Genet Resour Crop Evol (2026) 73:159
https://doi.org/10.1007/s10722-026-02787-8

RESEARCH ARTICLE

Genetic diversity and population structure of *Myrciaria dubia* from the Peruvian Amazon: implications for germplasm conservation and crop improvement

Eduardo Mejia de Loayza · **Guillain Estivals** · **Diana Castro-Ruiz** · **Werner Chota-Macuyama** · **Carlos Angulo-Chavez** · **Mike Corazon-Guivin** · **Ángel Rodriguez del Castillo** · **Jhon Alvarado Reategui** · **Carlos Angulo-Villacorta** · **Kember Mejia** · **Dennis Del Castillo Torres** · **Carmen García-Dávila**

Received: 26 November 2025 / Accepted: 13 March 2026 / Published online: 27 March 2026

Abstract *Myrciaria dubia* (camu-camu) is a shrubby fruit tree native to the continental Amazon whose fruits have been intensively harvested from wild stands, potentially reducing effective population sizes. We quantified genetic diversity and population structure across seven wild Peruvian Amazon populations and delineated river-basin genetic units to guide provenance-aware germplasm conservation and breeding. We genotyped 254 individuals from the Napo, Ucayali, Nanay, Tahuayo, Putumayo, Tigre, and Curaray basins using six polymorphic microsatellite loci. Overall, 48 alleles were detected. Observed heterozygosity (0.149–0.483) was generally lower than expected heterozygosity (0.220–0.531), and population-level inbreeding coefficients ($F_{IS} = -0.038$–0.560) indicated significant heterozygote deficits in Napo, Curaray, and Tahuayo. The Putumayo

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10722-026-02787-8.

E. Mejia de Loayza (✉) · C. Angulo-Villacorta · D. Del Castillo Torres
Dirección de Investigación en Manejo Integral del Bosque y Servicios Ecosistémicos, Instituto de Investigaciones de La Amazonía Peruana, Iquitos, Peru
e-mail: edu_345@hotmail.com; edu.mejia@ufrj.br

C. Angulo-Villacorta
e-mail: cangulovillacorta@gmail.com

D. Del Castillo Torres
e-mail: ddelcastillo@iiap.gob.pe

E. Mejia de Loayza
Departamento de Vertebrados, Museu Nacional, Universidade Federal do Rio de Janeiro, Quinta da Boa Vista , do Rio de Janeiro, São Cristóvão 20940-040, Brazil

G. Estivals · D. Castro-Ruiz · W. Chota-Macuyama · C. Angulo-Chavez · J. Alvarado Reategui · C. García-Dávila
Dirección de Investigación en Ecosistemas Acuáticos Amazónicos, Instituto de Investigaciones de La Amazonía Peruana, Iquitos, Peru
e-mail: ichtyos3134@hotmail.fr

D. Castro-Ruiz
e-mail: dcastro@iiap.gob.pe

W. Chota-Macuyama
e-mail: wernerchotam@yahoo.com

C. Angulo-Chavez
e-mail: cangulo@iiap.gob.pe

J. Alvarado Reategui
e-mail: jalvarado@iiap.gob.pe

C. García-Dávila
e-mail: cgarcia@iiap.gob.pe

M. Corazon-Guivin
Laboratorio de Biología y Genética Molecular, Universidad Nacional de San Martín, Tarapoto, Peru
e-mail: macorazong@unsm.edu.pe

Á. Rodriguez del Castillo
San Ramón Agricultural Experimental Station, Instituto Nacional de Innovación Agraria, Yurimaguas, Peru
e-mail: arodriguez@inia.gob.pe

population harbored nine private alleles, representing a unique genetic reservoir. Pairwise differentiation was substantial ($F_{ST}=0.093–0.660$; Nei's distance = 0.068–1.734), with the strongest divergence between Tigre and Ucayali. Neighbor-joining, Bayesian assignment, and Discriminant Analysis of Principal Components (DAPC) initially supported three major genetic units and highlighted Putumayo as genetically isolated; additionally, hierarchical STRUCTURE analyses resolved eight clusters, and DAPC distinguished seven population-specific groups. Analysis of molecular variance attributed 56.5% of the variation within individuals and 34.7% among populations. Mantel tests supported isolation by distance based on straight-line geographic distances ($r=0.53$–0.56; $p \leq 0.017$), whereas river-network distances were not significant. Overall, the data indicate a geographically structured genetic architecture shaped by dispersal limitation and basin-scale differentiation, supporting three provenance units for germplasm banking and breeding: (i) Napo–Ucayali–Nanay–Tahuayo, (ii) Tigre–Curaray, and (iii) Putumayo.

Keywords Conservation genetics · Gene flow, Population structure · Microsatellite markers · Molecular markers

Introduction

The Peruvian Amazon covers approximately 61% of the national territory and is dominated by lush forests that serve as essential sources of food and income for numerous rural communities (Gonzáles 2007). In this region, at least 60 tree and shrub species are commercialized, with more than 60% harvested directly from wild populations, the most prominent being *Mauritia flexuosa* L.f. (aguaje), *Myrciaria dubia* McVaugh (camu-camu), *Euterpe precatoria* Mart. (huasaí), and *Oenocarpus bataua* Mart. (ungurahui) in Amazonian ecosystems (Vásquez and Gentry 1989; Marshall et al. 2003; Shackleton and Shackleton 2004; Gonzáles 2007; Shanley 2011; Alvarado et al. 2020). Conservation of genetic resources and crop improvement depend on maintaining standing genetic variation in wild populations and on complementary ex situ collections that preserve distinct provenances and enable germplasm exchange among breeding programs (Castro et al. 2024). Sustaining these forest-harvested resources requires management that balances use and conservation, because market demand can reduce wild population sizes and erode existing genetic variation (Penn 2004, 2006; Ticktin 2004; Ticktin and Shackleton 2011).

Myrciaria dubia is a shrubby fruit-bearing species of the family Myrtaceae. It grows naturally in seasonally flooded soils along rivers and lakes in Peru, Colombia, Brazil, and Venezuela (Pinedo et al. 2001; Rodrigues and Marx 2006; Borges et al. 2014; Martin et al. 2014). The largest natural populations have been reported in Peru, particularly along the Napo, Nanay, Ucayali, Marañón, and Tigre rivers (Pinedo et al. 2001; Rodrigues and Marx 2006; Hernández and Barrera 2014; Castro et al. 2024). In Peruvian floodplains, natural stands of *Myrciaria dubia* can be exceptionally abundant and productive, with yields reported at the hectare scale, underscoring the feasibility and importance of managing and conserving wild populations in situ (Peters and Hammond 1990). Camu camu fruits are rich in antioxidants and have antimicrobial and cell-regenerating properties, attributed to their exceptionally high ascorbic acid content (877 to 3079 mg/100 g of pulp) and phenolic compounds such as ellagitannins, ellagic acid, quercetin glycosides, syringic acid, and myricetin (Schmidt et al. 2010; Correa et al. 2011; Fujita et al. 2015; Arellano-Acuña et al. 2016; Castro et al. 2018; Azevedo et al. 2019). Due to its nutritional potential, *M. dubia* has garnered significant attention in national and international markets, resulting in high harvesting pressure on wild populations, particularly since the 1990s, when families and communities in the Ucayali, Napo, and Tigre river basins began exporting fruit (Penn 2004, 2006). In 1994 alone, Japan, France, and the United States were the primary importers, with exports reaching 500 tons of pulp annually from Iquitos and Pucallpa (Penn 2004, 2006; Akter et al. 2011). In response, the Peruvian government launched reforestation initiatives and, in 1996, the National Camu-Camu Program (PNCC) to promote the species' domestication. However, most plantations under the PNCC were unsuccessful because implementation

---

K. Mejia
Dirección de Investigación en Diversidad Biológica Terrestre Amazónica, Instituto de Investigaciones de La Amazonía Peruana, Iquitos, Peru
e-mail: kmejia@iiap.gob.pe

was rushed and the promised technical support for site preparation was rarely delivered, leading many growers to establish plantings on suboptimal floodplain sites prone to prolonged inundation and riverbank erosion (Penn 2006), and wild populations remain the primary source of fruit. Overharvesting and habitat degradation have been associated with local declines in wild population sizes, threatening the conservation of the species and its genetic diversity (Penn 2004, 2006; Martin et al. 2014).

Several scientific research on *Myrciaria dubia*, has primarily focused on cultivation practices, reproductive biology, and fruit composition (Alvarado 1969; Gutiérrez-Ruiz 1969; Yuyama et al. 2002; Castro et al. 2012, 2013, 2018; Azevedo et al. 2019). In contrast, the genetic diversity and population structure of wild populations have remained relatively understudied (García-Dávila et al. 2008; Šmíd et al. 2017). Early molecular evidence for genetic structuring in *M. dubia* also comes from ISSR data in the Brazilian Amazon; Nunes et al. (2017) analyzed 10 populations (Rio Branco basin, Roraima, Brazil) using 14 ISSR primers and recovered three major genetic groups, indicating substantial differentiation among populations (35% of variation among populations). Although based on dominant markers and a geographically localized sampling design, these results support the existence of detectable population structure in the species and highlight the value of expanding population-level sampling across the Amazon. For instance, Šmíd et al. (2017) analyzed both cultivated and wild populations in the Peruvian Amazon using microsatellite markers, revealing high genetic diversity but also signs of inbreeding, particularly among isolated wild populations. More recently, Castro et al. (2024) developed new microsatellite loci and evaluated a large germplasm collection from eight river basins, providing valuable insights into the genetic structure of ex-situ accessions and emphasizing the importance of maintaining genetically diverse germplasm collections. Together, these studies substantially advance the SSR-based characterization of *M. dubia*; however, they also indicate the need for basin-replicated analyses based exclusively on natural stands to define provenance units under contemporary landscape and hydrological connectivity.

Despite these important contributions, basin-replicated studies focused exclusively on natural populations across broader geographic scales remain scarce. In floodplain ecosystems, hydrological connectivity can facilitate dispersal and gene flow among riverine stands, but the strength of this effect may vary across basins and dispersal pathways; coupled with harvest pressure and landscape change, this can generate spatially heterogeneous genetic structure. Therefore, quantifying genetic variation within and among wild populations is essential to infer habitat connectivity, evaluate disturbance effects, and define provenance units that complement ex situ collections and inform in situ conservation and sustainable management (Hajjar & Hodgkin 2007). Understanding genetic variability in wild populations of *Myrciaria dubia* is essential for developing improved cultivars through breeding programs targeting traits such as fruit yield and vitamin C content. Here, we aim to assess genetic diversity and population structure in seven geographically distinct wild populations of *Myrciaria dubia* from the Peruvian Amazon using six microsatellite loci and delineate river-basin genetic units to guide provenance-explicit germplasm banking and crop-improvement programs.

Materials and methods

Biological sampling, DNA extraction and microsatellites amplification

Biological material (leaf tissue) was collected from 254 individuals belonging to seven natural populations of *Myrciaria dubia* in the department of Loreto, Peruvian Amazon (Fig. 1). Sampling was designed to capture as much of the species' natural genetic diversity as possible and to delimit river-basin genetic units where large natural populations of *M. dubia* have been recorded (Pinedo et al. 2001; Rodrigues and Marx 2006; Hernández and Barrera 2014). Field sampling was conducted by IIAP (a public research institute attached to Peru's Ministry of Environment, MINAM), following applicable institutional and national procedures for research sampling. Within each basin, individuals were selected at random within natural stands; to minimize non-independence due to spatial clustering, we avoided sampling multiple stems from the same shrub clump and prioritized well-separated individuals. Sampled plants were spaced approximately 50 m apart to reduce the likelihood of collecting close neighbors

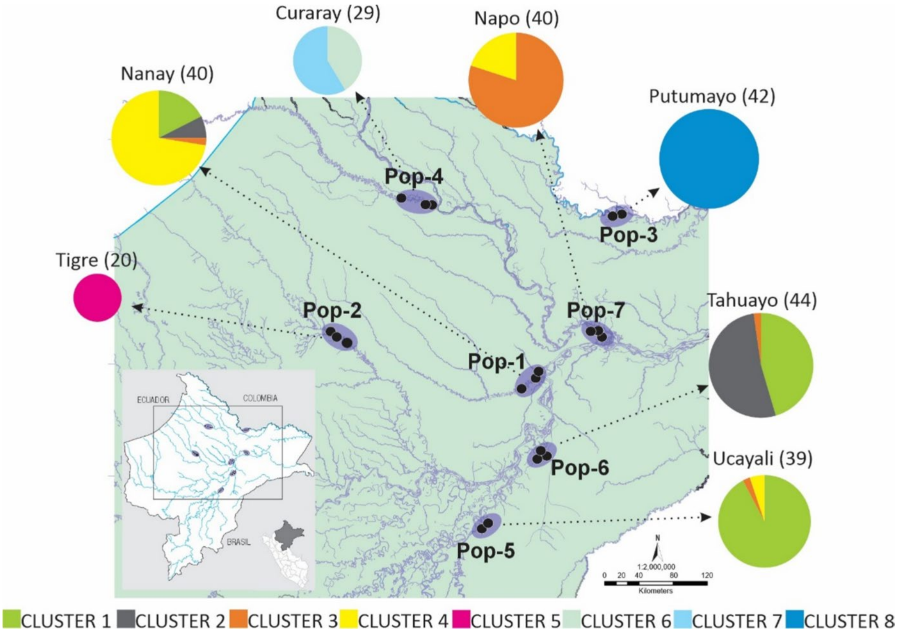

Fig. 1 Geographic distribution of the seven wild *Myrciaria dubia* populations in the Peruvian Amazon (black circles are multiple sampling points within each basin). Adjacent pie charts display the proportion of individual ancestry assigned to each of the eight genetic clusters inferred by STRUCTURE ($K=8$). Pie chart area is proportional to the number of genotyped individuals (n), and the value in parentheses next to each population name indicates the sample size (n)

and potential relatives, a precaution warranted by the partial self-pollination reported for the species (Maués and Couturier 2002). The study area includes populations from the Napo (73º19′15’’ / 3º50′39’’), Ucayali (73º43′18’’ / 4º56′23’’), Nanay (73º37′45’’ / 3º42′60’’), Tahuayo (73º19′15’’ / 3º50′39’’), Putumayo (72º37′45’’ / 2º26′22’’), Tigre (74º56′54’’ / 3º23′70’’), and Curaray (74º21′24’’ / 2º16′12’’) river basins.

Total genomic DNA was extracted from 0.1 g of dried leaf tissue using the CTAB method of Doyle & Doyle (1987). For genotyping, we amplified six EST-SSR loci of *Myrciaria dubia* (Rojas et al. 2008; Supplementary Table 1). The six markers were amplified in two multiplex reactions of three SSRs respectively to avoid spectral overlap and size-range conflicts: multiplex C2: MD07 (6-FAM), MD04 (HEX), MD02 (NED) and multiplex C3: MD05 (6-FAM), MD08 (HEX), MD06 (NED). Each multiplex PCR was performed in a final volume of 10 µL containing 1 µL of genomic DNA (10 ng/µL), 1.0 µL of Taq DNA polymerase (5 U/µL), 2 µL reaction buffer (5×), 1.2 µL of MgCl$_2$ (25 mM), 0.2 µL of dNTPs (10 mM each), and primers (0.15 µL of the forward primer [10 µM], 0.15 µL of the reverse primer [10 µM]) and the volume was completed to 10 µL with nuclease-free water (Invitrogen). Cycling conditions were initial denaturation at 95 °C for 1 min; 30 cycles of denaturation at 94 °C for 30 s, annealing at the multiplex-specific temperature (59 °C for multiplex C2; 52 °C for multiplex C3) for 1 min, and extension at 72 °C for 1 min; followed by a final extension at 72 °C for 40 min.

PCR products were denatured at 95 °C for 5 min and sized by capillary electrophoresis on an ABI 3130XL Genetic Analyzer (Applied Biosystems) using GeneScan 500 ROX as the internal size standard. Injection mixes contained 1.0 µL of PCR product, 8.7 µL of Hi-Di formamide (Applied Biosystems, Thermo Fisher Scientific), and 0.3 µL of GeneScan 500 ROX Size Standard (Applied Biosystems, Thermo Fisher Scientific). Allele sizes were registered using Peak Scanner v1.0 (Applied Biosystems). For allele binning, fragment lengths were first recorded as base-pair values relative to the ROX-500 ladder (raw fragment lengths in Supplementary Table S2). We then binned alleles to the nearest valid integer class consistent with each locus's core repeat motif (2-bp or 3-bp steps, as appropriate) and microsatellite locus with ≥ 2 failed amplifications were coded as missing for that locus (NA; -9 in STRUCTURE). The integer-binned genotype matrix used in all population-genetic analyses is provided in Supplementary Table S3. Additionally, a supplementary table reporting fragment lengths rounded to the nearest integer is provided (Supplementary Table S4).

Data analysis

Genetic diversity within and among populations

We used PowerMarker v3.25 (Liu and Muse 2005) to calculate locus-wise diversity indices across all 254 individuals: number of alleles (A), polymorphic information content (PIC), observed heterozygosity (Ho), and expected heterozygosity (He). For each locus, we additionally report the locus-wise heterozygote deficit as f = 1 – Ho/He. Population-wise diversity indices (N, At, P, Ho, He) were computed for each river-basin population. Population-level inbreeding coefficients (F_{IS}) were estimated in GENETIX (Belkhir et al. 2004) and reported as point estimates. We assessed statistical support for departures from Hardy–Weinberg expectations by testing each population's F_{IS} with a permutation procedure in GENETIX (1,000 replicates); p-values were calculated as the proportion of permuted F_{IS} values ≥ the observed F_{IS}, and populations with $p < 0.05$ were considered to depart significantly from Hardy–Weinberg expectations due to inbreeding and/or substructure. Pairwise genetic differentiation among populations was estimated as F_{ST} using the Weir & Cockerham (1984) estimator (θ) implemented in GENETIX.

Null alleles and Hardy–Weinberg equilibrium

To evaluate whether heterozygote deficits could be influenced by null alleles, we performed a locus-by-population screening based on heterozygote deficiency using a Brookfield-type estimator, $r = \left(H_e - H_o\right)/\left(H_e + H_o\right)$, truncated at zero. In addition, Hardy–Weinberg equilibrium (HWE) was tested for each locus within each population using an exact test with 10,000 Monte Carlo permutations; *p*-values were adjusted for multiple testing using the false discovery rate (FDR) procedure. These analyses were performed in the R package pegas (Paradis 2010).

Population structure

To visualize genetic relationships among river-basin populations, a neighbor-joining dendrogram based on Nei's genetic distances (Nei 1978) was generated in PowerMarker version 3.25 (Liu and Muse 2005), with 1000 bootstrap replicates. Allele frequency data were formatted and prepared using the PHYLIP package (Felsenstein 1989–2006). Population structure was analyzed using STRUCTURE version 2.3.3 (Pritchard et al. 2000) to infer the number of genetic clusters without a priori information about the geographic origin of individuals. The admixture model was applied, assuming that individuals may have mixed ancestry (Falush et al. 2003). The burn-in period was set to 10,000 iterations, followed by 100,000 Markov Chain Monte Carlo (MCMC) repetitions. Simulations were conducted for K values ranging from 1 to 7, with 10 replicates per K. The most likely number of genetic clusters was determined using the ΔK method of Evanno et al. (2005). For each STRUCTURE run (global and subset analyses), we evaluated K using ΔK (Evanno et al. 2005) and inspected consistency across replicate runs; we adopted a stepwise approach: after the initial run on all individuals, each major cluster identified was subjected to an independent STRUCTURE analysis (same parameters) to explore internal substructure. This recursive procedure was repeated until STRUCTURE no longer detected further subdivision within a given cluster. At each step individuals were assigned to the cluster with the highest membership coefficient (*q*-value). The analysis was concluded once STRUCTURE failed to detect further sub-structuring within the reanalyzed clusters. Subsets were not further subdivided when replicate runs did not support an additional stable partition (e.g., inconsistent solutions across replicates and/or ΔK not indicating clear structure).

A Discriminant Analysis of Principal Components (DAPC) was carried out using the adegenet package (Jombart 2008; Jombart et al. 2010) in R (R Core Team) to evaluate population structure without assuming Hardy–Weinberg equilibrium. The optimal number of genetic clusters was determined based on the lowest Bayesian Information Criterion (BIC). In addition, the number of genetic clusters was further corroborated using DAPC, an effective method for visualizing population structure and identifying the most influential factors contributing to variation among populations. The analysis retained 20 principal components, and the first two discriminant axes explained the total discriminant variance. Additionally, an Analysis of Molecular Variance (AMOVA; Excoffier et al. 1992) was performed using the poppr package with 999 permutations (Kamvar et al. 2014). Two analyses were conducted: one based solely on the geographic populations and another using groups defined by the NJ clustering and the DAPC results. This dual approach allowed us to assess genetic variation both at the population level and among genetically inferred clusters. Finally, to test isolation by distance, we assessed the association between genetic differentiation and geographic separation using Mantel tests. Pairwise genetic distances among populations were computed as Rousset's distance (Rousset 1997) $a = F_{ST}/\left(1 - F_{ST}\right)$ from the pairwise F_{ST} matrix (Supplementary Table S5). Geographic distances (Supplementary Table S6) were measured in two ways: (i) straight-line distances between population localities (km) and (ii) river-network distances following the main navigable fluvial route between localities (km). We performed Mantel tests (Pearson correlation) between the genetic-distance matrix and each geographic-distance matrix using 9999 permutations. Because isolation-by-distance relationships often become more linear after transformation of distance, we ran Mantel tests using both raw distances and log10-transformed distances (log10[km + 1]). We further evaluated whether genetic structure was better explained by straight-line distance or river-network distance using partial Mantel tests, correlating genetic distance with one geographic matrix while controlling for the other (9999 permutations). All Mantel and partial Mantel tests were conducted in R using the vegan package (Dixon 2003).

Results

Genetic diversity of alleles within and among populations

A total of 254 individuals of *Myrciaria dubia* from seven wild populations were genotyped using six polymorphic microsatellite markers (Fig. 1; See Microsatellite Genotypes in Table S2 and S3.). Forty-eight alleles were detected (mean = eight alleles per locus; Table 1). The number of alleles per locus ranged from 4 (MD07, MD04, MD02) to 18 (MD05). Observed heterozygosity (Ho) values ranged from 0.241 to 0.464 and were consistently lower than the expected heterozygosity (He) at all loci (Table 1). Locus-wise heterozygote deficit ranged from 0.042 (MD08) to 0.250 (MD02) (Table 1).

Table 1 Genetic diversity indices for six microsatellite loci genotyped in 254 *Myrciaria dubia* individuals

Locus	A	He	Ho	f	PIC
MD07	4	0.392	0.341	0.130	0.566
MD04	4	0.293	0.266	0.094	0.430
MD02	4	0.322	0.241	0.250	0.373
MD05	18	0.575	0.464	0.193	0.837
MD08	9	0.370	0.355	0.042	0.620
MD06	9	0.516	0.448	0.133	0.683
Mean	8	0.411	0.352	0.130	0.585

A: number of alleles per locus; He: expected heterozygosity; Ho: observed heterozygosity; f: locus-wise heterozygote deficit; PIC: polymorphic information content

At a population level, expected heterozygosity (Table 2) was highest in Napo population (He = 0.531) and lowest in Tigre (He = 0.220), while observed heterozygosity (Ho) ranged from 0.149 (Curaray) to 0.483 (Nanay; Table 2). Except for Ucayali, all populations exhibited lower observed heterozygosity than expected heterozygosity (Table 2); heterozygote deficits were more pronounced in Napo and Tahuayo and were strongest in Curaray (He = 0.340 vs Ho = 0.149), consistent with the population-level F_{IS} estimates (Table 2) and permutation tests (Supplementary Table S7). The number of alleles varied significantly among geographic populations, being highest in Putumayo (31 alleles) and Nanay (27 alleles) and lowest in Tigre (13 alleles), Curaray (16 alleles), and Ucayali (16 alleles). The Putumayo population displayed nine private alleles representing 29% of the alleles found in that population (Table 2).

Null-allele screening suggested generally low-to-moderate values across loci (Mean by locus ranged from 0.032 to 0.177), with the highest mean at MD05 (0.177) and the lowest at MD08 (0.032). Across populations, Curaray showed the highest average null-allele signal (mean $r = 0.420$), whereas other populations were lower (0.028–0.108). Exact HWE tests (FDR-corrected) identified significant departures in a subset of locus-by-population combinations, concentrated mainly in Curaray (3 loci), and to a lesser extent in Tahuayo (1 locus) and Tigre (1 locus).

Genetic differentiation and population structure

Pairwise F_{ST} and Nei's distances (Table 3) reveal that the Tahuayo and Ucayali populations are the most

Table 2 Genetic diversity indices for seven wild *Myrciaria dubia* populations in the Peruvian Amazon based on six microsatellite loci

Populations	N	At	P	He	Ho	F_{IS}
Napo	40	22	-	0.531	0.442	0.168*
Ucayali	39	16	-	0.346	0.359	-0.038
Nanay	40	27	-	0.495	0.483	0.023
Putumayo	42	31	9	0.504	0.476	0.054
Tigre	20	13	-	0.220	0.200	0.088
Curaray	29	16	-	0.340	0.149	0.560*
Tahuayo	44	18	-	0.444	0.356	0.199*

Significance of F_{IS} was assessed by permutation in GENETIX (1000 replicates; See supplementary Table S7)

N: number of individuals genotyped; At: total number of alleles; P: number of private alleles; He: expected heterozygosity; Ho: observed heterozygosity; F_{IS}: within-population inbreeding coefficient

$^{*}p < 0.001$

Table 3 Pairwise genetic differentiation index F_{ST} (upper triangle) and Nei's genetic distance (lower triangle) among seven wild *Myrciaria dubia* populations

Population	Curaray	Nanay	Napo	Putumayo	Tahuayo	Tigre	Ucayali
Curaray	–	0.451*	0.437*	0.492*	0.498*	0.464*	0.585*
Nanay	0.955	–	0.117*	0.234*	0.131*	0.518*	0.139*
Napo	0.955	0.151	–	0.256*	0.194*	0.495*	0.212*
Putumayo	0.955	0.461	0.461	–	0.170*	0.537*	0.289*
Tahuayo	1.071	0.143	0.256	0.204	–	0.570*	0.093*
Tigre	0.955	1.094	1.094	1.094	1.360	–	0.661*
Ucayali	0.955	0.151	0.225	0.461	0.068	1.094	–

*denote F_{ST} values significantly different from zero ($p < 0.001$)

similar ($D = 0.068$; $F_{ST} = 0.093$), consistent with low differentiation that may reflect recent shared ancestry and/or ongoing connectivity. In contrast, the Tigre and Ucayali populations are the most divergent ($D = 1.734$; $F_{ST} = 0.660$). Populations from the Napo, Ucayali, Nanay, and Tahuayo basins (Amazon-basin populations) exhibit lower genetic differentiation among themselves than with other populations (Table 3), whereas those from Tigre and Curaray, located in upper minor basins, were the most differentiated. All pairwise differences were statistically significant (Table 3). Permutation tests indicated significant heterozygote deficits ($p < 0.001$) in Napo, Curaray, and Tahuayo, whereas Tigre, Putumayo, Nanay, and Ucayali did not depart significantly from Hardy–Weinberg expectations (Supplementary Table S7).

Neighbour-Joining (NJ) dendogram (Fig. 2) grouped the seven populations into three principal genetic units: i) Napo, Ucayali, Tahuayo, and Nanay, (ii) Tigre and Curaray, and iii) Putumayo (Fig. 2). Although most bootstrap support values were low, the group formed by the Tigre population was moderately supported (bootstrap = 77%), indicating a significant level of divergence from the other populations. In contrast, the remaining populations grouped into two additional groups, consistent with weaker differentiation among them and/or more recent shared ancestry (Fig. 2).

Bayesian clustering analysis using STRUCTURE revealed that the 254 individuals could be partitioned into three primary clusters ($K = 3$) (Fig. 3I). In this scheme, Cluster A included the Napo, Ucayali, Nanay, and Tahuayo populations; Cluster B corresponded solely to the Putumayo population; and Cluster C comprised the Tigre and Curaray populations. These clusters corroborate the three genetic units suggested by the NJ tree. Further sub-analyses (Fig. 3II-IV) within Clusters A and C revealed additional sub-structuring. Within Cluster A, the individuals were divided into four distinct subclusters, and a similar subdivision was observed within Cluster C

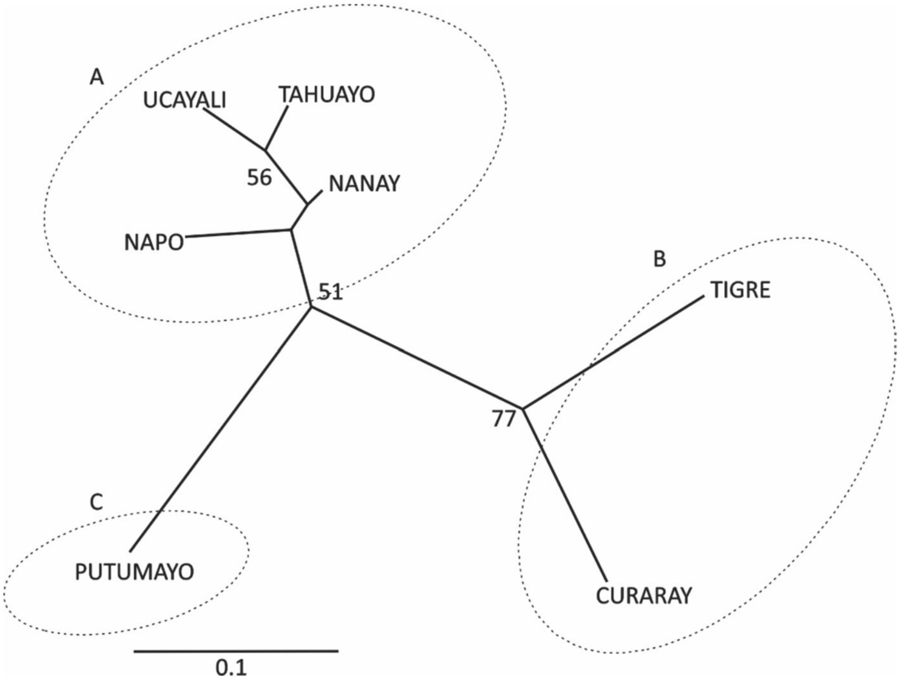

Fig. 2 Neighbour-joining dendrogram of seven wild *Myrciaria dubia* populations based on Nei's genetic distances calculated from six microsatellite loci. Bootstrap support values (1000 pseudoreplicates) are shown at each node. The dendogram resolves three main genetic units: **A** Napo, Ucayali, Tahuayo, and Nanay (top-left group), **B** Tigre + Curaray (right-hand group). **C** Putumayo (bottom-left group)

(Fig. 3II–IV). These sub-analyses were motivated by significant departures from Hardy–Weinberg expectations (permutation tests on F_{IS}; $p < 0.05$), which are consistent with heterozygote deficits due to non-random mating and/or cryptic substructure. The pie charts in Fig. 1 summarize the proportional membership of each population in the eight final STRUCTURE clusters (C1–C8), highlighting admixture in Napo–Ucayali–Nanay–Tahuayo and the near-homogeneous composition of Putumayo, Tigre, and Curaray. Final clusters comprise: four clusters containing some mixed individuals from Napo, Ucayali, Nanay, and Tahuayo (Clusters 1–4), while individuals from Tigre (Cluster 5), Curaray (Clusters 6 and 7), and Putumayo (Cluster 8) formed homogeneous groups (see clusters; Fig. 1).

Discriminant Analysis of Principal Components (DAPC) retained six discriminant functions, with the first two axes (FD1 and FD2) explaining 42.93% and 30.01% of the total discriminant variance, respectively (72.94% cumulatively; Fig. 4). Global assignment probabilities were high (98.8%), although variation existed among populations (97.3% for Group 1, 90.9% for Group 2, and 100% for Groups 3–7). The scatterplot of FD1 versus FD2 reveals seven distinct groups corresponding to the seven populations analyzed. Notably, group 3, composed of specimens from the Putumayo River, is separated and genetically isolated from the others (Fig. 4). In addition, populations 2 and 4 form a distinct group, although some individuals appear at intermediate positions between them, suggesting moderate genetic differentiation but also a degree of connectivity. In contrast, groups 1, 5, 6, and 7 (Amazonas basin) show substantial overlap, with groups 1, 5, and 6 being almost entirely superimposed; this pattern suggests gene flow or a recent shared ancestry among these populations (Fig. 4).

The first discriminant function separates group 3 along the negative axis, while the second function distinguishes group 2 from group 4 (Supplementary Table S8). The remaining groups occupy a central overlapping area (See supplementary Table S8 for mean discriminant scores and standard deviations in each group). It is worth noting that the relationship between populations and genetic groups was the same as those identified by the dendrogram of genetic distance (Fig. 2).

The AMOVA revealed contrasting patterns in the distribution of genetic variation in *Myrciaria dubia* (Table 4). At the population level (Table 4a), 56.45% of the variation occurred within individuals, 34.73% among populations, and 8.82% among individuals within populations, indicating significant genetic differentiation among geographic populations despite

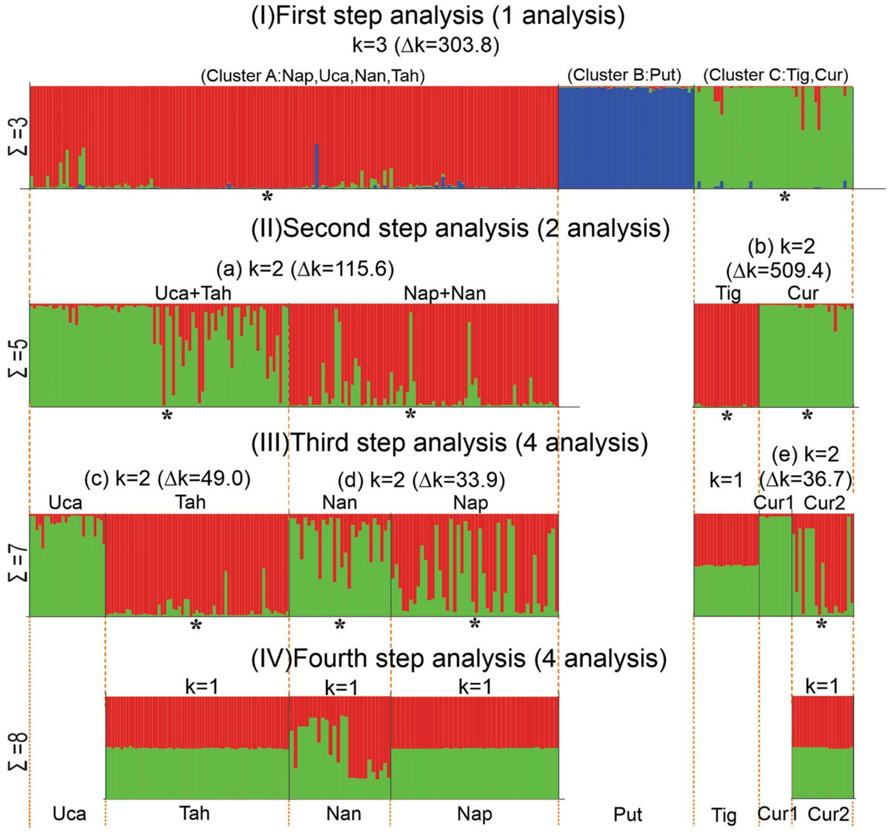

Fig. 3 Illustration of the four successive clustering-step analyses (I–IV) in STRUCTURE for *254 Myrciaria dubia* individuals. Histograms and the optimal K for each step were selected based on the highest ΔK following Evanno et al. (2005). (I) Global analysis identifying three primary genetic clusters ($K=3$). (II) Sub-analyses of the two main clusters from step I (a: $K=2$; $\Delta K=115.6$; and b: $K=2$; $\Delta K=509.4$) revealing additional structure. (III–IV) Additional sub-analyses within the resulting groups (c–e; each $K=2$; ΔK values shown above plots) to resolve finer-scale structure. The final assignment identifies eight genetic clusters ($K=8$), shown at the bottom, with population abbreviations beside the corresponding barplot blocks: Nap = Napo, Uca = Ucayali, Nan = Nanay, Tah = Tahuayo, Put = Putumayo, Tig = Tigre, Cur = Curaray. Asterisks below clusters indicate significant deviation from panmixia ($p<0.05$). Σ indicates the cumulative number of clusters identified across steps. Colors in intermediate-step panels are assigned independently and are not directly comparable across steps

the predominance of within-individual variation. When populations were grouped according to NJ-inferred genetic units (Table 4b), the variance components were slightly redistributed: 33.95% of the total variation was attributed to differences among NJ groups, 16.16% to variation among samples within groups, and 49.89% remained within individuals. These findings validate the groups and reinforce the congruence between the NJ and DAPC results.

Mantel tests supported isolation by distance when geographic separation was measured as straight-line distance (raw distances: $r=0.5285$, $p=0.0165$; log10-transformed distances: $r=0.5636$, $p=0.0087$; Fig. 5). In contrast, river-network distances were not significantly correlated with genetic distances (raw: $r=0.0145$, $p=0.3677$; log10: $r=0.1896$, $p=0.2528$). Partial Mantel tests indicated that straight-line distance remained significant after controlling for river-network distance ($r=0.5603$, $p=0.0081$), whereas river-network distance did not explain genetic structure after controlling for straight-line distance ($r=-0.1751$, $p=0.6308$).

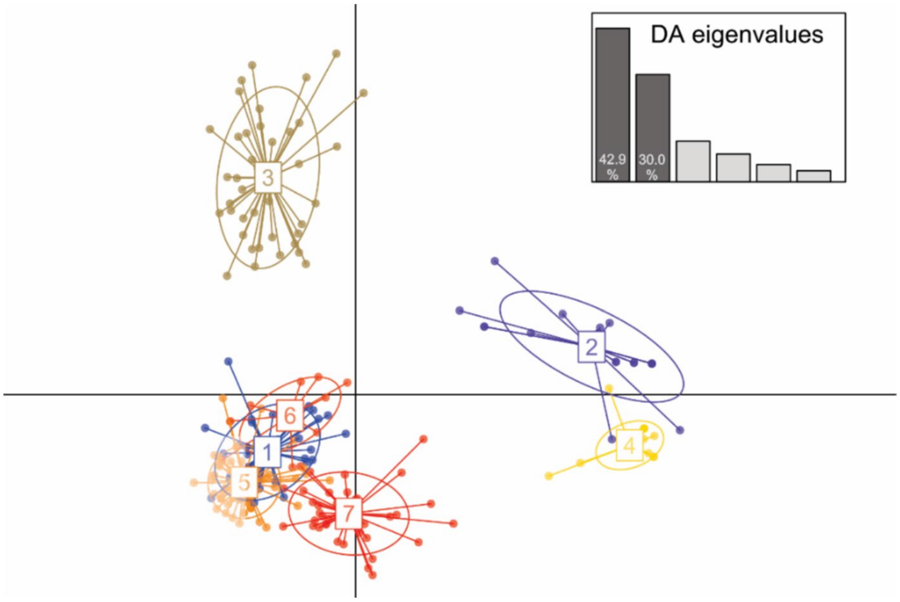

Fig. 4 Discriminant Analysis of Principal Components (DAPC) scatterplot for 254 *Myrciaria dubia* individuals. Individuals are plotted according to the first two discriminant functions (FD1 and FD2), which explain 72.94% of the total discriminant variance. Colors and labels (1–7) denote the seven a priori geographic populations (1: Nanay; 2: Tigre; 3: Putumayo; 4: Curaray; 5: Ucayali; 6: Tahuayo; 7: Napo); with 95% inertia ellipses drawn around each group. The inset barplot shows the eigenvalues for all discriminant axes. See Supplementary Table S8 for mean FD1 and FD2 scores (± SD) by population

Table 4 Analysis of molecular variance (AMOVA) for wild *Myrciaria dubia* populations, partitioned at two hierarchical levels

Source	df	SS	MS	Est. Var	%
(a) Geographic population level					
Among populations	6	614.063	102.344	1.381	34.73
Between samples within populations	247	727.571	2.946	0.351	8.82
Within samples	254	570.000	2.244	2.244	56.45
Total	507	1911.634	–	3.976	100
(b) NJ cluster level					
Among NJ groups	2	413.555	206.778	1.527	33.95
Between samples within groups	251	928.079	3.698	0.727	16.16
Within samples	254	570.000	2.244	2.244	49.89
Total	507	1911.634	–	4.498	100

In the poppr output, “samples” correspond to diploid individuals; “within samples” represents within-individual variance (heterozygosity) (a) among geographic populations and (b) among NJ-inferred genetic clusters

df, degrees of freedom; SS, sum of squares; MS, mean square; Est. Var., estimated variance component; %, percentage of total variance

Discussion

Genetic diversity within and among populations

Across seven wild *Myrciaria dubia* populations, microsatellite data indicate moderate-to-high genetic variability, in line with previous studies (Rojas et al. 2011; Šmíd et al. 2017; Castro et al. 2024). A total of 48 alleles were detected across six SSR loci (mean = 8.0 alleles per locus), providing a broad allelic state space for population-level inference. Because marker resolution depends on both allele number and allele-frequency distributions, summary metrics such as expected heterozygosity and the polymorphism information content (PIC) are widely used to evaluate the discriminatory ability of codominant markers and to prioritize the most informative loci (Nei 1978; Botstein et al. 1980; Serrote et al. 2020). In our dataset, PIC values ranged from 0.373 (MD02) to 0.837 (MD05), with an overall mean of 0.585 (Table 1), indicating moderate-to-high marker informativeness and supporting the suitability of this SSR panel for detecting allele-frequency heterogeneity among river basins (Botstein et al. 1980; Serrote et al. 2020). The precision of microsatellite-based inference depends not only on the number of loci but also on the total number of independent alleles surveyed; simulation analyses show that comparable accuracy in genetic distance estimation can be achieved with fewer loci when they collectively capture sufficient allelic diversity (Foulley and Hill 1999; Kalinowski 2002). Moreover, power simulations indicate that multi-allelic microsatellite datasets can provide substantial power to detect population divergence even at low levels of differentiation under commonly used sampling designs, with resolution influenced by how alleles are distributed across loci (Ryman et al. 2006). Empirical resampling further shows that increasing microsatellite numbers improves the stability of genetic distances and the robustness of distance-based phylograms, even when locus increases are moderate (Koskinen et al. 2004). Conversely, locus-reduction approaches indicate that a relatively small “core set” of highly polymorphic loci can still recover the major population structure and support high assignment accuracy, particularly when marker informativeness is high (Arthofer et al. 2018). Thus, although the total number of alleles detected here was lower than the 91 alleles reported by Šmíd et al. (2017) and the 313 alleles identified by Castro et al. (2024), these differences are expected given variation in marker number, sampling intensity, and geographic scope among studies, and the moderate-to-high PIC values observed here indicate that our SSR panel retains substantial information content for detecting diversity and population structure in wild stands.

Observed heterozygosity (Ho) values were lower than the expected heterozygosity (He) at all loci (Table 1), indicating a heterozygote deficit. This pattern is consistent with that reported by Rojas et al. (2011), who found high He values (0.691–0.903) in eight loci for camu-camu accessions, and with the results of Šmíd et al. (2017), which also showed lower Ho (0.137–0.527) compared to higher He (0.218–0.680). At population level, Šmíd et al. (2017) reported a consistent deficit of heterozygotes in both wild and cultivated camu-camu (Ho = 0.347–0.404 vs. He = 0.506–0.516). Likewise, Castro et al. (2024) reported observed heterozygosity between 0.468 and 0.644 and a mean fixation index of $F_{IS} = 0.276$ across river basins, reflecting pronounced heterozygote deficiency. Our results mirror this pattern, showing a systematic heterozygote deficit across loci, and population-level permutation tests (Supplementary Table S7) revealed significant heterozygote deficits (local inbreeding or substructure) in three of the seven populations: Napo ($F_{IS} = 0.168$), Curaray (0.560) and Tahuayo (0.199; $p < 0.001$), while the other four populations did not depart significantly from Hardy–Weinberg expectations. Notably, Ucayali showed Ho slightly exceeding He (negative F_{IS}), which indicates no heterozygote deficit under our marker set; however, this pattern alone does not demonstrate panmixia or contemporary gene flow and should be interpreted conservatively as consistency with Hardy–Weinberg expectations given sampling variance and the loci analyzed. These elevated inbreeding coefficients likely result from one or more factors: (i) a contribution of null alleles at a subset of loci, as suggested by the null-allele screening with highest signals in Curaray and at locus MD05 (Chapuis and Estoup 2007; Liu et al. 2013), and consistent with the fact that FDR-corrected HWE departures were concentrated mainly in Curaray, while most locus-by-population tests were not significant in other basins; (ii) cryptic reproductive substructure (Wahlund effect), which is associated with departures from Hardy–Weinberg expectations and heterozygote deficits (Nei 1978; Liu et al. 2013), (iii) partial self-pollination or nonrandom mating, and (iv) sampling of closely related individuals particularly when accessions are collected within restricted stands (Rossetto et al. 1999; Ferreira-Ramos et al. 2014; Lu et al. 2018).

Heterozygote deficit has been attributed to factors such as partial selfing, reproductive substructure, or residual sampling non-independence such as inadvertent sampling of relatives due to spatial clustering within stands (Brondani et al. 1998; Rossetto et al. 1999; Payn et al. 2008). Maués and Couturier (2002) analyzed the reproductive system of camu-camu and observed a temporal overlap between stigma receptivity and pollen viability, which can allow self-pollination, reporting early-morning anthesis and intense visitation by stingless bees (Meliponini), which they considered the main legitimate pollinators. Such a pollination context, together with floral-phase overlap, can increase geitonogamy and within-stand mating (Peters and Vásquez 1987). In large and continuous stands such as Putumayo, greater mate availability and pollinator-mediated pollen movement are expected to increase realized outcrossing and reduce biparental inbreeding, which helps maintain high allelic diversity and private alleles over time. In contrast, in smaller or more isolated populations such as Tigre or Curaray, inbreeding due to crosses among closely related individuals may explain the low levels of genetic diversity, a situation similar to that reported in studies of *Eugenia dysenterica* and *Eugenia uniflora*, which exhibit high inter-population variation but low overall genetic diversity, possibly due to reduced population sizes and limited pollen and seed dispersal (Salgueiro et al. 2004; Trindade and Chaves 2005; Ferreira-Ramos et al. 2014; Lu et al. 2018).

Population genetic structure

The high genetic polymorphism in populations such as Putumayo and Nanay (Table 2) are consistent with demographic and landscape processes that promote the long-term retention of genetic variation, including relatively large effective population size, habitat continuity, and basin connectivity, which counteract allele loss by genetic drift (Nei et al. 1975; Frankham 1996). In Putumayo, the occurrence of private alleles may reflect limited recent immigration and localized allele-frequency differentiation in comparatively less disturbed stands, particularly where individuals were sampled across extensive floodplain areas distant from human settlements (Rojas et al. 2011; Castro et al. 2024; Slatkin 1985). In Nanay, its central position within the Amazon system plausibly facilitates exchange with neighboring basins through hydrological connectivity, which can help maintain allelic diversity via gene flow (Stevens et al. 2007; Rojas et al. 2011). On the other hand, the Tigre, Curaray, and Ucayali populations showed comparatively low levels of polymorphism, which may reflect reductions in effective population size associated with intense harvesting pressure on natural stands by local communities in Jenaro Herrera, Mazán, and Tigre since 1994 (Akter et al. 2011). Importantly, repeated fruit extraction can reduce seed availability and recruitment over time, lowering population growth and effective population size (Ticktin 2004; Martin et al. 2014), and these demographic effects may be amplified in floodplain systems where recruitment is constrained by the flood pulse (Junk et al. 1989). In the Peruvian Amazon, harvesting pressure and habitat degradation have been associated with reduced allelic number and heterozygosity in camu-camu accessions and related contexts (Ferreira-Ramos et al. 2014; Šmíd et al. 2017; Castro et al. 2024), and similar processes could contribute to reduced diversity in Tigre and Curaray. Notably, Ucayali was the only population where Ho exceeded He (Table 2), suggesting either local demographic effects, recent admixture, or stochastic variation in allele frequencies given the modest number of loci.

Across our analyses, NJ, STRUCTURE, and DAPC consistently recovered three primary genetic units (Figs. 2–4), whereas the hierarchical STRUCTURE workflow resolved finer sub-structuring into eight clusters (Fig. 3). Differences in the number of inferred clusters among studies (Šmíd et al. 2017; Castro et al. 2024) are expected and likely reflect methodological and biological contrasts, including marker-panel resolution (number of SSR loci), sampling design (wild stands vs. cultivated/ex situ accessions), and the degree of provenance mixing. Nevertheless, a consistent qualitative pattern emerges. Šmíd et al. (2017) explored $K = 2–5$ and repeatedly recovered Putumayo populations as a distinct genetic group, with additional river-system subdivision at higher K. In Castro et al. (2024), the use of a larger SSR panel (16 loci) and ex situ germplasm accessions produced a more complex clustering solution (optimal $K \approx 10$) and extensive admixture among most basins; nevertheless, Putumayo remained the most distinctive unit (high private-allele content and low connectivity with other basins). Our wild-population dataset mirrors this broad pattern: Putumayo was consistently distinct across NJ/STRUCTURE/DAPC, whereas Amazon-connected basins show closer relationships and finer substructure when analyzed at higher resolution.

Although a basin-wide phylogeographic reconstruction for *Myrciaria dubia* is still lacking, broader phylogenetic and biogeographic syntheses of Myrtaceae (tribe Myrteae) provide context for basin-structured differentiation in the Neotropics, where geographically structured lineages are often interpreted considering historical changes in connectivity and landscape evolution (Thornhill et al. 2015; Vasconcelos et al. 2017; Wagner 2025). Mantel tests supported a basin-structured isolation-by-distance (IBD) signal: correlations were significant when geographic separation was measured as straight-line distance (Fig. 5), indicating that genetic differentiation increases with spatial separation among river basins, whereas river-network distance was not significantly associated with genetic distance. This pattern of IBD under straight-line distance, but not along the river network, is consistent with reports for other Myrtaceae (e.g., *Eugenia dysenterica*; Zucchi et al. 2003). Populations from the Napo, Ucayali, Nanay, and Tahuayo basins also largely overlapped in ordination/clustering, consistent with higher connectivity and/or more recent shared ancestry; consistent with this, several pairwise F_{ST} comparisons among these Amazon-connected basins were among the lowest observed (e.g., Tahuayo–Ucayali, $F_{ST} = 0.093$; Nanay–Napo, $F_{ST} = 0.117$; Nanay–Tahuayo, $F_{ST} = 0.131$; Table 2), whereas comparisons involving Tigre and Curaray were much higher (up to $F_{ST} = 0.661$), indicating restricted effective gene flow. This broad signal (higher connectivity among Amazon-connected basins and differentiation involving Tigre/Curaray) was also recovered in the ex situ dataset of Castro et al. (2024), although numerical F_{ST} values are not directly comparable given differences in loci number and sampling design. AMOVA indicated that most genetic variation occurred within individuals (56.45%), with additional variation among individuals within populations (8.82%) and a substantial fraction among populations (34.73%), supporting pronounced geographic differentiation. The concordant signal of among-population differentiation across summary statistics and clustering/ordination approaches is consistent with expectations that multi-allelic microsatellite datasets can yield high power to detect allele-frequency divergence even at modest differentiation levels (Ryman et al. 2006). These findings are broadly comparable to those reported by

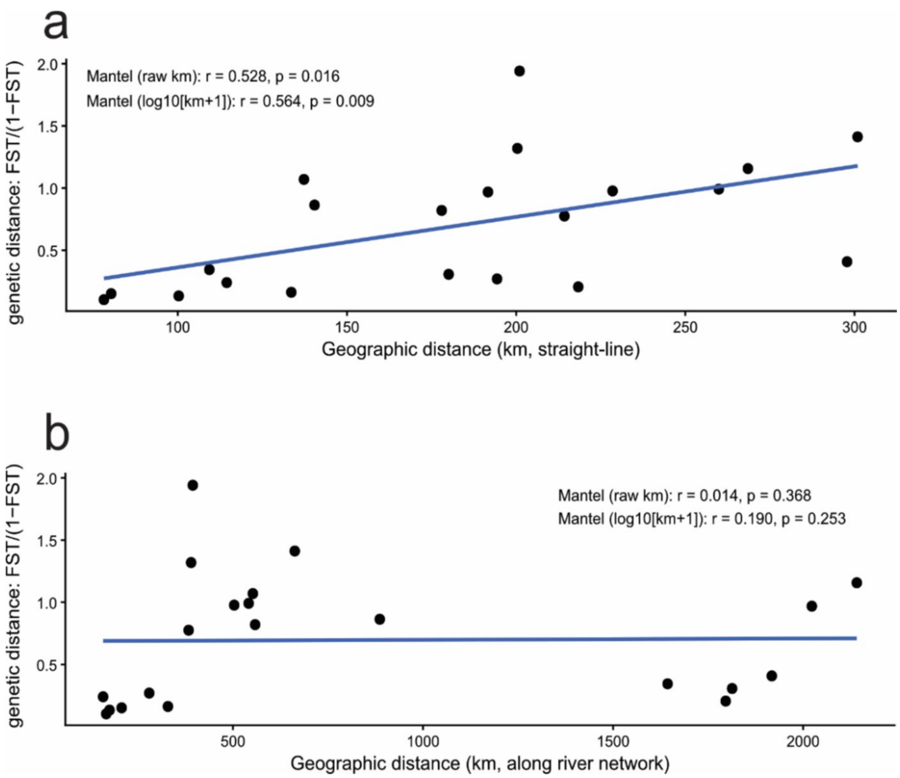

Fig. 5 Discriminant Analysis of Principal Components (DAPC) scatterplot for 254 *Myrciaria dubia* individuals. Individuals are plotted according to the first two discriminant functions (FD1 and FD2), which explain 72.94% of the total discriminant variance. Colors and labels (1–7) denote the seven a priori geographic populations (1: Nanay; 2: Tigre; 3: Putumayo; 4: Curaray; 5: Ucayali; 6: Tahuayo; 7: Napo); with 95% inertia ellipses drawn around each group. The inset barplot shows the eigenvalues for all discriminant axes. See Supplementary Table S8 for mean FD1 and FD2 scores (± SD) by population

Castro et al. (2024) for ex situ camu-camu germplasm banks, which also showed that most genetic variation is retained within sampling units (73% within accessions and 86% within river basins). Ex situ collections often assemble plant material collected from multiple origins in a single site for evaluation and improvement programs (e.g., progenies sourced from different institutions/origins and established together in a genetic test; Pinedo-Panduro et al. 2020). In contrast, wild floodplain populations are shaped more directly by spatially structured dispersal and strong environmental gradients associated with the Amazonian flood-pulse regime (hydrological connectivity, inundation duration, and flood dynamics; Sioli 1984; Junk et al. 1989). At the same time, basin-wide disturbances can modify floodplain habitats and connectivity: land-cover change and drought-related fire dynamics can affect forest structure and regeneration, hydropower development can alter river flow regimes and floodplain connectivity, and mining expansion can drive rapid habitat degradation in river corridors (Swenson et al. 2011; Winemiller et al. 2016; Aragão et al. 2018).

Based on this study of wild populations from the northern Peruvian Amazon, camu-camu populations are structured primarily at the river-basin scale influenced by three main factors: (i) reduced effective population size, potentially linked to high harvesting pressure and habitat degradation in some river systems, combined with the potential for self-pollination via geitonogamy and an optional outcrossing reproductive system (Peters and Vásquez 1987; Villachica 1996; Maués and Couturier 2002), which can affect within-population diversity and increase local inbreeding or substructure; (ii) hydrological connectivity and dispersal along floodplain networks, including seed transport by river currents and by frugivorous fishes (ichthyochory) such as gamitana (*Colossoma macropomum*), can increase connectivity and exchange among nearby riverine stands while still generating spatially structured genetic patterns (Villachica 1996; Kubitzki and Ziburski 1994; Anderson et al. 2009; Horn et al. 2011); (iii) spatial isolation among river basins, together with increasing habitat disturbance, may reduce connectivity and contribute to persistent differentiation (Zarfl et al. 2015; Winemiller et al. 2016; Aragão et al. 2018; Rudke et al. 2020). Given that extraction from flooded forests and associated aquatic environments is central to household economies in riverine communities of the Peruvian Amazon (Gram et al. 2001), conservation and management measures should be designed to reduce genetic erosion while remaining compatible with local livelihood dependence. Finally, the genetic affinity between Tigre and Curaray may reflect past connectivity, potentially linked to historical hydrological connections (water bodies that have shifted or disappeared over time) and/or dispersal by terrestrial vertebrates (e.g., ungulates or other mammals; García-Dávila et al. 2008). Together, these results provide a basin-scale baseline for interpreting wild population structure in northern Peru and for informing provenance-aware conservation and breeding decisions.

Conclusions and outlook

Wild populations of *Myrciaria dubia* in the Peruvian Amazon retain substantial genetic diversity but are strongly structured at the river-basin scale. The concordant signal from differentiation metrics (F_{ST}), clustering (STRUCTURE/NJ), ordination (DAPC), and AMOVA indicates that these populations represent distinct units for germplasm conservation and the design of breeding strategies, rather than interchangeable sources of planting material. Putumayo represents a distinct genetic reservoir and should be prioritized for both in situ protection and ex situ representation in regional germplasm banks. From a crop-improvement perspective, these river-basin units provide a practical framework for provenance-based parent selection and crossing schemes: selecting unrelated parents within a basin can help limit inbreeding risk in seed orchards and breeding nurseries, whereas controlled crosses among differentiated basins can broaden the breeding base while maintaining provenance identity.

Recent ex situ efforts in Peru further underscore both the value and the limitations of relying solely on collections. The INIA germplasm bank and its optimized core subset captured most allelic diversity while reducing the number of maintained genotypes (Castro et al. 2024), demonstrating that ex situ conservation can be highly efficient. However, Castro et al. (2024) also reported extensive admixture among most river basins, whereas Putumayo accessions were comparatively genetically homogeneous, indicating that provenance mixing can obscure natural structure and that some basins represent distinctive genetic resources. Together with our basin-structured signal in wild populations, this supports a provenance-explicit strategy in which ex situ holdings complement rather than replace in situ conservation of the river-basin units that maintain evolutionary processes and locally distinctive lineages.

Based on these results, we recommend: (i) maintaining provenance identity when establishing or enriching germplasm collections, ensuring that each major genetic unit is represented to preserve evolutionary potential; (ii) implementing provenance-aware seed collection and exchange at appropriate spatial scales to maintain genetic integrity while supporting connectivity; and (iii) prioritizing monitoring and, where warranted, targeted enrichment or ex situ safeguarding for populations showing reduced diversity and/or significant heterozygote deficits (e.g., Curaray, Napo, Tahuayo), where local inbreeding or substructure may reduce adaptive capacity. For crop improvement, the observed river-basin structure implies that parent selection and crossing designs should be provenance-aware: combining material from differentiated genetic units may broaden the breeding base, whereas unstructured mixing may dilute distinctive lineages and complicate the maintenance of locally adapted genetic combinations. Accordingly, breeding, restoration, and productive plantings should explicitly track basin provenance. Any germplasm exchange among breeding programs should be provenance-explicit and fully traceable. Where exchange would entail transboundary use of Peruvian *M. dubia* genetic resources for research or breeding, access is not automatic and is subject to access and benefit-sharing procedures (prior informed consent and mutually agreed terms) under the Nagoya Protocol and the applicable Andean/Peruvian framework (Andean Community Decision 391; Supreme Decree No. 019–2021-MINAM).

Beyond germplasm actions, conserving these river-basin units will require addressing landscape-scale pressures that increasingly affect camu-camu floodplain habitats across Amazonia, including deforestation-associated fire, mining expansion, and hydropower infrastructure (Zarfl et al. 2015; Winemiller et al. 2016; Rudke et al. 2020; Freitas et al. 2022). Such stressors can fragment riparian and várzea environments, reduce regeneration, and compromise the connectivity that sustains genetic diversity and adaptive capacity. Conservation planning should integrate provenance-explicit genetic units with monitoring of habitat loss and disturbance, prioritizing basins where unique lineages coincide with heightened vulnerability (Castro et al. 2024; Villachica 1996).

Acknowledgements This work received partial financial support from the Proyecto de Innovación Científica y Competitividad para el Agro Peruano (INCAGRO), sub-project “Mejoramiento genético de Myrciaria dubia (camu-camu) para sistemas de producción en suelos inundables”. We also thank the Instituto de Investigaciones de la Amazonía Peruana (IIAP) for institutional, logistical, and administrative support. We are grateful to Nicolas Hubert (IRD) for constructive comments and suggestions that improved the manuscript.

Author contribution All authors contributed to the study. Conceptualization: E. Mejia de Loayza, C. García-Dávila. Methodology: E. Mejia de Loayza, D. Castro-Ruiz, W. Chota-Macuyama. Investigation: all authors. Formal analysis: E. Mejia de Loayza, G. Estivals, C. García-Dávila. Data curation: E. Mejia de Loayza, G. Estivals. Writing – original draft: E. Mejia de Loayza. Writing – review & editing: all authors. Supervision: C. García-Dávila, D. Del-Castillo Torres, K. Mejia. All authors reviewed and approved the manuscript.

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).

Data availability All data supporting the findings of this study are available within the paper and its Supplementary Material. Microsatellite locus information (including primer sequences and original reference) is provided in Supplementary Table S1, and the microsatellite genotype datasets are provided in Supplementary Tables S2–S4.

Declarations

Conflict of interest The authors declare no competing interests.

References

Akter MS, Oh S, Eun JB, Ahmed M (2011) Nutritional compositions and health promoting phytochemicals of camu-camu (*Myrciaria dubia*) fruit: a review. Food Res Int 44(7):1728–1732. https://doi.org/10.1016/j.foodres.2011.03.045

Alvarado M (1969) Posibilidades del cultivo del camu-camu (*Myrciaria dubia*) en el Perú. Pontificia Universidad Católica del Perú, Lima, Perú, Tesis de maestría

Alvarado LF, Macedo VD, Peña PP, Díaz RP, Mejía K (2020) Estructura poblacional de *Mauritia flexuosa* y *Oenocarpus bataua* en tres comunidades de la cuenca alta del Putumayo, frontera Perú-Colombia. Cienc Amaz Puyo 8(2):151–166. https://doi.org/10.22386/ca.v8i2.294

Anderson JT, Saldaña Rojas J, Flecker AS (2009) High-quality seed dispersal by fruit-eating fishes in Amazonian floodplain habitats. Oecologia 161:279–290. https://doi.org/10.1007/s00442-009-1371-4

Aragão LE, Anderson LO, Fonseca MG, Rosan TM, Vedovato LB et al (2018) 21st century drought-related fires counteract the decline of Amazon deforestation carbon emissions. Nat Commun 9(1):536. https://doi.org/10.1038/s41467-017-02771-y

Arellano-Acuña E, Rojas-Zavaleta I, Paucar-Menacho LM (2016) Camu-camu (*Myrciaria dubia*): fruta tropical de excelentes propiedades funcionales que ayudan a mejorar la calidad de vida. Sci Agropecu 7(4):433–443. https://doi.org/10.17268/sci.agropecu.2016.04.08

Arthofer W, Heussler C, Krapf P, Schlick-Steiner BC, Steiner FM (2018) Identifying the minimum number of microsatellite loci needed to assess population genetic structure: a case study in fly culturing. Fly 12(1):13–22. https://doi.org/10.1080/19336934.2017.1396400

Azevedo L, de Araujo Ribeiro PF, de Carvalho Oliveira JA, Correia MG, Ramos FM, de Oliveira EB et al (2019) Camu-camu (*Myrciaria dubia*) from commercial cultivation has higher levels of bioactive compounds than native cultivation (Amazon Forest) and presents antimutagenic effects in vivo. J Sci Food Agric 99(2):624–631. https://doi.org/10.1002/jsfa.9224

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Montpellier, France: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II

Borges LL, Conceição EC, Silveira D (2014) Active compounds and medicinal properties of *Myrciaria* genus. Food Chem 153:224–233. https://doi.org/10.1016/j.foodchem.2013.12.064

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32(3):314

Brondani R, Brondani C, Tarchini R et al (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theor Appl Genet 97:816–827. https://doi.org/10.1007/s001220050961

Castro JC, Ruiz MC, Saavedra RR, Correa SI (2012) Aislamiento de ADN genómico de *Myrciaria dubia* (HBK) "camu camu" apropiado para análisis moleculares. Cienc Amaz Puyo 2:7–16. https://doi.org/10.22386/ca.v2i1.19

Castro JC, Gutiérrez F, Acuña C, Cerdeira LA, Tapullima A, Cobos M et al (2013) Variación del contenido de vitamina C y antocianinas en *Myrciaria dubia* "camu camu." Rev Soc Quím Perú 79(4):319–330

Castro JC, Vasquez-Guizado SJ, Vigil BE, Ascue F, Rojas-Villa N, Paredes JD et al (2024) Development and application of microsatellite markers for genetic diversity assessment and construction of a core collection of *Myrciaria dubia* (Kunth) Mcvaugh germplasm from the Peruvian Amazon. For 15(11):1873. https://doi.org/10.3390/f15111873

Castro JC, Maddox JD, Cobos M, Imán SA (2018) Myrciaria dubia "Camu Camu". fruit: health-promoting phytochemicals and functional genomic characteristics. In: Soneji and Nagesawara-Rao (ed) Breeding and health benefits of fruit and nut crops. InTech, London, pp 85–116

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24(3):621–631. https://doi.org/10.1093/molbev/msl191

Correa SI, Zamudio LB, Solís VS, Cruz CO (2011) Contenido de vitamina C en frutos de camu camu *Myrciaria dubia* (HBK) Mc Vaugh, en cuatro estados de maduración, procedentes de la Colección de Germoplasma del INIA Loreto, Perú. Sci Agropecu 2(3):123–130. https://doi.org/10.17268/sci.agropecu.2011.03.01

Dixon P (2003) VEGAN, a package of R functions for community ecology. J Veg Sci 14(6):927–930

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical bulletin

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131(2):479–491. https://doi.org/10.1093/genetics/131.2.479

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4):1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Felsenstein J (1989–2006) PHYLIP (phylogeny inference package). http://evolution.genetics.washington.edu/phylip.html

Ferreira-Ramos R, Accoroni KA, Rossi A, Guidugli MC, Mestriner MA, Martinez CA, Alzate-Marin AL (2014) Genetic diversity assessment for *Eugenia uniflora* L., *E. pyriformis* Cambess., *E. brasiliensis* Lam. and *E. francavilleana* O. Berg neotropical tree species (Myrtaceae) with heterologous SSR markers. Genet Resour Crop Evol 61:267–272. https://doi.org/10.1007/s10722-013-0028-7

Foulley J, Hill WG (1999) On the precision of estimation of genetic distance. Genet Sel Evol 31:457–464

Frankham R (1996) Relationship of genetic variation to population size in wildlife. Conserv Biol 10(6):1500–1508. https://doi.org/10.1046/j.1523-1739.1996.10061500.x

Freitas CE, de Mereles M, A, Pereira D, Siqueira-Souza FK, et al (2022) Death by a thousand cuts: Small local dams can produce large regional impacts in the Brazilian Legal Amazon. Environ Sci and Policy 136:447–452

Fujita A, Sarkar D, Wu S, Kennelly E, Shetty K, Genovese M (2015) Evaluation of phenolic-linked bioactives of camu-camu (*Myrciaria dubia* Mc. Vaugh) for antihyperglycemia, antihypertension, antimicrobial properties and cellular rejuvenation. Food Res Int 77(2):194–203. https://doi.org/10.1016/j.foodres.2015.07.009

García-Dávila C, Corazón-Guivin M, Castro-Ruiz D, Chota-Macuyama W, Rodríguez A, Delgado-Vásquez C, Renno J (2008) Variabilidad genética de cinco poblaciones naturales de camu-camu (*Myrciaria dubia* HBK Mc. Vaugh) de la Amazonía Peruana, evaluadas mediante DALP. Folia Amaz 17(1–2):91–98. https://doi.org/10.24841/fa.v17i1-2.271

Gonzáles C (2007) Frutales nativos amazónicos: patrimonio alimenticio de la humanidad. Instituto de Investigaciones de la Amazonía Peruana, Iquitos, Perú

Gram S, Kvist LP, Cáseres A (2001) The economic importance of products extracted from Amazonian flood plain forests. Ambio 30(6):365–368. https://doi.org/10.1579/0044-7447-30.6.365

Gutiérrez-Ruiz A (1969) Especies frutales nativas de la selva del Perú: estudio botánico y de propagación por semillas. Universidad Nacional Agrária La Molina, Lima, Perú, Tesis

Hajjar R, Hodgkin T (2007) The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 156:1–13. https://doi.org/10.1007/s10681-007-9363-0

Hernández M, Barrera J (2014) Organización social para el aprovechamiento sostenible del camu-camu (*Myrciaria dúbia* (Kunth) McVaugh) en Tarapacá, departamento del Amazonas, Colombia. Editorial Legis SA, Colombia

Horn MH, Correa SB, Parolin P, Pollux BJA, Anderson JT, Lucas C (2011) Seed dispersal by fishes in tropical and temperate fresh waters: the growing evidence. Acta Oecol 37(6):561–577. https://doi.org/10.1016/j.actao.2011.06.004

https://doi.org/10.57098/SciRevs.Biology.4.2.3

Jombart T (2008) Adegenet: an R package for the multivariate analysis of genetic markers. Bioinform 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:1–15. https://doi.org/10.1186/1471-2156-11-94

Junk WJ, Bayley PB, Sparks RE (1989) The flood pulse concept in river-floodplain systems. Can Spec Publ Fish Aquat Sci 106(1):110–127

Kalinowski ST (2002) How many alleles per locus should be used to estimate genetic distances? Heredity 88(1):62–65. https://doi.org/10.1038/sj.hdy.6800009

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

Koskinen MT, Hirvonen H, Landry PA, Primmer CR (2004) The benefits of increasing the number of microsatellites utilized in genetic population studies: an empirical perspective. Hereditas 141(1):61–67. https://doi.org/10.1111/j.1601-5223.2004.01804.x

Kubitzki K, Ziburski A (1994) Seed dispersal in flood plain forests of Amazonia. Biotropica. https://doi.org/10.2307/2389108

Liu K, Muse S (2005) Powermarker: an integrated analysis environment for genetic marker analysis. Bioinform 21:2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Liu W, Kang M, Tian H, Huang H (2013) A range wide geographic pattern of genetic diversity and population structure of *Castanea mollissima* populations inferred from nuclear and chloroplast microsatellites. Tree Genet Genomes 9(4):975–987. https://doi.org/10.1007/s11295-013-0610-3

Lu W, Arnold RJ, Zhang L, Luo J (2018) Genetic diversity and structure through three cycles of a *Eucalyptus urophylla* ST Blake breeding program. Forests 9(7):372. https://doi.org/10.3390/f9070372

Marshall E, Newton A, Schreckenberg K (2003) Commercialization of non-timber forest products: first steps in analyzing the factors influencing success. Int for Rev 5(2):128–137. https://doi.org/10.1505/IFOR.5.2.128.17410

Martin M, Peters C, Ashton M (2014) Revisiting camu-camu (*Myrciaria dubia*): twenty-seven years of fruit collection and flooding at an oxbow lake in Peruvian Amazonia. Econ Bot 68(2):169–176

Maués M, Couturier G (2002) Biologia floral e fenologia reprodutiva do camu-camu (Myrciaria dubia (H.B.K.) McVaugh, Myrtaceae) no Estado Pará, Brasil. Rev Bras Bot 25(4):441–448. https://doi.org/10.1590/S0100-84042002012000008

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet 89(3):583–590. https://doi.org/10.1093/genetics/89.3.583

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evol. https://doi.org/10.1111/j.1558-5646.1975.tb00807.x

Nunes CF, Setotaw TA, Pasqual M, Chagas EA, Santos EG, dos Santos DN, Lima CGB, de Almeida Cançado GM (2017). *Myrciaria dubia*, an Amazonian fruit: population structure and its implications for germplasm conservation and genetic improvement. Genetics and Molecular Research.

Paradis E (2010) Pegas: an R package for population genetics with an integrated–modular approach. Bioinform 26(3):419–420. https://doi.org/10.1093/bioinformatics/btp696

Payn K, Dvorak W, Janse B, Myburg A (2008) Microsatellite diversity and genetic structure of the commercially important tropical tree species *Eucalyptus urophylla*, endemic to seven islands in eastern Indonesia. Tree Genet Genomes 4(3):519–530. https://doi.org/10.1007/s11295-007-0128-7

Penn JW (2006) The cultivation of camu camu (*Myrciaria dubia*): a tree planting programme in the Peruvian Amazon. For Trees Livelihoods 16(1):85–101. https://doi.org/10.1080/14728028.2006.9752547

Penn JW (2004) Another boom for Amazonia? Examining the socioeconomic and environmental implications of the new camu camu industry in Peru. Dissertation, University of Florida

Peters CM, Hammond EJ (1990) Fruits from the Flooded Forests of Peruvian Amazonia: Yield Estimates for Natural Populations of Three Promising Species. Adv in Econ Bot 8:159–176

Peters C, Vasquez A (1987) Estudios ecológicos de camu-camu (*Myrciaria dubia*). I. Producción de frutos en poblaciones naturales. Acta Amaz 17:161–188. https://doi.org/10.1590/1809-43921987171174

Pinedo M, Riva R, Rengifo E, Delgado C, Villacrez J, Gonzales A, Inga H, López A, Farroñay R, Vega R, Linares C (2001) Sistema de Producción de Camu-Camu en Restinga, Instituto de Investigaciones de la Amazonia Peruana. Programa Manejo Integral del Bosque y Servicios Ambientales – PROBOSQUES. Loreto, Perú

Pinedo-Panduro M, Alves-Chagas E, Paredes-Davila E, Abanto-Rodriguez C et al (2020) Individual selection in a genetic test with 43 advanced progenies of camu-camu *Myrciaria dubia* Kunth (McVaugh) in floodable plot near Iquitos, Peruvian Amazon. Sci Agropecu 11(3):345–355. https://doi.org/10.17268/sci.agropecu.2020.03.06

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):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, Austria.

Rodrigues RB, Marx F (2006) Camu camu [*Myrciaria dubia* (H.B.K.) McVaugh]: a promising fruit from the Amazon Basin. Nutri 30(9):376–381

Rojas S, Rodrigues D, Lima M, Fhilo A (2008) Desenvolvimento e mapeamento de microssatélites gênicos (EST-SSRs) de camu-camu (*Myrciaria dubia* [H.B.K.] McVaugh). Rev Corpoica Cienc Tecnol Agropecu 9(1):14–21. https://doi.org/10.21930/rcta.vol9_num1_art:100

Rojas S, Yuyama K, Clement C, Nagao E (2011) Diversidade genética em acessos do banco de germoplasma de camu-camu (*Myrciaria dubia* [HBK] McVaugh) do INPA usando marcadores microssatélites (EST-SSR). Ciência y Tecnologia Agropecuária 12(1):51–64. https://doi.org/10.21930/rcta.vol12_num1_art:215

Rossetto M, Slade R, Baverstock P, Henry R, Lee L (1999) Microsatellite variation and assessment of genetic structure in tea tree *Melaleuca alternifolia* -Myrtacea. Mol Ecol 8:633–643. https://doi.org/10.1046/j.1365-294x.1999.00622.x

Rousset F (1997) Genetic differentiation and estimation of gene flow from f-statistics under isolation by distance. Genetics 145(4):1219–1228. https://doi.org/10.1093/genetics/145.4.1219

Rudke AP, Sikora de Souza VA, Santos AMD, Freitas Xavier AC et al (2020) Impact of mining activities on areas of environmental protection in the southwest of the Amazon: a GIS- and remote sensing-based assessment. J Environ Manage 263:110392. https://doi.org/10.1016/j.jenvman.2020.110392

Ryman N, Palm S, André C, Carvalho GR, Dahlgren TG, Jorde PE et al (2006) Power for detecting genetic divergence: differences between statistical methods and marker loci. Mol Ecol 15(8):2031–2045. https://doi.org/10.1111/j.1365-294X.2006.02839.x

Salgueiro F, Felix D, Caldas J, Margis-Pinheiro M, Margis R (2004) Even population differentiation for maternal and biparental gene markers in *Eugenia uniflora*, a widely distributed species from the Brazilian coastal Atlantic rain forest. Divers Distrib 10(3):201–210. https://doi.org/10.1111/j.1366-9516.2004.00078.x

Schmidt A, Lajolo F, Genovese M (2010) Chemical composition and antioxidant/antidiabetic potential of Brazilian native fruits and commercial frozen pulps. J Agric Food Chem 58(8):4666–4674. https://doi.org/10.1021/jf903875u

Serrote CML, Reiniger LRS, Silva KB, dos Santos Rabaiolli SM, Stefanel CM (2020) Determining the polymorphism information content of a molecular marker. Gene 726:144175. https://doi.org/10.1016/j.gene.2019.144175

Shackleton C, Shackleton S (2004) The importance of non-timber forest products in rural livelihood security and as safety nets: a review of evidence from South Africa. S Afr J Sci 100:658–664

Shanley P (2011) Fruit trees and useful plants in Amazonian life. Food and Agriculture Organization of the United Nations (FAO), the Center for International Forestry Research and People and Plants International, Rome, Italy

Sioli H (1984) The Amazon and its main affluents: Hydrography, morphology of the river courses, and river types. In: Sioli H (ed) The Amazon Monographiae Biologicae, vol 56. Springer, Dordrecht, p 127

Slatkin M (1985) Rare alleles as indicators of gene flow. Evolution 39(1):53–65. https://doi.org/10.1111/j.1558-5646.1985.tb04079.x

Šmíd J, Kalousová M, Mandák B, Houška J, Chládová A, Pinedo M, Lojka B (2017) Morphological and genetic diversity of camu-camu [*Myrciaria dubia* (Kunth) McVaugh] in the Peruvian Amazon. PLoS ONE. https://doi.org/10.1371/journal.pone.0179886

Stevens L, Salomon B, Sun G (2007) Microsatellite variability and heterozygote excess in *Elymus trachycaulus* populations from British Columbia in Canada. Biochem Syst Ecol 35(11):725–736. https://doi.org/10.1016/j.bse.2007.05.017

Swenson JJ, Carter CE, Domec JC, Delgado CI (2011) Gold mining in the Peruvian Amazon: global prices, deforestation, and mercury imports. PLoS ONE 6(4):e18875. https://doi.org/10.1371/journal.pone.0018875

Thornhill AH, Ho SY, Külheim C, Crisp MD (2015) Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. Mol Phylogenet Evol 93:29–43. https://doi.org/10.1016/j.ympev.2015.07.007

Ticktin T (2004) The ecological implications of harvesting non-timber forest products. J Appl Ecol 41(1):11–21. https://doi.org/10.1111/j.1365-2664.2004.00859.x

Ticktin T, Shackleton C (2011) Harvesting Non-timber Forest Products Sustainably: Opportunities and Challenges. In: Shackleton S, Shackleton C, Shanley P (eds) Non-Timber Forest Products in the Global Context Tropical Forestry, vol 7. Springer, Berlin, pp 149–169

Trindade M, Chaves L (2005) Genetic structure of natural *Eugenia dysenterica* DC (Myrtaceae) populations in northeastern Goiás, Brazil, accessed by morphological traits and RAPD markers. Genet Mol Biol 28(3):407–413. https://doi.org/10.1590/S1415-47572005000300013

Vasconcelos TN, Proença CE, Ahmad B, Aguilar DS, Aguilar R, Amorim BS et al (2017) Myrteae phylogeny, calibration, biogeography and diversification patterns: increased understanding in the most species rich tribe of Myrtaceae. Mol Phylogenet Evol 109:113–137

Vásquez R, Gentry A (1989) Use and misuse of forest-harvested fruits in the Iquitos area. Conserv Biol 3(4):350–361. https://doi.org/10.1111/j.1523-1739.1989.tb00241.x

Villachica LH (1996) El cultivo del camu camu (*Myrciaria dubia* HBK Mc Vaugh) en la Amazonía Peruana (No. F01–50). Tratado de Cooperación Amazonia, Lima, Perú

Wagner MA (2025) Current taxonomic and systematic knowledge on South American species of Myrtaceae: a review. Sci Rev Biol 4(2):21–29

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358–1370. https://doi.org/10.2307/2408641

Winemiller KO, McIntyre PB, Castello L, Fluet-Chouinard E, Giarrizzo T et al (2016) Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. Sci 351(6269):128–129

Yuyama K, Aguiar J, Yuyama LK (2002) Camu-camu: um fruto fantástico como fonte de vitamina C1. Acta Amaz 32(1):169–174. https://doi.org/10.1590/1809-43922002321174

Zarfl C, Lumsdon AE, Berlekamp J et al (2015) A global boom in hydropower dam construction. Aquat Sci 77:161–170. https://doi.org/10.1007/s00027-014-0377-0

Zucchi M, Brondani R, Pinheiro J, Chaves L, Coelho A, Vencovsky R (2003) Genetic structure and gene flow in *Eugenia dysenterica* DC in the Brazilian Cerrado utilizing SSR markers. Genet Mol Biol 26:449–457. https://doi.org/10.1590/S1415-47572003000400008