population relationships among *C. sylvestris* varieties in different ecosystems of Brazil, a similar pattern of genetic diversity was found in almost all populations. Thus, as measured by genetic diversity, the forest restoration plantations studied here were as conserved as forest remnants.

Casearia sylvestris is a pioneer species that is very abundant in disturbed fragments of the region and is highly dispersed by bats and birds, which may promote an intense gene flow between restoration plantations and remnants at the landscape level. Otálora et al. (2011) [24], studying the diversity and genetic structure of Mediterranean forests, suggest that as long as habitat quality and minimum connectivity remain among fragments, forest fragments can preserve the genetic diversity of original forests. In this context, forest restoration plantations can improve landscape connectivity and facilitate gene flow through seeds and pollen, which can mitigate the deleterious effects of inbreeding. In spite of the favorable results for genetic conservation, the studied populations had high F_{IS} values. Despite the mixed mating system of *C. sylvestris*, the differences found for these estimates might be related to the overwhelming level of forest fragmentation, as found in other studies conducted in fragmented landscapes, such as those of Vranckx et al. (2012) [61] and Dick et al. (2003) [62].

For the TT natural remnant, which is a highly fragmented forest, we expect that the reduction of population size in this type of forest and the isolation of the trees in the fragment contributed to the increase in F_{IS}. Andrianoelina et al. (2006) [63], studying the effects of fragmentation on the genetic diversity of an arboreal species in Madagascar, suggest that an increase of F_{IS} in fragmented populations can be explained by high rates of self-fertilization caused by the isolation of adult individuals after population fragmentation.

The results for the general F_{ST} and the pairwise F_{ST} estimates among populations also suggested that the genetic diversity in natural remnants and forest restoration plantations was not structured through population subdivision. The DAPC results are consistent with the F_{ST} estimates and reinforce the potential gene flow among the studied forests and with other fragments of the region. Martins et al. (2016) [64] conducted an important study on the role of small natural remnants of Atlantic Forest for conserving genetic diversity of the tree *Copaifera langsdorffii*, and these authors suggest that small forest fragments preserve the genetic diversity that existed before fragmentation. Therefore, forest fragments are crucial for *in situ* species conservation and the maintenance of long-term genetic diversity in human-modified tropical landscapes. Moreover, by establishing restoration plantations with high genetic diversity, the risk of extinctions of tropical trees threatened by reproductive isolation in fragmented landscapes can be reduced.

To explain the absence of spatial genetic structure, we suggest that seed and pollen dispersal in these populations might play a role [65]. This result is consistent with previous discussions of genetic differentiation and structure for *C. sylvestris* populations in Brazilian Atlantic Forest fragments [66]. Populations of *C. sylvestris*, which is self-compatible, may be experiencing gene flow within and among populations in such a way that genetic diversity is uniformly distributed in space.

Sato et al. (2006) [67], studying the effects of gene flow on SGS for a riparian canopy tree species in Japan, suggested extensive gene flow occurs within and among populations. Furthermore, they argue that high levels of gene flow can increase the local effective size of a population and maintain genetic diversity. Bittencourt and Sebbenn (2007) [68] also suggest that pollen dispersal into forest fragments promotes genetic variation, reduces kinship and inbreeding, and increases the effective population size.

Therefore, we suggest that efforts are required to not only conserve large natural remnants but also the small forest fragments that are most of the remaining remnants in the Brazilian Atlantic Forest [2]. These remnants preserve the genetic diversity that might be important in the establishment and maintenance of new populations under future climate change scenarios. Additionally, we emphasize the importance of forest restoration as a way to connect forest fragments across the landscape, which will contribute to the long-term resilience of tropical forests embedded within human-modified landscapes in the current scenario of global environmental change.

Supporting information

S1 Appendix. Distribution and relative size of the fragments of Atlantic Forest used in the study of diversity and genetic structure of *Casearia sylvestris* Sw. populations of natural remnants and forest restorations.
(PDF)

S2 Appendix. Kinship correlograms (Loiselle et al., 1995) for *Casearia sylvestris* populations from restoration plantations and natural remnants in the Atlantic Forest of southeastern Brazil.
(PDF)

S1 Dataset. Coordinates of all *Casearia sylvestris* Sw. samples used in this study.
(XLSX)

S2 Dataset. SNP data of *Casearia sylvestris* Sw. samples used in this study.
(VCF)

Author Contributions

Conceptualization: João Paulo Gomes Viana, Patricia Sanae Sujii, Marcelo Mattos Cavallari, Pedro H. S. Brancalion, Julian Catchen.

Data curation: João Paulo Gomes Viana, Julian Catchen.

Formal analysis: João Paulo Gomes Viana, Marcos Vinícius Bohrer Monteiro Siqueira, Patricia Sanae Sujii, Julian Catchen.

Funding acquisition: Anete Pereira de Souza, Maria I. Zucchi.

Investigation: João Paulo Gomes Viana, Marcos Vinícius Bohrer Monteiro Siqueira, Fabiano Lucas Araujo, Carolina Grando, Patricia Sanae Sujii, Ellida de Aguiar Silvestre, Mariana Novello, José Baldin Pinheiro, Marcelo Mattos Cavallari, Pedro H. S. Brancalion, Ricardo Ribeiro Rodrigues, Anete Pereira de Souza, Julian Catchen, Maria I. Zucchi.

Methodology: João Paulo Gomes Viana, Marcos Vinícius Bohrer Monteiro Siqueira, Fabiano Lucas Araujo, Carolina Grando, Patricia Sanae Sujii, Ellida de Aguiar Silvestre, Mariana Novello, José Baldin Pinheiro, Julian Catchen.

Project administration: Pedro H. S. Brancalion, Ricardo Ribeiro Rodrigues, Maria I. Zucchi.

Resources: José Baldin Pinheiro, Marcelo Mattos Cavallari, Anete Pereira de Souza.

Software: João Paulo Gomes Viana, Julian Catchen.

Supervision: José Baldin Pinheiro, Marcelo Mattos Cavallari, Pedro H. S. Brancalion, Ricardo Ribeiro Rodrigues, Anete Pereira de Souza, Julian Catchen, Maria I. Zucchi.

Validation: Pedro H. S. Brancalion.

Visualization: Pedro H. S. Brancalion.

Writing – original draft: João Paulo Gomes Viana, Marcos Vinícius Bohrer Monteiro Siqueira, Patricia Sanae Sujii, Marcelo Mattos Cavallari, Pedro H. S. Brancalion, Ricardo Ribeiro Rodrigues, Julian Catchen, Maria I. Zucchi.

References

1. Chazdon RL. Beyond deforesttation: Restoring forest and ecosystem services on degraded lands. Science. 2008; 320: 1458–1460. https://doi.org/10.1126/science.1155365 PMID: 18556551

2. Rodrigues RR, Lima RAF, Gandolfi S, Nave AG. On the restoration of high diversity forests: 30 years of experience in the Brazilian Atlantic Forest. Biol Conserv. 2009; 142: 1242–1251. https://doi.org/10.1016/j.biocon.2008.12.008

3. Holl KD. Restoring tropical forests from the bottom up. Science. 2017; 355: 455–456. https://doi.org/10.1126/science.aam5432 PMID: 28154033

4. Lamb D. Regreening the bare hills. 1st ed. Dordrecht: Springer Netherlands; 2011. 550p. https://doi.org/10.1007/978-90-481-9870-2

5. Holl KD, Aide TM. When and where to actively restore ecosystems? For Ecol Manage. 2011; 261: 1558–1563. https://doi.org/10.1016/j.foreco.2010.07.004

6. Letcher SG, Chazdon RL. Rapid recovey of biomass, species richness and species composition in a forest chronosequence in Northeastern Costa Rica. Biotropica. 2009; 41: 608–617. https://doi.org/10.1111/j.1744-7429.2009.00517.x

7. de Rezende CL, Uezu A, Scarano FR, Araujo DSD. Atlantic Forest spontaneous regeneration at landscape scale. Biodivers Conserv. 2015; 24: 2255–2272. https://doi.org/10.1007/s10531-015-0980-y

8. Kimball S, Lulow M, Sorenson Q, Balazs K, Fang YC, Davis SJ, et al. Cost-effective ecological restoration. Restor Ecol. 2015; 23: 800–810. https://doi.org/10.1111/rec.12261

9. Brancalion PHS, Viani RAG, Aronson J, Rodrigues RR, Nave AG. Improving Planting Stocks for the Brazilian Atlantic Forest Restoration through Community-Based Seed Harvesting Strategies. Restor Ecol. 2012; 20: 704–711. https://doi.org/10.1111/j.1526-100X.2011.00839.x

10. Arroyo-Rodríguez V, Melo FPL, Martínez-Ramos M, Bongers F, Chazdon RL, Meave JA, et al. Multiple successional pathways in human-modified tropical landscapes: new insights from forest succession, forest fragmentation and landscape ecology research. Biol Rev. 2017; 92: 326–340. https://doi.org/10.1111/brv.12231 PMID: 26537849

11. Thomas E, Jalonen R, Loo J, Boshier D, Gallo L, Cavers S, et al. Genetic considerations in ecosystem restoration using native tree species. For Ecol Manage. 2014; 333: 66–75. https://doi.org/10.1016/j.foreco.2014.07.015

12. Mijangos JL, Pacioni C, Spencer PBS, Craig MD. Contribution of genetics to ecological restoration. Mol Ecol. 2015; 24: 22–37. https://doi.org/10.1111/mec.12995 PMID: 25377524

13. Frankham R, Ballou JD, Briscoe DA. Introduction to Conservation Genetics. 2nd ed. Cambridge: Cambridge University Press; 2010. 644p.

14. Bell G, Collins S. Adaptation, extinction and global change. Evol Appl. 2008; 1: 3–16. https://doi.org/10.1111/j.1752-4571.2007.00011.x PMID: 25567487

15. Reed DH, Frankham R. Correlation between Fitness and Genetic Diversity. Conserv Biol. 2003; 17: 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

16. Kahilainen A, Puurtinen M, Kotiaho JS. Conservation implications of species-genetic diversity correlations. Glob Ecol Conserv. 2014; 2: 315–323. https://doi.org/10.1016/j.gecco.2014.10.013

17. Toro MA, Caballero A. Characterization and conservation of genetic diversity in subdivided populations. Philos Trans R Soc B Biol Sci. 2005; 360: 1367–1378. https://doi.org/10.1098/rstb.2005.1680 PMID: 16048780

18. Ganzhorn SM, Perez-Sweeney B, Thomas WW, Gaiotto FA, Lewis JD. Effects of fragmentation on density and population genetics of a threatened tree species in a biodiversity hotspot. Endanger Species Res. 2015; 26: 189–199. https://doi.org/10.3354/esr00645

19. Schaberg PG, DeHayes DH, Hawley GJ, Nijensohn SE. Anthropogenic alterations of genetic diversity within tree populations: Implications for forest ecosystem resilience. For Ecol Manage. 2008; 256: 855–862. https://doi.org/10.1016/j.foreco.2008.06.038

20. Kettenring KM, Mercer KL, Reinhardt Adams C, Hines J. Application of genetic diversity-ecosystem function research to ecological restoration. J Appl Ecol. 2014; 51: 339–348. https://doi.org/10.1111/1365-2664.12202

21. Ritchie AL, Krauss SL. A Genetic assessment of ecological restoration success in banksia attenuata. Restor Ecol. 2012; 20: 441–449. https://doi.org/10.1111/j.1526-100X.2011.00791.x

22. Kettle CJ, Ennos RA, Jaffré T, Gardner M, Hollingsworth PM. Cryptic genetic bottlenecks during restoration of an endangered tropical conifer. Biol Conserv. 2008; 141: 1953–1961. https://doi.org/10.1016/j.biocon.2008.05.008

23. Burgarella C, Navascués M, Soto Á, Lora Á, Fici S. Narrow genetic base in forest restoration with holm oak (Quercus ilex L.) in Sicily. Ann For Sci. 2007; 64: 757–763. https://doi.org/10.1051/forest:2007055

24. Otálora MG, Martínez I, Belinchón R, Widmer I, Aragón G, Escudero A, et al. Remnants fragments preserve genetic diversity of the old forest lichen Lobaria pulmonaria in a fragmented Mediterranean mountain forest. Biodivers Conserv. 2011; 20: 1239–1254. https://doi.org/10.1007/s10531-011-0025-0

25. Lloyd H, Marsden SJ. Between-Patch Bird Movements within a High-Andean Polylepis Woodland/Matrix Landscape: Implications for Habitat Restoration. Restor Ecol. 2011; 19: 74–82. https://doi.org/10.1111/j.1526-100X.2009.00542.x

26. Leidner AK, Haddad NM. Combining Measures of Dispersal to Identify Conservation Strategies in Fragmented Landscapes. Conserv Biol. 2011; 25: 1022–1031. https://doi.org/10.1111/j.1523-1739.2011.01720.x PMID: 21848964

27. Sezen UU. Genetic Consequences of Tropical Second-Growth Forest Regeneration. Science. 2005; 307: 891–891. https://doi.org/10.1126/science.1105034 PMID: 15705843

28. Etter PD, Johnson E. RAD Paired-End Sequencing for Local De Novo Assembly and SNP Discovery in Non-model Organisms. In: Methods in molecular biology (Clifton, NJ). 2012. p. 135–151. https://doi.org/10.1007/978-1-61779-870-2_9 PMID: 22665280

29. Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet. 2003; 4: 981–994. https://doi.org/10.1038/nrg1226 PMID: 14631358

30. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011; 6: 1–10. https://doi.org/10.1371/journal.pone.0019379 PMID: 21573248

31. Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. Harnessing genomics for delineating conservation units. Trends Ecol Evol. 2012; 27: 489–496. https://doi.org/10.1016/j.tree.2012.05.012 PMID: 22727017

32. Tabarelli M, Villani JP, Mantovani W. Aspectos da sucessão secundária em trecho da Floresta Atlântica no Parque Estadual da Serra do Mar, SP. Rev do Inst Florestal, São Paulo [Internet]. 1993; 5: 95–112. Available from: http://www.iflorestal.sp.gov.br/RIF/RevistaIF/RIF5-1/RIF5-1_99-112.pdf

33. Silva AF Da, Oliveira RV De, Santos NRL, Paula A De. Composição florística e grupos ecológicos das espécies de um trecho de floresta semidecídua submontana da Fazenda São Geraldo, Viçosa-MG. Rev Árvore. 2003; 27: 311–319. https://doi.org/10.1590/S0100-67622003000300006

34. Marmontel CVF, Rodrigues VA, Martins E Martins T, Starzynski R, de Carvalho JL. Caracterização da vegetação secundária do bioma mata atlântica com base em sua posição na paisagem. Biosci J. 2013; 29: 2042–2052.

35. DIAS MC, VIEIRA AOS, NAKAJIMA JN, PIMENTA JA, LOBO PC. Composição florística e fitossociologia do componente arbóreo das florestas ciliares do rio Iapó, na bacia do rio Tibagi, Tibagi, PR. Rev Bras Botânica. 1998; 21: 183–195. https://doi.org/10.1590/S0100-84041998000200011

36. Sleumer HO. Flacourtiaceae. In: Flora Neotropica. New York: The New York Botanical Garden; 1980.

37. Ramalho M. Stingless bees and mass flowering trees in the canopy of Atlantic Forest: a tight relationship. Acta Bot Brasilica. 2004; 18: 37–47. https://doi.org/10.1590/S0102-33062004000100005

38. Barbosa AAA. Biologia reprodutiva de uma comunidade de Campo Sujo, Uberlandia/MG. D. Sc. Thesis. Universidade Estadual de Campinas; 1997. Available from: http://repositorio.unicamp.br/jspui/handle/REPOSIP/316111

39. Wanderley M das GL, Shepherd GJ, Giulietti AM, Melhem TS, Kameyama C, Bittrich V. Flora Fanerogâmica do Estado de São Paulo. Vol. 2. 2002. 386p. Available from: http://botanica.sp.gov.br/ffesp_online/

40. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. Biodiversity hotspots for conservation priorities. Nature. 2000; 403: 853–858. https://doi.org/10.1038/35002501 PMID: 10706275

41. SOS Mata Atlantica, INPE. Atlas dos remanescentes florestais da Mata Atlântica período 2013–2014. Instituto Nacional de Pesquisas Espaciais. 2015. Available from: http://mapas.sosma.org.br/site_media/download/atlas_2013-2014_relatorio_tecnico_2015.pdf

42. Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FJ, Hirota MM. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. Biol Conserv. 2009; 142: 1141–1153. https://doi.org/10.1016/j.biocon.2009.02.021

43. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987; 19: 11–15.

44. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: An analysis tool set for population genomics. Mol Ecol. 2013; 22: 3124–3140. https://doi.org/10.1111/mec.12354 PMID: 23701397

45. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: A workbench to detect molecular adaptation based on a F st -outlier method. BMC Bioinformatics. 2008; 9: 1–5.

46. Paquette SR. PopGenKit: Useful functions for (batch) file conversion and data resampling in microsatellite datasets. 2012. Available from: https://cran.r-project.org/package=PopGenKit

47. Goudet J, Jombart T. hierfstat: Estimation and Tests of Hierarchical F-Statistics. 2015. Available from: https://cran.r-project.org/package=hierfstat

48. Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. 2016. Available from: https://cran.r-project.org/package=VennDiagram

49. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available from: https://www.r-project.org

50. Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol. 2013; 4: 782–788. https://doi.org/10.1111/2041-210X.12067

51. Jombart T. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24: 1403–1405. https://doi.org/10.1093/bioinformatics/btn129 PMID: 18397895

52. Jombart T, Ahmed I. adegenet 1.3–1: New tools for the analysis of genome-wide SNP data. Bioinformatics. 2011; 27: 3070–3071. https://doi.org/10.1093/bioinformatics/btr521 PMID: 21926124

53. Loiselle BA, Sork VL, Nason J, Graham C. Spatial Genetic Structure of a Tropical Understory Shrub, Psychotria officinalis (Rubiaceae). Am J Bot. 1995; 82: 1420–1425. Available from: http://www.jstor.org/stable/2445869

54. Vekemans X, Hardy OJ. New insights from fine-scale spatial genetic structure analyses in plant populations. Mol Ecol. 2004; 13: 921–935. https://doi.org/10.1046/j.1365-294X.2004.02076.x PMID: 15012766

55. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2002; 2: 618–620. https://doi.org/10.1046/j.1471-8286.2002.00305.x

56. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet. 2016; 17: 81–92. https://doi.org/10.1038/nrg.2015.28 PMID: 26729255

57. Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA. Genotyping-by-sequencing in ecological and conservation genomics. Mol Ecol. 2013; 22: 2841–2847. https://doi.org/10.1111/mec.12350 PMID: 23711105

58. Greenbaum G, Templeton AR, Zarmi Y, Bar-David S. Allelic richness following population founding events—A stochastic modeling framework incorporating gene flow and genetic drift. PLoS One. 2014; 9: 1–23. https://doi.org/10.1371/journal.pone.0115203 PMID: 25526062

59. Caballero A, García-Dorado A. Allelic diversity and its implications for the rate of adaptation. Genetics. 2013; 195: 1373–1384. https://doi.org/10.1534/genetics.113.158410 PMID: 24121776

60. Cavallari MM, Gimenes MA, Billot C, Torres RB, Zucchi MI, Cavalheiro AJ, et al. Population genetic relationships between Casearia sylvestris (Salicaceae) varieties occurring sympatrically and

allopatrically in different ecosystems in south-east Brazil. Ann Bot. 2010; 106: 627–636. https://doi.org/10.1093/aob/mcq151 PMID: 20699231

61. Vranckx G, Jacquemyn H, Muys B, Honnay O. Meta-Analysis of Susceptibility of Woody Plants to Loss of Genetic Diversity through Habitat Fragmentation. Conserv Biol. 2012; 26: 228–237. https://doi.org/10.1111/j.1523-1739.2011.01778.x PMID: 22044646

62. Dick CW, Etchelecu G, Austerlitz F. Pollen dispersal of tropical trees (Dinizia excelsa: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. Mol Ecol. 2003; 12: 753–764. https://doi.org/10.1046/j.1365-294X.2003.01760.x PMID: 12675830

63. Andrianoelina O, Rakotondraoelina H, Ramamonjisoa L, Maley J, Danthu P, Bouvet JM. Genetic diversity of Dalbergia monticola (Fabaceae) an endangered tree species in the fragmented oriental forest of Madagascar. Biodivers Conserv. 2006; 15: 1109–1128. https://doi.org/10.1007/978-1-4020-5208-8_4

64. Martins K, Kimura RK, Francisconi AF, Gezan S, Kainer K, Christianini A V. The role of very small fragments in conserving genetic diversity of a common tree in a hyper fragmented Brazilian Atlantic forest landscape. Conserv Genet. 2016; 17: 509–520. https://doi.org/10.1007/s10592-015-0800-7

65. Hamrick JL, Nason JD. Consequence of dispersal in plants. In: Rhodes OE, Ronald KC, Smith MH, editors. Population Dynamics in Ecological Space and Time. Chicago; 1996. p. 203–235.

66. Araujo FL, Siqueira MVBM, Grando C, Viana JPG, Pinheiro JB, Alves-Pereira A, et al. Genetic diversity of Casearia sylvestris populations in remnants of the Atlantic Forest. Genet Mol Res. 2017; 16: 4393–4396. https://doi.org/10.4238/gmr16019105 PMID: 28128404

67. Sato T, Isagi Y, Sakio H, Osumi K, Goto S. Effect of gene flow on spatial genetic structure in the riparian canopy tree Cercidiphyllum japonicum revealed by microsatellite analysis. Heredity. 2006; 96: 79–84. https://doi.org/10.1038/sj.hdy.6800748 PMID: 16304606

68. Bittencourt JVM, Sebbenn a M. Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree Araucaria angustifolia in southern Brazil. Heredity. 2007; 99: 580–591. https://doi.org/10.1038/sj.hdy.6801019 PMID: 17928810