

Genetic structure and diversity of populations of polyploid *Tibouchina pulchra* Cogn. (Melastomataceae) under different environmental conditions in extremes of an elevational gradient

Vinícius L. G. Brito^{1,2,3} · Gustavo M. Mori^{3,4,5} · Bianca B. Z. Vigna^{3,6} ·
Marianne Azevedo-Silva⁷ · Anete P. Souza^{3,8} · Marlies Sazima⁸

Received: 29 November 2015 / Revised: 3 September 2016 / Accepted: 26 September 2016 / Published online: 7 October 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The genetic structure and diversity of plants may change significantly in an elevational gradient because different elevations regulate different ecological conditions. Several factors may influence genetic variation, such as mutations, selection, genetic drift, and gene flow. The aim of the present study was to evaluate the genetic structure and diversity of populations of *Tibouchina pulchra* Cogn. (Melastomataceae) trees in two extremes of an elevational gradient experiencing different environmental conditions. Nine polymorphic microsatellite loci were used to measure the genetic diversity of 14 adult populations, whose structure was evaluated using frequentist and Bayesian analyses. We also carried out progeny structure and paternity analyses comparing the number of fathers of each progeny and the probability of the progeny genotypes to be the result of selfing in order to evaluate the possible current processes leading to such genetic structure. Genetic structure analyses indicated the existence of genetic differentiation between populations in adults and progenies,

but with a contact interface between them. The population from the higher region showed smaller genetic diversity when compared to the population at the lower region. However, the pollen variability delivered to the stigmas at the higher region was not different from that of the lower region. These results may be explained by the dynamics of gene flow mediated by pollen, especially by the different amounts of pollination events in each region, as well as local adaptation, distribution, and reproduction characteristics of *T. pulchra*.

Keywords Polyploidy · Pollen flow · Atlantic rainforest · Serra do Mar · Manacá-da-Serra

Introduction

Plant populations are not a set of genotypes randomly structured in time and space. In general, there is a genetic

Communicated by D. Grattapaglia

Electronic supplementary material The online version of this article (doi:10.1007/s11295-016-1059-y) contains supplementary material, which is available to authorized users.

✉ Vinícius L. G. Brito
viniciusduartina@gmail.com

¹ Institute of Biology, Federal University of Uberlândia, Uberlândia, Minas Gerais, Brazil

² Programa de Pós Graduação em Biologia Vegetal, Laboratório de Biossistemática, Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, Brazil

³ Center for Molecular Biology and Genetic Engineering (CBMEG), University of Campinas, Campinas, São Paulo, Brazil

⁴ São Paulo Agency for Agribusiness Technology, Piracicaba, São Paulo, Brazil

⁵ Present address: Biosciences Institute - Coastal Campus - São Paulo State University - UNESP, São Vicente, São Paulo, Brazil

⁶ EMBRAPA Southeast Livestock, Brazilian Agricultural Research Corporation, São Carlos, São Paulo, Brazil

⁷ Programa de Pós Graduação em Ecologia, Department of Animal Biology, Institute of Biology, University of Campinas, Campinas, São Paulo, Brazil

⁸ Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, Brazil

organization among geographically distinct populations, within local plant groups or even in the progeny of individuals (Loveless and Hamrick 1984). In this sense, several factors may influence the amount of genetic structure and diversity of plant populations, such as mutations, selection, genetic drift, and gene flow, which may be considered the most important factor in large populations (Loveless and Hamrick 1984; Slatkin 1987). In plants, the gene flow may occur only in two life history moments: pollen and seed dispersal. Therefore, pollination and seed dispersal dynamics have direct implications on the genetic structure of populations, and both processes have been considered the causes of intraspecific genetic diversity in most of angiosperm species and, consequently, the substrate for the high species diversity on this plant group (Bawa 1995).

The lack of gene flow can decrease the population effective size favoring genetic drift, decreasing genetic diversity in small populations, and possibly leading to inbreeding depression (Ellstrand 1992; Charlesworth and Charlesworth 1987). In turn, such loss of genetic diversity may limit the potential of natural populations to adapt to a changing world (Willi et al. 2006; Byars et al. 2009). By contrast, high levels of genetic diversity, generated by the movement of pollen and/or seeds and the consequent exchange of genes between populations, enable a great number of genotypic combinations, increasing the evolutionary potential of the species due to the greater adaptation capacities to possible environmental changes (Sebbenn et al. 2000).

Different ecological conditions, in particular habitats along elevational gradients (e.g., pollinator availability, soil quality, temperature), may influence significantly the genetic structure and diversity of plants by local adaptation and drift (Thomas et al. 2001; Semagn et al. 2000; Yan et al. 2009; Misiewicz and Fine 2014). Populations established in different elevations may also present asynchrony in their flowering and fruiting phenology, which may also favor genetic structure (Jordano and Godoy 2000). In this sense, it has been demonstrated for several species of plants occurring along an elevational gradient that the genetic diversity can increase, decrease, or even be constant along the altitude (Yan et al. 2009; Ohsawa and Ide 2008; Zhao et al. 2006; Semagn et al. 2000; Reis et al. 2015). Some of these studies suggest that the restricted gene flow between populations with strong natural selection influence the genetic diversity of plants in such places (Thomas et al. 2001; Slatkin 1987). Therefore, plants that occur in elevational gradients and rely on pollinator activity for their reproduction are ideal models to understand the role of pollen flow on genetic structure and diversity of populations, and the Melastomataceae family presents some species that match these biological features.

Melastomataceae comprehends 4200–4500 species in total and they are well represented in tropical and subtropical ecosystems of the Americas, where around 3000 of these species are found (Renner 1993). There are 68 genera in Brazil, with estimations of over 1300 species (Goldenberg et al. 2012; Renner 1993). Their flowers are habitually pollinated by bees

which collect pollen from the often poricidal tubular anthers. During a flower visit, the bees embrace the anthers and perform vibration movements with their wing muscles. Pollen grains are expelled and adhere to the bee body, a place frequently touched by the flower stigmas. This type of pollination is called buzz pollination and is frequently found in the genus *Tibouchina* (Buchmann 1983; Brito and Sazima 2012). Consequently, in such systems, gene flow by pollen is almost restricted to the movement of the pollinators due to pronounced flower herkogamy (Renner 1989). On the other hand, because *Tibouchina* produces capsular fruits, seed dispersal in this genus is almost restricted to autochory, that is, with no apparent dispersal vector (Silveira et al. 2013). Such dispersal system would possibly reduce the effects of seed dispersal on gene flow. Estimating genetic structure and diversity of *T. pulchra* Cogn. (Melastomataceae) is particularly challenging because there is evidence that this species is polyploid (Brito et al. 2010) making the definition of the allelic constitution of a specific locus a methodological issue which, in most of the cases, precludes the use of allele frequency-based methods (Dufresne et al. 2014). Therefore, specific analytical tools are required to minimize eventual biases and circumventing the complexity that arises from genotype ambiguity of partial heterozygotes.

The aim of this study is to evaluate the genetic structure and diversity of *T. pulchra* populations, whose gene flow is basically driven by pollinators' movement, occurring at two extremes of an elevational gradient of a tropical forest, considering the polyploid nature of its genome. As the individuals of the higher region in the Atlantic Rainforest (Núcleo Santa Virgínia of Parque Estadual da Serra do Mar—NSV) experience different climate and ecological conditions (Reis et al. 2015) and are more pollinator-limited than populations occurring at the lower region (Núcleo Picinguaba of the same conservation unit—NP) (Brito and Sazima 2012), we expect to find the following: (1) genetic differentiation of populations between regions and (2) lower genetic diversity in the populations occurring at the higher region. If the processes generating such genetic structure and diversity pattern is the low pollen diversity delivered to the stigmas at NSV as compared to NP due to pollinator limitation and higher selfing rates, we also expect that (3) in comparison to NP, in NSV, the number of fathers and the probability of selfing in each progeny would be smaller and greater, respectively.

Materials and methods

Study area: location and characterization

The study was developed at the Parque Estadual da Serra do Mar, at NSV and NP regions, both situated in the northern region of São Paulo state coast (Fig. 1). The NSV is located in the region of scarps and reverse faults of the Serra do Mar

mountain range, on the Paraitinga-Paraibuna plateau, close to the city of São Luís do Paraitinga—SP, between altitudes that vary between 870 and 1110 m (Tabarelli and Mantovani 1999). The vegetation is classified as montane Atlantic rainforest (Veloso et al. 1991), and the regional climate is subtropical humid without a dry season and with temperate summer (Alvares et al. 2014). Average annual precipitation is 2180 mm, and December, January, and February are the most humid months, while June, July, and August are the less humid months. Average precipitation is never inferior to 60 mm per month (Tabarelli and Mantovani 1999). The NP is located on the coastal plain, in the municipality of Ubatuba—SP, covering the coastline and the foothills of the Serra do Mar mountain range. The vegetation in the area is classified as dense rainforest (Veloso et al. 1991). The climate is tropical rainy and presents a super humid season from October to April, with an average precipitation superior to 200 mm per month and a drier season from May to September, with an average monthly precipitation of 100 mm (Morellato et al. 2000). The average annual rainfall is 2100 mm and the average annual temperature is around 22 °C (Bencke and Morellato 2002).

Studied species

The *Tibouchina* genus (Melastomataceae) is amply distributed in the Neotropics and has the greatest number of species within the group of capsular fruit Melastomataceae, with around 240 species (Renner 1993). It is found from East Mexico to the Northeast Argentina and Paraguay (Todzia and Almeida 1991; Guimarães and Martins 1997). *T. pulchra* Cong., commonly known as “manacá-da-serra” in Brazil, is a pioneer polyploid tree species that has flowers displaying two colors: new white flowers presenting a pleasant and weak perfume characteristic to melittophilous flowers (sensu Faegri and Van der Pijl 1979) and old pink flowers that are scentless and rarely receive visits (Pereira et al. 2011; Brito and Sazima 2012, Brito et al. 2015). The new flowers possess five large stamens and five small ones, presenting poricidal anthers. Pollen grains are the only reward offered by flowers and it is removed only by buzz pollination (Buchmann 1983). *T. pulchra* is self-compatible, but does not self-pollinate, being their pollinators necessary for seed production. Its fruits are capsular and they set over than 3000 seeds, which are dispersed autochorically and germinate after 15 days (Brito and Sazima 2012).

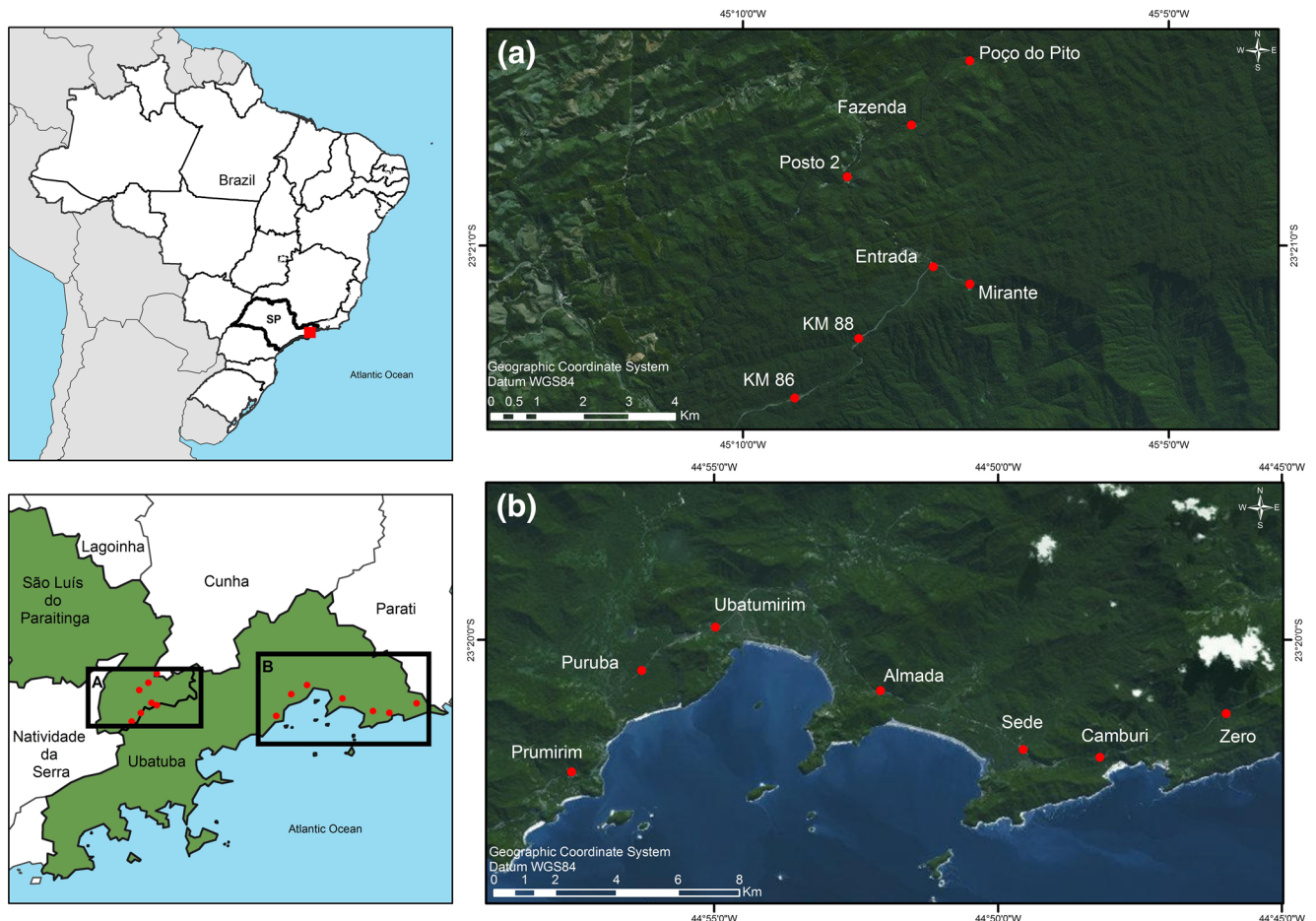


Fig. 1 Map with population locations for *Tibouchina pulchra* in the Parque Estadual da Serra do Mar—SP. **a** Populations in the upper region at Núcleo Santa Virgínia (NSV). **b** Populations in lower region at Núcleo Picinguaba (NP). Each population was composed of 15 adult individuals

Despite the unknown ploidy level of *T. pulchra*, its polyploid nature has already been observed (Brito et al. 2010). Almeda and Chuang (1992) pointed out that tetraploidy is the most common condition in the genus *Tibouchina*, followed by diploidy and hexaploidy.

DNA extraction, amplification, and genotyping

Young leaves from 210 adult individuals (105 individuals from NSV and NP each) distributed in 14 populations between the two regions were collected using sterile scissors and stored in the Biofreezer at -80°C . As *T. pulchra* is self-dispersed, we chose individuals at least 10 m apart from each other to avoid oversampling relatives. We sampled individuals occurring in seven distinct populations about 2 km apart from each other within each region (Table 1, Fig. 1a, b). Subsequently, during the fruiting season, we chose some of these adults to collect one or two fruits assigning the region, the population, and the mother for each one, carrying to distribute the sampling in all the cardinal directions. From each fruit collected, we germinated their seeds and sampled from ten to 21 progeny individuals when the seedlings reached a size enough to perform deoxyribonucleic acid (DNA) extraction (Table 1).

Total genomic DNA was extracted from the leaf tissues using the DNeasy Plant Mini Kit (QIAGEN). Out of the 12 microsatellites previously developed for *T. pulchra* (Brito

et al. 2010), we only used the markers (TP01, TP03, TP05, TP12, TP20, TP23, TP25, TP27, TP29) that were reliable for the entire sample set. The loci were amplified using polymerase chain reaction (PCR) carried out in a final volume of 10 μL containing 1.5 ng of template DNA, $1\times$ PCR buffer, 3-mM magnesium chloride, 0.2 μM of each dNTP, 0.1 μM of each primer, 0.1 μM of 700 or 800 nm Infrared Dyes (Li-Cor Biosciences, Lincoln, NE, USA) and 1 U Taq DNA polymerase. We followed three touchdown protocols (Don et al. 1991), according to the thermo-cycling conditions: 94°C for 4 min; $10\times$ [94°C for 45 s, 60°C ($-0.5^{\circ}\text{C}/\text{cycle}$) or 57°C ($-0.5^{\circ}\text{C}/\text{cycle}$) or 63°C ($-0.5^{\circ}\text{C}/\text{cycle}$) for 1 min and 72°C for 1 min 15 s]; $25\times$ [94°C for 45 s, 50°C for 1 min and 72°C for 1 min 15 s]; and 72°C for 10 min. In each forward primer, a M13 tail (5'- CACGACGTTGTAACGAC - 3') was added at its 5' end (Schuelke 2000), which enabled the amplified microsatellite fragments to be scored on 4300 DNA Analyzer (Li-Cor Biosciences, Lincoln, NE, USA). Amplification products sizes were determined using SagaTM software (Li-Cor) and used for further analyses. As the majority of individuals presented a maximum of four bands per marker, a typical tetraploid pattern, we considered *T. pulchra* to be tetraploid in all analyses. Therefore, we excluded the phenotype information when it presented more than four bands per locus (only three individuals at one microsatellite). The presence of more than two alleles per

Table 1 Adult and progeny location and sample size (n) of progeny of *Tibouchina pulchra* situated in two regions: Núcleo Santa Virgínia (NSV) and Núcleo Picinguaba (NP), in the Parque Estadual da Serra do Mar

Area	Population	Location (latitude/longitude)	Designation of adults with collected fruits	Fruits from each adult (n) ^a	Seedlings from each fruit (n) ^b
NSV	Entrada	$-23.35/-45.13$	S146 + S157	2 + 2	(11 + 12) + (20 + 20)
NSV	Fazenda	$-23.33/-45.13$	S198 + S195	2 + 1	(20 + 20) + (15)
NSV	Km 88	$-23.37/-45.14$	S138	1	(14)
NSV	Km 86	$-23.38/-45.16$	S129 + S120	2 + 1	(20 + 20) + (20)
NSV	Mirante	$-23.36/-45.12$	S101 + S113	2 + 1	(21 + 21) + (21)
NSV	Poço do Pito	$-23.34/-45.12$	S183 + S189	1 + 1	(20) + (11)
NSV	Posto II	$-23.34/-45.15$	S169	1	(20)
NP	Almada	$-23.35/-44.87$	P184 + P188	1 + 2	(10) + (20 + 14)
NP	Camburi	$-23.37/-44.80$	P175	1	(20)
NP	Promirim	$-23.37/-44.96$	P101	2	(20 + 20)
NP	Puruba	$-23.35/-44.94$	P121 + P129	2 + 1	(20 + 19) + (20)
NP	Sede	$-23.36/-44.83$	P202	1	(16)
NP	Ubatumirim	$-23.33/-44.92$	P139 + P136	2 + 2	(20 + 20) + (20 + 17)
NP	Zero	$-23.35/-44.77$	NA	NA	NA

In each population, we collected total DNA from 15 adult individuals

NA not applicable

^a Each number corresponds to the number of fruits collected from the same adults from the previous column, respectively; + signal is used to separate different adult origins

^b Each number corresponds to the number of seedlings collected from each fruit from the previous column, respectively; parenthesis indicates different adult origins and + signal indicates different fruit origins

individual associated with the complex inheritance patterns in polyploids does not allow for the direct calculation of allele frequency and impedes the utilization of traditional methods to measure genetic variability (De Silva et al. 2005). Some methods were developed to deal with partial heterozygote's genotype ambiguity (for example, Falush et al. 2007) and, alternatively, polymorphic markers can be utilized as dominant markers in ecological and evolutionary studies involving polyploids (Byrne et al. 2008; Caddah et al. 2009; Dufresne et al. 2014). In this sense, in AMOVA, DAPC, Mantel, and COLONY analyses (see below), we converted each phenotype at a k-allele codominant locus to diploid phenotypes at k dominant "loci" (Wang and Scribner 2014). On the other hand, we used the same polymorphic markers as codominant with genotype ambiguity in partial heterozygotes in STRUCTURE and Fdash analyses (see below).

Genetic structure of adults and progeny

Analysis of molecular variance (AMOVA) (Excoffier et al. 1992) was performed using the ade4 1.6–2 R package (Dray et al. 2007) to know how genetic variation is partitioned between regions (NSV and NP), among populations within regions (the 14 sampling localities), and within each populations of adults. The same approach was employed to partition the genetic variation between regions (NSV and NP), among progenies within regions (from the 21 sampling mother trees), and within each progeny from the same mother. The significance of the proportion of variation at each category was obtained by a Monte Carlo test with 1000 permutations.

We also tested the correlation between genetic and geographic distances among adult individuals using Mantel tests (Legendre and Legendre 1998). The genetic distance was calculated by the Lynch distance, which is two times the number of bands that two phenotypes have in common, divided by the total number of bands that the two genotypes have (Lynch 1990). The significance of correlations between both distances considering all the adults and the adults in each region was estimated with 10,000 permutations of the rows and columns of the geographic distance matrix.

The genetic relationship among all adult individuals, as well as among progeny individuals, and genetic clustering was accessed through a discriminant analysis on principal components (DAPC) (Jombart et al. 2010) implemented in adegenet R package (Jombart 2008). DAPC maximizes the variance between groups while minimizes the variance within groups, using as the input the principal components (PCs) generated from the genotypes of individuals (Jombart et al. 2010). As the number of retained PCs for the analysis can bias the fit of DAPC, we chose the number of principal components using the a-score function, which is the difference between the proportion of successful reassignment of the analysis (observed discrimination) and values obtained using

random groups (random discrimination). We ran this analysis considering the regions and the populations within each region as groups for adults and the regions and the mothers within each region as groups for the progeny.

The evaluation of the genetic structure of the adult populations and progeny was also made through Bayesian inference performed with the software STRUCTURE 2.2.3 considering an admixture model with correlated allele frequency (Pritchard et al. 2000; Falush et al. 2003) accounting for genotype ambiguity for codominant markers in polyploids (Falush et al. 2007). The individuals were assigned to possible genetic groups, K, varying from two to ten without a priori definition of populations and disregarding recessive alleles. Fifty independent Markov Chain Monte Carlo (MCMC) runs were carried out with 500,000 iterations following a burn-in period of 500,000 iterations for each value of the number of clusters (K). The best estimate of K was calculated from the ad hoc statistic delta K, proposed by Evanno et al. (2005). The software CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) was used to deal with label switching and multimodality issues.

Genetic diversity of adult populations

The following estimators of the genetic diversity on each locus in adult populations occurring at the two regions were calculated using the FDASH program (Obbard et al. 2006): total number of alleles, average number of unshared alleles between pairs of individuals, average number of different alleles carried by each individual, total number of phenotypes, average Shannon-Weaver phenotype diversity, and average number of shared alleles between pairs of individuals. These estimations were calculated through the average number of alleles by which pairs of individuals differ. These estimates are ideal to allopolyploids (Obbard et al. 2006), as seems to be the case of *T. pulchra*. An analysis of variance (ANOVA) was performed at 95 % confidence interval using each estimator as response and the regions, populations, and locus as factors in R 2.15.0 software (<http://www.r-project.org/>).

Paternity analyses

Following Wang and Scribner (2014), we estimated the number of fathers of each progeny within a fruit and the probability of such progeny genotypes to be a result of a selfing event using the software COLONY version 2.0.6.1 (Jones and Wang 2010). For each progeny (fruit), we set the following parameters given the nature of *T. pulchra* and the experimental design: mating system I: monogamous male and polygamous female; mating system II: with inbreeding and without clones; species: monoecious and diploid (following Wang and Scribner 2014); analysis method: full likelihood (FL). All the other parameters were set to default. As *T. pulchra* is

self-compatible (Brito and Sazima 2012), we stipulated a probability of 0.5 to the mother genotype be a father candidate of the progeny. From the estimation of the number of fathers per fruit, we adjusted a generalized linear mixed-effect model with the Poisson distribution family in R, considering the number of fathers per fruit as the response, the region as fixed factor, and the population and mother as random factors. The same rationale was applied to analyze the selfing probability, but we used a beta distribution since it better describes the variation of values between 0 and 1 (Cribari-Neto and Zeileis 2010). In this analysis, we considered the selfing probability as the response, the region as fixed factor, and the population, mother, and fruit as random factors.

Results

Although leaves from 105 adults were collected in each region, one adult from NP region was lost, resulting in a total of 209 trees and 562 progeny individuals genotyped. The progenies were collected from 31 fruits in 21 adults (mothers) from different populations (Table 1). The total number of bands per locus varied from four to 25, while the number of bands per individual per locus varied from one to four, a multiband pattern typical of polyploid organisms. Six of the nine loci presented more than 90 % of the individuals with disomic-inheritance band pattern with one or two band per individual, while three loci presented polysomic-inheritance multiband pattern, with more than 10 % of the genotyped individuals with more than two bands.

Genetic structure of adults and progeny

AMOVA indicated that most of the variation (91.55 %) is within adult populations, while 5.94 % varies among populations and 2.51 % between regions (Table 2). The Monte Carlo test indicated that this variation within populations is lower than the expected by chance, whereas it is greater than the expected by chance among populations and between regions ($p < 0.001$ in each test). When considering the progeny, AMOVA indicated the same pattern, showing the most of variation within progeny populations (65.56 %), and 31.67 % of

variation among progenies and 2.77 % between regions. The Monte Carlo test was also significant ($p < 0.001$) in each test and the variation was lower than the expected by chance within population, while it was greater than the expected by chance among progeny and between regions (Table 3).

We observed a significant positive correlation between genetic and geographic distances when considering all adult individuals ($r = 0.17$, mantel- $p < 0.001$). However, such correlation holds true for the lower region ($r = 0.13$, mantel- $p < 0.001$), whereas it was only marginally significant at the higher region ($r = 0.05$, mantel- $p = 0.04$).

The DAPC discriminated 88 % of the adult individuals from NSV and 86 % of the individuals from NP, considering the region as groups and a retention of 24 PCs given by the a-score (Fig. 2a). However, 26 % of total individuals were associated with both regions indicating the existence of a contact interface between the populations from higher and lower regions. A similar result was found using the populations as groups and a retention of 21 PCs as suggested by the a-score: the populations from each region grouped together but with a contact interface between them (Supplementary Fig. 1). Considering the progeny, DAPC discriminated 97 % of progeny individuals from NSV and 95 % of progeny individuals from NP, considering the region as groups and a retention of 44 PCs given by the a-score (Fig. 2b). Only 8 % of progeny individuals were associated with both regions. When considering the progeny as groups, the progenies of each region became evident only on the axis of the third discriminant function after a retention of 12 PCs given by the a-score (Supplementary Fig. 2).

The allocation of adult individuals to populations utilizing the model-based clustering showed that the most likely K is three, based on Evanno et al. (2005) delta K statistics and not considering the information about the location of each individual (Fig. 3). Despite all genetic clusters are well represented in each individual, it is noteworthy that the majority of the individuals occurring in a given region has a great probability to share a similar cluster in accordance with their occurrence at each region (blue: NSV; red: NP—Fig. 3). A similar result was found for the progeny individuals (Fig. 4). A similar result was found for the progeny individuals (Supplementary Fig. 3) for which $K = 2$ and $K = 8$ scenarios were also likely based on

Table 2 Results of the analysis of molecular variance (AMOVA) for populations of *Tibouchina pulchra* situated in two regions: Núcleo Santa Virgínia (NSV) and Núcleo Picinguaba (NP), in the Serra do Mar State Park

Variation source	Df	Sum of squares	Variation percentage	Variation as compared to a random process	p value
Within populations	195	1552.23	91.55	Less	<0.001
Among populations within regions	12	188.01	5.94	Greater	<0.001
Between regions	1	38.45	2.51	Greater	<0.001
Total	297	2597.59	100.00		

Df degrees of freedom

Table 3 Results of the analysis of molecular variance (AMOVA) for the progenies of *Tibouchina pulchra* situated in two regions: Núcleo Santa Virginia (NSV) and Núcleo Picinguaba (NP), in the Serra do Mar State Park

Variation source	Df	Sum of squares	Variation percentage	Variation as compared to a random process	<i>p</i> value
Within progeny	541	2911.82	65.56	Less	<0.001
Among progenies within regions	19	1399.34	31.67	Greater	<0.001
Between regions	1	151.02	2.77	Greater	<0.001
Total	561	4462.18	100.00		

Df degrees of freedom

Evanno et al. (2005) delta K statistics. In this case, the majority of the progeny individuals also shares a common genetic cluster in accordance with their occurrence at each region (blue: NSV; red: NP—Supplementary Fig. 3).

Genetic diversity of adult populations

The genetic diversity estimators calculated by the FDASH program indicated that the individuals located at NSV region form a group genetically less diverse than the individuals occurring at NP (Supplementary Table 1, Fig. 5). Only the average number of shared alleles between pairs of individuals was statistically not different between regions, but all the other metrics were lower at the higher region (Supplementary Table 1, Fig. 5). The genetic diversity indexes were not structured by the populations except for the total number of alleles (Supplementary Table 1). Conversely, there was a difference among loci in all genetic diversity indexes (Supplementary Table 1).

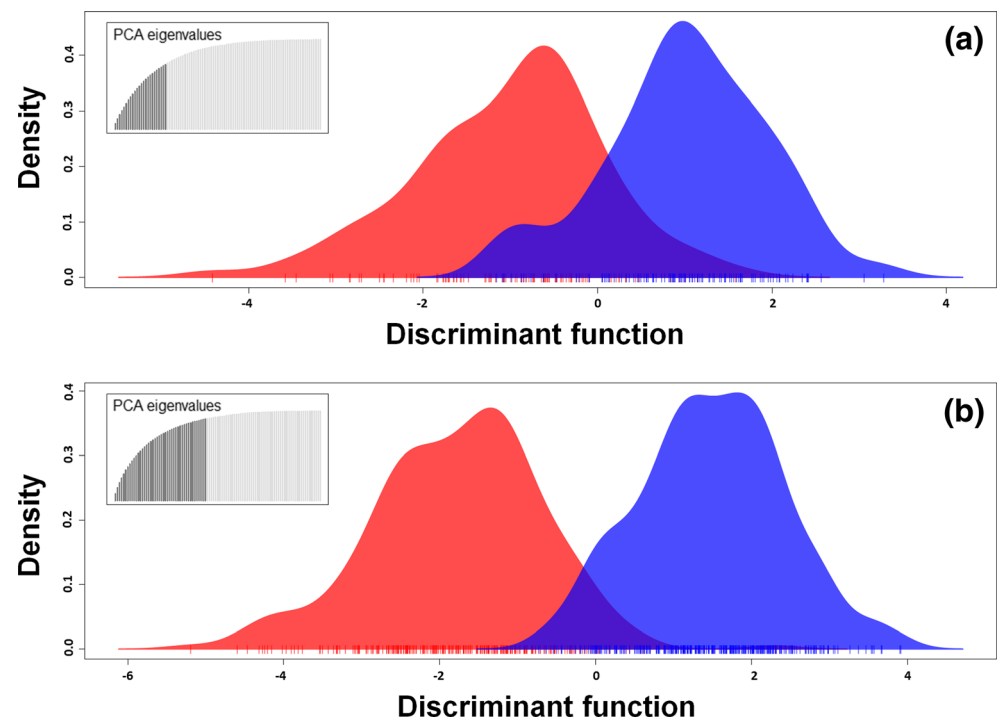
Paternity analyses

The paternity analysis indicated that the number of fathers of each progeny (fruit) is not different between regions (Fig. 6a; $z = -0.052$; $p > 0.05$). The same result was found for the probability of the progeny genotypes to be a result of a selfing event (Fig. 6b; $t = -0.364$; $p > 0.05$). On average, each fruit is fathered by two to three adults and there is an average probability of 39.53 % of each progeny individual to be a result of a selfing event.

Discussion

In general and in agreement with our first expectation, our results indicated that there is a weak but significant genetic structure among adult populations of *T. pulchra* occupying the higher and lower areas. However, just a small amount of the genetic difference between these groups is related to their

Fig. 2 Discriminant analysis on principal components (DAPC) for **a** adult and **b** progeny individuals of *Tibouchina pulchra* considering 24 and 44 PCs retained by the a-score, respectively. Blue area=Santa Virginia (higher region, NSV); red area=Picinguaba (lower region, NP). The upper graph shows the retained PCs by the a-score



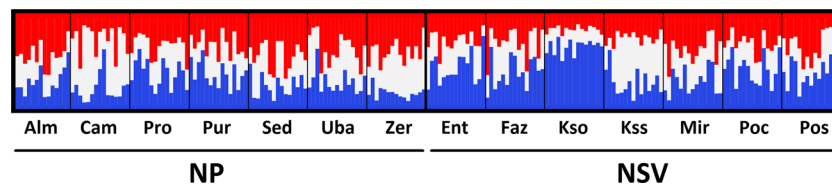


Fig. 3 Analysis performed by STRUCTURE 2.2.3 software for adult individuals of *Tibouchina pulchra* occurring in Santa Virgínia (higher region, NSV) and Picinguaba (lower region, NP). Groups were obtained under admixture model and allele frequency correlated between populations without a priori definition of populations. The best K , $K = 3$, is shown. Each color represents a group (K). Bars that present three colors indicate individuals that belong to more than one group. The size of each

color bar represents the probability of the individual to belong to that group. In this graph, the individuals are grouped in accordance with the population in which they were collected. Population names: Alm—Almada; Cam—Camburi; Pro—Promirim; Pur—Puruba; Sed—Sede; Uba—Ubatumirim; Zer—Zero; Ent—Entrada; Faz—Fazenda; Kso—Km88; Kss—Km86; Mir—Mirante; Poc—Poço do Pito; Pos—Posto II

geographic distances which indicate that such genetic structure is mostly related to other processes. This genetic structure was also found for the progeny array. Moreover, adult populations occupying the higher area present lower genetic diversity than populations occurring at the lower area, as we expected. However, differently from our third expectation, such difference cannot be attributed to a smaller number of adults fathering the progenies or a higher proportion of selfing at the higher regions, which indicate that pollen diversity delivered to stigmas is not the cause of the lower genetic diversity.

Specifically, AMOVA results indicated that there is a low genetic difference among adult populations within regions and between the regions formed by adult individuals occurring at the higher and lower regions. However, these differences were greater than expected by chance, suggesting that despite the high diversity within populations, there is genetic structure even in fine scales. The same pattern was found for progeny populations at the same regions. These results were confirmed by the DAPC that assigned the majority of the adult and progeny individuals to their actual region, even when we consider populations and mothers as groups, respectively. The STRUCTURE analysis also confirmed this trend showing that the majority of individuals occurring in each region share majorly the same allelic pool. In this sense, we could observe that, despite the occurrence of *T. pulchra* across the entire

elevation range, there is a genetic differentiation between the elevation extremes we studied. On the other hand, our analyses also suggest a contact interface between individuals occurring at both regions which is expected due to this continuous distribution to the species high reproductive and germinative capacities (Zaia and Takaki 1998) and also to the many possible reproductive events during the plant lifetime (Brito and Sazima 2012).

Differentiation between populations can be the product of two non-mutually exclusive processes influencing gene flow: isolation by distance or habitat differentiation (Bockelmann et al. 2003). Our analysis suggested that isolation by distance is not a predominant force driving genetic differentiation among and within regions. On the other hand, drastic habitat differences even in small scale, such as temperature, soil, luminosity, and pollinator limitation may restrict the establishment of seedlings and limit gene flow, leading to the differentiation of NSV and NP populations and such differences are observed in the altitudinal gradient studied (Joly et al. 2012; Brito and Sazima 2012).

T. pulchra is a pioneer tree species, occurring densely in a great area of distribution (Lorenzi 1992); its fruits produce thousands of seeds, dispersed autochorically and 10 to 30 % are able to germinate in natural conditions (Brito, non-published data). At Serra do Mar mountain range, the

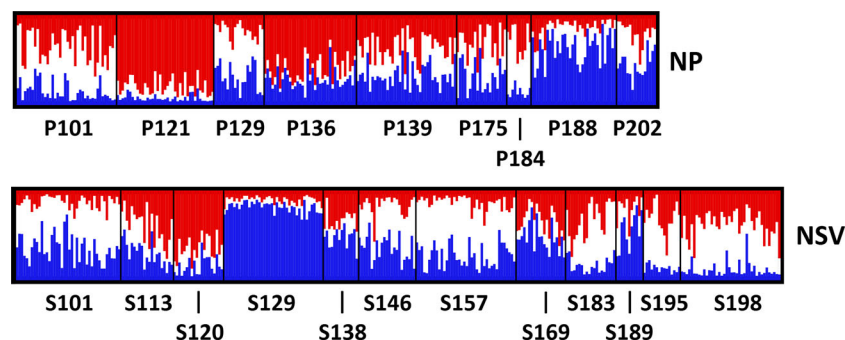


Fig. 4 Analysis performed by STRUCTURE 2.2.3 software for progeny individuals of *Tibouchina pulchra* occurring in Santa Virgínia (higher region, NSV) and Picinguaba (lower region, NP). Groups were obtained under admixture model and allele frequency correlated between populations without a priori definition of populations. The best K , $K = 3$,

is shown. Each color represents a group (K). Bars that present three colors indicate individuals that belong to more than one group. The size of each color bar represents the probability of the individual to belong to that group. In this graph, the individuals are grouped in accordance with the mother in which they were collected (Table 1)

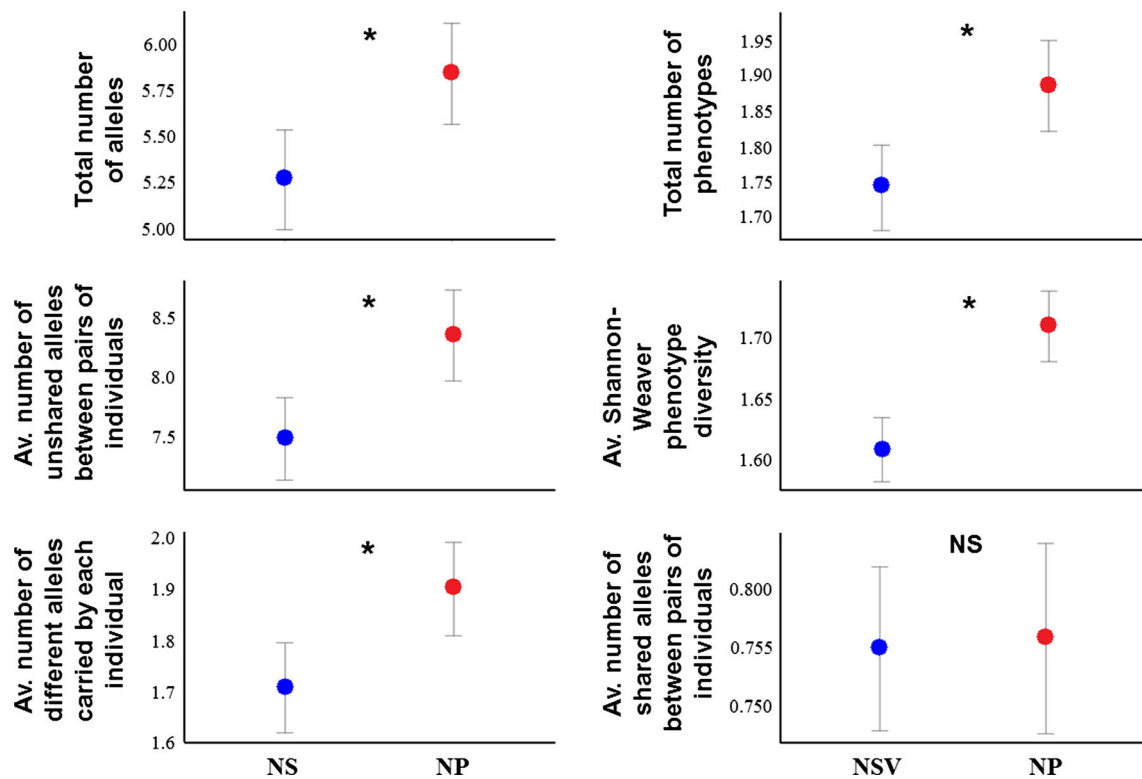


Fig. 5 Comparison between genetic diversity indexes for *Tibouchina pulchra* occurring in higher (Santa Virginia—NSV) and lower (Picinguaba—NP) regions. Asterisk indicates significance at 95 % confidence interval. NS non-significant

distribution of this species is continuous along the entire elevation range, occurring from the restinga (sandbank) region to the more elevated regions (Lorenzi 1992). However, the availability of pollinators varies drastically between both regions: at the more elevated region (NSV), the frequency of bees capable of extracting pollen from poricidal anthers is up to 200 times smaller than in the lower region (NP) (Brito and Sazima 2012). This implies different pollen flow dynamics within the populations, according to the altitude. Moreover, these populations have phenological displacement in flowering and fruiting (Brito and Sazima 2012), and it may also increase the genetic differentiation between populations

from both regions (Hall et al. 1994). Such reproductive phenological difference may lead to partial reproductive isolation which favor adaptation to different habitats and neutral genetic differentiation, as has been shown for other forest trees (Kraj and Sztorc 2009; Shi et al. 2011). In this sense, the weak genetic structure between populations at NSV and NP regions may be due to both limited gene flow by pollen and seed dispersal and by the adaptation to particular environmental factors.

Genetic differentiation between populations on an elevational gradient has been demonstrated for several tree species and the amount of variation between regions is even

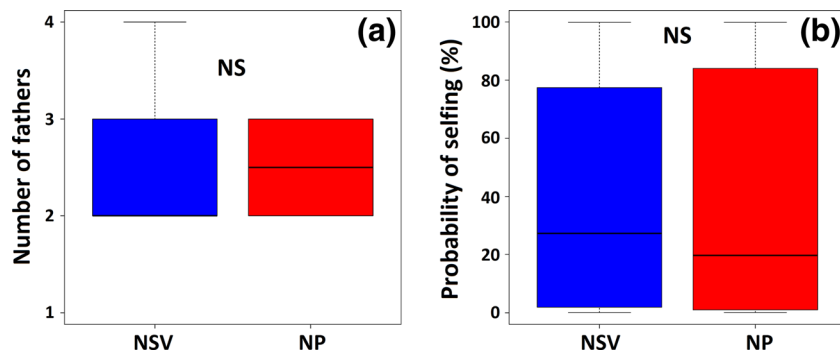


Fig. 6 Boxplot showing **a** the number of fathers of each progeny within a fruit and **b** the probability of such progeny genotypes to be a result a selfing event in *Tibouchina pulchra* occurring in higher (Santa Virginia—NSV) and lower (Picinguaba—NP) regions. Boxplots show

outliers as circles, the extreme of the lower and upper whisker as bars, the lower and upper hinges as the box extremes and the median as the darker line. NS non-significant. Blue Santa Virginia (higher region, NSV); red Picinguaba (lower region, NP)

larger than the one showed here (e.g., Reis et al. 2015; Shi et al. 2011; Ohsawa et al. 2007). In such cases, and also in other studies involving plants occurring in altitudinal gradients, diversity can increase, decrease, or even be constant with the altitude (Yan et al. 2009; Zhao et al. 2006; Semagn et al. 2000). Specifically, the pattern of lower genetic diversity at higher regions, as we found for *T. pulchra*, was already described and attributed to genetic drift and bottlenecks occurring during range expansion from lowlands or clonal reproduction (Ohsawa and Ide 2008; Quiroga and Premoli 2007; Williams and Arnold 2001). *T. pulchra* does not reproduce vegetatively and, in our study, we cannot exclude neutral processes, as genetic drift and bottlenecks, as the causes of genetic structure along the elevational gradient. However, relevant ecological conditions, such as pollinator limitation and phenological segregation, may also lead to similar genetic structure and diversity patterns expected for intense genetic drift and bottlenecks, and this would be even stronger in plants exclusively pollinated by insects (Reis et al. 2015).

Studies have already reported a pollinator limitation at higher regions in Andean and Alpine sites (Arroyo et al. 1985; Totland 1993), and the same was seemed for *T. pulchra* occurring at the Serra do Mar mountain range (Brito and Sazima 2012). Nonetheless, our progeny analysis indicated that the number of fathers per progeny (fruit), as well as the probability of the progeny genotypes to be a result of a selfing event, is not different between regions. This result refuted our former hypothesis that low pollen variability delivered by bees to the stigmas is the main cause of the lower diversity at the higher region. Despite this result, the idea that the dynamic of pollen flow influences genetic diversity within the populations in each region still holds. If pollen has a better dispersal capacity than seeds, as seems to be the case for *T. pulchra* and most tropical trees (Loveless and Hamrick 1984), gene dispersal will depend almost exclusively on the latter component (Hardy et al. 2006). Therefore, we speculate that the pattern of genetic diversity found in *T. pulchra* populations is related to the different amount of pollination events within each region. The lack of pollinators and the consequent pollen limitation observed in the highest region associated with the dependence on pollinators due to pronounced herkogamy lead to a lower fruiting in higher areas and a consequent low input of new variants in this region. Even with greater flowering and greater germination rate of seeds produced by cross-pollination in NSV (Brito and Sazima 2012), the reduced number of bees does not ensure a mixture of allelic pools, as seems to happen in NP, decreasing the effective size of such populations. The smaller values of genetic diversity measured by many indexes in NSV reinforce this idea. The reduced genetic variability within these populations could lead to inbreeding and decrease of the adaptive capacity (Willi et al. 2006). However, to test the occurrence of inbreeding depression (Ellstrand 1992; Charlesworth and Charlesworth 1987)

in these populations, studies with more comprehensive collections should be done in the elevated regions of Serra do Mar.

Our study showed that *T. pulchra* is a polyploid tree species whose populations occurring at two extreme regions of the elevational gradient of the Serra do Mar mountain range are genetically different and that the populations occurring at the higher site (NSV) presented lower genetic diversity. Despite the lack of genetic diversity differences among the populations within regions in the majority of indexes considered, we also found genetic structure at this level, indicating that some processes not evaluated by this study, such as landscape heterogeneity within region, are also determining these patterns, and this should be considered in further studies. Our progeny analysis corroborated this pattern and also indicated that the variability of pollen grains delivered to stigmas is not the major process influencing the genetic diversity contrast of *T. pulchra* in both regions. However, we suggest that the amount of pollination events, which is much smaller at NSV and decreases the effective size of such populations, reduces the input of new mutations in this site. To specifically test this hypothesis, we encourage a further study comparing the structure and diversity of *T. pulchra* seedlings in both regions.

Acknowledgments The authors kindly thank the help of our colleagues in the field and lab: Cristiano Silva, Marcelo M. Egea, Rafael S. Oliveira, and Fernanda Piccolo. We also thank Paulo E. Oliveira and two anonymous reviewers for the review of an early version of the manuscript and BSc, Giovana Maranhão Bettiol for helping with Fig. 1. We thank São Paulo Research Foundation (FAPESP) for the scholarship provided to V. L. G. B. (grant #2010/51494-5) and for the project financial support (grants #2008/52197-4 and #2012/50425-5). A. P. S. and M. S. received research fellowship from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). This work was supported by CNPq (grants #131969/2008-0 to V. L. G. B. and #302452/2008-7 to M. S.) and by the FAPESP as part of the Thematic Project Functional Gradient (grant #03/12595-7), within the BIOTA/FAPESP Program—The Biodiversity Virtual Institute (<http://www.biota.org.br>).

Data archiving statement Data would be deposited in the Dryad repository (<http://datadryad.org/>) after acceptance for review.

References

- Almeda F, Chuang TI (1992) Chromosome numbers and their systematic significance in some Mexican Melastomataceae. *Syst Bot* 17(4): 583–593
- Alvares CA, Stape JL, Sentelhas PC, Golçalves JLM, Sparovek G (2014) Köppen's climate classification map for Brazil. *Meteorol Z* 22(6): 711–728
- Arroyo MTK, Armesto JJ, Primack RB (1985) Community studies in pollination ecology in the high temperate Andes of central Chile. II. Effect of temperature on visitation rates and pollination possibilities. *Plant Syst Evol* 149:187–203
- Bawa KS (1995) Pollination, seed dispersal and diversification of angiosperms. *Trends Ecol Evol* 10(8):311–312
- Bencke CC, Morellato PLC (2002) Comparação de dois métodos de avaliação da fenologia de plantas, sua interpretação e representação. *Rev Bras Bot* 25(3):269–275

- Bockelmann AC, Reusch TBH, Bijlsma R, Bakker JP (2003) Habitat differentiation vs. isolation-by-distance: the genetic population structure of *Elymus athericus* in European salt marshes. *Mol Ecol* 12(2):505–515
- Brito VLG, Sazima M (2012) *Tibouchina pulchra* (Melastomataceae): reproductive biology of a tree species at two sites of an elevational gradient in the Atlantic rainforest in Brazil. *Plant Syst Evol* 298: 1271–1279
- Brito VLG, Vigna BBZ, Souza AP (2010) Characterization of 12 microsatellite loci from an enriched genomic library in polyploid *Tibouchina pulchra* Cogn. (Melastomataceae). *Conserv Genet Resour* 2(1):193–196
- Brito VLG, Weynans K, Sazima M, Lunau K (2015) Trees as huge flowers and flowers as oversized floral guides: the role of floral color change and retention of old flowers in *Tibouchina pulchra*. *Front Plant Sci* 6:362
- Buchmann SL (1983) Buzz pollination in angiosperms. In: Jones CE, Little RJ (eds) *Handbook of experimental pollination biology*. Van Nostrand Reinhold, New York, pp. 73–113
- Byars SG, Parsons Y, Hoffmann AA (2009) Effect of altitude on the genetic structure of an Alpine grass, *Poa hiemata*. *Ann Bot* 103: 885–899
- Byrne M, Hankinson M, Sampson JF, Stankowski S (2008) Microsatellite markers isolated from a polyploid saltbush, *Atriplex nummularia* Lindl. (Chenopodiaceae). *Mol Ecol Resour* 8(6):1426–1428
- Caddah MK, Campos T, Sforça DA et al (2009) Microsatellite markers isolated from polyploid *Kielmeyera coriacea* Mart. and Zucc. (Clusiaceae) from an enriched genomic library. *Conserv Genet* 10(5): 1533–1535
- Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Syst* 18:237–268
- Cribari-Neto F, Zeileis A (2010) Beta regression in R. *J Stat Softw* 34(2): 1–24
- De Silva HN, Hall AJ, Rikkerink E, McNeilage MA, Fraser LG (2005) Estimation of allele frequencies in polyploids under certain patterns of inheritance. *Heredity* 95(4):327–334
- Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) Touchdown™ PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* 19(14):4008
- Dray S, Dufour AB, Chessel D (2007) The ade4 package-II: two-table and K-table methods. *R News* 7(2):47–52
- Dufresne F, Stift M, Vergilino R, Mable BK (2014) Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Mol Ecol* 23:40–69
- Ellstrand NC (1992) Gene flow by pollen: implications for plant conservation genetics. *Oikos* 63(1):77–86
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8):2611–2620
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131(2):479–491
- Faegri K, Van der Pijl L (1979) *The principles of pollination ecology*. Pergamon Press, Oxford
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4):1567–1587
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null allele. *Mol Ecol Notes* 7:574–578
- Goldenberg R, Baumgratz JFA, Souza MLDR (2012) Taxonomia de Melastomataceae no Brasil: retrospectiva, perspectivas e chave de identificação para os gêneros. *Rodriguésia* 63:145–161
- Guimarães PJF, Martins AB (1997) *Tibouchina* sect. *Pleroma* (D.Don) Cogn. (Melastomataceae) no estado de São Paulo. *Rev Bras Bot* 2(1):11–33
- Hall P, Chase MR, Bawa KS (1994) Low genetic variation but high population differentiation in a common tropical forest tree species. *Conserv Biol* 8(2):471–482
- Hardy OJ, Maggia L, Bandou E et al (2006) Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Mol Ecol* 15:559–571
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806
- Joly CA, Assis MA, Bernacci LC et al (2012) Florística e fitossociologia em parcelas permanentes da Mata Atlântica do sudeste do Brasil ao longo de um gradiente altitudinal. *Biota Neotrop* 12(1):123–145
- Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94
- Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* 10:551–555
- Jordano P, Godoy JA (2000) RAPD variation and population genetic structure in *Prunus mahaleb* (Rosaceae), an animal-dispersed tree. *Mol Ecol* 9:1293–1305
- Kraj W, Sztorec A (2009) Genetic structure and variability of phenological forms in the European beech (*Fagus sylvatica* L.). *Ann For Sci* 66: 203–209
- Legendre P, Legendre L (1998) *Numerical ecology*. Elsevier Science B.V, Amsterdam
- Lorenzi H (1992) *Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil*. Editora Plantarum Ltda, Nova Odessa
- Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annu Rev Ecol Syst* 15(1):65–95
- Lynch M (1990) The similarity index and DNA fingerprinting. *Mol Biol Evol* 7:478–484
- Misiewicz TM, Fine PVA (2014) Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subseratum* (Burseraceae). *Mol Ecol* 23:2543–2558
- Morellato LP, Talora DC, Takahasi A, Bencke CC, Romera EC, Zipparro VB (2000) Phenology of Atlantic rainforest trees: a comparative study. *Biotropica* 32(4b):811–823
- Obbard DJ, Harris SA, Pannell JR (2006) Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Heredity* 97(4):296–303
- Ohsawa T, Ide Y (2008) Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Glob Ecol Biogeogr* 17:152–163
- Ohsawa T, Tsuda Y, Saito Y, Sawada H, Ide Y (2007) Altitudinal genetic diversity and differentiation of *Quercus crispula* in the Chichibu Mountains, Central Japan. *Int J Plant Sci* 168(3):333–340
- Pereira AC, Silva JB, Goldenberg R, Melo GA, Varassin IG (2011) Flower color change accelerated by bee pollination in *Tibouchina* (Melastomataceae). *Flora* 206:491–497
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Quiroga MP, Premoli AC (2007) Genetic patterns in *Podocarpus parlatorei* reveal the long-term persistence of coldtolerant elements in the southern Yungas. *J Biogeogr* 34:447–455
- Reis TS, Ciampi-Guillardi M, Bajay MM, Souza AP, Santos FAM (2015) Elevation as a barrier: genetic structure for an Atlantic rain forest tree (*Bathysa australis*) in the Serra do Mar mountain range, SE Brazil. *Ecol Evol* 5(9):1919–1931

- Renner SS (1989) A survey of reproductive biology in neotropical Melastomataceae and Memecylaceae. *Ann Mo Bot Gard* 76:469–518
- Renner SS (1993) Phylogeny and classification of the Melastomataceae and Memecylaceae. *Nord J Bot* 13:519–540
- Schuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. *Nat Biotechnol* 18:233–234
- Sebbenn AM, Siqueira ACMF, Gurgel GLMA, Angerami EMRA (2000) Variabilidade genética e interação genótipo x locais em jequitibá-rosa—*Cariniana legalis* (Mart.) O. Ktze. *Rev Inst Florest* 12(1):13–23
- Semagn K, Bjornstad A, Stedje B, Bekele E (2000) Comparison of multivariate methods for the analysis of genetic resources and adaptation in *Phytolacca dodecandra* using RAPD. *Theor Appl Genet* 101(7):1145–1154
- Shi MM, Michalski SG, Chen XY, Durka W (2011) Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*. *PLoS One* 6(6):e21302
- Silveira FAO, Fernandes GW, Lemos-Filho JP (2013) Seed and seedling ecophysiology of neotropical Melastomataceae: implications for conservation and restoration of savannas and rainforests. *Ann Mo Bot Gard* 99(1):82–99
- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* 236(4803):787–792
- Tabarelli M, Mantovani W (1999) A riqueza da floresta Atlântica de encosta no estado de São Paulo (Brasil). *R Bras. Biol* 59(2):239–250
- Thomas G, Joseph L, Varghese G, Kalyanaraman S, Kuriachan P (2001) Analysis of phenotypic and genetic variations among populations of *Oryza malampuzhaensis* show evidence of altitude-dependent genetic changes. *Can. J Botany* 79:1090–1098
- Todzia CA, Almeida F (1991) A revision of *Tibouchina* sect. *Lepidota* (Melastomataceae: Tibouchinae). *Proc Calif Acad Sci* 47(6):175–206
- Totland Ø (1993) Pollination in alpine Norway: flowering phenology, insect visitors, and visitation rates in two plant communities. *Can. J Botany* 71:1072–1079
- Veloso HP, Rangel-Filho ALR, Lima JCA (1991) Classificação da vegetação brasileira adaptada a um sistema universal. IBGE, Rio de Janeiro
- Wang J, Scribner KT (2014) Parentage and sibship inference from markers in polyploids. *Mol Ecol Resour* 14:541–553
- Willi Y, Buskirk JV, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *Annu Rev Ecol Evol S* 37:433–458
- Williams JH, Arnold ML (2001) Sources of genetic structure in the woody perennial *Betula occidentalis*. *Int J Plant Sci* 162:1097–1109
- Yan XB, Guo YX, Zhao C, Liu FY, BR L (2009) Intra-population genetic diversity of two wheatgrass species along altitude gradients on the Qinghai-Tibetan Plateau: its implication for conservation and utilization. *Conserv Genet* 10(2):359–367
- Zaia JE, Takaki M (1998) Estudo da germinação de sementes de espécies arbóreas pioneiras: *Tibouchina pulchra* Cong. e *Tibouchina granulosa* Cong. (Melastomataceae). *Acta Bot Bras* 12(3):221–229
- Zhao N, Gao Y, Wang J, Ren A, Xu H (2006) RAPD diversity of *Stipa grandis* populations and its relationship with some ecological factors. *Acta Ecol Sin* 26(5):1312–1318