



Genome-wide SNPs reveal adaptation and population structure in *Syagrus romanzoffiana* (Arecaceae) from southern South America

Kauanne Karolline Moreno Martins¹ · Matheus Scaketti² · Ana Flávia Francisconi² · Igor Araújo Santos de Carvalho² · Caroline Bertocco Garcia² · Thiago Deomar Ludwig² · Enéas Ricardo Konzen³ · Maria Imaculada Zucchi^{1,2,4}

Received: 14 October 2025 / Accepted: 15 April 2026 / Published online: 14 May 2026
© The Author(s) 2026

Abstract

Palms play crucial ecological and economic roles, and *Syagrus romanzoffiana* is a widely distributed species in South America. Despite its ecological importance and economic potential, population-level genetic studies of this palm remain limited. Here, we assessed genetic diversity, population structure, and signatures of local adaptation across distinct vegetation types in Brazil and Paraguay using 24,859 single nucleotide polymorphisms (SNPs) derived from genotyping-by-sequencing (GBS). We analyzed 91 individuals from 8 populations, revealing high genetic diversity within populations and moderate differentiation among them. Observed heterozygosity was generally high, with some populations showing an excess of heterozygotes, consistent with outcrossing and gene flow. Pairwise F_{ST} and AMOVA confirmed that most variation is maintained within populations, while genetic structure analyses highlighted the influence of vegetation type on differentiation, with *restinga* populations displaying distinct genetic signatures. Despite geographic proximity, some populations exhibited greater divergence associated with local environmental conditions, supporting isolation by environment. Environmental association analyses (LFMM and RDA) identified SNPs associated with temperature and precipitation, indicating adaptive genomic differentiation across contrasting habitats. Together, these findings demonstrate how habitat type and climate jointly shape neutral and adaptive genetic variation in *Syagrus romanzoffiana*, providing a genomic baseline relevant for conservation, restoration, and seed sourcing strategies under ongoing habitat fragmentation and climate change.

Keywords Neotropical palms₁ · Genetic diversity₂ · Population genomics₃ · Molecular markers₄ · Conservation₅ · Local adaptation₆

Introduction

Understanding how genetic diversity and population structure are distributed across heterogeneous environments is a central question in conservation genetics, with direct implications for restoration, management, and the long-term persistence of wild populations (Helyar et al. 2011; Brumfield et al. 2003). Across heterogeneous landscapes, populations may diverge not only through neutral processes such as genetic drift and restricted gene flow, but also through environmentally driven selection, a distinction increasingly recognized as critical for both basic and applied conservation science (Allendorf et al. 2010). While analyses based on neutral genetic markers provide insights into demographic history and connectivity, they are limited in their ability to detect adaptive responses to local environmental conditions.

✉ Kauanne Karolline Moreno Martins
k228341@dac.unicamp.br

✉ Maria Imaculada Zucchi
mizucchi@sp.gov.br

¹ Biology Institute, State University of Campinas – UNICAMP, Campinas, São Paulo, Brazil

² Department of Genetics, University of São Paulo – ESALQ/USP, Piracicaba, São Paulo, Brazil

³ Center for Limnological, Coastal and Marine Studies, North Coastal Campus, Federal University of Rio Grande do Sul – UFRGS, Imbé, Rio Grande do Sul, Brazil

⁴ Secretariat of Agriculture and Food Supply of São Paulo State, APTA, UPDR-Piracicaba, São Paulo, Brazil

Identifying genomic loci underlying local adaptation is therefore essential to disentangle demographic from selective processes, particularly in widespread species distributed across contrasting habitats. For species used in ecological restoration and seed sourcing programs, accounting for adaptive genetic variation is especially relevant: translocating individuals without considering local adaptation may compromise the long-term success of restoration efforts under ongoing habitat fragmentation and climate change (Hoffmann and Sgrò 2011; Hemstrom et al. 2024). Genome-wide approaches such as genotyping-by-sequencing (GBS) now provide the resolution needed to simultaneously characterize neutral population structure and scan for signatures of selection across thousands of loci, making them powerful tools for addressing these questions in non-model organisms (Hohenlohe et al. 2010; Helyar et al. 2011).

Palm species (Arecaceae), with more than 2,600 recognized species worldwide, have provided a wide range of services to humankind over centuries (Cámara-Leret et al. 2017; Eiserhardt et al. 2011; Levis et al. 2017). Many are ecological keystone species, with large numbers of animals depending on them for fruit and flower resources to survive (Onstein et al. 2017). Moreover, palms supply non-timber forest products such as food, medicine, fibers, oils, and building materials, supporting both traditional communities and modern industries (Sen and Samanta 2015; Colombo et al. 2018; Balslev et al. 2011).

In palms, environmental factors such as temperature and precipitation are expected to play a central role in shaping species distributions and adaptive limits, particularly at regional and continental scales (Eiserhardt et al. 2011). At the southern limits of palm distribution, low winter temperatures and seasonal water availability can impose strong selective pressures, influencing survival, growth, and reproductive success. Experimental and field-based studies in related palm species have demonstrated that tolerance to cold and frost can directly constrain distributional ranges and population persistence (Gatti et al. 2008). Therefore, climatic gradients provide a biologically meaningful axis along which adaptive genetic variation may be detected. Other environmental factors, such as soil properties and hydrology, are expected to act at finer spatial scales and were not explicitly modeled in this regional analysis.

Syagrus romanzoffiana (Cham.) Glassman, commonly known as jerivá, is a widespread palm native to southern and eastern South America, occurring in Brazil, Paraguay, northern Argentina, and Uruguay (Lorenzi 2010). In addition to its ecological importance, *S. romanzoffiana* is widely cultivated in landscaping and urban forestry and supports rural communities through products derived from fruits, leaves, stems, and seeds (Zona and Henderson 1989; Scarriot 2015). Despite its broad distribution and socioeconomic

relevance, population-level genomic studies of this species remain scarce.

Across its southern distribution, *S. romanzoffiana* occupies environments that differ markedly in climate, vegetation structure, soil characteristics, and landscape history, including Atlantic Rain Forest, Atlantic Seasonal Forest, and coastal *restinga* formations. *Restinga* environments are characterized by sandy soils, high salinity, and strong wind exposure, whereas inland forests exhibit more stable soils, higher canopy cover, and distinct temperature and precipitation regimes (Lorenzi 2010; Scarano 2009; Henderson et al. 1995). Such environmental heterogeneity provides a natural framework to test whether genetic differentiation in this species reflects not only geographic distance but also adaptation to contrasting habitats, consistent with isolation by environment (IBE) (Wang and Bradburd 2014).

In this study, we focused on the southern range of *S. romanzoffiana*, sampling populations in southern Brazil and Paraguay to establish a regional genomic baseline. We hypothesized that (i) population genetic structure would reflect both habitat type and geographic context, and (ii) climatic variables, particularly temperature and precipitation, would explain patterns of adaptive genomic variation across vegetation types. To test these hypotheses, we used genome-wide single nucleotide polymorphisms (SNPs) obtained through genotyping-by-sequencing (GBS) to (i) assess genetic diversity and population structure and (ii) identify genomic regions associated with environmental variables. As the first SNP-based population genomic study of *S. romanzoffiana*, this work provides a high-resolution baseline for integrating neutral and adaptive genetic variation into conservation and restoration planning in Atlantic Forest and *restinga* ecosystems.

Materials and methods

Plant material

We sampled leaf tissues from 91 adult and reproductive individuals of *Syagrus romanzoffiana* from different locations representing three vegetation types in which they are found in southern South America: Atlantic Rain Forest, Atlantic Seasonal Forest, and *restinga* (Henderson et al. 1995; Lorenzi 2010). The samples were obtained from wild population remnants only, and included Linha Lucena-Cecília, Travessa, and Herval (Venâncio Aires, RS, Brazil; Atlantic Seasonal Forest); San Ramón (Concepción, Paraguay; Seasonal Forest); Pixirica (Morrinhos do Sul, RS, Brazil) and Borússia (Osório, RS, Brazil; Atlantic Rain Forest); and Condomínio Marítimo, (Cond. Marítimo) (Tramandaí, RS, Brazil) and Gravatá (Laguna, SC, Brazil; *restinga*) (Fig. 1).

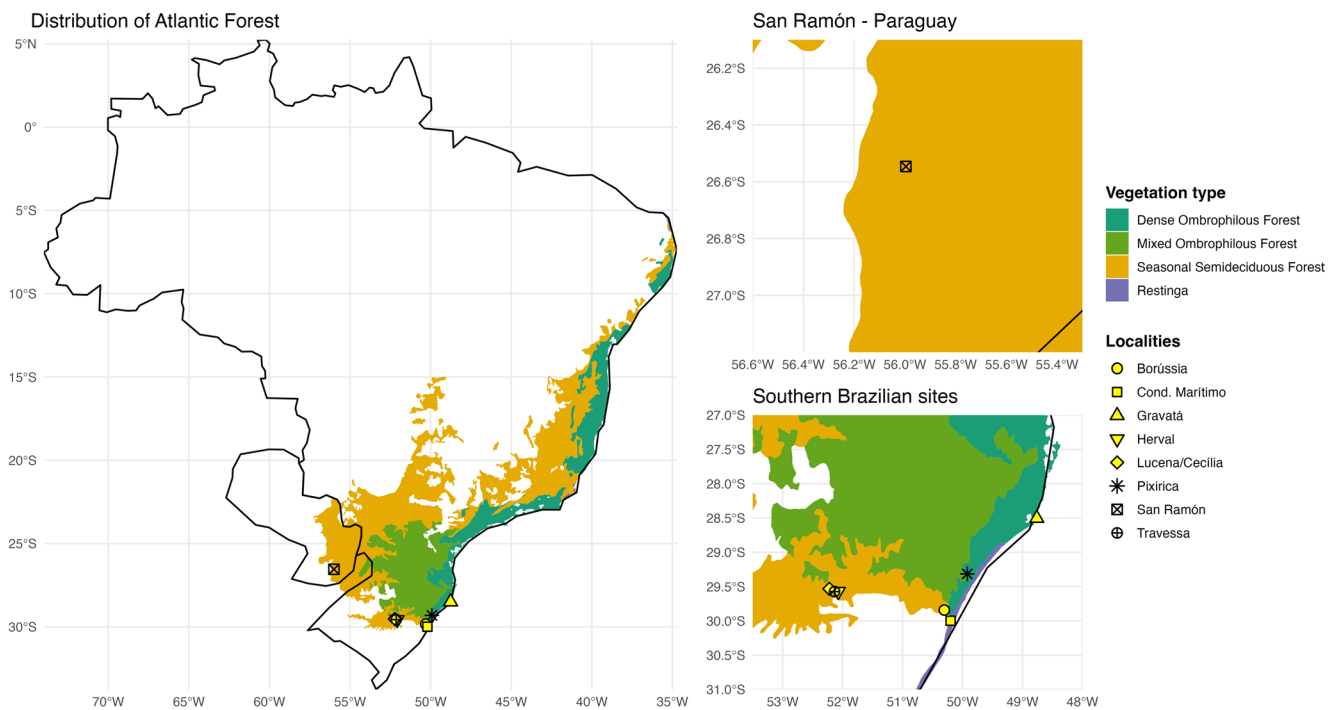


Fig. 1 Sampled populations of *Syagrus romanzoffiana* in Brazil and Paraguay. The main map shows the study region between Brazil and Paraguay; inset maps provide zoomed views of the sampling areas in Paraguay (upper right) and southern Brazil (bottom right). Each symbol represents a sampled population: Borússia – RS (dark green), Cond. Marítimo – RS (orange), Gravatá – SC (pale blue), Herval –

RS (pink), Lucena-Cecília – RS (light green), Pixirica – RS (yellow), Travessa – RS (gray), and San Ramón – Concepción, Paraguay (light brown). The vegetation types and distribution of the Atlantic Forest were delimited according to shape files obtained from the Ecoregions 2017 platform (Dinerstein et al., 2017)

Within each population, individuals were sampled with sufficient spatial separation to reduce the likelihood of sampling closely related individuals. The minimum geographic distance between sampled populations was approximately 6 km, while the maximum distance reached approximately 1,150 km, encompassing a broad regional spatial scale across southern Brazil and Paraguay. This spatial extent allows the evaluation of both geographic and environmental contributions to population genetic differentiation. Leaf samples were stored in silica gel and transported to the Department of Genetics at Luiz de Queiroz College of Agriculture – University of São Paulo – ESALQ/USP, for DNA extraction.

This study was approved by the Ethics Committee of the University of Campinas and conducted in accordance with Brazilian Ministry of the Environment (MMA) regulations and the National System for Genetic Heritage and Associated Traditional Knowledge (SisGen; registration A3A928).

DNA extraction

DNA was extracted from leaf tissue using a modified CTAB protocol (Doyle 1991) optimized for palms. DNA integrity

was checked on 1% agarose gels, and concentration was standardized to 30 ng/μL after quantification with a Qubit fluorometer (Thermo Fisher Scientific).

A detailed step-by-step description of the extraction protocol, including buffer composition, centrifugation conditions, and incubation times, is provided in Supplementary Information to facilitate reproducibility.

Genotyping-by-sequencing (GBS) library preparation

Genomic libraries were prepared using a double-digest GBS protocol (Poland et al. 2012) with modifications. Genomic DNA (30 ng/μL) was digested with PstI and MseI, followed by ligation of Illumina-compatible adapters with unique barcodes. Libraries were purified (QIAquick PCR Purification Kit, Qiagen), enriched by PCR with Illumina primers, and quantified by qPCR. Fragment size distribution was evaluated on a 2100 Bioanalyzer (Agilent Technologies), and sequencing was performed on the Illumina NextSeq 1000/2000 platform at Luiz de Queiroz College of Agriculture – University of São Paulo – (ESALQ-USP).

A complete description of adapter design, PCR conditions, and purification steps is provided in Supplementary Information to ensure reproducibility.

SNP discovery and filtering

Raw read quality was first assessed with *FastQC* (Andrews 2010). Demultiplexing and initial filtering were conducted in *Stacks v2.62* (Catchen et al. 2013) using the *process_radtags* module, to remove low-quality reads ($\text{Phred} < 20$), adapter contamination, uncalled bases (Ns), and sequences lacking restriction sites. *De novo* assembly was performed in *Stacks: ustacks* was run with a minimum depth of coverage of 4 ($-m\ 4$), allowing up to 5 mismatches between stacks within individuals ($-M\ 5$) (Boutet et al. 2016). A catalog of loci was built with *cstacks* (tolerance of 4 mismatches among individuals) (Boutet et al. 2016), followed by *tsv2bam* and genotyping in *gstacks*.

Variant filtering was conducted in *VCFTools* (Danecek et al. 2011). To reduce linkage disequilibrium, one SNP per locus was retained for structure analyses. Filters applied were per-locus missingness $\leq 20\%$, per-individual missingness $\leq 20\%$, $\text{MAF} \geq 0.05$, and Hardy–Weinberg equilibrium within populations ($p < 1 \times 10^{-6}$). Filtering choices followed the guidelines of Hemstrom et al. (2024).

Population genetic diversity

We estimated expected heterozygosity (H_s), observed heterozygosity (H_o), allelic richness (Ar) and fixation index (F_{IS}) using the *hierfstat* package (Goudet 2005). The number of alleles was calculated with *adegenet* package (Jombart and Ahmed 2011), and private alleles were identified using *poppr* package (Kamvar et al. 2014).

Population genetic structure

To characterize population genetic structure and disentangle the effects of geographic distance and environment, we combined complementary analytical approaches. Wright's F-statistics and AMOVA were used to quantify genetic differentiation among populations, Mantel tests were applied to evaluate isolation by distance, and multivariate and clustering methods (DAPC and sNMF) were used to explore population structure without explicit spatial assumptions.

Wright's F statistics (F_{ST} , F_{IT} , F_{IS}) and pairwise estimates of F_{ST} were estimated using *hierfstat* (Goudet 2005). Because the pairwise F_{ST} matrix is symmetric, the upper triangle of the matrix was used to display geographic Euclidean distances among populations. Geographic distances were calculated as Euclidean distances from population coordinates (latitude and longitude) using the *dist* function in R. Analyses of Molecular Variance (Excoffier et al. 1992) were conducted in *poppr* package (Kamvar et al. 2014), and the significance of the AMOVA was tested through 20,000 permutations, using population identity

as the hierarchical grouping factor. To minimize bias, loci with more than 5% missing data were excluded. A more stringent missing data threshold was applied for AMOVA to reduce potential bias in variance partitioning, whereas a less restrictive threshold was used in other analyses to retain statistical power for population structure and environmental association tests.

A Mantel test was performed with *ade4* (Mantel 1967) using 10,000 permutations testing for correlations between genetic and geographic distances matrices (Thioulouse et al. 2018). This analysis was used to assess isolation by distance (IBD), and its results were interpreted in contrast with environmental association analyses to evaluate the relative contributions of geographic and environmental factors to genetic differentiation. Population structure was further explored with discriminant analysis of principal components (DAPC) using *adegenet* (Jombart and Ahmed 2011). The optimal number of PCs was determined using α -score optimization, and results were visualized with scatterplots and barplots.

Environmental association and detection of candidate adaptive loci

Prior to environmental association analyses, genotype data were imputed to reduce potential bias associated with missing values. Although overall missing data per individual remained below 20%, a subset of loci exceeded 10% missingness. In such cases, genotype imputation is recommended to improve statistical power and stability in LFMM analyses (Frichot 2013). Imputation was performed using ancestry coefficients inferred from sNMF, with $K=4$, reflecting the main neutral genetic structure of the dataset. This ancestry-informed approach accounts for population structure while minimizing spurious genotype–environment associations.

Environmental association analyses were conducted with Latent Factor Mixed Models (LFMM) implemented in the *LEA* package (Frichot and François 2015). Twenty bioclimatic variables from *WorldClim v2* (Fick and Hijmans 2017) and elevation were considered. *WorldClim* variables were selected due to their widespread use in landscape genomics studies, facilitating comparison with previous work, and because their spatial resolution is appropriate for the regional scale of sampling considered here. To reduce collinearity among predictors, a Principal Component Analysis (PCA) was performed; the first two PCs explained $\sim 85\%$ of the variance (Figure S1, methodological). The most correlated variables were retained: bio06 (minimum temperature of the coldest month) and bio19 (precipitation of the coldest quarter), both commonly reported as drivers of local adaptation in tropical plants species (Eckert et al. 2009; Manel et al. 2010; De Kort et al. 2014).

These two variables were also used to calculate environmental distances among populations. Specifically, values of bio6 and bio19 were extracted for each population based on their geographic coordinates, and a pairwise environmental distance matrix was computed using Euclidean distance. A Mantel test was then performed using the *ade4* package (Mantel 1967), with 10,000 permutations, to evaluate the correlation between environmental and geographic distance matrices. This analysis aimed to assess the extent to which environmental variation covaries with geographic distance among populations. LFMM analyses were performed on the imputed genotype matrix using $K=4$ latent factors, as defined by cross-entropy optimization from preliminary sNMF analyses (10 repetitions, 200,000 iterations). Each LFMM run consisted of 10,000 iterations, including a burn-in of 5,000 iterations, with 10 replicates per environmental variable. P-values were adjusted using genomic inflation (λ) and the Benjamini–Hochberg false discovery rate procedure ($q < 0.05$) (François et al. 2016; Ahrens et al. 2019).

Candidate SNPs identified by LFMM were used in two downstream analyses: (i) a second sNMF analysis restricted to candidate loci, and (ii) Redundancy Analysis (RDA) implemented in the *vegan* package (Oksanen et al. 2024). In the RDA, SNP genotypes were treated as response variables, with bio06 and bio19 used as explanatory constraints. All analyses were performed in R v4.3.0 (R Core Team 2024).

Results

Genetic diversity in southern South America

After data filtering, a total of 24,859 SNP markers were retained for the *S. romanzoffiana*. The average percentage of missing data per individual was 6.11%, with a maximum of 14% (Figure S2A). The average sequencing coverage per individual was 15.25x (Figure S2B), and the mean coverage per locus was 13.68x, with no coverage exceeding 100x or falling below 10x per locus (Figure S3A). The SNP data

also revealed a higher number of transitions compared to transversions (Figure S3B).

Genetic diversity varies across sites (Table 1). Observed heterozygosity (H_o) and expected (H_s) heterozygosity were high across populations, with H_o exceeding H_s in six sites. Lucena-Cecilia was an exception, showing similar H_o and H_s values (0.169), while Travessa had a lower H_o than H_s . The highest H_o values were in San Ramón (0.205) and Gravatá (0.175), with the highest H_s also in San Ramón (0.203), followed by Lucena-Cecilia (0.169). San Ramón, Gravatá and Lucena-Cecilia had the greatest number of alleles (A , Table 1), with San Ramón and Lucena-Cecilia showing the highest allelic richness (Ar) at 1.567 and 1.493, respectively, followed by Gravatá, which also had a notable number of private alleles (Ap , Table 1). Allelic richness was generally higher in inland forest populations compared to *restinga* sites, with the exception of Gravatá, which also displayed high Ar values (1.473). The highest fixation index (f) showed variation among populations, with the highest value observed in Travessa (0.0145), while negative values in Gravatá (−0.0261), Herval (−0.0106), and Pixirica (−0.0230), indicated an excess of heterozygotes.

Population genetic structure in southern South America

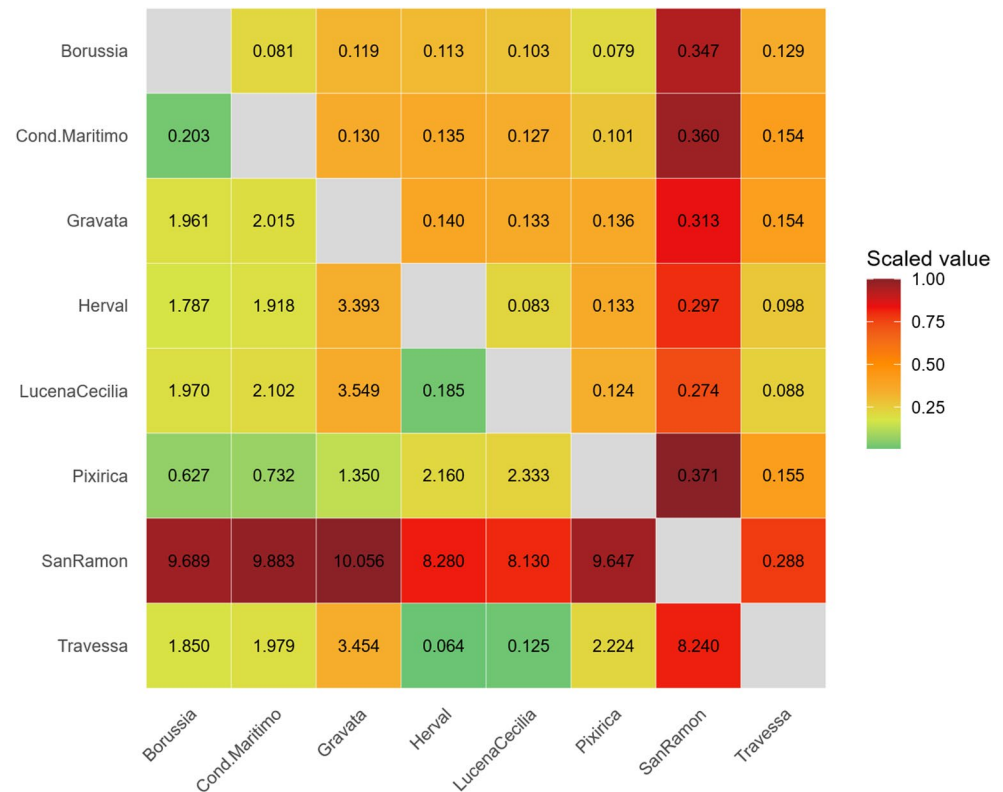
Based on 24,859 SNPs, Wright's F statistics revealed significant differentiation among populations. The estimates across the eight sampled *S. romanzoffiana* populations were: total inbreeding coefficient ($F_{IT} = 0.139$), within-population inbreeding coefficient ($F_{IS} = 0.021$), and genetic differentiation coefficient among populations ($F_{ST} = 0.122$). Pairwise F_{ST} analyses indicated that San Ramón was consistently the most divergent population, with values ranging from 0.274 (vs. Lucena-Cecilia) to 0.371 (vs. Pixirica), while the lowest differentiation was observed between Borússia and Pixirica ($F_{ST} = 0.079$), both located in Atlantic Rain Forest (Fig. 2). In contrast, Borússia and Cond. Marítimo, despite being geographically close (geographic Euclidean

Table 1 Population genomic diversity of *Syagrus romanzoffiana* based on 24,859 SNPs

Location	H_o	H_s	A	AR	Ap	f
Borússia	0.145	0.143	35,479	1.416	245	0.0015 (−0.0209; −0.0027)
Cond.Marítimo	0.140	0.138	36,125	1.388	653	0.0043 (−0.0238; −0.0067)
Gravatá	0.175	0.167	38,196	1.473	3291	−0.0261 (−0.0531; −0.0387)
Herval	0.163	0.159	36,631	1.453	364	−0.0106 (−0.03; −0.0135)
Lucena-Cecilia	0.169	0.169	37,555	1.493	329	0.0065 (−0.0098; 0.0062)
Pixirica	0.139	0.133	34,935	1.373	476	−0.0230 (−0.0541; −0.0354)
San Ramón	0.205	0.203	40,079	1.567	17,426	0.0021 (−0.0159; −0.0031)
Travessa	0.160	0.164	35,518	1.468	98	0.0145 (0.0128; 0.0313)

H_o = Observed heterozygosity, H_s = Expected heterozygosity, A = Number of alleles, Ar = Allelic richness, Ap = Private alleles, f = Fixation index (bootstraps). Locations in Brazil: Borússia – RS; CondMarítimo – RS, Herval – RS, Lucena-Cecilia – RS, Pixirica – RS, Travessa – RS and Gravatá – SC. Location in Paraguay: San Ramón – Concepción

Fig. 2 Heatmap (Galili et al. 2018) of pairwise genetic differentiation (F_{ST}) among 91 individuals of *Syagrus romanzoffiana* sampled across different locations, based on 24,859 SNPs. The lower triangle shows pairwise F_{ST} values, while the upper triangle shows geographic Euclidean distances calculated from population coordinates, allowing direct comparison between genetic and spatial distances



distance=0.203), showed markedly higher genetic differentiation ($F_{ST}=0.081$), reflecting the influence of habitat type (*restinga* vs. Atlantic Rain Forest) on genetic structure. The geographic Euclidean distances shown in the lower triangle of Fig. 2 further illustrate that San Ramón is substantially more distant from all Brazilian populations (ranging from 8.130 to 10.056), yet its elevated F_{ST} values cannot be attributed to geographic distance alone, as the non-significant Mantel test ($r=-0.009$, $p=0.141$) indicates no overall pattern of isolation by distance across the dataset.

These pairwise comparisons highlight both geographic and ecological influences on population structure. Consistently, AMOVA confirmed significant differentiation among populations, with 29.7% of the genetic variance distributed among populations and 70.3% within populations (Table 2). These findings indicate that most of the genetic diversity is maintained within populations, while moderate structure exists among them. This pattern may result from substantial gene flow within populations and ecological or geographic factors constraining gene flow between populations. However, the Mantel test showed a non-significant correlation between genetic and geographic distances ($r=-0.009$, $p=0.141$), indicating no evidence of isolation by distance.

Although San Ramón represents the most geographically distant population, the lack of isolation by distance indicates that its strong genetic differentiation is not explained solely

Table 2 Analyses of molecular variance (AMOVA) performed for eight populations of *Syagrus romanzoffiana* collected in the in Brazil and Paraguay

Source of Variation	df	MS	Est. Var.	PV (%)
Among populations (Among Locations)	7	6617.983	483.8844	29.7%
Within populations (Within Locations)	83	1145.607	1145.6073	70.3%
Total	90	1571.237	1629.4917	100%

p -value=0.001 (Estimated based on 20000 permutations). Df =Degrees of Freedom; MS =Mean Squares. *Est. Var.* = Variance component. *PV*=Percentage of Variation

by spatial separation, supporting the role of environmental factors in shaping divergence. Discriminant Analysis of Principal Components (DAPC) identified clear population structure. The first discriminant function (LD1) explained 19.22% of the variation and separated San Ramón from all other sites (Fig. 3A). The remaining populations showed partial overlap along LD1, reflecting more subtle differentiation. Membership probabilities (Fig. 3B) further supported these results, with San Ramón, representing a colder and more seasonal environment, formed a distinct cluster, while *restinga* populations (Cond. Maritimo and Gravata), characterized by sandy and coastal environments, tended to group together, and inland forest populations showed mixed assignment.

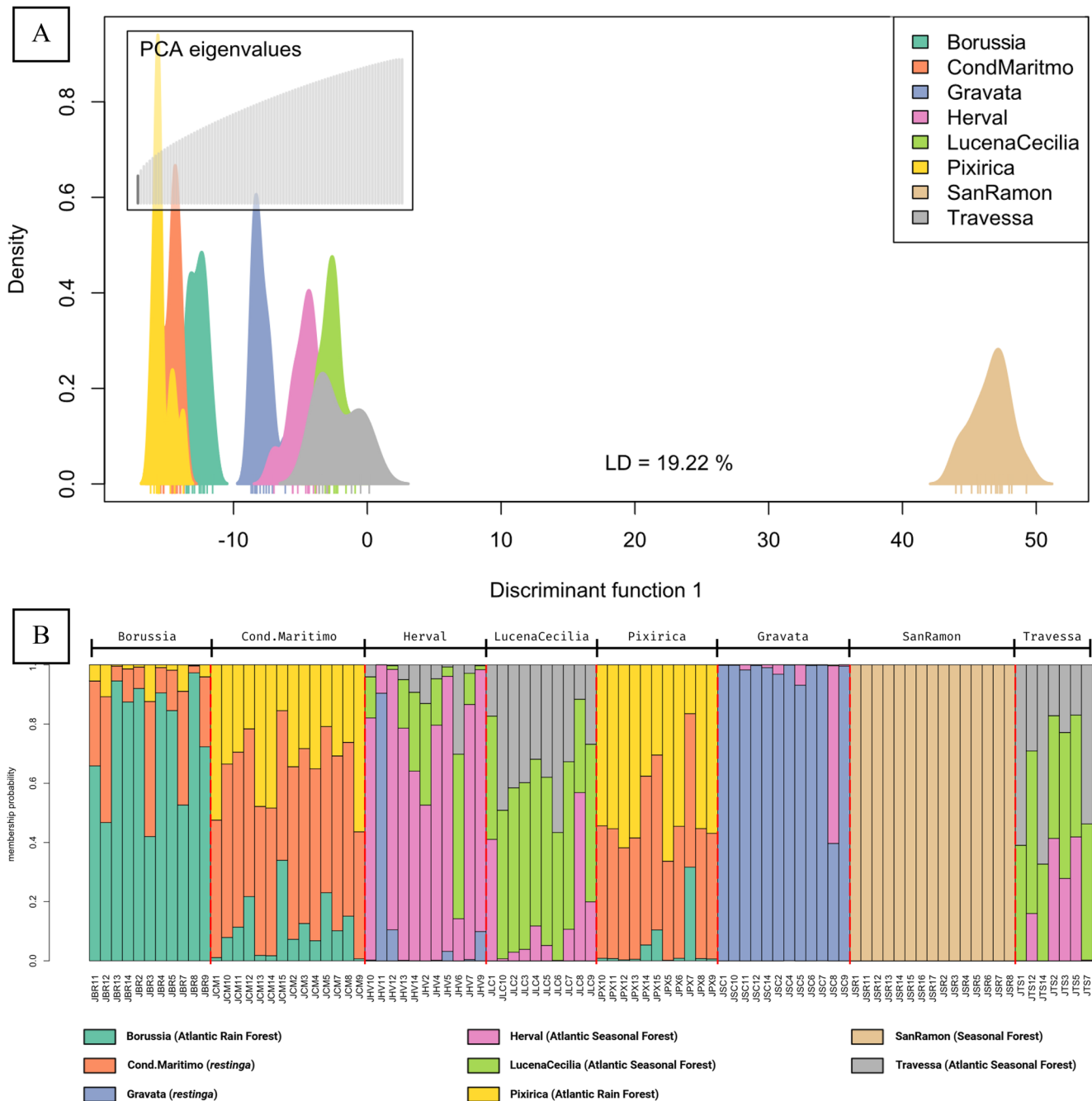


Fig. 3 (A) Scatter Plot of Discriminant Analysis of Principal Components for 91 individuals from 8 *Syagrus romanzoffiana* populations. (B) Membership barplot/compoplot with genetic groups

Detection of adaptive loci and environmental associations

To assess whether environmental variation covaries with spatial distance among populations, we tested the correlation between environmental and geographic distances using a Mantel test. Environmental distances were calculated using Euclidean distance based on the climatic variables identified

in the PCA as most relevant (bio06 – minimum temperature of the coldest month, and bio19 – precipitation of the coldest quarter). The Mantel test revealed a non-significant correlation between environmental and geographic distances ($r=0.28$, $p=0.13$), indicating that environmental differentiation among populations is not strongly structured by spatial distance. This result suggests that the environmental signals detected in subsequent genotype–environment association

analyses are unlikely to be driven solely by spatial covariance between environmental variables and geographic distance.

Environmental association analyses were conducted using LFMM on the imputed SNP dataset, incorporating population structure inferred from sNMF. This approach increased robustness to missing data while preserving the underlying neutral genetic structure. LFMM identified 4,621 SNPs significantly associated with bio06 (minimum temperature of the coldest month) and bio19 (precipitation of the coldest quarter) (Figures S4, 4).

A new sNMF analysis based exclusively on these candidate SNPs revealed a population structure with $K = 4$ clusters (Figure 5), indicating that environmental variables contribute to population differentiation beyond neutral processes. San Ramón and the two *restinga* (Gravatá and Cond. Marítimo) populations formed distinct genetic clusters, while inland forest populations showed higher levels of admixture, consistent with weaker environmental differentiation.

To further assess consistency among environmental associations, we quantified the overlap between candidate SNPs associated with bio06 and bio19 (Figure 6). A substantial proportion of loci were shared between both variables, indicating partially convergent genomic responses to temperature and precipitation gradients rather than independent or stochastic associations driven by population structure alone.



Fig. 6 Overlap between candidate SNPs identified by LFMM for bio06, minimum temperature of the coldest month, and bio19, precipitation of the coldest quarter, based on the imputed genotype dataset. Shared loci represent genomic regions jointly associated with temperature and precipitation gradients, indicating partially convergent adaptive responses to multiple climatic drivers in *Syagrus romanzoffiana*

Redundancy Analysis (RDA) based on candidate SNPs from the imputed LFMM confirmed strong genotype–environment associations (Fig. 7). The global model constrained by bio06 and bio19 was highly significant ($p < 0.001$), with an adjusted R^2 of 20%. Both variables contributed

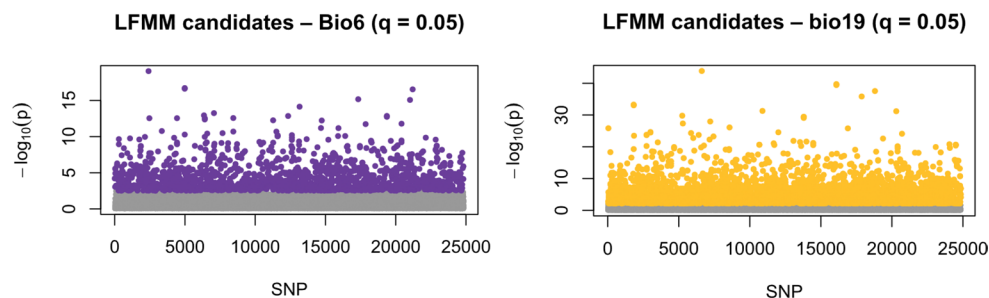


Fig. 4 Distribution of adjusted $-\log_{10}(p)$ values for the candidate SNPs under selection identified by the LFMM model, associated with the bioclimatic variables bio06 (minimum temperature of the coldest

month) and bio19 (precipitation of the coldest quarter). Purple and yellow dots represent loci with significant association ($q < 0.05$), indicating possible regions under adaptive selection

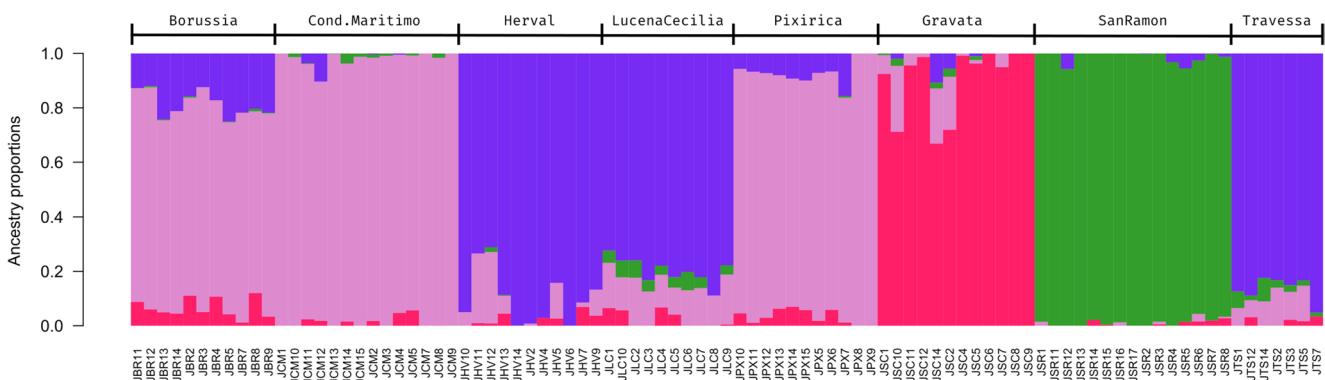
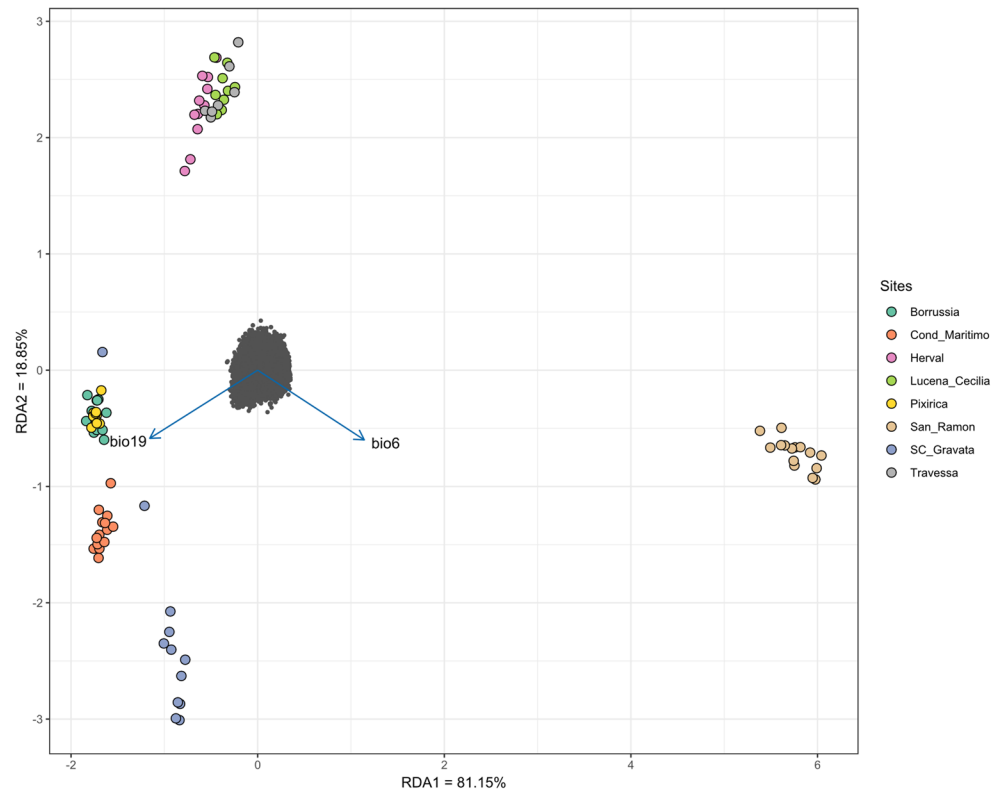


Fig. 5 Genetic structure inferred by sNMF using candidate adaptive SNPs identified by LFMM analysis ($q < 0.05$), with $K = 4$. Each bar represents an individual, and colors indicate ancestry proportions associated with adaptive genetic clusters

Fig. 7 Redundancy analysis (RDA) based on candidate SNPs identified by LFMM. Colored points represent individuals from each population of *Syagrus romanzoffiana*, while gray points represent candidate SNPs associated with environmental variables. Arrows indicate the bioclimatic variables bio06 (minimum temperature of the coldest month) and bio19 (precipitation of the coldest quarter), which explain genotypic variation among populations. The first two constrained axes explain nearly all of the adaptive variance captured by the candidate loci



significantly to the explained genetic variance, and the first two constrained axes explained nearly all of the adaptive variation captured by candidate loci (99.9%).

Individuals from San Ramón were positively associated with bio06, indicating adaptive differentiation related to colder minimum temperatures, while Cond. Marítimo, Pixirica, Borússia, and Gravata were more strongly associated with bio19, suggesting adaptive responses to precipitation seasonality across southern Atlantic Forest and *restinga* environments. Together, these results confirm that climatic variables related to temperature and precipitation are key drivers of adaptive genomic differentiation in *Syagrus romanzoffiana*.

Discussion

The genetic diversity estimates across *Syagrus romanzoffiana* populations varied among sites and were broadly consistent with values reported for other outcrossing neotropical palms assessed with SNP markers. Observed heterozygosity (H_O) exceeded expected heterozygosity (H_S) in five of eight populations, and H_O values ranged from 0.139 (Pixirica) to 0.205 (San Ramón), a pattern consistent with a predominantly outcrossing mating system and ongoing gene flow (Zona and Henderson 1989). For comparison, similar H_O ranges have been reported in SNP-based studies of congeneric species: *Euterpe oleracea* showed H_O values ranging from 0.185 to 0.260 across Amazonian populations

(Francisconi et al. 2025a) and *precatoria* also exhibited generally low fixation indices and heterozygote excess across sampled localities (Francisconi et al. 2025b), suggesting that outcrossing and gene flow are common mechanisms maintaining within-population diversity across neotropical palms. Populations with an excess of heterozygotes are generally expected to be more resilient to short-term environmental pressures, which may be relevant for persistence under ongoing habitat fragmentation and climate change (Hoffmann and Sgrò 2011).

In addition to heterozygosity, allelic richness (A_r) was particularly notable in San Ramón (1,567) and Lucena-Cecília (1,493), highlighting their genetic potential for adaptation and long-term evolutionary resilience (Pritchard et al. 2000). The high number of private alleles (A_p) found in Gravata (3,291) and San Ramón (17,426) emphasizes the genetic uniqueness of these populations, the former likely reflecting specific local adaptations to coastal *restinga* environments characterized by sandy soils, high salinity, and strong wind exposure (Zamith and Scarano 2006), and the latter reflecting its geographic distance from the other sampled sites. These patterns reinforce that habitat type and geographic context, rather than proximity alone, play central roles in shaping allelic diversity, as has been reported for other neotropical trees such as *Eugenia uniflora* (Vetö et al. 2023) and *Euterpe edulis* (Conte et al. 2008).

Fixation indices (f) were low and close to zero across all populations, ranging from -0.0261 (Gravata) to 0.0145

(Travessa), indicating that no population showed pronounced inbreeding or substantial excess of homozygosity. Negative f values in Gravatá, Pixirica, and Herval indicate a slight excess of heterozygotes, consistent with outcrossing and gene flow in populations occupying environmentally heterogeneous or well-connected habitats (Allendorf et al. 2010). The positive values observed in Travessa, Lucena-Cecília, Borússia, Cond. Marítimo, and San Ramón are close to zero and fall within confidence intervals overlapping zero (Table 1), suggesting no biologically meaningful departure from Hardy-Weinberg expectations. Together, these results suggest that, despite occurring in landscapes with varying degrees of fragmentation, the sampled populations of *S. romanzoffiana* have maintained largely panmictic mating dynamics, likely sustained by effective pollen and seed dispersal by insects, birds, and mammals (Zona and Henderson 1989).

Wright's F -statistics revealed moderate differentiation ($F_{ST} = 0.122$), indicating that while gene flow persists among populations, they remain sufficiently differentiated to maintain distinct genetic compositions (Wright 1931, 1949). AMOVA corroborated this finding, showing that most genetic variation occurs within populations (70.3%), consistent with expectations for long-lived, outcrossing tropical trees (Anderson 2001; Griffiths and Tavaré 1994), where large effective population sizes and extensive gene flow maintain diverse alleles within local gene pools.

Interestingly, the lack of a significant correlation between genetic and geographic distances in the Mantel test indicates that geographic separation is not the main driver of differentiation. This result indicates that geographic distance alone is insufficient to explain the observed genetic structure. Instead, ecological factors appear to play a greater role. Populations occupying contrasting vegetation types showed greater genetic differentiation than populations within the same habitat, even when geographically closer. For instance, Borússia and Pixirica, although ~90 km apart, were genetically more similar to each other than Borússia and Cond. Marítimo, which are separated by only ~20 km but differ in vegetation type (*restinga* vs. Atlantic Rain Forest). This contrast reinforces the importance of isolation by environment (IBE) over isolation by distance, highlighting the role of local ecological conditions and microclimates in shaping genetic divergence (Mantel 1967; Wang and Bradburd 2014).

The *restinga* environment occurs in sandy soils, prone to high salinity and much more windy sites. On the other hand, Pixirica and Borússia belong with a much humid forest, and a basaltic-derived soil. Moreover, Cond. Marítimo is located near the ocean, some 5 km from the beach. Geological studies of the area indicate that the sea level was approximately 2–5 m above its current level in a postglacial

marine transgression peak some 5 to 6 thousand years ago (Dillenburg et al. 2009), which may probably be incompatible with any occurrence of *S. romanzoffiana* in this area at the time. After that, a gradative regression to the current levels occurred (Barboza et al. 2021), probably accompanied by the expansion of the *restinga* vegetation and by the colonization of the area with *S. romanzoffiana*. Borússia and Pixirica occur in higher elevations and much older geological terrains, which probably allowed *S. romanzoffiana* to expand much earlier. This time difference in the setting of the populations along with the soil and climate conditions of each site is probably implicated in local adaptations at these sites.

Consistent with this interpretation, environmental association analyses provided compelling evidence for local adaptation in *S. romanzoffiana*. SNPs associated with climatic variables (bio06 and bio19) highlight the role of temperature and precipitation as selective pressures across vegetation types. Notably, adaptive genetic structure emphasized that selection could act independently of geographic distance or historical demography. Although historical processes undoubtedly contribute to background genetic structure, the consistency of genotype–environment associations across multiple analytical approaches supports a primary role of environmental selection. Similar patterns of environment-driven divergence have been documented in *Euterpe edulis* (Gaiotto et al. 2003), supporting the idea that neotropical palms frequently undergo local adaptation even across relatively small spatial scales. In this study, populations from San Ramón showed strong signals of adaptation to colder winters, while southern Brazilian populations, particularly those in *restinga* and coastal forest habitats, aligned with wetter and more variable rainfall regimes.

Collectively, these results indicate that environmental heterogeneity plays a central role in shaping genetic differentiation in *Syagrus romanzoffiana*. The absence of isolation by distance, combined with consistent genotype–environment associations, supports isolation by environment as a key process structuring populations across vegetation types. Although demographic history and genetic drift were not explicitly modeled, the concordance among LFMM, sNMF, and RDA analyses indicates that environmental selection contributes measurably to genomic differentiation beyond neutral structure. Future studies incorporating broader sampling and genomic resources will be important to further disentangle the relative contributions of drift and selection.

Conservation implications

These results provide relevant insights for the conservation and management of *Syagrus romanzoffiana*. Although this species is widespread and not currently threatened, genetic

variation is unevenly distributed across habitats, with distinct genetic signatures associated with *restingas*, inland forests, and seasonal environments. These findings highlight the importance of conserving ecologically distinct habitats that harbor locally adapted genetic variants. Given the frequent use of *S. romanzoffiana* in restoration, landscaping, and sustainable use programs, incorporating habitat-specific genetic information into seed sourcing and restoration planning may enhance long-term population resilience under ongoing habitat fragmentation and climate change.

Conclusion

Our results show that *Syagrus romanzoffiana* maintains high levels of genetic diversity within populations and moderate differentiation among populations across southern South America. Genetic structure was not explained by geographic distance, as indicated by the absence of isolation by distance, but instead reflected contrasts among vegetation types and climatic conditions. Populations from distinct habitats, particularly *restingas* and inland forests, were more genetically differentiated than geographically closer populations occupying similar environments.

Environmental association analyses identified loci associated with minimum temperature and precipitation, supporting the hypothesis that climatic variables contribute to adaptive genomic differentiation in this species. These adaptive signals were detected independently of spatial distance, indicating that environmental selection plays a measurable role alongside gene flow and demographic history.

Together, these findings provide a regional genomic baseline for *S. romanzoffiana* and demonstrate that habitat-specific environmental conditions shape both neutral and adaptive genetic variation. This information is directly relevant for restoration and management, as it indicates that seed sourcing strategies should prioritize matching source populations to local vegetation types and climatic regimes, particularly in *restinga* and fragmented Atlantic Forest landscapes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10592-026-01788-5>.

Acknowledgements We thank Prof. Guilherme Dubal dos Santos Seger and Ms. Francisco Agustin Vergara for his assistance in guiding and harvesting leaf samples from populations on a few sites of this study.

Author contributions K.K.M.M. First author, conceptualization, performed the experiment, analyzed the results, data curation, original draft writing, review & editing; M.S. Performed the experiment, analyzed the results; A.F.F. Analyzed the results, review & editing; I.A.S.de C. Developing of GBS plates & DNA quantification; C.B.G. Developing of GBS plates; T.D.L. Responsible for submitting the data in the NCBI's GenBank repository and preparing the data avail-

ability information; E.R.K. Collected the samples, responsible for the taxonomic identification of the plant material, data curation, review & editing; M.I.Z. Funding, conceptualization, data curation, review & editing.

Funding The Article Processing Charge (APC) for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614). We want to thank CAPES (Coordination for the Improvement of Higher Education Personnel) for the promotion granted to those involved. This study was supported by grants, including São Paulo Research Foundation (2011/50296-8), (2021/10319-0) and the National Council for Scientific and Technological Development (CNPq—313417/2023-7). The authors declare that they have no conflict of interest.

Data availability The plant material used in this study was taxonomically identified by Dr. Enéas Ricardo Konzen, an expert in the taxonomy of palms of the genus *Syagrus*. No voucher specimens were deposited in a public herbarium; however, the corresponding genetic data generated from these samples are publicly available in NCBI's GenBank repository (<https://www.ncbi.nlm.nih.gov/>) (<https://www.ncbi.nlm.nih.gov/>) under the accession number *Syagrus romanzoffiana** (PRJNA1295587; <http://www.ncbi.nlm.nih.gov/bioproject/1295587>) (<http://www.ncbi.nlm.nih.gov/bioproject/1295587>)).

Declarations

Ethics approval and consent to participate This study was approved by the Ethics Committee of the University of Campinas and conducted in accordance with Brazilian Ministry of the Environment (MMA) regulations and the National System for Genetic Heritage and Associated Traditional Knowledge (SisGen; registration A3A9281).

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahrens CW, Byrne M, Rymer PD (2019) Standing genomic variation within coding and regulatory regions contributes to the adaptive capacity to climate in a foundation tree species. *Mol Ecol* 28:2502–2516. <https://doi.org/10.1111/mec.15092>
- Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nat Rev Genet* 11:697–709. <https://doi.org/10.1038/nrg2844>
- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32–46. <https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x>

- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Balslev H, Kahn F, Millan B et al (2011) Species Diversity and Growth Forms in Tropical American Palm Communities. *Bot Rev* 77:381–425. <https://doi.org/10.1007/s12229-011-9084-x>
- Barboza EG, Dillenburg SR, do, Nascimento-Ritter M et al (2021) Holocene sea-level changes in Southern Brazil based on high-resolution Radar Stratigraphy. *Geosciences* 11:326. <https://doi.org/10.3390/geosciences11080326>
- Boutet G, Alves-Carvalho S, Falque M et al (2016) SNP discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea RIL population. *BMC Genomics* 17:121. <https://doi.org/10.1186/s12864-016-2447-2>
- Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol Evol* 18:249–256. [https://doi.org/10.1016/S0169-5347\(03\)00018-1](https://doi.org/10.1016/S0169-5347(03)00018-1)
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22:3124–3140. <https://doi.org/10.1111/mec.12354>
- Câmara-Leret R, Fortuna MA, Albuixech J et al (2017) Fundamental species traits explain provisioning services of tropical American palms. *Nat Plants* 3:16220. <https://doi.org/10.1038/nplants.2016.220>
- Colombo CA, Berton LHC, Diaz BG, Ferrari RA (2018) Macauba: a promising tropical palm for the production of vegetable oil. *OCL* 25:D108. <https://doi.org/10.1051/ocl/2017038>
- Conte R, Reis MS, Mantovani A, Vencovsky R (2008) Genetic Structure and Mating System of *Euterpe edulis* Mart. Populations: A Comparative Analysis Using Microsatellite and Allozyme Markers. *J Hered* 99:476–482. <https://doi.org/10.1093/jhered/esn055>
- Danecek P, Auton A, Abecasis G et al (2011) The variant call format and VCFtools. *Bioinformatics* 27:2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- De Kort H, Vandepitte K, Mergeay J et al (2014) Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Mol Ecol* 23:4709–4721. <https://doi.org/10.1111/mec.12813>
- Dillenburg SR, Barboza EG, Tomazelli LJ et al (2009) The Holocene Coastal Barriers of Rio Grande do Sul. *Geology and Geomorphology of Holocene Coastal Barriers of Brazil. Lecture Notes in Earth Sciences*, vol 107. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-44771-9_3
- Dinerstein E, Olson D, Joshi A (2017) An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* 67:534–545. <https://doi.org/10.1093/biosci/bix014>
- Doyle J (1991) DNA protocols for plants. In: Hewitt GM, Johnston AWB, Young JPW (eds) *Molecular Techniques in Taxonomy*. Springer, Berlin, pp 283–293. https://doi.org/10.1007/978-3-642-83962-7_18
- Eckert AJ, van Heerwaarden J, Wegrzyn JL et al (2009) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics* 182:1289–1302. <https://doi.org/10.1534/genetics.109.102350>
- Eisnerhardt WL, Svenning JC, Kissling WD, Balslev H (2011) Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Ann Bot* 108:1391–1416. <https://doi.org/10.1093/aob/mcr146>
- 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* 131:479–491. <https://doi.org/10.1093/genetics/131.2.479>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315. <https://doi.org/10.1002/joc.5086>
- Francisconi AF, Scaketti M, Morales-Marroquín JA, de Carvalho IAS et al (2025a) Genomic consequences of açai extraction in the Amazon: insights into selective pressures, genomic diversity, and population structure. *Front Plant Sci* 16:1688760. <https://doi.org/10.3389/fpls.2025.1688760>
- Francisconi AF, Scaketti M, Morales-Marroquín JA et al (2025b) High genomic structuring in *Euterpe precatoria* in the Brazilian Amazônia and new implications for restoration and conservation efforts. *Sci Rep* 16:1525. <https://doi.org/10.1038/s41598-025-31277-7>
- François O, Martins H, Caye K, Schoville SD (2016) Controlling false discoveries in genome scans for selection. *Mol Ecol* 25:454–469. <https://doi.org/10.1111/mec.13513>
- Frichot E (2013) A short manual for LFMM (command-line version). Available at: <https://www.academia.edu/76894852>
- Frichot E, François O (2015) LEA: an R package for landscape and ecological association studies. *Methods Ecol Evol* 6:925–929. <https://doi.org/10.1111/2041-210X.12382>
- Gaiotto FA, Grattapaglia D, Vencovsky R (2003) Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart). *J Hered* 94:399–406. <https://doi.org/10.1093/jhered/esg087>
- Galili T, O'Callaghan A, Sidi J, Sievert C (2018) heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 34:1600–1602. <https://doi.org/10.1093/bioinformatics/btx657>
- Gatti MG, Campanello PI, Montti LF, Goldstein G (2008) Frost resistance in the tropical palm *Euterpe edulis* and its pattern of distribution in the Atlantic Forest of Argentina. *Ecol Manag* 256:633–640. <https://doi.org/10.1016/j.foreco.2008.05.012>
- Goudet J (2005) hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* 5:184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Griffiths RC, Tavaré S (1994) Ancestral inference in population genetics. *Statist Sci* 9:307–319. <https://doi.org/10.1214/ss/1177010378>
- Helyar SJ, Hemmer-Hansen J, Bekkevold D et al (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Mol Ecol Resour* 11:123–136. <https://doi.org/10.1111/j.1755-0998.2010.02943.x>
- Hemstrom K, Xuereb A, Prunier J et al (2024) Next-generation data filtering in the genomics era. *Nat Rev Genet* 25:750–767. <https://doi.org/10.1038/s41576-024-00738-6>
- Henderson A, Galeano G, Bernal R (1995) Field guide to the palms of the Americas. Princeton University Press, Princeton. <https://doi.org/10.2307/j.ctvczzzd>
- Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature* 470:479–485. <https://doi.org/10.1038/nature09670>
- Hohenlohe PA, Bassham S, Etter PD et al (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* 6:e1000862. <https://doi.org/10.1371/journal.pgen.1000862>
- Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281. <https://doi.org/10.7717/peerj.281>
- Levis C, Costa FRC, Bongers F et al (2017) Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science* 355:925–931. <https://doi.org/10.1126/science.aal0157>
- Lorenzi H (2010) Flora Brasileira: Arecaceae (Palmeiras). Instituto Plantarum, Nova Odessa

- Manel S, Joost S, Epperson BK et al (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Mol Ecol* 19:3760–3772. <https://doi.org/10.1111/j.1365-294X.2010.04717.x>
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220
- Oksanen J, Simpson GL, Blanchet FG et al (2024) *vegan*: Community Ecology Package. R package version 2.6-8. <https://CRAN.R-project.org/package=vegan>
- Onstein RE, Baker WJ, Couvreur TLP et al (2017) Frugivory-related traits promote speciation of tropical palms. *Nat Ecol Evol* 1:1903–1911. <https://doi.org/10.1038/s41559-017-0348-7>
- Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7:e32253. <https://doi.org/10.1371/journal.pone.0032253>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- R Core Team (2024) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.r-project.org/>
- Scarano FR (2009) Plant communities at the periphery of the Atlantic rain forest: rare-species bias and its risks for conservation. *Biol Conserv* 142:1201–1208. <https://doi.org/10.1016/j.biocon.2009.02.027>
- Scariot AO (2015) Palmeiras brasileiras: botânica, ecologia, usos e conservação. In: Lopes R, Oliveira MSP, Cavallari MM, Barbieri RL, Conceição LDHCS (ed) *Palmeiras nativas do Brasil*, Embrapa, Brasília, pp 15–33
- Sen T, Samanta SK (2015) Medicinal plants, human health and biodiversity: a broad review. *Adv Biochem Eng Biotechnol* 147:59–110. https://doi.org/10.1007/10_2014_273
- Thioulouse J, Dray S, Dufour AB, Siberchicot A, Jombart T, Pavoine S (2018) Multivariate analysis of ecological data with ade4. Springer, New York
- Vetö NM, Oliveira GCX, Lemos-Filho JP et al (2023) Population structure and signals of local adaptation in *Eugenia uniflora* (Myrtaceae), a widely distributed species in the Atlantic Forest. *Bot J Linn Soc* 201:100–113. <https://doi.org/10.1093/botlinnean/boac012>
- Wang JJ, Bradburd GS (2014) Isolation by environment. *Mol Ecol* 23:5649–5662. <https://doi.org/10.1111/mec.12938>
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159. <https://doi.org/10.1093/genetics/16.2.97>
- Wright S (1949) The genetical structure of populations. *Ann Eugen* 15:323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Zamith LR, Scarano FR (2006) Restoration of a *restinga* sandy coastal plain in Brazil: survival and growth of planted woody species. *Restor Ecol* 14:87–94. <https://doi.org/10.1111/j.1526-100X.2006.00108.x>
- Zona S, Henderson A (1989) A review of animal-mediated seed dispersal of palms. *Selbyana* 11:6–21

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.