

Article

Genomic Diversity and Structure of *Copaifera langsdorffii* Populations from a Transition Zone Between the Atlantic Forest and the Brazilian Savanna

Marcos Vínicius Bohrer Monteiro Siqueira ^{1,*} , Juliana Sanchez Carlos ², Wilson Orcini ², Miklos Maximiliano Bajay ³ , Karina Martins ⁴ , Arthur Tavares de Oliveira Melo ⁵ , Elizabeth Ann Veasey ⁶, Evandro Vagner Tambarussi ⁷  and Enéas Ricardo Konzen ⁸ 

- ¹ Universidade do Estado de Minas Gerais, Avenida Escócia, 1001, Cidade das Águas, Frutal 38202-436, MG, Brazil
 - ² Universidade do Sagrado Coração, Rua Irmã Arminda, 10-50—Jardim Brasil, Bauru 17011-160, SP, Brazil; jusanchez93@gmail.com (J.S.C.); wilson.orcini@gmail.com (W.O.)
 - ³ Centro de Educação Superior da Região Sul, Universidade do Estado de Santa Catarina, Rua Cel. Fernandes Martins, 270, Progresso, Laguna 88790-000, SC, Brazil; mmbajay@gmail.com
 - ⁴ Departamento de Biologia, Universidade Federal de São Carlos, Rod. João Leme dos Santos, km 110-SP-264, Bairro do Itinga, Sorocaba 18052-780, SP, Brazil; karimartins@yahoo.com
 - ⁵ Inova Genética LTDA, Fazenda Santa Barbara, Rodovia BR 251, KM 18, Cristalina 73850-000, GO, Brazil; arthurmelo10@gmail.com
 - ⁶ Departamento de Genética, Universidade de São Paulo (USP), Av. Pádua Dias 11, Piracicaba 13418-900, SP, Brazil; eann.veasey@gmail.com
 - ⁷ Departamento de Produção Vegetal, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Av. Universitária 3780, Botucatu 18610-034, SP, Brazil; evandro.tambarussi@unesp.br
 - ⁸ Centro de Estudos Costeiros, Limnológicos e Marinhos, Campus Litoral Norte, Universidade Federal do Rio Grande do Sul, Av. Tramandaí, 976, Imbé 95625-000, RS, Brazil; erkonzen@gmail.com
- * Correspondence: mvbsiqueira@gmail.com

Abstract

Copaifera langsdorffii is a neotropical tree widely distributed in the Brazilian Atlantic Forest and Brazilian Savanna. Population genetic analyses can identify the scale at which tree species are impacted by human activities and provide useful demographic information for management and conservation. Using a Restriction site Associated DNA Sequencing approach, we assessed the genomic variability of six *C. langsdorffii* population relicts in a transition zone between the Seasonal Atlantic Forest and Savanna biomes in Southeastern Brazil. We identified 2797 high-confidence SNP markers from six remnant populations, with 10 to 29 individuals per population, in a transition zone between the Seasonal Atlantic Forest and Savanna biomes in Southeastern Brazil. Observed heterozygosity values (0.197) were lower than expected heterozygosity (0.264) in all populations, indicating an excess of homozygotes. Differentiation among populations (F_{ST}) was low (0.023), but significant (0.007–0.044, c.i. 95%). A clear correlation was observed between geographic versus genetic distances, suggesting a pattern of isolation by distance. Bayesian inferences of population structure detected partial structuring due to the transition between the Atlantic Forest and the Brazilian Savanna, also suggested by spatial interpolation of ancestry coefficients. Through the analysis of F_{ST} outliers, 28 candidates for selection have been identified and may be associated with adaptation to these different phytophysiognomies. We conclude that the genetic variation found in these populations can be exploited in programs for the genetic conservation of the species.

Keywords: Atlantic rainforest; Brazilian trees; Cerrado; population structure; SNP markers



Academic Editor: Brian Traw

Received: 2 August 2025

Revised: 2 September 2025

Accepted: 10 September 2025

Published: 13 September 2025

Citation: Siqueira, M.V.B.M.; Carlos, J.S.; Orcini, W.; Bajay, M.M.; Martins, K.; Melo, A.T.d.O.; Veasey, E.A.; Tambarussi, E.V.; Konzen, E.R. Genomic Diversity and Structure of *Copaifera langsdorffii* Populations from a Transition Zone Between the Atlantic Forest and the Brazilian Savanna. *Plants* **2025**, *14*, 2858. <https://doi.org/10.3390/plants14182858>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the past 20 years, remarkable progress has been made towards understanding how the loss of biodiversity affects the functioning of ecosystems as well as human populations [1]. These losses have further implications as they can intensify collapses of ecosystem functions. The most recognized and potentially well-documented examples of biodiversity and ecosystem loss have been in the tropics, home to half of the world's species [2]. Biodiversity loss is largely driven by human activities such as hunting and habitat loss through deforestation, agricultural expansion, industrialization, and urbanization [3].

The Atlantic Rainforest, one of the world's major hotspots of biodiversity, originally covering approximately 1,500,000 km² from regions of eastern Brazil suffered a large loss of primary habitat due to continuous anthropogenic pressures and now is reduced to 11–16% of its original area, mostly as small and degraded remnants embedded in matrices dominated by pastures, agricultural lands, and urban areas [4]. The Brazilian savanna (Cerrado) has been the target of high human pressures in the last 50 years, leading to an unprecedented loss of biodiversity and deforestation rates [5,6]. It is the most bio-diverse and threatened savannah on the planet, with 48.2% of its original floral cover already lost, and it is being affected by an intense process of habitat fragmentation [7]. It includes a range of vegetation types, such as forests, savannahs, and grasslands, that together reach 12 thousand plant species, of which 44% are endemic [8].

The Neotropical forest-savannah, including Atlantic Rainforest and Cerrado biomes, shows a geographical discontinuity, along with specific soil and climate characteristics, frequently resulting in high levels of genetic structure between populations of endemic biota found on different forest physiognomies [9]. Those transition zones require appropriate conservation efforts [9]. Cerrado, with more than 200 million ha, shares ecological transition zones with four of the five Brazilian biomes: the Amazon (rainforest), Caatinga (semi-arid region), Pantanal (wetland), and Atlantic Forest (coastal forest) [10]. Cerrado and the Atlantic Forest are connected by a Seasonally Dry Tropical Forest (SDTF) in southeastern Brazil. SDTF is among the most devastated forest types because of its strategic location: low slopes and not far from the coastal zone, where agriculture and urban growth easily occur [11]. In addition, the SDTF is second among the most threatened centers of endemism in the Atlantic Forest, with only 7.1% of its original cover remaining and 6.8% of its remaining forest under protection [4–12].

Population genetics analysis, through molecular technologies, has been pivotal in understanding the diversity and vulnerability of many forest areas, as well as contributing to the development of conservation strategies [13,14]. A recent meta-analysis at global scale reveals that genetic diversity conservation needs urgent attention from conservationists [15], so more studies on neglected species need to be pursued and effective actions need to be designed after the studies. Recent advances in genomic methods, including next-generation sequencing, whole genome scans and gene-expression pattern analysis, allowed stepping up from a limited number of neutral markers to genome-wide estimates of functional genetic variation [16]. Restriction-site associated DNA sequencing (RADSeq) is a method [17] that enables the identification of several molecular markers, such as single-nucleotide polymorphisms (SNPs), even without a reference genome [18]. By selecting a certain number of RADSeq loci and based on the choice of different restriction enzymes, a large-scale SNP genotyping matrix can be obtained for a target species [18,19]. Thus, a more detailed understanding of population genomics is possible, allowing the development of effective strategies for the conservation of genetic resources [18].

Although genetic analyses have become an integral part of conservation biology, very few conservation genetics studies have been completely integrated into the development of conservation and management strategies and formal policies [20]. For example, much of the

genetic diversity of native Brazilian flora has not yet been studied [21], showing a large gap between the supply of plants and the few researchers assigned to them. Therefore, many more studies need to be conducted to appropriately characterize and provide a panorama of the genetic diversity of the species whose environments of origin are the tropics. *Copaifera langsdorffii* Desf., a key and widely distributed species in Atlantic Forest and Cerrado areas in Brazil [22], has been receiving special attention due to its medicinal properties, used not only in traditional, but also in modern medicine. Studies with *C. langsdorffii* to date have been performed with isozymes and microsatellite markers (SSR) [22–28], detecting patterns of diversity of natural populations, genetic structure of populations in fragmented areas, mating system and gene flow. These previous studies revealed important aspects of neutral genetic diversity but relied on a few molecular markers. In the present study, we moved forward to use a genomic approach based on reduced genomic representation to investigate patterns of adaptive genomic variation in *C. langsdorffii* in a transition zone between Atlantic Forest and Cerrado in Southeastern Brazil.

Using SNP markers detected through RADseq, we assessed the genetic variability and potential to conserve the genetic diversity of *C. langsdorffii* populations in two different biomes: the Brazilian Atlantic Forest and the Savanna of the Central-West Region of São Paulo state. Our specific goals were to: (1) characterize genetic diversity and population structure of *C. langsdorffii* in Brazilian Atlantic Forest, Cerrado and transition zones, and (2) identify markers under directional selection to assess if transition zones display any response to environmental stressors in populations from distinct phytophysiognomies.

2. Materials and Methods

2.1. Study Species

Copaifera langsdorffii, popularly known as Copaíba (Supplementary Material S2—ESM2), belongs to the Fabaceae family and Detarioideae subfamily [29]. According to [30], *Copaifera* has 72 described species, 16 of which are found exclusively in Brazil [31], and is distributed from the northern to the southernmost regions of the country [32]. The distribution of the species encompasses biomes such as the Amazon Rainforest, Atlantic Forest, and the riparian forests of the Cerrado region [22]. This tropical tree can reach 40 m in height and live up to 400 years. It is pollinated by small insects (*Apis mellifera*, *Trigona* sp., among others), and is predominantly outcrossed [32,33]. The fruits are a follicle type with a black seed surrounded by an orange aryl [31,34], traits that indicate that seed dispersal occurs mainly through animals, especially birds [34].

Natural populations have been gradually eliminated due to habitat degradation and habitat fragmentation. Currently, most populations are restricted to small remnants within forest fragments [25]. *Copaifera* sp. produces an oil-resin that is well described in the literature and may vary in quantity and quality among species [35]. Products derived from *Copaifera* spp. (copaíba oil-resin) are typically extracted via tapping into the trunk rather than logging, and are widely used in medicinal, cosmetic, and industrial applications. Although extraction does not entail felling, traditional and modern practices have been documented to be unsustainable when performed intensively over time, especially without long-term management protocols. Ethnobiological studies in Brazil report that historical extractivist economies were often destructive, even when extraction methods were ritualized or respectful [36]. The oil can be used to treat ulcers, sinusitis, inflammation of the skin, kidneys, and throat. It is also used in the perfume industry for its emollient properties, such as bactericidal and anti-inflammatory, in the manufacture of soaps, bath creams and foams, shampoos, conditioner creams, moisturizing and hair lotions [31,37]. Copaíba plays key ecological functions by contributing to forest structure with high phenotypic plastic-

ity [38], supporting biodiversity through pollination and seed dispersal interactions [39], and enhancing ecosystem services such as carbon storage and soil protection [40].

2.2. Sampling Sites

The study was conducted in the Central-West Region of São Paulo State, southeastern Brazil, in areas of Cerradão (a forest phytophysiology of Brazilian Savannah) and Semideciduous Seasonal Forest (an Atlantic Forest phytophysiology). These areas have a Cwa climate (humid temperate climate with dry winter), where summer is hot, with the air temperatures in the hottest month rising above 22 °C; the average rainfall is less than 60 mm in at least one of the months of the season [41]. The sampled areas involved secondary, regenerating, and ecotone forests, and six remaining fragments were selected and defined as populations, located in four municipalities, comprising public and private areas. A total of 95 adult trees were sampled. Population sampling took place between May and June 2015. During sample collection, reproductive individuals were prioritized, and a minimum distance of 10 m between them was maintained. In the municipality of Bauru, three population remnants were sampled: 1. Municipal Botanical Garden of Bauru (JB) with 321.71 hectares (ha) of extension (29 individuals sampled); 2. Campus of the São Paulo State University (UN) with 132 ha (20 individuals); and 3. Aimorés Forest (HA) with 5424.15 ha (11 individuals). In the municipality of Pederneiras, we sampled one population with 14 individuals at the State Forest (FP) with 430 ha. In the municipality of Agudos, a population was sampled with 11 individuals at the reserve of a private property (DU) with 2000 ha. In the municipality of Gália, 10 individuals were sampled at the Caetetus Ecological Station (CAE), with 2178.84 ha (Figure 1). The JB, UN, and FP fragments belong to areas where Savanna and Semideciduous Seasonal Forest vegetation occur, being important ecotones that were considered in this study. The other sampled areas (HA, DU, and CAE) comprise only the composition of Seasonal Semideciduous Forest, and DU stands out for being an area of natural regeneration. For details of *C. langsdorffii* populations sample collection, please refer to Supplementary Material (ESM1 and ESM2), which shows details on their location.

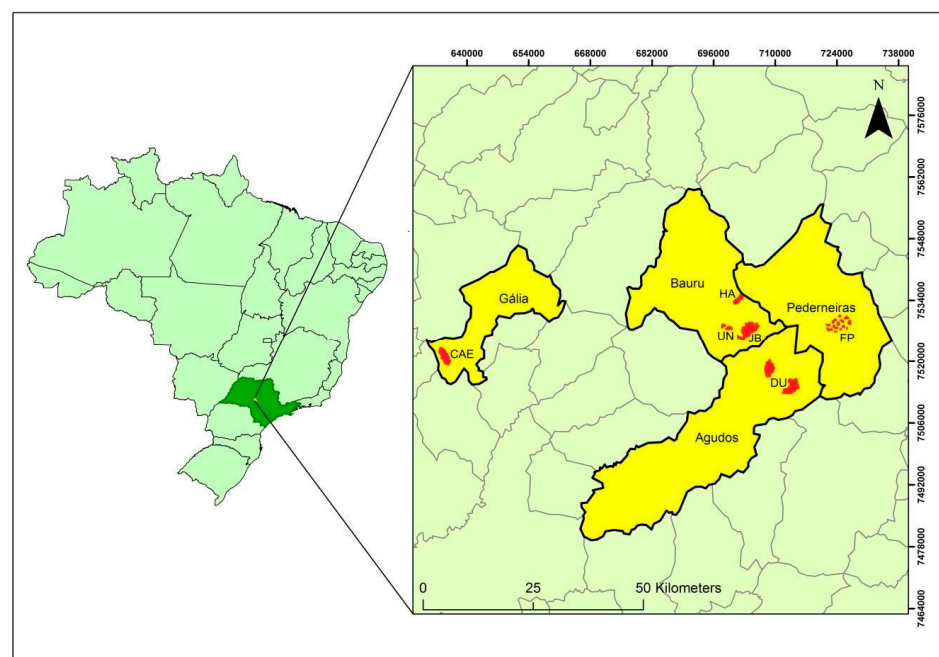


Figure 1. Map of Brazil, highlighting the state of São Paulo (dark green in left map) and the Central-West Region, showing the municipalities and collection sites (in yellow) of *Copaifera langsdorffii* populations: Bauru—JB, UN, HA; Pederneiras—FP; Agudos—DU and Gália—CAE.

Leaf samples of each individual of *C. langsдорffii* were immediately packed in liquid nitrogen and then stored at -80°C until DNA extraction. Measurements on tree height, diameter at breast height (DBH), MSL, and geographical coordinates were also gathered during sampling (Supplementary Material).

2.3. DNA Isolation and RADSeq Library Preparation

Total genomic DNA was extracted according to the protocol established by the DNeasy Plant Mini Kit manufacturer (QIAGEN, Hilden, Germany), generating high-quality samples according to quantification performed on NanoDrop (Spectrophotometer 2000, Thermo Scientific; Waltham, MA, USA) and 1.2% agarose gel, stained with ethidium bromide.

The sequencing was performed by SNPsaurus LLC (Eugene, OR, USA) using RADSeq [42] and genomic DNA was transformed into RADSeq-like libraries using selective primers. Genomic DNA was first fragmented with Nextera reagent (Illumina, Inc., San Diego, CA, USA), to which small adapters were attached at the end of each fragment. The DNA fragments were then amplified with one of the primers, combining the adapter and extending nine arbitrary nucleotides in genomic DNA with the selective sequence. Thus, only fragments starting with a sequence that can be hybridized by the selective primer sequence have been efficiently amplified. The resulting fragments were fixed at the selective end, presenting random lengths depending on the initial Nextera fragmentation. In this sense, DNA amplified from a particular locus may appear in fragments of different sizes, so careful selection of the library size is not necessary. The RADSeq libraries were sequenced on a HiSeq 4000 (University of Oregon) using a custom sequencing primer that includes the selective sequence so that the entire read is used for genotyping.

2.4. SNP Calling and Genotyping

The SNPs were called using custom scripts that created a *de novo* reference from abundant reads and then mapped all the reads to the reference with an alignment identity threshold of 95% [42,43]. Genotype calling was performed using Samtools [44] and bcftools [45] and filtered using Plink 1.9 [46].

Data were filtered excluding non-polymorphic SNP markers or those with more than 10% missing data and SNP loci with minor allele frequencies (MAF) of less than 1%. We also excluded samples with more than 25% missing data. In total, 95 individuals remained for population genomic analyses.

2.5. Data Analysis

2.5.1. Genetic Diversity and Structure

Observed (H_O) and expected (H_E) heterozygosities were estimated using both '*hierfstat*' [47] and '*adegenet*' [48] packages in R [49]. F statistics, allelic richness, and the respective confidence intervals, based on 1000 bootstrap resamples, were estimated using the R '*diversity*' package version 1.9.90 [50].

A matrix of Nei's (1972) [51] genetic distances was generated and used for constructing a dendrogram through the unweighted pair group method with arithmetic mean (UPGMA) approach [52], using '*poppr*' [53]. The consistency of the nodes was estimated through 1000 bootstrap replications.

Genetic structure was analyzed with a Bayesian approach using Structure v. 2.3.3 [54] under the mixture model, delimiting K groups from 1 to 10 with 10 replicates, a 200,000 burn-in period followed by 500,000 MCMC iterations. K was chosen according to the ΔK method of [55] using Structure Harvester v. 0.6.7 [56]. The genetic structure of populations was also assessed using discriminant analysis of principal components (dAPC) with '*adegenet*' in R [47] and principal component analysis (PCA) with '*ape*' [57], '*ade4*' [58], and '*ggplot2*' [59] packages in R. The coancestry coefficients were also calculated from '*tess3r*', using the

tess3 function [60]. Spatial interpolation of the coancestry coefficients was computed and displayed with ‘*tess3r*’, along with ‘*raster*’ [61] and ‘*rworldmap*’ [62] packages from R.

2.5.2. F_{ST} Outlier Analysis

The discovery of the candidate sites for the divergent selection (F_{ST} outliers) was conducted with the FDIST2 method [63], in which the outliers are the F_{ST} estimates that lie above the 99% threshold of a simulated distribution of neutral F_{ST} values. This method was implemented in the Lositan software v1.0 [64], and the statistical significance of F_{ST} outliers was considered using the false discovery rate (FDR) strategy to correct for multiple tests [65].

Estimates of genetic diversity parameters were compared between the total set of markers and the non-neutral loci to estimate the effect of potential adaptive loci on populations from the Semideciduous Seasonal Forest and the savanna environments.

3. Results

3.1. Genetic Diversity and Structure

The molecular characterization resulted in 2797 SNP markers, most of which had observed heterozygosity values (H_O) lower than the expected heterozygosity values (H_E) in all six fragment populations (Table 1). Genetic diversity indices revealed slight differences among all populations; however, the JB population showed the highest values for all the genetic statistics, followed by FP. The inbreeding coefficients (F_{IS}) were all significant with an average of 0.209 (0.158–0.251, confidence interval at 95%) (Table 1).

Table 1. Genetic diversity estimates for six populations of *Copaifera langsdorffii* in an ecotone region between Atlantic Forest and Brazilian savanna in southeastern Brazil.

Population	N	H_O	H_E	F_{IS}
JB	29	0.205	0.257	0.203 *
UN	20	0.189	0.247	0.235 *
HA	11	0.187	0.244	0.235 *
FP	14	0.203	0.245	0.173 *
DU	11	0.178	0.238	0.251 *
CAE	10	0.196	0.233	0.158 *
Mean	-	0.197	0.264	0.209

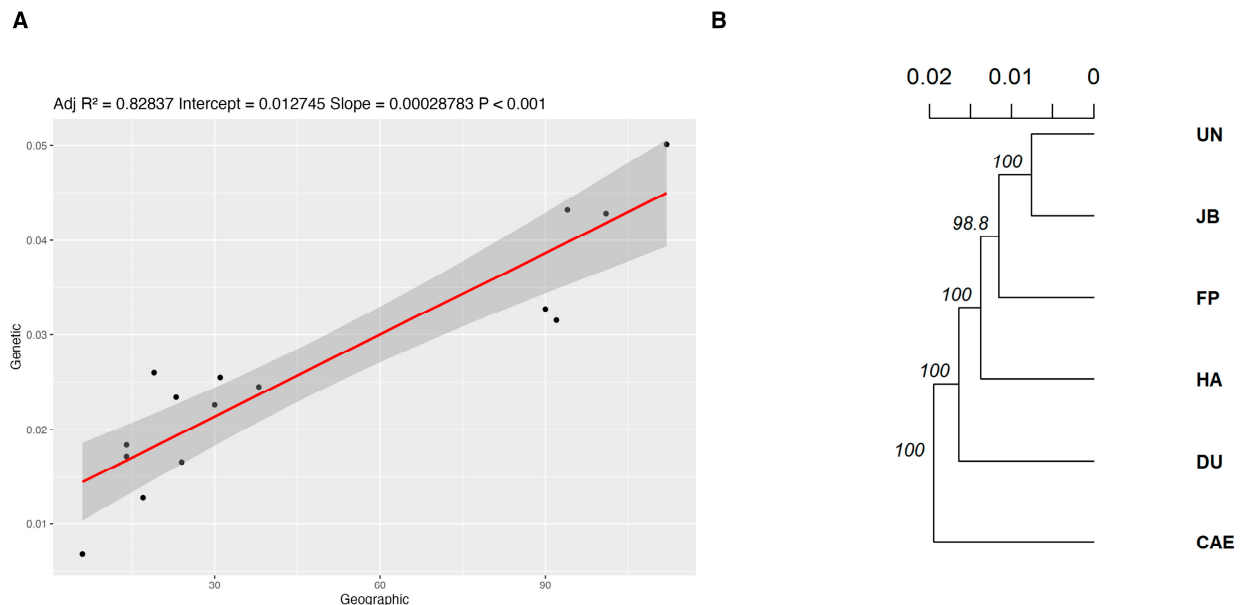
N: number of samples, H_O : observed heterozygosity, H_E : expected heterozygosity, F_{IS} : inbreeding coefficient (* $p < 0.05$). JB, UN, HA, FP, DU, and CAE are the acronyms for population names (see methods).

Population differentiation was low ($F_{ST} = 0.023$), but significant (0.007–0.044, c.i 95%). Pairwise F_{ST} suggested a pattern of isolation by distance, with the most isolated population (CAE, approximately 90–112 km away from the other populations) showing higher and significant F_{ST} estimates (Table 2). The hierarchy of the genotypes in the UPGMA dendrogram (Figure 2A) and the high correlation between the genetic distance and the geographic distance, as inferred from a linear model (Figure 2B), corroborate the tendency of isolation by distance found in pairwise F_{ST} (Table 2) data.

Through principal component analysis, the two first components explained only 4.66% of the total genetic variability found in 95 genotypes (Figure 3). The samples from CAE and JB were more dispersed from the others (Figure 3A). The individuals collected in areas of Seasonal Semideciduous Forest (in the graph marked as Forest) and the individuals collected in the Savanna area (in the graph marked as Savanna), were marked by different colors (Figure 3B). However, a partial overlap was observed between the two groups.

Table 2. Pairwise F_{ST} estimated for *Copaifera langsdorffii* populations (below diagonal) and pairwise geographic distances (km) (above diagonal).

Pop.	JB	UN	HA	FP	DU	CAE
JB	—	6	14	23	19	94
UN	0.0068	—	17	24	14	90
HA	0.0172	0.0128	—	31	30	92
FP	0.0234 *	0.0166	0.0255	—	38	112
DU	0.0260	0.0184	0.0226	0.0245	—	101
CAE	0.0432 *	0.0327 *	0.0315 *	0.0501 *	0.0428	—

* $p < 0.05$.**Figure 2.** (A) UPGMA dendrogram obtained by UPGMA method for six populations of *Copaifera langsdorffii* from the Central-West Region of São Paulo state, southeastern Brazil. The scale indicates Nei's distance (1972) values and node values are the supporting bootstraps. (B) Linear model fitting Nei's genetic distances with the corresponding geographic distances. The correlation coefficient associated with the model is $r = 0.9101$.

Despite their overlap, a trend of separation between forest and savanna ecosystems can be observed, for example, in JB, where samples belonging to the forest are mostly separated from the savanna (compare graphs A and B in Figure 4). The isolation of distance of CAE, however, is also clear from the comparison (Figure 3A). Given those clues, further Bayesian analysis showed the highest ΔK estimation in a population structure of $K = 2$, with most of group I (red) belonging to the genotypes collected in the Forest ecosystem, while most of group II (green) included the genotypes collected in the Savanna (Figure 4A). This pattern, however, was not clear, as a mixture of ancestry between both groups was also observed. After spatial interpolation of the ancestry coefficients, using TESS structure analysis, an overall trend of separation between Savanna and Forest was also detected (Figure 4B). Once more, the site of CAE shows a genetic separation even from the other samples of the Forest group, suggesting more genetic divergence from the other sites.

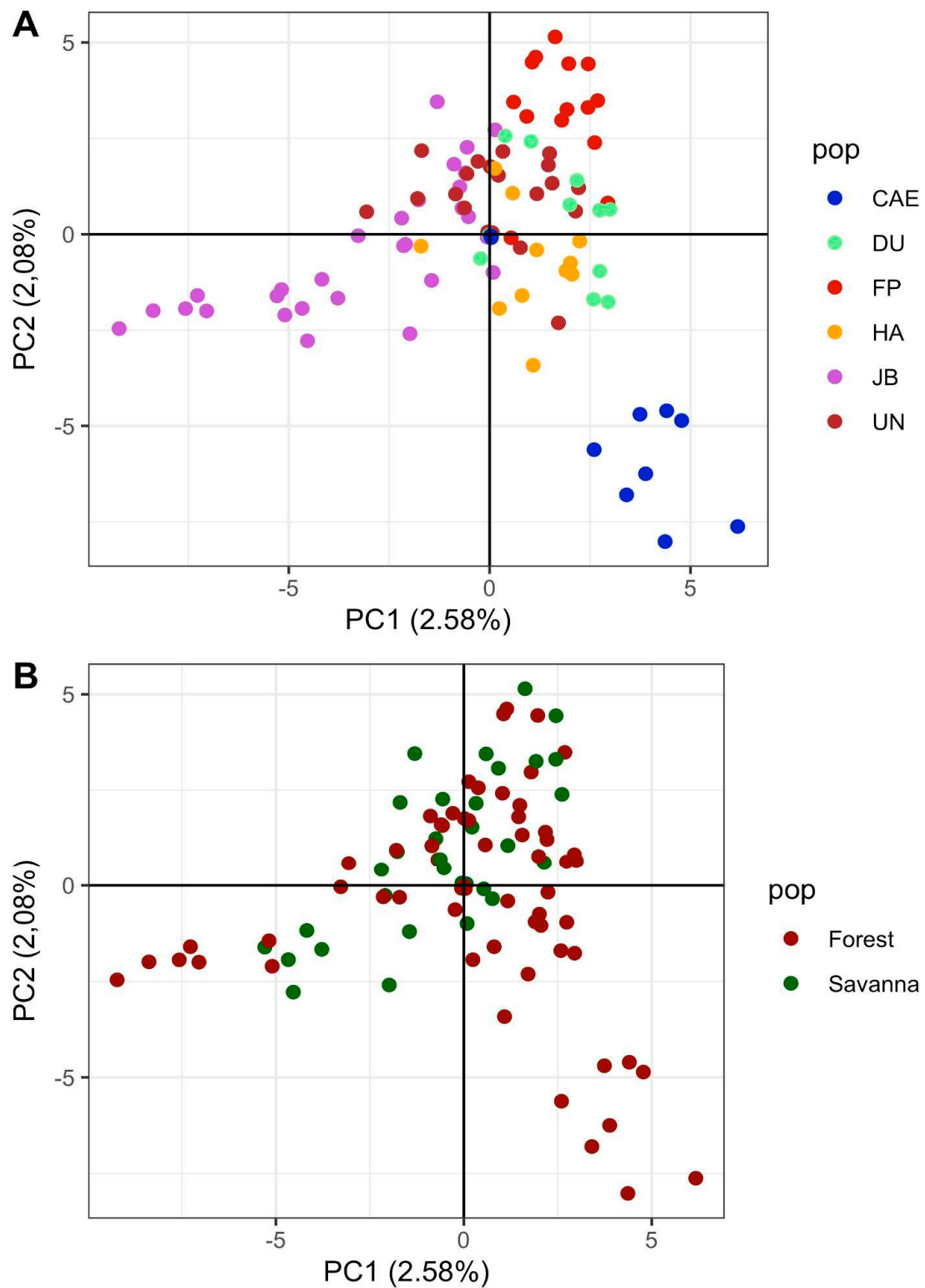


Figure 3. Principal component analysis of SNP markers obtained from *Copaifera langsdorffii* individuals sampled from sites within the Seasonal Semideciduous Forest and Brazilian savanna (Cerrado) at different collection sites. (A) Scatter plot colored by forest type. (B) Scatter plot colored based on collection site.

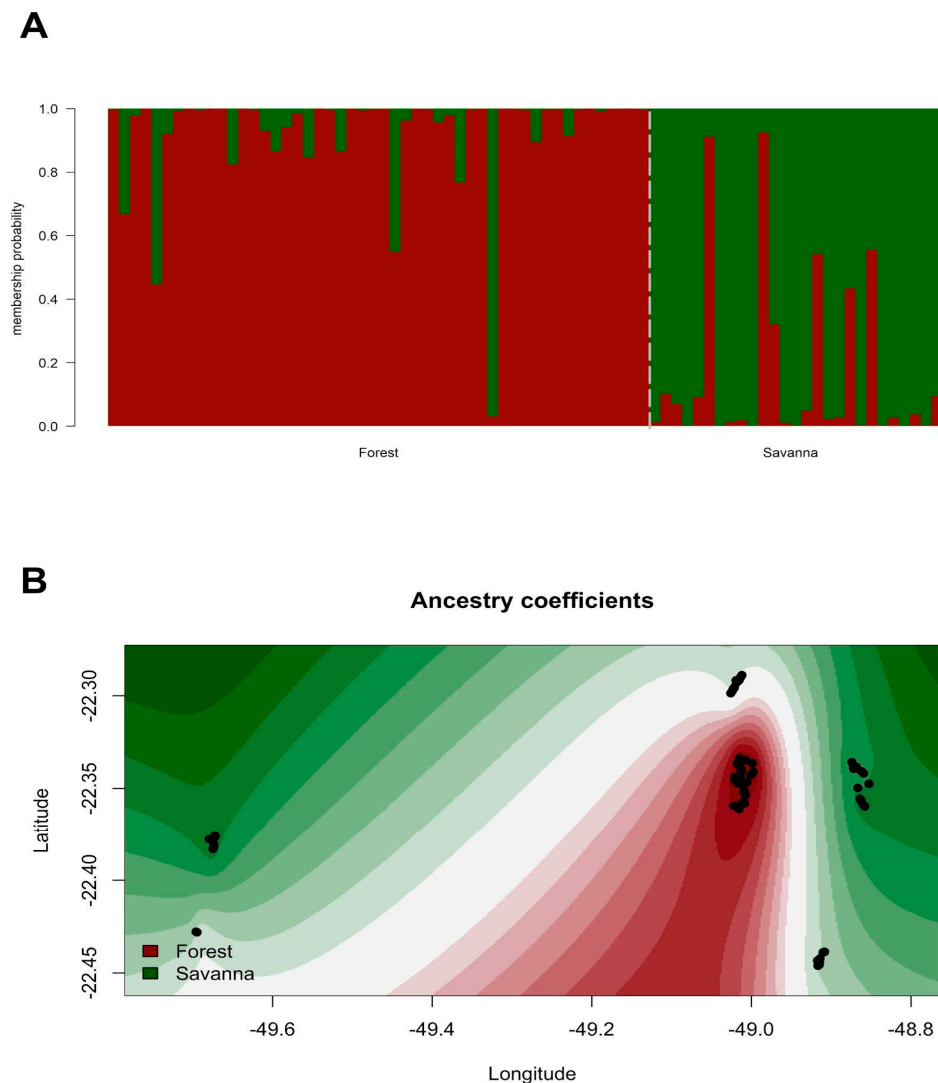


Figure 4. Genetic structure of *Copaifera langsdorffii* sampled from sites with Seasonal Semideciduous Forest and Brazilian savanna (Cerrado). (A) Genetic structure display based on $K = 2$ obtained with Structure. (B) Spatial interpolation of ancestry coefficients as obtained with Tess.

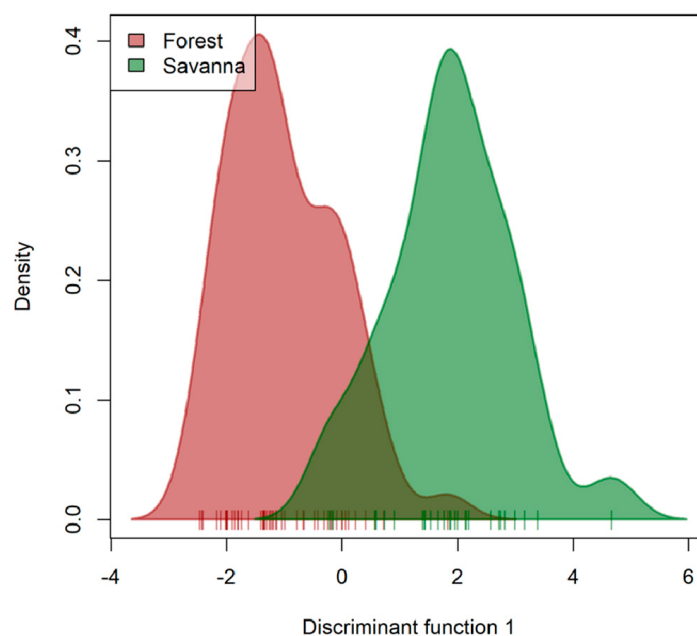
3.2. Outlier F_{ST} Analysis

Twenty-eight outlier loci were detected with Fdist2 analysis, which were significant after the FDR correction (Supplementary Material S1). All were confirmed as the most significant loci to separate the two groups by dAPC. Of these loci, 21 are monomorphic in the Savanna samples, indicating that they may be relevant in adapting the species to this environment. The dAPC analysis discriminated the same 28 loci as the SNPs with the greatest contribution to the differentiation of the clusters. The density graph (Figure 5A) shows that the samples are better grouped according to this contrast of environments using the 28 outlier loci. This dispersion pattern is very evident in the dAPC scatter plot (Figure 5B).

From our previous classification into two groups of ancestry, the Forest and Savanna groups, no clear differences in the genetic diversity estimates were observed (Figure 6). Here, we also estimated the allelic richness between the two groups, revealing that the coefficient of inbreeding is lower, on average, in the Savanna than in the Forest group. The allelic richness is also lower in the Savanna, on average. That difference is not clear, however, from the analysis of all SNP markers. When only the outliers are compared for

the estimates, the inbreeding is indeed lower in the Savanna, and the allelic richness is significantly lower in the Savanna than in the Forest group (Figure 6).

A



B

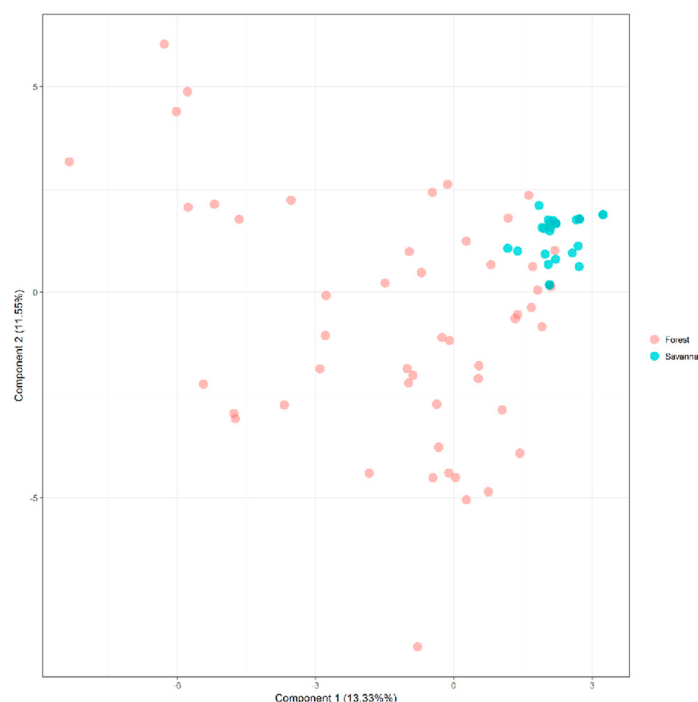


Figure 5. dAPC clustering of outlier SNP loci as determined based on the genotyping-by-sequencing of *Copaifera langsdorffii* populations sampled from two forest types: Season semideciduous forest and Brazilian savanna (Cerrado). **(A)** Individual density graph according to the defined groups (Cluster 1—Seasonal Semideciduous Forest; and Cluster 2—Savanna), along the first discriminant function obtained by DAPC, with cutoff = 0.001. **(B)** Scatter plot based on 28 outlier loci collected from six fragments in the Central-West Region of São Paulo State, Brazil.

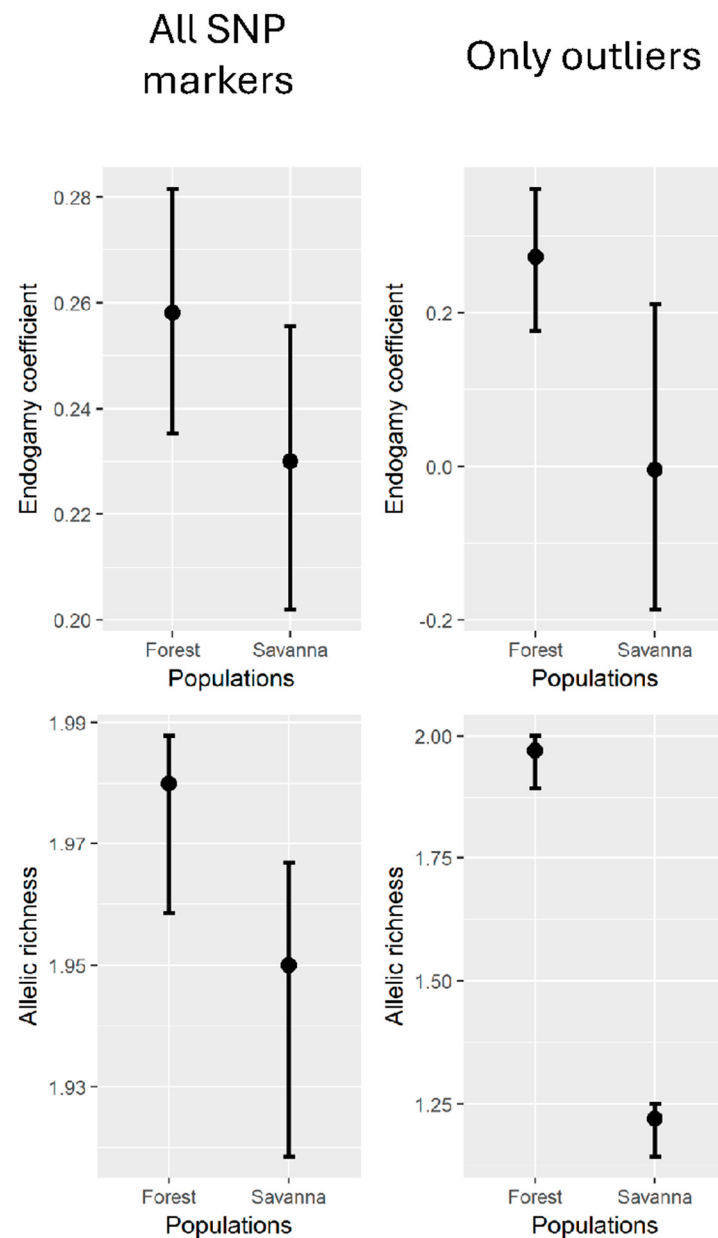


Figure 6. Inbreeding coefficient and allelic richness based on SNP markers obtained after genotyping-by-sequencing with samples of *Copaifera langsdorffii* collected from two forest types: seasonal semideciduous forest and Cerrado. The left column shows estimates with all SNP loci, while the right graph estimates are based on the outlier loci only. Vertical lines stand for $\pm 95\%$ CI.

4. Discussion

Our study revealed that all the sampled sites had an excess of homozygotes, whether belonging to the Forest or the Savanna group. Moreover, an important result revealed by our genotyping and sampling is a weak but noticeable genetic structure between the Forest and Savanna groups. Despite the spatial structure due to the isolation by distance, a trend of structuring by ancestry between the two groups was detected based on Bayesian inferences. Population structuring between Forest and Savanna was more prominent based on 28 outlier loci that were detected from the overall set of SNP markers, 21 of them in homozygosity in the samples belonging to the Savanna.

These 21 monomorphic loci in Savanna populations were identified based on the results of the *FST* outlier analysis. Using FDIST2 and confirmed by DAPC (Discriminant Analysis of Principal Components), we identified 28 candidate loci putatively under selec-

tion. Of these, 21 loci were found to be fixed (monomorphic) in the Savanna populations while remaining polymorphic in Forest populations.

This pattern suggests that selective pressures specific to the Savanna environment, such as nutrient-poor soils, water stress, and recurrent fire regimes, may have favored certain alleles, leading to the loss of polymorphism at these loci. In population genomics, loci that deviate significantly from neutral expectations (outliers) and show habitat-specific fixation are widely considered candidates for local adaptation, as they likely reflect directional or purifying selection eliminating maladaptive variants [23,63,66,67].

Similar approaches in forest trees have demonstrated that outlier SNPs fixed in contrasting environments often carry signatures of adaptive divergence [68,69]. Therefore, our interpretation is consistent with the broader framework of evolutionary and landscape genomics, where fixed outlier loci are considered indicative of adaptation to local edaphoclimatic conditions.

4.1. Excess of Homozygotes in All Sites

The excessive levels of homozygotes detected in each site are in line with most studies performed with different types of molecular markers, such as isozymes [28] and microsatellites [22,26,27,39,70]. The first recorded genetic study with natural populations of *C. langsdorffii* was that of [71], who identified an excess of homozygotes in two riparian areas based on allozyme markers. Anthropogenic intervention over the years may have directly affected the dispersers of the species and, consequently, generated an increase in homozygotes among these populations. Ref. [28] found, based on isoenzymatic markers, that 60 *C. langsdorffii* individuals in three populations (Cerrado, Atlantic Forest and Riparian Zone) had high observed and expected heterozygosity when compared to most species, both tropical and temperate. According to the authors, the Savanna and Semideciduous Seasonal Forest populations expressed a tendency for an excess of homozygotes. Exceptionally, they were more heterozygous in the Riparian Zone, contrary to what is commonly observed for other vegetation formations. Similar data were obtained by [26] analyzing 80 *C. langsdorffii* individuals based on six SSR loci, finding $H_o = 0.715$ and $H_e = 0.882$, on average, for the three forest areas collected, also indicating an excess of homozygotes. An excess of heterozygosity suggests inbreeding, the presence of null alleles, or the occurrence of the Wahlund effect. As *C. langsdorffii* has hermaphroditic flowers, the excessive homozygosity could be explained by selfing and mating among relatives [70].

4.2. Outcrossing Rates May Be High, but There Is Also Isolation by Distance

Despite the current excess of homozygotes, the F_{ST} values observed in the sampled sites of *C. langsdorffii* suggest high levels of outcrossing [72], which may result from historical gene flow among sites on a geological timescale. A similar study with *Copaifera reticulata* Ducke, in the Amazon region, using SSR markers, concluded that most of the genetic variation was found within the collection areas, with weak genetic differentiation between areas [73]. Similar data were observed with the tropical species *Protium spruceanum* Benth Engl [74] and *Prunus africana* (Hook. f.) [75], indicating low population genetic structure for both. Low F_{ST} values are found for most tree species (Hamrick and Godt 1996) [76]. This can be explained by a combination of factors throughout the life of these plants, specifically their breeding system [77]. Ref. [78] pointed out that the reasons for these low F_{ST} values may be because tropical trees have low but significant rates of self-fertilization and bi-parental crossing, as well as genetic drift due to the low population densities of most species.

A strong association between genetic divergence and geographic distances was observed for all six sites. Based on the environmental landscape and in a seed dispersal study [70] with the same species in the same area, four sampled sites are completely iso-

lated from each other (HA, CAE, FP, and DU) (Figure 1). Even with the sites JB and UN considered as two distinct populations, there is evidence that in the recent past, they were a unique large forest fragment. Ref. [27] also observed a strong association between genetic and geographic distances for *C. langsdorffii* in two experimental stations from the Forest Institute (Assis and Itirapina), located at Pedregulho Conservation Unit and in a private property at Brotas municipality, both in São Paulo State.

Spatial Genetic Structure (SGS) analysis for the *C. langsdorffii* populations suggested isolation by distance. Similar results were found in previous studies [22,27], indicating that spatial isolation is a product of forest fragmentation, which decreases the effective population size and consequently the genetic diversity. Ref. [70] highlighted seeds of *C. langsdorffii* generally have short-distance dispersal, and the isolation by distance seems due to restricted pollen and seed flow among populations. The SGS characterization, along with effective population size and effective pollen donor size estimations and models of both pollen and seed dispersals, are fundamental parameters to be considered to set up efficient strategies of tree conservation.

4.3. Population Structure Reflects the Transition Between Forests and Savannas

The Bayesian hierarchical clustering analysis showed that the six sampled sites were structured into two larger groups ($K = 2$), suggesting a distinction between the genotypes belonging to the Savanna and the Semideciduous Seasonal Forest. The spatial interpolation of the ancestry coefficients also revealed a trend of separation between the Savanna and the Forest. Because the Bauru region is an important transition area [79], the analysis of genetic structure corroborates the initial idea that there is a transition between these two forest formations. On the one hand, we observed two distinct vegetation structures on the JB population, illustrating here an important ecotone, and certainly this disposition is manifested by the separation of its genotypes (Figures 3 and 4). In addition, the DU population, whose forest area is a continuous part of the JB Semideciduous Seasonal Forest (Figures 3 and 4), presents its genotypes isolated from most of the JB genotypes sampled in Savanna areas. According to [80,81], stretches today covered by the Savanna vegetation correspond to areas where savannahs and/or countryside physiognomies predominated in the past, according to the descriptions of [82], in a study on the native vegetation of Bauru. According to [81], the comparison between old and current aerial images of the vegetation allows us to verify a change from countryside and savannah physiognomies to forest physiognomy.

4.4. Outlier Loci Analysis Highlights Putative Local Adaptations

Out of 28 candidate outlier loci pointed out by the neutrality test, 21 were monomorphic loci, leading to the hypothesis that the fixed loci have important evolutionary significance in the study areas. One hypothesis is that most of these loci may be associated with important alleles for adaptation to the savanna, and genotypes with polymorphism in such regions are negatively selected. It is likely that part of the 21 outlier loci that were fixed in these areas may be related to the adaptation to the edaphoclimatic conditions found in that environment.

One of the goals of evolutionary biology studies is to determine the influences of positive and purifying selection as well as neutral forces that shape the landscape genetic variation and natural populations that span large climate gradients, offering an ideal opportunity to investigate these evolutionary forces [23,67]. However, population genomic studies with tropical trees aiming to understand the speciation processes are still incipient [83], specifically in transition areas of vegetative physiognomies. Due to the strong anthropogenic pressures in the region of the present study, understanding adaptive genetic responses to climate change is one of the main challenges for the preservation of local

biological diversity. Similar studies have addressed genetic and ecological variables to understand changes in character expression or allelic frequencies over a series of contiguous populations within a given geographical distance.

Recent studies have sought to identify genes associated with geographic gradients, mostly in temperate regions. Meger et al. [68] identified 781 SNPs in 389 phenology-related genes associated with geographic, climatic, and phenotypic gradients in natural populations of *Quercus robur*, providing evidence of local adaptation to current and future climate conditions. Similarly, Kremer et al. [69] identified 38 outlier SNPs through whole-genome screening that showed strong geographic structure, allowing for the discrimination of four Western European white oak species and highlighting signatures of spatial genetic differentiation.

Our results corroborate those observed by [66], who suggest genetic differentiation in *Erythrina falcata* through a small spatial scale analysis, despite the increasingly reduced gene flow. The authors hypothesized that heterogeneous environments may cause molecular divergence, possibly associated with local adaptation, a situation similar to what was observed in the transition area in the Central-West Region of São Paulo, between the two phytophysiognomies studied.

Our findings point to several directions for future research that can strengthen conservation strategies for *C. langsdorffii*. The detection of outlier loci potentially associated with local adaptation highlights the need for functional studies to validate their role under contrasting phytophysiognomies. Landscape genomic approaches that integrate environmental variables may further elucidate adaptive processes and support the definition of seed transfer zones resilient to climate change. In addition, coupling demographic and genomic data will be essential to refine estimates of effective population size and dispersal, improving strategies to maintain genetic connectivity. Finally, the establishment of long-term monitoring and the integration of in situ and ex situ conservation initiatives will be crucial to safeguard both neutral and adaptive components of genetic diversity in fragmented landscapes.

5. Conclusions

Our study revealed that all population relicts sampled have an excessive level of homozygotes, which may be due to inbreeding or other underlying factors that need to be addressed for the proper design of conservation practices and management. A clear pattern of isolation by distance was detected from our study, with high correlation between geographic and genetic distances. Therefore, there is spatial genetic structure among the population relicts. Moreover, Bayesian inferences obtained by two algorithms, STRUCTURE and TESS, showed population structure of the ancestry coefficients. Spatial interpolation of the ancestry coefficients supported a trend of separation by the phytophysiognomy, the Atlantic Forest, or the Brazilian Savanna. The transition zone between these biomes partially influences population structuring.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants14182858/s1>. Supplementary Material S1: Excel spreadsheet; Supplementary Material S2: A—Jardim Botânico (JB), B—UNESP-Bauru (UN), C—Horto Aimorés (HA), D—Floresta Pederneiras (FP), E—Duratex (DU), F—E.E. Caetetus (CA), G—Copaifera langsdorffii leaves, H—Copaifera langsdorffii trunk

Author Contributions: M.V.B.M.S. and J.S.C. conceived the research; M.V.B.M.S., J.S.C. and W.O. performed the experiments; M.V.B.M.S., J.S.C. and M.M.B. analyzed the data; M.V.B.M.S., J.S.C., M.M.B., K.M., A.T.d.O.M., E.A.V., E.V.T. and E.R.K. wrote the manuscript; All authors contributed to editing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the São Paulo Research Foundation (FAPESP—Grant 2011/06756-4). The financial support for publication was provided by the Minas Gerais State University (Edital PROPPG N° 02/2025—PROPUBLIC/UEMG).

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: MVBMS thanks the many students who contributed to the collections as well as the institutions that facilitated the fieldwork. JSC thanks FAPESP for the scholarship. We are grateful to the three anonymous reviewers for their constructive feedback, insightful suggestions, and careful corrections, which greatly improved the quality of this manuscript.

Conflicts of Interest: The authors did not receive support from any organization for the submitted work. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no competing interests to declare that are relevant to the content of this article. Author A.T.d.O.M. was employed by the company Inova Genética. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Clare, E.; Garcia, C.; Bolan, S. Ecosystem functioning during biodiversity loss and recovery. *Oikos* **2024**, *2024*, e10154. [[CrossRef](#)]
2. Flores, B.M.; Montoya, E.; Sakschewski, B.; Nascimento, N.; Staal, A. Critical transitions in the Amazon forest system. *Nature* **2023**, *624*, 65–72. [[CrossRef](#)]
3. Jaureguiberry, P.; Titeux, N.; Wiemers, M.; Bowler, D.E.; Coscieme, L.; Golden, A.S.; Guerra, C.A.; Jacob, U.; Takahashi, Y.; Settele, J.; et al. The direct drivers of recent global anthropogenic biodiversity loss. *Sci. Adv.* **2022**, *8*, eabm9982. [[CrossRef](#)] [[PubMed](#)]
4. de Lima, R.A.F.; Oliveira, A.A.; Pitta, G.R.; Gasper, A.L.; Vibrans, A.C.; Chave, J.; Steege, H.; Prado, P. The erosion of biodiversity and biomass in the Atlantic Forest biodiversity hotspot. *Nat. Commun.* **2020**, *11*, 6347. [[CrossRef](#)] [[PubMed](#)]
5. Alencar, A.; Shimbo, J.Z.; Lenti, F.; Marques, C.B.; Zimbres, B.; Rosa, M.; Arruda, V.; Castro, I.; Ribeiro, J.P.F.M.; Varela, V.; et al. Mapping Three Decades of Changes in the Brazilian Savanna Native Vegetation Using Landsat Data Processed in the Google Earth Engine Platform. *Remote Sens.* **2020**, *12*, 924. [[CrossRef](#)]
6. Aguiar, B.I.; Sebbenn, A.M.; Tarazi, R.; Vogado, N.A.; Morellato, L.P.C.; Tambarussi, E.V.; Moreno, M.A.; Pereira, L.C.S.M.; Montibeller, C.; Ferraz, E.M.; et al. Phenology, Seed Germination, and Genetics Explains the Reproductive Strategies of *Diospyros lasiocalyx* (Mart.) B. Wall. *Trop. Plant Biol.* **2020**, *13*, 23–35. [[CrossRef](#)]
7. Rabeschini, G.; Persson, U.M.; West, C.; Kastner, T. Choosing fit-for-purpose biodiversity impact indicators for agriculture in the Brazilian Cerrado ecoregion. *Nat. Commun.* **2025**, *16*, 1799. [[CrossRef](#)]
8. Resende, F.M.; Cimon-Morin, J.; Poulin, M.; Meyer, L.; Loyola, R. Consequences of delaying actions for safeguarding ecosystem services in the Brazilian Cerrado. *Biol. Conserv.* **2019**, *234*, 90–99. [[CrossRef](#)]
9. Muniz, A.C.; Pimenta, R.J.G.; Cruz, M.V.; Rodrigues, J.G.; Buzatti, R.S.O.; Heuertz, M.; Lemos-Filho, J.P.; Lovato, M.B. Hybrid zone of a tree in a Cerrado/Atlantic Forest ecotone as a hotspot of genetic diversity and conservation. *Ecol. Evol.* **2022**, *22*, e8540. [[CrossRef](#)]
10. Sano, E.E.; Rodrigues, A.A.; Martins, E.S.; Bettiol, G.M.; Bustamante, M.M.C.; Bezerra, A.S.; Couto, A.F.; Vasconcelos, V.; Schüler, J.; Bolfe, E.L. Cerrado ecoregions: A spatial framework to assess and prioritize Brazilian savanna environmental diversity for conservation. *J. Environ. Manag.* **2019**, *232*, 818–828. [[CrossRef](#)]
11. Ramos, E.A.; Nuvoloni, F.M.; Lopes, E.R.N. Landscape Transformations and loss of Atlantic Forests: Challenges for conservation. *J. Nat. Conserv.* **2022**, *66*, 126152. [[CrossRef](#)]
12. Nery, E.K.; Caddah, M.K.; Santos, M.F.; Nogueira, A. The evolution of ecological specialization underlies plant endemism in the Atlantic Forest. *Ann. Bot.* **2023**, *131*, 921–940. [[CrossRef](#)]
13. Konzen, E.R. Towards conservation strategies for forest tree endangered species: The meaning of population genetic statistics. *Adv. For. Sci.* **2014**, *1*, 45–51. [[CrossRef](#)]
14. Feng, J.; Dan, X.; Cui, Y.; Gong, Y.; Peng, M.; Sang, Y.; Ingvarsson, P.K.; Wang, J. Integrating evolutionary genomics of forest trees to inform future tree breeding amid rapid climate change. *Plant Commun.* **2024**, *5*, 101044. [[CrossRef](#)] [[PubMed](#)]
15. Shaw, R.E.; Farquharson, K.A.; Bruford, M.W.; Coates, D.J.; Elliott, C.P.; Mergeay, J.; Ottewell, K.M.; Segelbacher, G.; Hoban, S.; Hvilsom, C.; et al. Global meta-analysis shows action is needed to halt genetic diversity loss. *Nature* **2025**, *638*, 704–710. [[CrossRef](#)]
16. Satam, H.; Joshi, K.; Mangrolia, U.; Waghoo, S.; Zaidi, G.; Rawool, S.; Thakare, R.P.; Banday, S.; Mishra, A.K.; Das, G.; et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* **2023**, *13*, 997. [[CrossRef](#)]

17. Miller, M.R.; Dunham, J.P.; Amores, A.; Cresko, W.A.; Johnson, E.A. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* **2007**, *17*, 240–248. [\[CrossRef\]](#)
18. Davey, J.W.; Blaxter, M.L. RADSeq: Next-generation population genetics. *Brief. Funct. Genom.* **2010**, *9*, 416–423. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Herrera, S.; Reyes-Herrera, P.H.; Shank, T.M. Predicting RAD-seq Marker Numbers across the Eukaryotic Tree of Life. *Genome Biol. Evol.* **2015**, *3*, 3207–3225. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Napolitano, C.; Clavijo, C.; Rojas-Bonzi, V.; Miño, C.I.; González-Maya, J.F.; Bou, N.; Giraldo, A.; Martino, A.; Miyaki, C.Y.; Aguirre, L.F.; et al. Understanding the conservation-genetics gap in Latin America: Challenges and opportunities to integrate genetics into conservation practices. *Front. Genet.* **2024**, *8*, 1425531. [\[CrossRef\]](#)
21. Lira, C.F. Neutral Genetic Diversity of Brazilian Native Flora: Current Approaches and Gaps. *Environ. Earth Sci. Proc.* **2024**, *31*, 7. [\[CrossRef\]](#)
22. Martins, K.; Santos, J.D.; Gaiotto, F.A.; Moreno, M.A.; Kageyama, P.Y. Estrutura genética populacional de *Copaifera langsdorffii* Desf. (Leguminosae-Caesalpinioideae) em fragmentos florestais no Pontal do Paranapanema, SP, Brasil. *Rev. Bras. Bot.* **2008**, *31*, 61–69. [\[CrossRef\]](#)
23. Carvalho, D.; Ingvarsson, P.K.; Joseph, J.; Suter, L.; Sedivy, C.; Macaya-Sanz, D.; Cottrell, J.; Heinze, B.; Schanzer, I.; Lexer, C. Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Mol. Ecol.* **2010**, *19*, 1638–1650. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Tarazi, R.; Sebbenn, A.M.; Mollinari, M.; Vencovsky, R. Mendelian inheritance, linkage and linkage disequilibrium in microsatellite loci of *Copaifera langsdorffii* Desf. *Conserv. Genet. Resour.* **2010**, *2*, 201–204. [\[CrossRef\]](#)
25. Manoel, R.O.; Cardin, L.T.; Moreira, J.P.; Silva, E.C.B.; Senna, S.N.; Kubota, T.Y.K.; Freitas, M.L.M.; Moraes, M.L.T.; Sebbenn, A.M. Sistema de reprodução, parentesco e tamanho efetivo em sementes de polinização livre de populações fragmentadas de *Copaifera langsdorffii* Desf. por análise de locos microssatélites. *Sci. For.* **2012**, *40*, 145–155.
26. Gonela, A.; Sebbenn, A.M.; Soriani, H.H.; Mestriner, M.A.; Martinez, C.A.; Alzate-Marin, A.L. Genetic diversity and mating system of *Copaifera langsdorffii* (Leguminosae/Caesalpinioideae). *Genet. Mol. Res.* **2013**, *12*, 569–580. [\[CrossRef\]](#)
27. Antikeira, L.M.O.R.; Souza, R.G.V.D.C.; Bajay, M.M.; Kageyama, P.Y. Genetic structure and diversity of *Copaifera langsdorffii* Desf. in Cerrado fragments of the São Paulo State, Brazil. *Rev. Árvore* **2014**, *38*, 667–675. [\[CrossRef\]](#)
28. Carvalho, D.; Oliveira, A.F. Genetic structure of *Copaifera langsdorffii* Desf. natural populations. *Cerne* **2015**, *10*, 137–153.
29. Azani, N.; Babineau, M.; Bailey, C.D.; Banks, H.; Barbosa, A.R.; Pinto, R.B.; Boatwright, J.S.; Borges, L.M.; Brown, G.K.; Bruneau, A.; et al. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny The Legume Phylogeny Working Group (LPWG). *Taxon* **2017**, *66*, 44–77. [\[CrossRef\]](#)
30. *Index Kewensis Supplement XX*; Clarendon Press: Oxford, UK, 1996.
31. Veiga Junior, V.; Pinto, A.C. O gênero *Copaifera* L. *Química Nova* **2002**, *25*, 273–286. [\[CrossRef\]](#)
32. Carvalho, P.E.R. *Espécies Arbóreas Brasileiras*; Embrapa Informação Tecnológica: Brasília, Brazil, 2003; Volume 1.
33. Oliveira, A.F.; Carvalho, D.; Rosado, S.C.S. Taxa de cruzamento e sistema reprodutivo de uma população natural de *Copaifera langsdorffii* Desf. na região de Lavras (MG) por meio de isoenzimas. *Rev. Bras. Bot.* **2002**, *25*, 331–338. [\[CrossRef\]](#)
34. Pedroni, F.; Sanchez, M.; Santos, F.A.M. Fenologia da copaíba (*Copaifera langsdorffii* Desf.-Leguminosae, Caesalpinioideae) em uma floresta semidecídua no sudeste do Brasil. *Rev. Bras. Bot.* **2002**, *25*, 183–194. [\[CrossRef\]](#)
35. Oliveira, E.C.P.; Lameira, O.A.; Zoghbi, M.G.B. Identificação da época de coleta do óleo-resina de copaíba (*Copaifera* spp.) no município de Moju, PA. *Rev. Bras. Plantas Med.* **2006**, *8*, 14–23.
36. DeVore, J. The mind of the copaíba tree: Notes on extractivism, animism and ontology from Southern Bahia. *J. R. Anthropol. Inst.* **2017**, *23*, 234–252. [\[CrossRef\]](#)
37. Almeida, S.D.; Proença, C.E.B.; Sano, S.M.; Ribeiro, J.F. *Cerrado: Espécies Vegetais Úteis*; Embrapa-CPAC: Planaltina, Brazil, 1998.
38. Siqueira, M.V.B.M.; Silvério, G.H.; Carlos, J.S.; Toledo, J.A.M.; da Silva, C.J.; de Paula-Souza, J.; Galastri, N.A. Phenotypic plasticity in *Copaifera langsdorffii* Desf. *An. Acad. Bras. Ciências* **2023**, *95*, e20220048.
39. Tarazi, R.; Sebbenn, A.; Kageyama, P.; Vencovsky, R. Long-distance dispersal in a fire- and livestock-protected savanna. *Ecol. Evol.* **2013**, *3*, 1003–1015. [\[CrossRef\]](#)
40. Trindade, R.; da Silva, J.K.; Setzer, W.N. *Copaifera* of the Neotropics: A Review of the Phytochemistry and Pharmacology. *Int. J. Mol. Sci.* **2018**, *19*, 1511. [\[CrossRef\]](#)
41. Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; de Moraes Gonçalves, J.L.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorol. Z.* **2013**, *22*, 711–728. [\[CrossRef\]](#)
42. Russello, M.A.; Waterhouse, M.D.; Etter, P.D.; Johnson, E.A. From promise to practice: Pairing non-invasive sampling with genomics in conservation. *PeerJ* **2015**, *3*, e1106. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Bushnell, B.; Rood, J.; Singer, E. BBMerge—Accurate paired shotgun read merging via overlap. *PLoS ONE* **2017**, *12*, e0185056. [\[CrossRef\]](#)

44. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
45. Danecek, P.; McCarthy, A.S. BCFtools/csq: Haplotype-aware variant consequences. *Bioinformatics* **2017**, *33*, 2037–2039. [CrossRef]
46. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]
47. Goudet, J. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Resour.* **2005**, *5*, 184–186. [CrossRef]
48. Jombart, T. ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef] [PubMed]
49. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: <https://www.R-project.org/> (accessed on 22 August 2023).
50. Keenan, K.; McGinnity, P.; Cross, T.F.; Crozier, W.W.; Prodöhl, P.A. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **2013**, *4*, 782–788. [CrossRef]
51. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292. [CrossRef]
52. Sokal, R.R.; Michener, C.D. A statistical method for evaluating systematic relationships. *Kans. Univ. Sci. Bull.* **1958**, *38*, 1409–1438.
53. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2014**, *2*, e281. [CrossRef]
54. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
55. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software Structure: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef] [PubMed]
56. Earl, D.A.; VonHoldt, B.M. Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [CrossRef]
57. Paradis, E.; Claude, J.; Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289–290. [CrossRef]
58. Dray, S.; Dufour, A.B. The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **2007**, *22*, 1–20. [CrossRef]
59. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2010.
60. Caye, K.; Deist, T.M.; Martins, H.; Michel, O.; François, O. TESS3: Fast inference of spatial population structure and genome scans for selection. *Mol. Ecol. Resour.* **2016**, *16*, 540–548. [CrossRef]
61. Hijmans, R.J.; van Etten, J.; Cheng, J.; Mattiuzzi, M.; Sumner, M.; Greenberg, J.A.; Lamigueiro, O.P.; Bevan, A.; Racine, E.B.; Shortridge, A. Raster: Geographic Data Analysis and Modeling. R Package Version 2.5-2. 2015. Available online: <https://CRAN.R-project.org/package=raster> (accessed on 7 June 2019).
62. South, A. rworldmap: A new R package for mapping global data. *R J.* **2011**, *3*, 35–43. [CrossRef]
63. Beaumont, M.A.; Nichols, R.A. Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B Biol. Sci.* **1996**, *263*, 1619–1626. [CrossRef]
64. Antão, T.; Lopes, A.; Lopes, R.J.; Beja-Pereira, A.; Luikart, G. LOSITAN: A workbench to detect molecular adaptation based on a F_{ST} -outlier method. *BMC Bioinform.* **2008**, *9*, 323. [CrossRef]
65. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [CrossRef]
66. Audigeos, D.; Brousseau, L.; Traissac, S.; Scotti-Saintagne, C.; Scotti, I. Molecular divergence in tropical tree populations occupying environmental mosaics. *J. Evol. Biol.* **2013**, *26*, 529–544. [CrossRef] [PubMed]
67. Agashe, D.; Sane, M.; Singhal, S. Revisiting the Role of Genetic Variation in Adaptation. *Am. Nat.* **2023**, *202*, 486–502. [CrossRef]
68. Meger, J.; Ulaszewski, B.; Chmura, D.J.; Burczyk, J. Signatures of local adaptation to current and future climate in phenology-related genes in natural populations of *Quercus robur*. *BMC Genom.* **2024**, *25*, 28. [CrossRef]
69. Kremer, A.; Delcamp, A.; Lesur, I.; Wagner, S.; Rellstab, C.; Guichoux, E.; Leroy, T. Whole-genome screening for near-diagnostic genetic markers for four western European white oak species identification. *Ann. For. Sci.* **2024**, *81*, 21. [CrossRef]
70. Sebbenn, A.M.; Carvalho, A.C.M.; Freitas, M.L.M.; Moraes, S.M.B.; Gaino, A.P.S.C.; Silva, J.M.; Jolivet, C.; Moraes, M.L.T. Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. *Heredity* **2011**, *106*, 134–145. [CrossRef]
71. Pinto, S.I.C.; Souza, A.M.; Carvalho, D. Variabilidade genética por iso-enzimas em populações de *Copaifera langsdorffii* Desf. em dois fragmentos de mata ciliar. *Sci. For.* **2004**, *65*, 40–48.
72. Wright, S. The interpretation of population structure by f-statistics with special regard to systems of mating. *Evolution* **1965**, *19*, 395–420. [CrossRef]
73. Silva, E.S.; Mathias, C.; Lima, M.C.F.; Veiga Junior, V.F.; Rodrigues, D.P.; Clement, C.R. Análise físico-química do óleo-resina e variabilidade genética de copaíba na Floresta Nacional do Tapajós. *Pesqui. Agropecu. Bras.* **2012**, *47*, 1621–1628. [CrossRef]

74. Vieira, F.A.; Carvalho, D. Genetic differentiation of the neotropical tree species *Protium spruceanum* (Benth.) Engler (Burseraceae) between fragments and vegetation corridors in Brazilian Atlantic forest. *Acta Bot. Bras.* **2009**, *23*, 1180–1185. [[CrossRef](#)]
75. Farwig, N.; Braun, C.; Böhning-Gaese, K. Human disturbance reduces genetic diversity of an endangered tropical tree, *Prunus africana* (Rosaceae). *Conserv. Genet.* **2008**, *9*, 317–326. [[CrossRef](#)]
76. Hamrick, J.L.; Godt, M.J.W. Effects of life history traits on genetic diversity in plant species. *Philos. Trans. R. Soc. B Biol. Sci.* **1996**, *351*, 1291–1298. [[CrossRef](#)]
77. Jordano, P.; Godoy, J.Á. RAPD variation and population genetic structure in *Prunus mahaleb* (Rosaceae), an animal-dispersed tree. *Mol. Ecol.* **2000**, *9*, 1293–1305. [[CrossRef](#)]
78. Dick, C.W.; Hardy, O.J.; Jones, F.A.; Petit, R.J. Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Trop. Plant Biol.* **2008**, *1*, 20–33. [[CrossRef](#)]
79. Cavassan, O.; Weiser, V.L. Vascular flora of the cerrado of Bauru-SP. *Biota Neotrop.* **2015**, *15*, e20140093. [[CrossRef](#)]
80. Faraco, A.G. Composição Florística e Fitossociologia de Uma Área de Cerrado na Reserva Legal do Campus de Bauru da UNESP-SP. Repositório Institucional UNESP, São Paulo, Brazil. 2007. Available online: <https://repositorio.unesp.br/handle/11449/88136> (accessed on 12 August 2019).
81. Cavassan, O. Bauru: Terra de cerrado ou floresta? *Ciênc. Geogr.* **2013**, *17*, 46–54.
82. Serra Filho, R.; Cavalli, A.C.; Guillaumon, J.R.; Chiarini, J.V.; Nogueira, F.P.; Ivancko, C.M.A.M.; Barbieri, J.L.; Donzeli, P.L.; Coelho, A.G.S.; Bittencourt, I. Levantamento da cobertura natural e do reflorestamento no cerrado estado de São Paulo. *Bol. Técnico Do Inst. Florest.* **1974**, *11*, 1–5.
83. Beheregaray, L.B.; Cooke, G.M.; Chao, N.L.; Landguth, E.L. Ecological speciation in the tropics: Insights from comparative genetic studies in Amazonia. *Front. Genet.* **2014**, *5*, 477. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

ID	Local	GPS	MSL (m)	DAP (cm)	Hight (m)
JB1	Jardim Botânico	22° 20' 12" S 49° 01' 04" W	572	182	16
JB2	Jardim Botânico	22° 20' 11" S 49° 01' 02" W	553	50	9
JB3	Jardim Botânico	22° 20' 10" S 49° 01' 02" W	554	118	10
JB4	Jardim Botânico	22° 20' 11" S 49° 01' 00" W	544	71	11
JB5	Jardim Botânico	22° 20' 10" S 49° 01' 00" W	544	72	11
JB6	Jardim Botânico	22° 20' 11" S 49° 00' 58" W	526	60	12
JB7	Jardim Botânico	22° 20' 10" S 49° 00' 59" W	555	155	10
JB8	Jardim Botânico	22° 20' 01" S 49° 00' 54" W	516	123	13
JB9	Jardim Botânico	22° 20' 02" S 49° 00' 52" W	523	92	14
JB10	Jardim Botânico	22° 20' 03" S 49° 00' 53" W	527	138	15
JB11	Jardim Botânico	22° 20' 23" S 49° 00' 47" W	528	90	4
JB12	Jardim Botânico	22° 20' 22" S 49° 00' 47" W	526	100	10
JB13	Jardim Botânico	22° 20' 23" S 49° 00' 47" W	531	75	7
JB14	Jardim Botânico	22° 20' 23" S 49° 00' 46" W	530	115	14
JB15	Jardim Botânico	22° 20' 24" S 49° 00' 46" W	545	140	13
JB16	Jardim Botânico	22° 20' 25" S 49° 00' 47" W	533	180	14
JB17	Jardim Botânico	22° 20' 26" S 49° 00' 47" W	533	37	6
JB18	Jardim Botânico	22° 20' 26" S 49° 00' 46" W	547	112	15
JB20	Jardim Botânico	22° 20' 36" S 49° 00' 51" W	514	108	16
JB21	Jardim Botânico	22° 20' 42" S 49° 00' 49" W	547	40	10
JB22	Jardim Botânico	22° 20' 46" S 49° 00' 51" W	549	69	12
JB23	Jardim Botânico	22° 20' 49" S 49° 00' 52" W	541	58	11
JB24	Jardim Botânico	22° 20' 43" S 49° 00' 53" W	534	128	16
JB25	Jardim Botânico	22° 20' 41" S 49° 00' 53" W	540	101,7	16
JB41	Jardim Botânico	22° 20' 11" S 48° 59' 55" W	581	105,5	16
JB42	Jardim Botânico	22° 20' 28" S 48° 59' 52" W	577	151	18
JB43	Jardim Botânico	22° 20' 34" S 49° 00' 04" W	586	38	4
JB44	Jardim Botânico	22° 20' 48" S 49° 00' 21" W	597	38	6
JB45	Jardim Botânico	22° 20' 05" S 49° 00' 26" W	600	25,67	3
UN26	UNESP-Bauru	22° 20' 50" S 49° 01' 01" W	561	108	12
UN28	UNESP-Bauru	22° 20' 47" S 49° 01' 03" W	567	44	7
UN29	UNESP-Bauru	22° 20' 46" S 49° 01' 05" W	572	53	10
UN30	UNESP-Bauru	22° 20' 41" S 49° 01' 06" W	561	29,67	5
UN31	UNESP-Bauru	22° 20' 43" S 49° 01' 08" W	574	99	10
UN32	UNESP-Bauru	22° 20' 43" S 49° 01' 09" W	565	44	12
UN33	UNESP-Bauru	22° 20' 43" S 49° 01' 01" W	567	64	12
UN34	UNESP-Bauru	22° 20' 44" S 49° 01' 13" W	573	64	13
UN36	UNESP-Bauru	22° 20' 42" S 49° 01' 13" W	568	97	16
UN37	UNESP-Bauru	22° 20' 42" S 49° 01' 13" W	563	52	12
UN38	UNESP-Bauru	22° 20' 43" S 49° 01' 16" W	562	44	8
UN39	UNESP-Bauru	22° 20' 40" S 49° 01' 15" W	553	104	10
UN40	UNESP-Bauru	22° 20' 38" S 49° 01' 14" W	549	14,5	4
UN46	UNESP-Bauru	22° 21' 10" S 49° 00' 31" W	603	171	20
UN47	UNESP-Bauru	22° 21' 04" S 49° 00' 34" W	601	73	5
UN48	UNESP-Bauru	22° 21' 13" S 49° 00' 27" W	603	99	9
UN49	UNESP-Bauru	22° 21' 29" S 49° 00' 31" W	615	104	10
UN50	UNESP-Bauru	22° 21' 28" S 49° 00' 43" W	615	60,5	7
UN51	UNESP-Bauru	22° 21' 40" S 49° 00' 56" W	609	168	15
UN52	UNESP-Bauru	22° 21' 34" S 49° 01' 20" W	610	123	8

HA53	Horto Aimorés	22° 17' 55" S 49° 01' 33" W	520	121	10
HA54	Horto Aimorés	22° 17' 51" S 49° 01' 27" W	521	54	9
HA55.1	Horto Aimorés	22° 17' 47" S 49° 01' 22" W	524	79	6
HA55.2	Horto Aimorés	22° 17' 47" S 49° 01' 22" W	524	79	6
HA56	Horto Aimorés	22° 17' 44" S 49° 01' 17" W	521	99	7
HA57	Horto Aimorés	22° 17' 35" S 49° 01' 08" W	528	95	12
HA58	Horto Aimorés	22° 17' 30" S 49° 01' 09" W	537	70	12
HA59	Horto Aimorés	22° 17' 31" S 49° 00' 59" W	550	150	15
HA60	Horto Aimorés	22° 17' 30" S 49° 00' 56" W	553	200	15
HA61	Horto Aimorés	22° 17' 24" S 49° 00' 49" W	556	94	10
HA62	Horto Aimorés	22° 17' 20" S 49° 00' 43" W	558	94	8
FP63	Floresta Pederneiras	22° 20' 22" S 48° 52' 19" W	201	129	15
FP64	Floresta Pederneiras	22° 20' 17" S 48° 52' 20" W	250	62	10
FP65	Floresta Pederneiras	22° 20' 09" S 48° 52' 27" W	519	81	10
FP66	Floresta Pederneiras	22° 20' 19" S 48° 52' 16" W	579	67,5	15
FP67	Floresta Pederneiras	22° 20' 18" S 48° 52' 07" W	577	61	16
FP68	Floresta Pederneiras	22° 20' 22" S 48° 52' 01" W	558	67,5	14
FP69	Floresta Pederneiras	22° 20' 27" S 48° 51' 42" W	545	68	10
FP70	Floresta Pederneiras	22° 00' 31" S 48° 51' 34" W	536	212	20
FP71	Floresta Pederneiras	22° 21' 35" S 48° 51' 29" W	621	80	6
FP72	Floresta Pederneiras	22° 21' 31" S 48° 51' 37" W	575	44	10
FP73	Floresta Pederneiras	22° 21' 23" S 48° 51' 46" W	581	87,67	5
FP74	Floresta Pederneiras	22° 21' 20" S 48° 51' 50" W	578	62,67	5
FP75	Floresta Pederneiras	22° 20' 59" S 48° 52' 00" W	575	71,5	13
FP76	Floresta Pederneiras	22° 20' 51" S 48° 51' 10" W	559	202	24
DU77	Duratex	22° 06' 46" S 48° 55' 00" W	578	27	4
DU78	Duratex	22° 26' 44" S 48° 54' 58" W	584	50	6
DU79	Duratex	22° 26' 44" S 48° 54' 54" W	580	52	7
DU80	Duratex	22° 26' 41" S 48° 54' 54" W	581	26	5
DU81	Duratex	22° 26' 38" S 48° 54' 53" W	577	58	10
DU82	Duratex	22° 26' 32" S 48° 54' 51" W	580	52	9
DU83	Duratex	22° 26' 20" S 48° 54' 35" W	568	86	11
DU84	Duratex	22° 26' 20" S 48° 54' 37" W	566	69	7
DU85	Duratex	22° 26' 19" S 48° 54' 30" W	564	47	8
DU86	Duratex	22° 26' 36" S 48° 55' 02" W	583	60	20
DU87	Duratex	22° 26' 37" S 48° 55' 01" W	577	106	18
CAE88	E.E. Caetetus	22° 25' 41" S 49° 41' 41" W	557	221,5	10
CAE89	E.E. Caetetus	22° 25' 40" S 49° 41' 45" W	548	206,5	18
CAE90	E.E. Caetetus	22° 22' 39" S 49° 40' 27" W	679	25	6
CAE91	E.E. Caetetus	22° 22' 40" S 49° 40' 46" W	638	50	14
CAE92	E.E. Caetetus	22° 22' 34" S 49° 40' 21" W	670	65	12
CAE93	E.E. Caetetus	22° 22' 34" S 49° 40' 18" W	680	65,5	7
CAE94	E.E. Caetetus	22° 22' 50" S 49° 40' 25" W	675	80	3
CAE96	E.E. Caetetus	22° 22' 54" S 49° 40' 25" W	679	41	7
CAE97	E.E. Caetetus	22° 22' 59" S 49° 40' 29" W	674	26	8
CAE98	E.E. Caetetus	22° 22' 58" S 49° 40' 28" W	679	44	12

[illegible]

[illegible]

Supplementary Material 2

A – Jardim Botânico (JB), B – UNESP-Bauru (UN), C - Horto Aimorés (HA), D - Floresta Pederneiras (FP), E - Duratex (DU), F - E.E. Caetetus (CA), G - *Copaifera langsdorffii* leaves, H - *Copaifera langsdorffii* trunk

