forests

MDPI

Article

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

Juan C. Castro [1,2,*], Stalin J. Vasquez-Guizado [3], Bianca E. Vigil [4], Francisco Ascue [5], Naysha Rojas-Villa [6], Jae D. Paredes [1], Marianela Cobos [1,2], Carlos G. Castro [1], Daniel E. Motta [1], Pedro M. Adrianzén [1,2], Sixto A. Imán [7,8] and J. Dylan Maddox [9,10,*]

1 Specialized Unit of Biotechnology Research Laboratory, Natural Resources Research Center of UNAP, Universidad Nacional de la Amazonía Peruana (UNAP), Iquitos 16001, Peru; jaedianap@gmail.com (J.D.P.); marianela.cobos@unapiquitos.edu.pe (M.C.); castrocoboscarlosgilbertox13@gmail.com (C.G.C.); dmottas17@gmail.com (D.E.M.); pedro.adrianzen@unapiquitos.edu.pe (P.M.A.)
2 Academic Department of Biomedical Sciences and Biotechnology, Faculty of Biological Sciences, Universidad Nacional de la Amazonía Peruana (UNAP), Iquitos 16001, Peru
3 Postgraduate School, Faculty of Biological Sciences, Universidad Nacional de San Agustín de Arequipa, Arequipa 04001, Peru; svasquezg@unsa.edu.pe
4 Plant Biotechnology Laboratory, FORESLAB S.A.C., Madre de Dios 20150, Peru; bianca.vigil@foreslab.com
5 Postgraduate Unit, Universidad de Ingeniería y Tecnología, Lima 15047, Peru; francisco.ascue@utec.edu.pe
6 International Center of Research for Sustainability, Universidad Nacional de Cañete, Cañete 15701, Peru; n_rojas@undc.edu.pe or naysharojas.v@gmail.com
7 San Roque Agricultural Experiment Station, Direction of Genetic Resources and Biotechnology, Instituto Nacional de Innovación Agraria (INIA), Iquitos 16001, Peru; siman@inia.gob.pe
8 Academic Department of Ecology and Conservation, Faculty of Forestry Sciences, Universidad Nacional de la Amazonía Peruana (UNAP), Iquitos 16001, Peru
9 Pritzker Laboratory for Molecular Systematics and Evolution, Field Museum of Natural History, 1400 S. Lake Shore Drive, Chicago, IL 60605, USA
10 Environmental Sciences, American Public University System, Charles Town, WV 25414, USA
* Correspondence: juan.castro@unapiquitos.edu.pe (J.C.C.); dmaddox@fieldmuseum.org (J.D.M.); Tel.: +51-965-017215 (J.C.C.); +1-(217)-390-5081 (J.D.M.)

Abstract: The Amazonian shrub *Myrciaria dubia* (camu-camu) produces vitamin C-rich fruits of growing commercial interest. However, sustainable utilization requires assessment and protection of the genetic diversity of the available germplasm. This study aimed to develop and apply microsatellite markers to assess genetic diversity and construct a core collection of *M. dubia* germplasm from the Peruvian Amazon. Sixteen polymorphic microsatellite loci were developed using an enrichment approach. The evaluation of 336 genotypes from 43 accessions of the germplasm bank, originating from eight river basins, was conducted using these newly developed markers. Genetic diversity parameters, including observed and expected heterozygosity, were calculated. Analysis of molecular variance (AMOVA) was performed to assess the distribution of genetic variation within and among accessions and river basins. Bayesian clustering analysis was employed to infer population structure. A core collection was constructed to maximize allelic richness. High genetic diversity was observed, with heterozygosity values ranging from 0.468 to 0.644 (observed) and 0.684 to 0.817 (expected) at the river basin level. AMOVA indicated significant genetic variation within (73–86%) compared to among (14–27%) accessions and river basins. Bayesian clustering detected ten genetic clusters, with several degrees of admixture among river basins, except for the genetically homogeneous Putumayo River basin. A core collection comprising 84 plant genotypes (25% of the full collection) was established, representing 90.82% of the overall allelic diversity. These results have important implications for *M. dubia* conservation strategies and breeding programs, in demonstrating a need for genetic connectivity between populations but preserving unique genetic resources in isolated basins. These results validate the expected levels of diversity and population subdivision in a crop and stress

check for updates

Citation: Castro, J.C.; Vasquez-Guizado, S.J.; Vigil, B.E.; Ascue, F.; Rojas-Villa, N.; Paredes, J.D.; Cobos, M.; Castro, C.G.; Motta, D.E.; Adrianzén, P.M.; et al. 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. *Forests* **2024**, *15*, 1873. https://doi.org/10.3390/f15111873

Academic Editors: Víctor Rincón, Javier Velázquez Saornil and Derya Gulcin

Received: 31 July 2024
Revised: 20 August 2024
Accepted: 20 August 2024
Published: 25 October 2024

the need to secure genetically diverse germplasms, underscoring the importance of thorough genetic characterization for ex situ germplasm management.

Keywords: Amazonian crop; ex situ conservation; core collection; molecular markers; plant genetic resources; simple sequence repeats

1. Introduction

Myrciaria dubia (Kunth) McVaugh, commonly known as camu-camu, is a shrub native to the Amazon rainforest [1]. It is a diploid Myrtaceae species, with 2n = 22 chromosomes [2]. Its complex reproductive biology includes hermaphrodite flowers and a mainly allogamous mating system. The latter is ensured by longistyly, a form of heterostyly where the style is extended beyond the stamen filaments. This is the main morphological feature that promotes cross-pollination between different plants, carried out almost entirely by stingless bees, mainly *Melipona fuscopilosa* Moure & Kerr and *Scaptotrigona postica* Latreille & Moure [3,4]. Though allogamy is the primary reproduction mode, self-compatibility is manifest in this plant to a certain extent. Geitonogamy—pollen transfer from one flower to another on the same plant—is possible, although less frequent [4]. A mixed mating strategy like this would combine the benefits of cross-pollination with potential self-fertilization, therefore helping to contribute to the genetic diversity and adaptability of the *M. dubia* populations in their native habitats. Investigations into its genetic structure showed that there were moderate-to-high levels of genetic diversity within populations, while significant differentiation among populations from different river basins in Amazonia was indicated [5,6].

M. dubia is an underutilized crop that produces one of the richest natural sources of vitamin C, as its fruits have up to 7355 mg L-ascorbic acid per 100 g of their pulp [7,8]. Beyond its exceptionally high vitamin C content, camu-camu fruits are abundant in bioactive compounds with antioxidant and anti-inflammatory properties, including anthocyanins, flavonoids, and phenolic acids [9–12]. The growing demand for natural sources of antioxidants and functional foods has sparked interest in domesticating and commercializing this species [13,14].

While the nutritional and economic potential of *M. dubia* has driven interest in domestication, long-term conservation and sustainable utilization of this genetic resource requires securing representative germplasm ex situ. Establishing ex situ field gene banks and seed banks is an important strategy, but maximizing captured genetic diversity remains a major challenge [15,16]. Many ex situ collections, however, have been founded based on comprehensive sampling in the native distribution area of plant species. This can give a better representation of the gene pool and mitigate genetic erosion in successive generations [17,18]. Questions relating to these concerns are discussed in several publications about ex situ conservation and genetic diversity [19–24]. Thus, comprehensive assessments of genetic diversity in available ex situ germplasms become very critical in optimizing management practices, identifying unique accessions, and delineating subsets that maximize diversity for breeding and long-term conservation [15,25]. For *M. dubia*, however, genetic diversity assessments within and among ex situ collections have been limited.

Microsatellites, or simple sequence repeats (SSRs) offer a powerful tool for genetic analysis in plant species. These highly polymorphic codominant molecular markers have been widely used to assess genetic diversity, population structure, gene flow, and mating systems [26]. While several studies have successfully developed and applied SSR markers for assessing the genetic diversity of various Myrtaceae species [27–31], only a few have been published for *M. dubia* [5,6], potentially limiting comprehensive germplasm characterization.

Given the wide geographic distribution of *M. dubia* across the Amazon basin, we hypothesize that the ex situ germplasm collection assembled from multiple regions in

the Peruvian Amazon harbors substantial genetic diversity. However, potential sampling biases may result in an unrepresentative collection with underrepresented or missing genotypes. To address this, we aim to develop informative microsatellite markers for robust quantification of genetic diversity parameters and the creation of a rationally selected core collection with maximum allelic richness. Such a core collection will be an important resource for germplasm management, conservation, and breeding of this species.

This study represents the first comprehensive assessment of genetic diversity and core collection development for *M. dubia* germplasm from the Peruvian Amazon. Our objectives were to (1) develop a set of novel polymorphic SSR markers for *M. dubia* suitable for characterizing the germplasm bank, (2) evaluate the genetic diversity and population structure of an ex situ *M. dubia* germplasm collection assembled from multiple locations across the Peruvian Amazon, and (3) construct a core collection that maximizes genetic diversity within this collection to facilitate conservation and breeding efforts.

The results of this study confirmed expected levels of diversity and population structure, ensuring genetically diverse germplasm accessions and demonstrating the importance of effective genetic characterization for ex situ germplasm management. This nutritionally and economically promising Amazonian crop species will also benefit from the newly developed microsatellite markers and the genetically characterized core collection. These will be invaluable genomic resources for future research, conservation efforts, and genetic improvement initiatives.

2. Materials and Methods

2.1. Plant Materials

A total of 336 genotypes of *M. dubia* were obtained from the ex situ germplasm bank maintained by the National Institute of Agrarian Innovation (INIA) in Peru. The germplasm bank, established in 1988 (36 years ago), contains exclusively diploid plants with a chromosomal count of 2n = 22. These genotypes represent 43 accessions originating from eight major river basins (Nanay, Itaya, Napo, Ucayali, Putumayo, Curaray, Tigre, and Amazonas) across the Loreto region of the Peruvian Amazon (Figure 1, Table S1). Young leaf samples were collected from each plant genotype and immediately transported to the laboratory facilities at the Natural Resources Research Center of UNAP (CIRNA) in containers with dry ice to preserve the tissue integrity. Leaf samples were stored at −20 °C until further processing.

2.2. DNA Isolation and Analysis

Genomic DNA was isolated from young leaf tissues following a cetyltrimethylammonium bromide (CTAB) extraction protocol with minor modifications [32]. Approximately 100 mg of fresh leaf material was ground to a fine powder in liquid nitrogen using a pre-chilled mortar and pestle. The powdered tissue was then incubated in pre-warmed CTAB extraction buffer supplemented with 2% β-mercaptoethanol at 65 °C for 60 min with intermittent mixing. Cell debris was removed by centrifugation, and the supernatant was extracted with a chloroform:isoamyl alcohol solution (24:1). Nucleic acids were precipitated from the aqueous phase by adding ice-cold isopropanol and centrifugation. The resulting DNA pellet was washed twice with 70% ethanol, air-dried briefly, and resuspended in 1X TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) for long-term storage. The integrity and purity of the extracted DNA were evaluated using a combination of agarose gel electrophoresis and UV–Vis spectrophotometry. Agarose gel electrophoresis was performed by loading an aliquot of each DNA sample onto a 0.8% agarose gel containing 0.5 µg/mL ethidium bromide [33]. After electrophoresis in 1X TBE buffer, the gels were visualized under UV illumination to assess DNA integrity based on the presence of a prominent high-molecular-weight band without significant smearing or degradation. DNA purity was determined by measuring the absorbance ratios at 260/280 and 260/230 nm using a NanoDrop 2000c UV–Vis spectrophotometer (Thermo Fisher Scientific, Inc., Pittsburgh, PA, USA). Only DNA samples exhibiting an A260/280 ratio between 1.8–2.0 and an A260/230

ratio greater than 1.8, indicative of high purity with minimal protein/polysaccharide contamination, were considered suitable for downstream applications and selected for microsatellite genotyping.

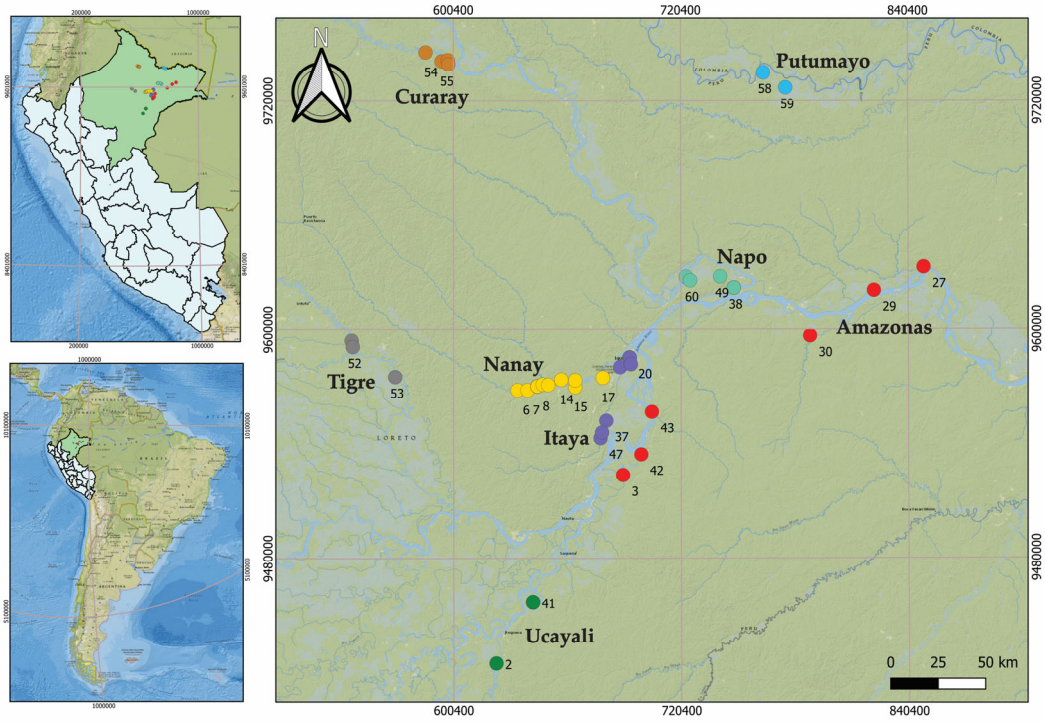

Figure 1. Geographical locations of the eight river basins in the Peruvian Amazon where INIA germplasm bank accessions of *M. dubia* were collected from wild populations. Each river basin is represented by a different color, with the Cuaray River basin in orange, the Tigre River basin in gray, the Itaya River basin in purple, the Nanay River basin in yellow, the Napo River basin in turquoise, the Amazonas River basin in red, the Putumayo River basin in blue, and the Ucayali River basin in green. The circles with numbers correspond to the accession numbers for the samples collected from those sites.

2.3. Isolation of Microsatellite DNA Loci and Analysis

SSR markers were isolated from the genomic DNA of a single *M. dubia* individual following a described enrichment protocol [34] that resulted in the identification of 16 polymorphic SSR markers that were used in this genetic diversity assessment study. Detailed information on these 16 SSR markers, including their GenBank accession numbers, primer sequences, annealing temperatures (Tm), repeat motif sequences, size ranges, number of alleles, and polymorphic information content (PIC) values, is provided in Supplementary Table S2.

Plants were genotyped following Schuelke [35] with all forward primers tagged with M13-tails (5′-TGTAAAACGACGGCCAGT-3′) to incorporate fluorescently labeled M13 primers. Reverse primers were developed with a "pigtail" (5′-GTGTCTT-3′) on the 5′ end to facilitate the addition of an adenine to all PCR products. SSR markers were amplified individually in 10 µL reactions with the following final concentrations: ≈25 ng of genomic DNA, 1 U Taq DNA polymerase, 1x PCR buffer (10 mM Tris-HCL, 50 mM KCL), 0.5 mM dNTPs, 2.0–2.5 mM $MgCl_2$, 1x BSA, 0.16 µM fluorescently labeled M13 primer (VIC, PET, NED, or 6-FAM), 0.04 µM forward primer, and 0.16 µM reverse primer. The thermal profile for all PCRs consisted of an initial denaturation step at 94 °C for 4 min, followed by 30 cycles at 94 °C for 15 s, 62 °C for 15 s, 72 °C for 45 s, followed by 8 cycles at 94 °C for 15 s, 53 °C for 15 s, 72 °C for 45 s, and a final extension step for 10 min at 72 °C. PCR reactions (1 µL) were combined with 8 µL formamide and 1 µL of a custom standard [36] and run on an

ABI 3730xl DNA analyzer (Applied Biosystems, Foster City, CA, USA) at the Pritzker DNA Laboratory for Molecular Systematics and Evolution (Field Museum, Chicago, IL, USA). Alleles were visualized and called using the microsatellite plugin v1.4.7 in Geneious Prime v2024.0.5 [37].

2.4. Data Analysis

2.4.1. Analysis of the Genetic Diversity and Population Structure

Standard genetic diversity parameters, including the number of alleles (*Na*), the effective number of alleles (*Ne*), Shannon's information index (*I*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), and Wright's fixation index (F_{IS}), were calculated using GenAlEx v6.5 [38]. Analysis of molecular variance (AMOVA) was implemented in GenAlEx to partition genetic variability within and among populations [39]. Pairwise genetic differentiation (F_{ST}) and gene flow (*Nm*) among river basins were estimated using GenAlEx. *Nm* values among river basins were determined based on the equation [40,41]:

$$Nm = 0.25\frac{(1 - F_{ST})}{F_{ST}} \quad (1)$$

Clustering based on Bayesian analysis was performed using the number of clusters (K), in the "find.cluster" function of the R package adegenet v1.3.1 [42,43], which does not take Hardy–Weinberg equilibrium into account. The basic algorithm and extensions to the method were previously described [44–47]. Twenty independent runs were performed for each K value ranging from 1 to 34, with a burn-in of 100,000 iterations and 500,000 Markov Chain Monte Carlo repetitions [44–47]. The best value of K is the first break in the curve that corresponds to the lowest value of the Bayesian Information Criterion (BIC) [48].

The number of genetic clusters was further corroborated using discriminant analysis of principal components (DAPC), an effective method for visualizing population structure that enables us to ascertain the most influential factors contributing to variation among the populations under investigation. To accomplish this, the R package adegenet v1.3.1 [42,43] was used, which comprised 200 principal components and a priori grouping of the number of river basins.

To visualize the genetic relationships among populations of river basins, a neighbor-joining tree based on Nei's genetic distances [49,50] was constructed. A distance matrix was first generated using the R package StAMPP v1.6.3 [51] and a tree was built with 1000 bootstrap replicates using the R package Poppr v2.9.6 [52,53].

2.4.2. Construction of the Core Collection

To capture the maximum allelic diversity in a representative core collection, Core-Hunter3 v3.2.2 [54] was used to select a core subset of 84 genotypes (25% of the full collection) based on a maximization strategy using Modified Rogers distance [55]. Allelic richness was calculated for the full collection and the core subset using the R package hierfstat v0.5-11 [56] to evaluate allelic representation in the core collection.

All statistical analyses were performed using R v4.4.0 [57]. Figures were generated (see scripts in https://github.com/FranciscoAscue/diversity-core-collection-myrciaria-dubia) using R packages ggplot2 v3.3.6 [58,59], pheatmap v1.0.12 [60], and SRplot [61].

3. Results

3.1. Genetic Diversity Parameters

The *M. dubia* germplasm bank, consisting of 43 accessions obtained from eight river basins (Amazonas, Curaray, Itaya, Nanay, Napo, Putumayo, Tigre, and Ucayali) in the Peruvian Amazon, revealed substantial variation in genetic parameters, indicating a rich and complex genetic structure within this species. In total, 336 genotypes were examined using 16 polymorphic SSR markers. All SSR markers were polymorphic, with the number of alleles per locus ranging from 10 to 28, resulting in 313 alleles. The polymorphic information content ranged from 0.66 to 0.94 (Figure 2, Table S2).

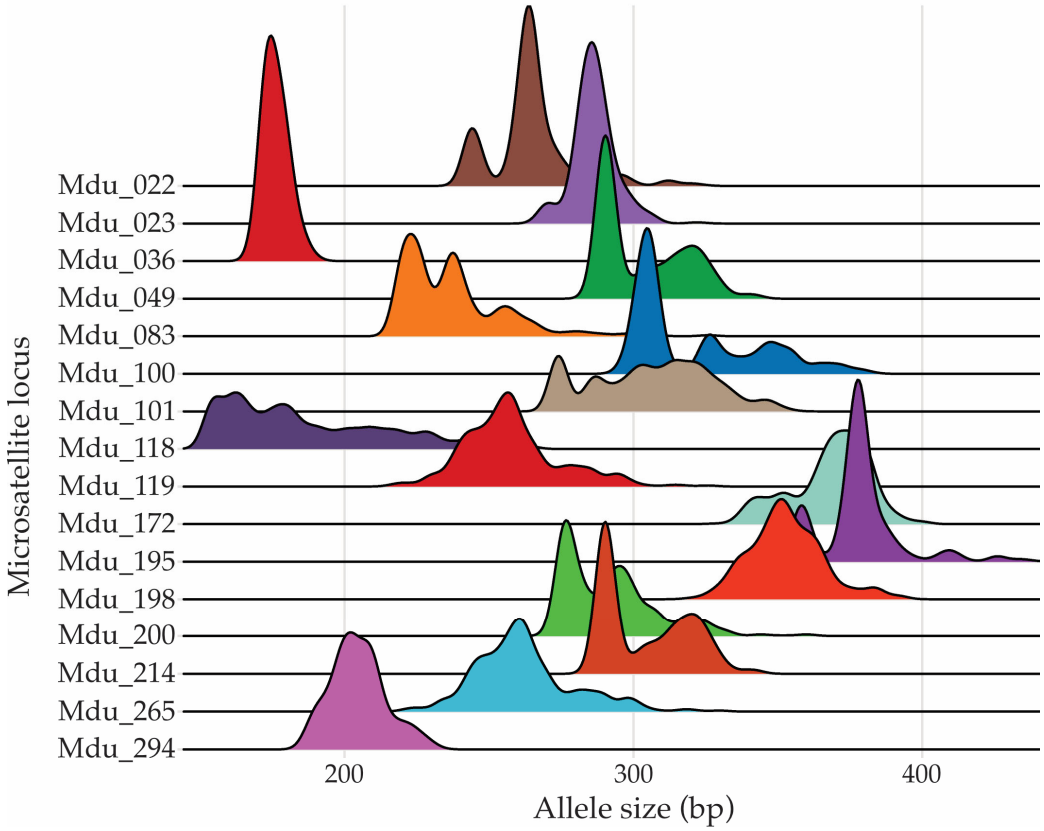

Figure 2. Allele size distribution of sixteen polymorphic SSR markers used for assessing genetic diversity and constructing the core collection of the *M. dubia* germplasm bank from the Peruvian Amazon. Each color represents the size variation and relative abundance of a specific SSR marker.

At the accession level (Table S3), the *Na* varied from 1.875 to 6.688, with an average of 5.438 ± 1.140, while the *Ne* ranged from 1.395 to 4.721, with an average of 3.899 ± 0.813. Shannon's diversity index, which measures allelic richness and evenness, ranged from 0.397 to 1.623, with an average of 1.381 ± 0.287. The *Ho* ranged from 0.104 to 0.854, with an average of 0.604 ± 0.152, whereas the *He* varied from 0.236 to 0.760, with an average of 0.675 ± 0.119. The degree of inbreeding, as measured by the F_{IS}, ranged from −0.259 to 0.511, with an average of 0.103 ± 0.177.

At the river basin level (Table 1), the *Na* per locus was 10.180 ± 1.849, with Nanay containing the highest (12.438) and Putumayo the lowest (6.813) number of alleles. The *Ne* ranged from 3.834 (Tigre) to 6.303 (Napo), with an average of 5.395 ± 0.947. Shannon's diversity index averaged 1.800 ± 0.203. The highest Shannon's diversity value was observed in Napo (2.006), whereas the lowest was noted in Putumayo (1.489). *Ho* showed an average of 0.551 ± 0.06; Amazonas presented the highest value, with 0.644, while Nanay had the lowest, with 0.468. The *He* was, in general, higher, averaging 0.757 ± 0.059, with Napo recording the maximum value of 0.817 and Tigre the minimum of 0.672. The F_{IS} averaged 0.276 ± 0.077, with values ranging from 0.160 in Putumayo to 0.371 in Itaya, showing a trend toward deficiency of heterozygotes in all populations. These results taken together show considerable genetic diversity within the *M. dubia* germplasm bank in the Peruvian Amazon, evidencing remarkable variation among river basins.

An evaluation of the exclusive genetic characteristics across the eight river basins revealed substantial differences in the prevalence of unique genetic variations within the germplasm bank of *M. dubia*. A total of 237 private alleles were identified across the river basins (Figure S2). Although it had the fewest number of individuals (19), the content of the Putumayo River basin was very high, comprising 91 private alleles. Not far behind was the Tigre River basin, with 70, which indicated a significant pool of diversity. The third

position was occupied by the Nanay River basin, with 31 private alleles, followed by the Curaray and Amazonas River basins, with 17 and 15 unique alleles, respectively. In sharp contrast, only 13 private alleles were found in the Itaya, Napo, and Ucayali River basins, indicating a general lack of regional genetic uniqueness among those populations.

Table 1. Assessment of genetic diversity parameters at the river basin level for *M. dubia* germplasm bank from the Peruvian Amazon using sixteen SSR markers.

River Basin	*N*	Genetic Diversity Parameters					
		Na	***Ne***	***I***	***Ho***	***He***	F_{IS}
Amazonas	50	11.375	6.154	1.939	0.644	0.797	0.191
Curaray	32	11.063	6.294	1.992	0.556	0.809	0.313
Itaya	53	10.625	5.847	1.910	0.507	0.803	0.371
Nanay	91	12.438	5.247	1.779	0.468	0.717	0.356
Napo	37	11.313	6.303	2.006	0.638	0.817	0.225
Putumayo	19	6.813	4.219	1.489	0.580	0.684	0.160
Tigre	24	8.375	3.834	1.520	0.473	0.672	0.306
Ucayali	30	9.438	5.262	1.765	0.546	0.754	0.284
Average/Total	336	10.180	5.395	1.800	0.551	0.757	0.276
Standard Deviation	-	1.849	0.947	0.203	0.068	0.059	0.077

Note: *N*, number of genotyped individuals, *Na*, number of alleles; *Ne*, number of effective alleles; *I*, Shannon's diversity index; *Ho*, observed heterozygosity; *He*, expected heterozygosity; F_{IS}, fixation index.

3.2. Analysis of Molecular Variance (AMOVA)

AMOVA showed very useful information for the distribution of genetic variation within the *M. dubia* germplasm bank (Table 2). At the accession level, it was observed that there was a greater degree of genetic variation within accessions than among accessions. In particular, 73% of the total genetic variance was attributed to within-accession variation, while 27% was attributed to among-accession variation. The estimated variance of within-accession components was high (13.821), as opposed to the estimated variance of among-accession components (5.200).

Table 2. Genetic variation among and within accessions and river basins of *M. dubia* germplasm bank from the Peruvian Amazon based on sixteen SSR markers and determined with AMOVA.

Source	*df*	SS	MS	Est. Var.	%
At Accession Level					
Among accessions	42	2284.679	54.397	5.200	27
Within accessions	293	4049.461	13.821	13.821	73
Total	335	6334.140		19.020	100
At River Basin Level					
Among river basins	7	904.018	129.145	2.785	14
Within river basins	328	5430.121	16.555	16.555	86
Total	335	6334.140		19.341	100

Abbreviations: *df*, degrees of freedom; SS, sum of squares; MS, mean of the squares; Est. Var., estimated variance of components; %, percentage of total variance contributed by each component.

Upon examining the genetic variance within the river basin, a more pronounced pattern was detected. Out of the total genetic variation, 86% was found within river basins, while just 14% occurred among different river basins. These findings suggest that *M. dubia* populations from various river basins in the Peruvian Amazon have significant genetic similarities. The estimated variation of components within river basins was much greater, at 16.555, compared to the variation among river basins, which was 2.785. This observation provides evidence for the hypothesis that local populations maintain a high level of genetic variation.

3.3. Analysis of Genetic Differentiation and Gene Flow among River Basins

Figure 3 shows the results for F_{ST} among the plants of the *M. dubia* germplasm bank derived from eight different river basins: Amazonas, Curaray, Itaya, Nanay, Napo, Putumayo, Tigre, and Ucayali. The pairwise F_{ST} values ranged from 0.018 to 0.166, thus showing different degrees of genetic differentiation among the river basins. The highest F_{ST} value was observed between Putumayo and Tigre, with an F_{ST} = 0.166. This would indicate very strong genetic differentiation between the two populations. Other high F_{ST} values were observed between Putumayo and Nanay, with an F_{ST} = 0.154; Putumayo and Ucayali, with an F_{ST} = 0.141; and Putumayo and Amazonas, with an F_{ST} = 0.130. In contrast, relatively low F_{ST} values were observed between Amazonas and the other river basins, ranging from 0.023 to 0.046. Likewise, for the Ucayali River basin, there were low F_{ST} values of 0.018 with Nanay, 0.023 with Amazonas, 0.037 with Itaya, and 0.042 with Napo, indicating closer genetic similarity among these populations. The overall F_{ST} values showed high genetic differentiation among the river basins; Putumayo and Tigre represent the largest differentiation in relation to other river basins.

The gene flow (*Nm*), measured as the number of migrants per generation between individuals of the same river basins, was estimated (Figure 3). The higher the value of *Nm*, the more intense the gene flow between populations, opposing genetic differentiation. The values of *Nm* ranged from 1.26 to 13.62 in this study, indicating a moderate gene flow among river basins. Itaya and Ucayali had the highest gene flow, with *Nm* = 13.62, indicating that there had been a high degree of genetic exchange between these two populations. Other high values of *Nm* were between Amazonas and Ucayali, with an *Nm* of 10.71; between Nanay and Amazonas, with an *Nm* of 9.47; and between Itaya and Amazonas, with an *Nm* of 8.97. In contrast, the river basins of Putumayo presented low values of *Nm* against all others. The lowest value was noticed between Putumayo and Tigre, where *Nm* = 1.26, pointing to a very low rate of gene flow between these populations. In the same way, low *Nm* values were recorded between Putumayo and Nanay, Putumayo and Ucayali, and Putumayo and Amazonas, with values of 1.37, 1.53, and 1.67, respectively. In general, the *Nm* values indicated that, in fact, there is fairly high gene flow between several river basins, mostly between Amazonas, Itaya, Nanay, and Ucayali, responsible for keeping genetic connectivity. However, low gene flow detected among Putumayo and with other river basins, particularly Tigre, indicated restricted genetic exchange that adds up to the genetic differentiation observed.

When the relationship between F_{ST} and *Nm* among the river basins was analyzed, an inverse relationship was observed, indicating that as gene flow increased, genetic differentiation decreased (Figure S1). The equation describing this relationship had a very high coefficient of determination (R^2 = 0.9995), suggesting an excellent fit of the model to the data. This suggests that the transfer of genes between various plant genotypes is crucial in minimizing the genetic variation among the river basins of *M. dubia* accessions from the Peruvian Amazon.

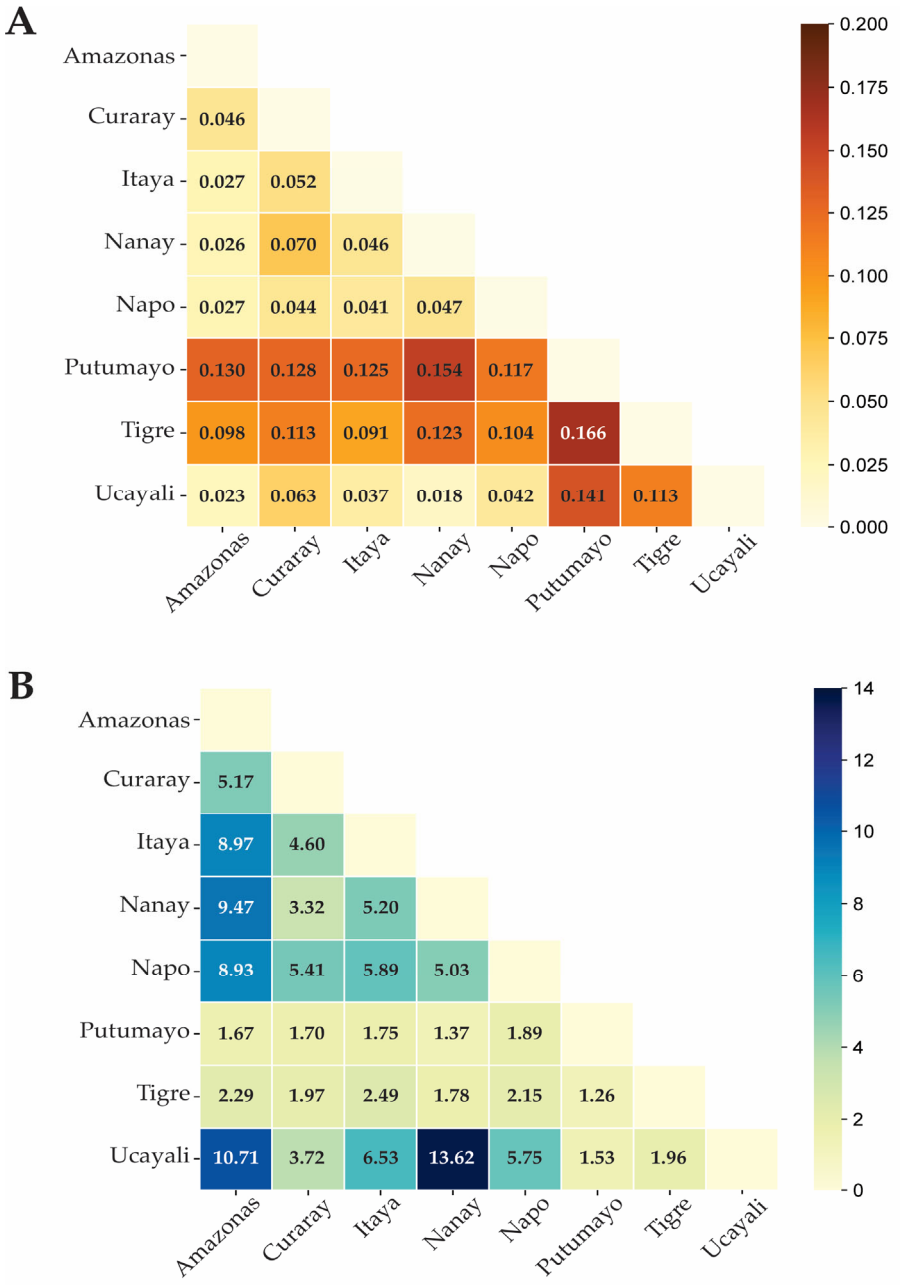

Figure 3. Pairwise comparisons of F_{ST} and *Nm* among *M. dubia* plants from the germplasm bank derived from eight river basins. (**A**) Heatmap of pairwise F_{ST} values among river basins. Higher F_{ST} values indicate greater genetic differentiation among populations. (**B**) Heatmap of pairwise *Nm* values among *M. dubia* plants from the germplasm bank derived from eight river basins, with higher *Nm* values indicating greater gene flow among populations.

3.4. Principal Component Discriminant Analysis (DAPC)

DAPC was conducted based on eight river basins, with 200 principal components representing 98% of the total variation. The scatterplots of DAPC are shown in Figure 4 and delimit genetic relationships among plant genotypes from different river basins.

The results indicated that the plant genotypes from the Putumayo, Curaray, and Tigre River basins could be separated from the rest to fall into two distinct main groups. Figure 3A clearly allows the separation to be seen in this scatter plot, underlining the different genetic profiles of these river basins. Even if the other remaining plant genotypes from the other river basins do not have complete differentiation, this already hinted at

substantial genetic connectivity among the other river basins, particularly the Itaya, Nanay, Amazonas, Napo, and Ucayali.

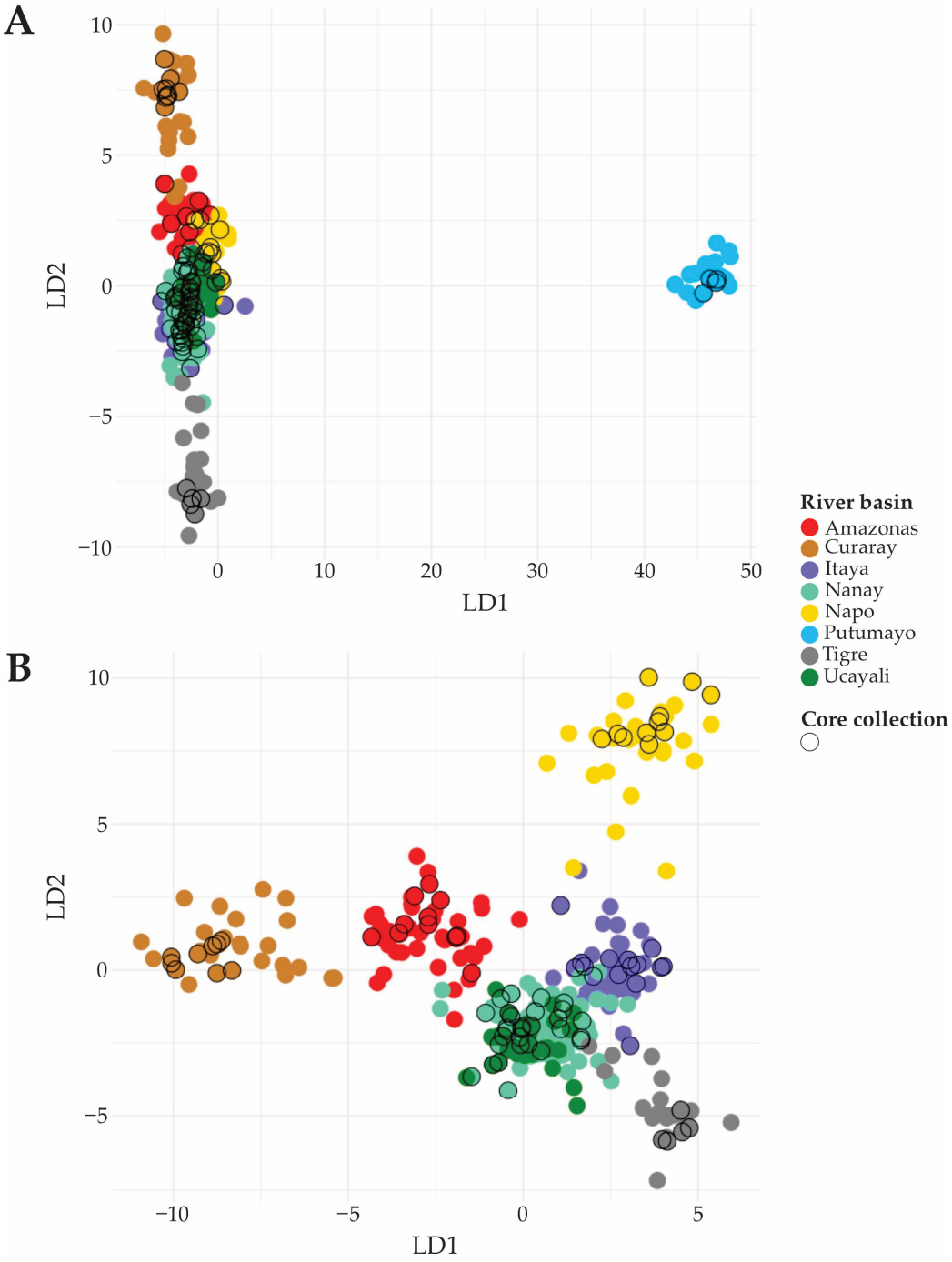

Figure 4. DAPC for plant genotypes of *M. dubia* from eight river basins of the germplasm bank, based on sixteen SSR markers. (**A**) DAPC scatterplot showing genetic differentiation among all eight river basins, highlighting the distinct separation of Putumayo, Curaray, and Tigre River basins. (**B**) DAPC scatterplot excluding the Putumayo River basin genotypes, revealing finer genetic structure among the remaining seven river basins.

To attain a greater level of resolution among plant genotypes belonging to the remaining river basins, we conducted an additional DAPC with the same parameters, excluding individuals from the Putumayo River basin (Figure 4B). This analysis revealed more sepa-

ration of the Tigre, Napo, and Curaray River basins. Although the Itaya, Nanay, Amazonas, and Ucayali River basins do not fully differentiate from one another, they do exhibit some distinctiveness, suggesting subtle genetic distinctions despite the overall genetic connectivity.

3.5. Bayesian Cluster Analysis

Bayesian cluster analysis revealed unique genetic structures and admixture patterns of the *M. dubia* accessions across the river basins (Figure 5). There was evidence that K = 10 was the optimal number of genetic clusters since, at this point, the Bayesian Information Criterion (BIC) had the lowest value; this is indicated by the red dot in Figure 5A. From K = 1 to K = 10, the BIC values decreased sharply before increasing again for higher K values. The pattern suggests that, up to ten clusters, model fit improves appreciably, after which additional clusters provide diminishing returns or even reduce model explanatory power.

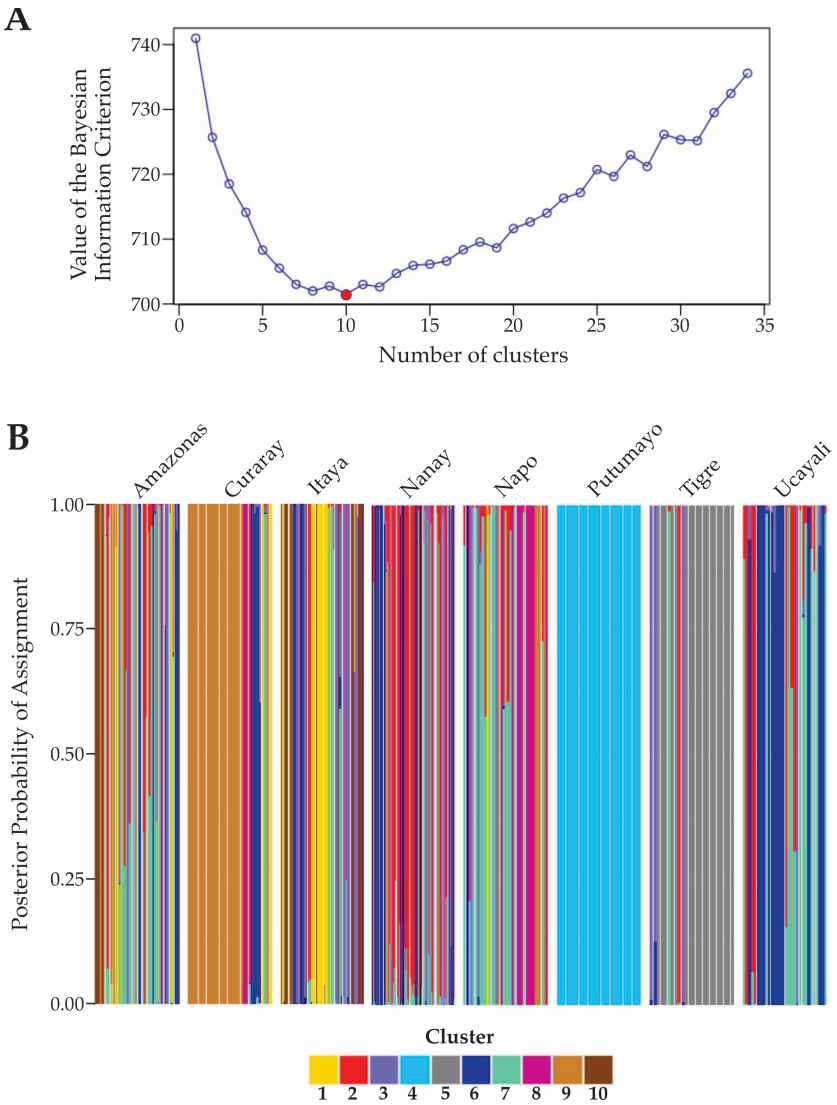

Figure 5. Genetic structure analysis of the *M. dubia* germplasm bank at the river basin level. (**A**) A graph showing the optimal number of genetic clusters (red dot) based on the Bayesian Information Criterion, (**B**) A bar plot representing the probability of assignment of individuals to different genetic clusters across the eight river basins. Each vertical bar represents an individual plant genotype from the respective river basin, while colors indicate the proportion of their genome from the different genetic clusters.

Figure 5B depicts the results, in bar plots, of the posterior probability of the assignment, for each river basin, of the individual plant genotypes to the ten genetic clusters identified. The results exhibited a complex genetic structure, with different degrees of admixture of the analyzed *M. dubia* germplasm accessions from eight river basins. Germplasm accessions from the Amazonas, Itaya, Nanay, Napo, and Ucayali River basins showed some degree of admixture; that is, plant genotypes were assigned to multiple genetic clusters. Those germplasm accessions derived from the Tigre and Curaray River basins displayed moderate grades of admixture, mostly being assigned to clusters 5 (gray) and 9 (light brown), respectively. In sharp contrast to this, germplasm accessions from the Putumayo River basin did not show any admixture, in which all of the plant genotypes are classified in genetic cluster 5 (light blue). These findings give evidence regarding the complex genetic structure and variable levels of admixture present in the eight river basins and their corresponding germplasm accessions of *M. dubia*.

3.6. Phylogenetic Relationship Among River Basins

Based on Nei's genetic distances, the neighbor-joining tree provided insights into the genetic relationships and divergence among germplasm accessions from the eight river basins in the study area, as depicted in Figure 6. Tree topology revealed two main genetic clusters that would suggest the existence of distinct evolutionary lineages. One large cluster comprised the Curaray, Napo, Amazonas, Ucayali, Nanay, and Itaya River basins, thus indicating their genetical proximity and therefore higher historical gene flow rates among these river basins. Their close genetic relationships among these six river basins agree with the high admixture levels observed in these populations. Larger genetic distances of the second major cluster, including that of the Putumayo and Tigre River basins from all others, suggest their independent genetic background and reduced admixture. The results underline the complex genetic structure and heterogeneous genetic connectivity among plant genotypes of *M. dubia* in the different river basins.

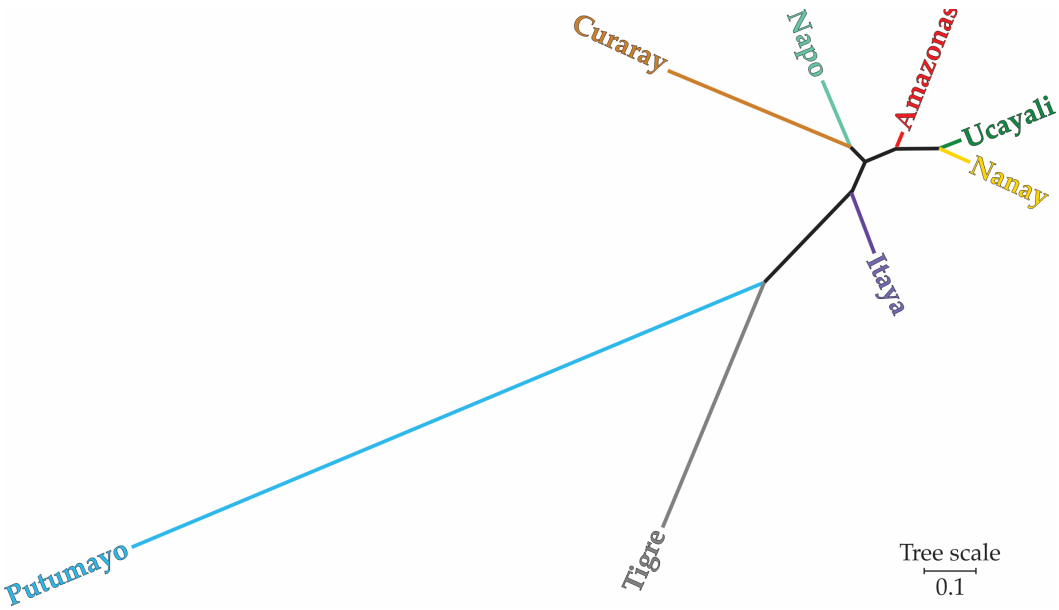

Figure 6. Neighbor-joining phylogenetic tree depicting the genetic relationships among *M. dubia* germplasm accessions from eight river basins. The tree was constructed based on Nei's genetic distances and was generated with 1000 bootstrap replicates.

3.7. Construction of the Core Collection

The core collection was constructed using a thorough and precise approach through the use of multipurpose core subset selection tools. These tools utilized local search algorithms

when generating subsets based on one or more distance and allelic richness metrics derived from the information given by the sixteen polymorphic SSR markers. This strategy ensured that the selected core collection would represent the entire germplasm bank and would be optimum in genetic diversity and allelic richness for the conservation of the general genetic structure of *M. dubia*. In the end, a total of 84 representative plant genotypes (25% of the full collection) were selected to fully cover genetic variation. The core collection was made up of genotypes from the eight river basins: Amazonas (13.10%), Curaray (10.71%), Itaya (19.05%), Nanay (25.00%), Napo (13.10%), Putumayo (4.76%), Tigre (5.95%), and Ucayali (8.33%). This distribution mirrors the proportional representation of each river basin in the core collection (Figure 7, Table S4).

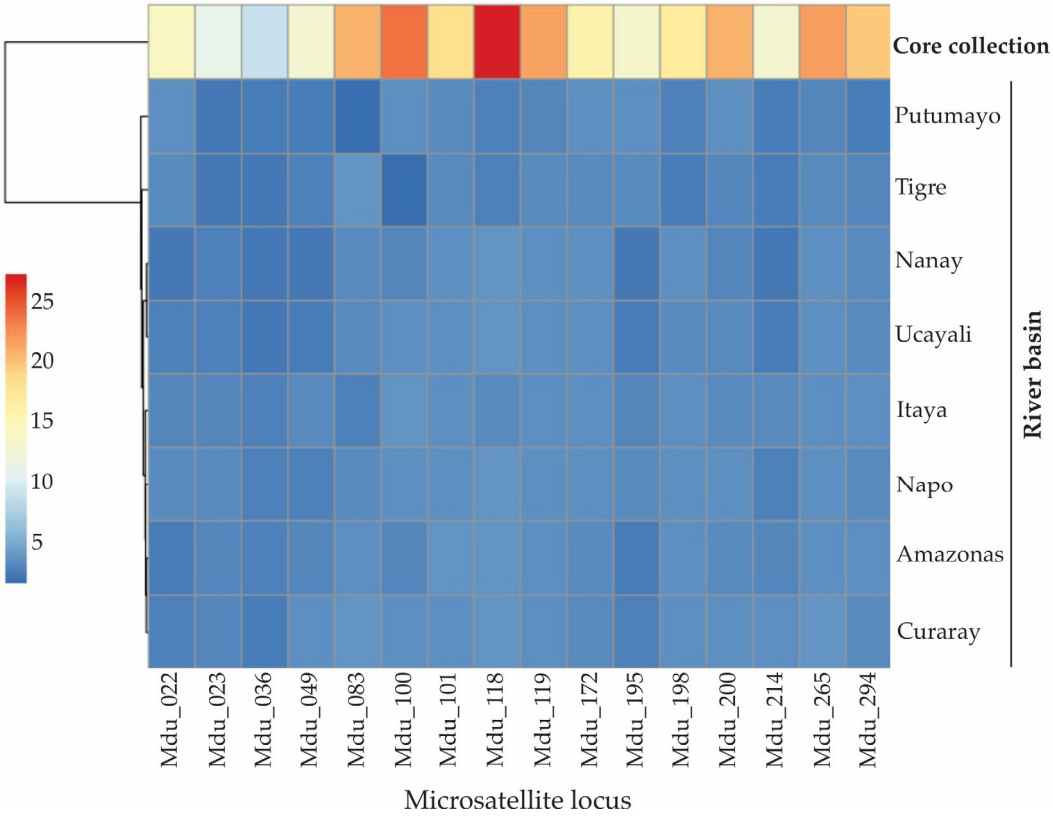

Figure 7. Heatmap depicting the allelic richness of a core collection from the *M. dubia* germplasm bank in the Peruvian Amazon based on sixteen polymorphic SSR markers.

The allele coverage in the core collection was high, at 90.82%, which in itself is indicative of substantial allelic richness. This allelic richness for the whole sample set and just the core collection is depicted graphically by the heatmap in Figure 7. A high value for allele coverage showed that this core collection had managed to pick up most of the genetic diversity in the larger germplasm bank.

4. Discussion

4.1. Genetic Diversity Parameters

Genetic diversity parameters estimated in the present study for the different river basins strongly support the need for comprehensive sampling of germplasm to ensure the conservation of *M. dubia* throughout the Amazonian region. The moderate-to-high genetic diversity found in *M. dubia* (Table 1) is consistent with its wide geographic distribution and outcrossing behavior. However, the presence of positive fixation indices and deficiencies of heterozygotes in several river basins may indicate inbreeding within the populations (Table 1).

These results are consistent with previous studies on germplasm collections of *M. dubia*. Studies on accessions from the Brazilian [5] and Peruvian [6] Amazon exhibited a similar pattern of heterozygote deficits and high fixation index values. The coherence observed across these studies suggests a common phenomenon of inbreeding in the populations of *M. dubia*. This may be a consequence of genetic drift, restricted gene flow, or population fragmentation. Some of these factors could generate local adaptation, similar to what has already been shown in other species, like *Eugenia dysenterica* DC [62].

The significant number of private alleles, particularly in accessions from the Putumayo, Tigre, and Nanay River basins (Figure S2), highlights the unique genetic diversity within these populations, emphasizing the importance of continued maintenance to maximize allelic richness of germplasm across diverse geographic sources. Such river basins might represent major reservoirs of distinct genetic diversity for *M. dubia* and hence may have to be targeted by focused conservation efforts to safeguard the adaptive potential of this species. Similar patterns have already been reported for other Amazonian fruit trees, such as *Theobroma grandiflorum* (Willd. ex Spreng.) K. Schum [63].

These results imply that targeted conservation strategies will be needed to preserve genetic diversity and minimize inbreeding in *M. dubia* populations across the Amazon basin. Preservation of the dissimilar genetic material found in particular river basins will be very instrumental to conserve the adaptive potential of this key species for future research and breeding programs.

4.2. Analysis of Molecular Variance (AMOVA)

AMOVA for the *M. dubia* germplasm bank identified important patterns of genetic diversity: more genetic variation within accessions than among accessions, and even more within river basins than among river basins (Table 2). This implies a reasonable maintenance of the genetic diversity within the local populations, suggesting that the *M. dubia* populations from different Peruvian Amazon basins share a substantial portion of their genes. These results are similar to previous reports on the genetic diversity of *M. dubia* from the Brazilian Amazon region, mostly from the active INPA germplasm bank [5]. Similar patterns have been described for other tropical tree species, such as *Bertholletia excelsa* Humb. & Bonpl. [64], *Theobroma cacao* L. [65], and *Euterpe edulis* Martius [66], in which high within-population diversity is attributed to outcrossing mating systems together with efficient mechanisms of gene flow. It could also be maintained by long-distance gene flow, possibly via seed dispersal through water or animal vectors, as had been reported for other Myrtaceae species [67].

Therefore, the genetic variation of *M. dubia* has some important implications for the appropriate design of conservation strategies. The high within-accession and within-river basin diversity suggests that collecting strategies focused on sampling multiple individuals at fewer locations might capture a large proportion of the species' genetic diversity. In fact, this strategy is recommended for other plant species of comparable genetic structure [16,20,68]. Nevertheless, a non-negligible amount of variation among accessions (27%) and river basins (14%) exists, indicating that sampling from several accessions and river basins is necessary for capturing all the existing genetic diversity. These results highlight the necessity for broad geographic representation in germplasm collections, as suggested by studies on other Amazonian fruit species [69].

These results enable the inference of genetic structure by a combination of historical and contemporary processes of population dynamics, gene flow, and local adaptation. Further studies on these factors, probably using landscape genetics methods [70], may provide insight into the evolutionary history and contemporary dynamics of *M. dubia* populations in the Peruvian Amazon.

Finally, several breeding program implications have been drawn from these results. A high genetic variation in germplasm banks is essential in breeding strategies to maximize genetic diversity in many studies across different plant species, including *Arabidopsis thaliana* (L.) Heynh. [71], *Melia dubia* Cav. [72], and other plant species [73]. This probably

induces heterosis effects [74]; thus, all of these studies underline the importance of genetic diversity within and among populations. Greater genetic variability due to the use of diverse germplasm will carry better traits for hybridization with maximum heterosis effects. It is by exploiting the existing genetic diversity through the inclusion of genetic material from different populations that breeders can develop superior plant varieties that possess desired traits, which is evidenced by various promising tropical fruit tree breeding programs in Latin America [75–77].

4.3. Analysis of Genetic Differentiation and Gene Flow Among River Basins

Pairwise F_{ST} values indicate varying degrees of genetic differentiation among the river basins (Figure 3). The strongest differentiation detected between Putumayo and the rest of the basins may indicate that this population has been relatively isolated or has undergone different selective pressures. Similar patterns of population differentiation have been reported in other Amazonian tree species, such as *Swietenia macrophylla* King, where genetic differentiation values were positively and significantly correlated to geographical distance under the isolation-by-distance model [78]. Higher F_{ST} values were reported for non-woody tropical plants that have mixed-mating systems and are pollinated by small insects [79]. Of the ecological factors tested, the latitudinal region explained the largest portion of variance, followed by the mode of pollination, mating system, and growth form, while the mode of seed dispersal did not significantly relate to genetic differentiation [80].

The gene flow patterns among river basins revealed a complex landscape of genetic connectivity and isolation. The high levels of gene flow observed between the Ucayali and the Nanay and Amazonas River basins, and between the Amazonas and the Napo, Nanay and Itaya River basins (Figure 3), suggest extensive historical gene exchange. This could be facilitated by interconnected river networks, animal- or human-mediated dispersal, and other ecological factors [79]. In contrast, the low-to-moderate gene flow estimates between some river basin pairs, such as Putumayo-Tigre and Putumayo-Nanay, may be explained by geographic barriers or poor dispersal mechanisms that could have caused stronger genetic isolation.

The inverse relationship between F_{ST} and Nm is a common pattern in plant population genetics, supporting the view that as gene flow increases, so genetic differentiation between populations decreases [80–82]. In the case of *M. dubia*, the high coefficient of determination ($R^2 = 0.9995$) emphasizes the fact that gene flow plays an essential role in the formation of genetic structure among these populations. Such a high correlation suggests that gene flow is homogenizing populations, reducing genetic differentiation by allowing the exchange of alleles across populations. Consequently, maintaining landscape connectivity becomes paramount for the long-term genetic health and viability of *M. dubia*. Connectivity enables gene flow, which preserves genetic diversity, enhances adaptive potential, and minimizes risks related to inbreeding and genetic drift. The result concurs with wider conservation principles in emphasizing that fragmented habitats and disrupted dispersal pathways mean populations become isolated, with reduced genetic variability and hence reduced resilience. This implies that maintaining habitats and ecological pathways is a prerequisite for sustaining stable and genetically diverse populations of *M. dubia* and related species.

4.4. Principal Component Discriminant Analysis (DAPC)

The DAPC results also revealed a complex genetic structure complexity in the *M. dubia* germplasm bank (Figure 4). In fact, the major clusters from the Putumayo, Curaray, and Tigre River basins were clearly separated from the remaining populations, indicating that they have particular genetic profiles that could be a consequence of isolation or adaptation to special environmental conditions. Geographically isolated populations of other Amazonian tree species such as *Bertholletia excelsa* Humb. & Bonpl. [83], evidenced similar genetic distinctiveness.

The weak structuring among the other basins (Itaya, Nanay, Amazonas, Napo, and Ucayali) is a signal of relatively high genetic connectivity among these populations. For

this species, the pattern of admixture observed suggests that mechanisms offering gene flow, like winds, water dispersion, or even animal-mediated pollination and seed dispersal, would be effective over long-distance genetic material exchange. This is therefore one of the most important factors for the maintenance of genetic diversity and the attainment of maximum adaptive potential in populations faced with environmental changes. Similar patterns of admixture were described for other broadly distributed Amazonian trees, such as *Ceiba pentandra* (L.) Gaertn, where high levels of gene flow were attributed to its broad distribution and effective pollination mechanisms [84]. This kind of dynamics has to be understood in conservation efforts because it underlines the need for the conservation of natural habitats and processes that allow gene flow. On the other hand, these guarantee a long-term viability of genetic diversity in such populations.

4.5. Bayesian Cluster Analysis

The Bayesian cluster analysis returning ten genetic clusters underlines the complex genetic structure of *M. dubia* within the studied river basins (Figure 5). The high admixture levels recorded in most populations, especially in Amazonas, Itaya, Nanay, Napo, and Ucayali, suggest historical and contemporary gene flow between these river basins. This pattern of admixture is similar to what has been described for other broadly distributed tree species in the Amazon, such as *Jacaranda copaia* (Aubl.) D. Don, where gene flow maintains genetic connectivity across large geographic areas [85]. Also, Zhang et al. [65], applying the Bayesian clustering method to a collection of *Theobroma cacao*, found an appreciable genetic structure conditioned by the river systems in the Peruvian Amazon.

That genetic homogeneity, observed in this particular case, in the Putumayo River basin, is striking and possibly unique in its mechanisms. On the one hand, this homogeneity would represent a founder effect where a new population was established by a few individuals with limited genetic variation. Otherwise, it could also be the result of a recent population bottleneck—some environmental event or human activity reduced population size and lost genetic diversity. It is in this region of Putumayo that strong adaptation to local conditions can explain the genetic uniformity observed. Distinct genetic clusters, which serve as relevant conservation units, have already been described for other Amazonian species. For example, in studies on *Theobroma cacao*, these genetically distinct populations have proved to possess some very useful traits valuable for conservation and breeding purposes [86]. Nevertheless, their preservation depends on their understanding and preservation of the integrity of the general genetic diversity and resilience of species in the Amazon basin.

4.6. Phylogenetic Relationship Among River Basins

The neighbor-joining tree, based on Nei's genetic distances, contains two principal genetic clusters that give clues about the evolutionary history of *M. dubia* from the Peruvian Amazon. A high degree of relatedness to genes among the Curaray, Napo, Amazonas, Ucayali, Nanay, and Itaya River basins suggests a common evolutionary history that may have been driven by historical and present-day gene flow. This may indicate similar selective pressures and continuing gene flow between these populations, which helps in the retention of overall genetic diversity, and hence, adaptability, of these river basins.

In sharp contrast, the clustering of the Putumayo and Tigre River basins suggests an independent evolutionary history. This could have resulted from geographic barriers, environmental gradients, or historic isolation, reducing gene flow and enhancing genetic differentiation. These results imply that the Putumayo and Tigre River basins are rich in privately held alleles and adaptations; therefore, they become very important targets for conservation. Such phylogenetic analyses will be important for unraveling the evolutionary history and the genetic structure of *M. dubia* in formulating effective conservation strategies. This ensures that both the genetic diversity within the interconnected populations and the unique genetic resources found in isolated populations are conserved.

The pattern of genetic relationships among populations from eight river basins in *M. dubia* aligns closely with other phylogeographic studies of Amazonian tree species. For example, in *Theobroma cacao*, the river systems clearly show a strong imprint on their current genetic structure [86]. The observed genetic structure in *M. dubia* likely results from an interplay of historical processes, including past climatic changes and geological events, combined with more recent anthropogenic influences.

4.7. Construction of the Core Collection

This is an important dimension for the conservation and utilization of *M. dubia* genetic resources: a core collection was built, representing 90.82% of the allelic diversity present in the entire germplasm bank (Figure 7, Table S4). At such a high level of allele capturing, the core collection represents the genetic diversity of the species, whereas the number of accessions to be maintained is greatly reduced. It offers an effective method of management and conservation and helps to ensure that important genetic variations are conserved with minimal input.

The methodologies that had been adapted for the construction of core collection using multipurpose core subset selection tools and local search algorithms had prior antecedents already applied with success in other crop species such as *Cunninghamia lanceolata* (Lamb.) Hook [87], *Ginkgo biloba* L. [88], *Larix decidua* Mill [89], *Phoebe zhennan* S. Lee [90], and *Pinus koraiensis* Siebold & Zucc. [91]. Such strategies maximize the capture of genetic diversity as part of the core collection, enhancing its representativeness and utility. The genotypes were distributed proportionally to the source river basin origin; the *M. dubia* core collection was, therefore, suitable for further breeding programs and genetic studies, allowing targeted research and development efforts to work on the discovery of beneficial traits and improvement of cultivars.

The establishment of such a core collection is therefore of great significance for an underutilized species like *M. dubia*, as it offers an avenue for efficient characterization, evaluation, and utilization of genetic resources. Core collections have contributed to the genetic enhancement of tropical crops including cassava, strawberry, and cacao. Optimization procedures based on genetic distances and phenotypic and genotypic data have kept significant genetic diversity in cassava core collections, and, as such, are useful for breeding programs and genomic investigations [92]. In strawberries, the integration of genomic and pedigree information in a core collection captured the maximum genetic variation in a small subset of genotypes for its future-proofing and facilitating haplotype reference panel development to be used for genotyping [93]. Finally, in cacao, the construction of the core collection was performed using 15 microsatellite loci by researchers who performed it with five different sampling algorithms [94]. As these data showed, the role of core collections is actually important in the genetic improvement of tropical plants. The well-curated and diversified genetic foundation of the *M. dubia* core collection will help in the selection of superior genotypes useful for cultivation and study, in species conservation and development.

5. Conclusions

The new microsatellite markers developed in this study enabled the quantification of genetic parameters, revealing high diversity among the accessions and river basins. Our analysis demonstrated the need for conserving diversity at several scales due to the significant genetic differentiation among accessions and river basins.

The complicated patterns of gene flow and genetic structure among river basins reflect the interplay of geographic, dispersal, and evolutionary processes in the formation of genetic diversity of *M. dubia* in the Amazonian landscape. Basins with a higher level of admixture and connectivity, like Amazonas, Curaray, Nanay, and Ucayali, most likely underwent historical gene exchange events that should be driven by the interconnection of river systems, and animal- or human-mediated dispersals. Other basins, such as Putumayo

and Tigre, appear to have been relatively more genetically isolated, perhaps because of geographical barriers or reduced effective levels of dispersal.

In the constructed core collection, the robust coverage of alleles and proportional basin representation will ensure that the total genetic diversity present in this larger germplasm bank is preserved. This approach to core collection development, with its emphasis on maximum allelic richness, provides a valuable resource for future germplasm management, conservation, and breeding efforts for *M. dubia*.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f15111873/s1, Figure S1: Relationship between genetic differentiation (F_{ST}) and gene flow (*Nm*) among plant genotypes belonging to the eight river basins in the *M. dubia* germplasm bank from the Peruvian Amazon; Figure S2: Distribution of private alleles across plant genotypes belonging to the eight river basins in the *M. dubia* germplasm bank from the Peruvian Amazon; Table S1: Comprehensive dataset of *M. dubia* germplasm bank from the Peruvian Amazon, including codes, river basin data, collection site, geographical information, and profiles of the sixteen microsatellite loci; Table S2: Information of the sixteen polymorphic microsatellite loci used for assessment the genetic diversity parameters and construction of the core collection of *M. dubia* germplasm bank from the Peruvian Amazon; Table S3: Assessment of genetic diversity parameters at the accession level for *M. dubia* germplasm bank from the Peruvian Amazon using sixteen microsatellite loci; Table S4: A core collection of *M. dubia* germplasm bank from the Peruvian Amazon obtained based on the sixteen polymorphic microsatellite loci.

Author Contributions: Conceptualization, J.C.C. and J.D.M.; methodology, J.D.P., C.G.C., D.E.M., S.A.I. and P.M.A.; validation, S.A.I., C.G.C. and P.M.A.; formal analysis, S.J.V.-G., B.E.V., F.A. and N.R.-V.; resources, J.C.C. and J.D.M.; writing—original draft preparation, S.J.V.-G. and J.C.C.; writing—review and editing, J.C.C. and J.D.M.; supervision, J.C.C.; project administration, M.C.; funding acquisition, J.C.C. and J.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: Funding was provided by the Universidad Nacional de la Amazonía Peruana (grant numbers: R.R. 0686-2015-UNAP and R.R. 0449-2024-UNAP), the American Public University System, the Grainger Bioinformatics Center, and the Pritzker Laboratory for Molecular Systematics and Evolution operated with support from the Pritzker Foundation.

Data Availability Statement: The data are included in the article or Supplemental Materials.

Acknowledgments: We thank K. Feldheim for assistance in the Pritzker Laboratory for Molecular Systematics and Evolution.

Conflicts of Interest: The authors declare no conflicts of interest. Bianca E. Vigil is the owner of FORESLAB S.A.C. Her company was not involved in this study, and there is no relevance between this research and her company.

References

1. Villachica, H. *El Cultivo Del Camu Camu (Myrciaria dubia H.B.K. McVaugh) En La Amazoni'a Peruana*, 1st ed.; Pro-Tempore, S., Ed.; Tratado de Cooperación Amazónica: Lima, Peru, 1996.
2. Uchiyama, H.; Koyama, T. Chromosomes of *Myrciaria dubia*, Myrtaceae. *Chromosome Inf. Serv.* **1993**, *54*, 16–17.
3. Maués, M.M.; Couturier, G. Biologia Floral e Fenologia Reprodutiva Do Camu-Camu (*Myrciaria dubia* (H.B.K.) McVaugh, Myrtaceae) No Estado Pará, Brasil. *Rev. Bras. Botânica* **2002**, *25*, 441–448. [CrossRef]
4. Peters, C.M.; Vasquez, A. Estudios ecológicos de Camu-Camu (*Myrciaria dubia*). I. Producción de frutos en poblaciones naturales. *Acta Amaz.* **1987**, *17*, 161–188. [CrossRef]
5. Rojas, S.; Clement, Y.; Nagao, E.O. Diversidade genética em acessos do banco de germoplasma de camu-camu (*Myrciaria dubia* [H.B.K.] McVaugh) do INPA usando marcadores microssatélites (EST-SSR). *Cienc. Tecnol. Agropecu.* **2011**, *12*, 51–64. [CrossRef]
6. Šmíd, J.; Kalousová, M.; Mandák, B.; Houška, J.; Chládová, A.; Pinedo, M.; Lojka, B. Morphological and Genetic Diversity of Camu-Camu [*Myrciaria dubia* (Kunth) McVaugh] in the Peruvian Amazon. *PLoS ONE* **2017**, *12*, e0179886. [CrossRef]
7. Bradfield, R.B.; Roca, A. Camu-Camu-a Fruit High in Ascorbic Acid. *J. Am. Diet. Assoc.* **1964**, *44*, 28–30. [CrossRef] [PubMed]
8. Imán Correa, S.; Bravo Zamudio, L.H.; Solís, V.S.; Cruz, C.O. Contenido de vitamina C en frutos de camu camu *Myrciaria dubia* (H.B.K) Mc Vaugh, en cuatro estados de maduración, procedentes de la Colección de Germoplasma del INIA Loreto, Perú. *Sci. Agropecu.* **2011**, *2*, 123–130. [CrossRef]

9. Akter, M.S.; Oh, S.; Eun, J.B.; Ahmed, M. Nutritional Compositions and Health Promoting Phytochemicals of Camu-Camu (*Myrciaria dubia*) Fruit: A Review. *Food Res. Int.* **2011**, *44*, 1728–1732. [CrossRef]
10. Akachi, T.; Shiina, Y.; Kawaguchi, T.; Kawagishi, H.; Morita, T.; Sugiyama, K. 1-Methylmalate from Camu-Camu (*Myrciaria dubia*) Suppressed D-Galactosamine-Induced Liver Injury in Rats. *Biosci. Biotechnol. Biochem.* **2010**, *74*, 573–578. [CrossRef]
11. Castro, J.C.; Maddox, J.D.; Cobos, M.; Imán, S.A. Myrciaria dubia "Camu Camu" Fruit: Health-Promoting Phytochemicals and Functional Genomic Characteristics. In *Breed. Health Benefits Fruit. Nut Crops*, 1st ed.; Soneji, S.R., Nageswara-Rao, M., Eds.; IntechOpen: London, UK, 2018; pp. 85–116. [CrossRef]
12. Azevedo, L.; de Araujo Ribeiro, P.F.; de Carvalho Oliveira, J.A.; Correia, M.G.; Ramos, F.M.; de Oliveira, E.B.; Barros, F.; Stringheta, P.C. 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.* **2019**, *99*, 624–631. [CrossRef]
13. Akinnifesi, F.K. *Indigenous Fruit Trees in the Tropics: Domestication, Utilization and Commercialization*, 1st ed.; CABI: Cambridge, MA, USA, 2008; ISBN 978-1-84593-110-0.
14. Penn, J. The Cultivation of Camu Camu (*Myrciaria dubia*): A Tree Planting Programme in the Peruvian Amazon. *For. Trees Livelihood* **2006**, *16*, 85–101. [CrossRef]
15. Khoury, C.; Laliberté, B.; Guarino, L. Trends in Ex Situ Conservation of Plant Genetic Resources: A Review of Global Crop and Regional Conservation Strategies. *Genet. Resour. Crop Evol.* **2010**, *57*, 625–639. [CrossRef]
16. Priyanka, V.; Kumar, R.; Dhaliwal, I.; Kaushik, P. Germplasm Conservation: Instrumental in Agricultural Biodiversity—A Review. *Sustainability* **2021**, *13*, 6743. [CrossRef]
17. Cohen, J.I.; Williams, J.T.; Plucknett, D.L.; Shands, H. Ex Situ Conservation of Plant Genetic Resources: Global Development and Environmental Concerns. *Science* **1991**, *253*, 866–872. [CrossRef]
18. Rajpurohit, D.; Jhang, T. In Situ and Ex Situ Conservation of Plant Genetic Resources and Traditional Knowledge. In *Plant Genetic Resources and Traditional Knowledge for Food Security*; Salgotra, R.K., Gupta, B.B., Eds.; Springer: Singapore, 2015; pp. 137–162. ISBN 978-981-10-0060-7.
19. Li, D.-Z.; Pritchard, H.W. The Science and Economics of Ex Situ Plant Conservation. *Trends Plant Sci.* **2009**, *14*, 614–621. [CrossRef]
20. Plucknett, D.L.; Smith, N.J.H.; Williams, J.T.; Anishetty, N.M. Crop Germplasm Conservation and Developing Countries. *Science* **1983**, *220*, 163–169. [CrossRef]
21. Griffith, M.P.; Cartwright, F.; Dosmann, M.; Fant, J.; Freid, E.; Havens, K.; Jestrow, B.; Kramer, A.T.; Magellan, T.M.; Meerow, A.W.; et al. Ex Situ Conservation of Large and Small Plant Populations Illustrates Limitations of Common Conservation Metrics. *Int. J. Plant Sci.* **2021**, *182*, 263–276. [CrossRef]
22. Volis, S.; Blecher, M. Quasi in Situ: A Bridge between Ex Situ and in Situ Conservation of Plants. *Biodivers. Conserv.* **2010**, *19*, 2441–2454. [CrossRef]
23. Wei, X.; Jiang, M. Meta-Analysis of Genetic Representativeness of Plant Populations under Ex Situ Conservation in Contrast to Wild Source Populations. *Conserv. Biol.* **2021**, *35*, 12–23. [CrossRef]
24. Engels, J.M.M.; Ebert, A.W. A Critical Review of the Current Global Ex Situ Conservation System for Plant Agrobiodiversity. I. History of the Development of the Global System in the Context of the Political/Legal Framework and Its Major Conservation Components. *Plants* **2021**, *10*, 1557. [CrossRef]
25. Ebert, A.W.; Engels, J.M.M. Plant Biodiversity and Genetic Resources Matter! *Plants* **2020**, *9*, 1706. [CrossRef] [PubMed]
26. Vieira, M.L.C.; Santini, L.; Diniz, A.L.; Munhoz, C.D.F. Microsatellite Markers: What They Mean and Why They Are So Useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [CrossRef] [PubMed]
27. Pérez, F.; Irarrázabal, C.C.; Cossio, M.; Peralta, G.; Segovia, R.; Bosshard, M.; Hinojosa, L.F. Microsatellite Markers for the Endangered Shrub *Myrceugenia rufa* (Myrtaceae) and Three Closely Related Species. *Conserv. Genet. Resour.* **2014**, *6*, 773–775. [CrossRef]
28. Miwa, M.; Tanaka, R.; Yamanoshita, T.; Norisada, M.; Kojima, K.; Hogetsu, T. Analysis of Clonal Structure of *Melaleuca cajuputi* (Myrtaceae) at a Barren Sandy Site in Thailand Using Microsatellite Polymorphism. *Trees* **2001**, *15*, 242–248. [CrossRef]
29. Albaladejo, R.G.; Sebastiani, F.; González-Martínez, S.C.; González-Varo, J.P.; Vendramin, G.G.; Aparicio, A. Isolation of Microsatellite Markers for the Common Mediterranean Shrub *Myrtus communis* (Myrtaceae). *Am. J. Bot.* **2010**, *97*, e23–e25. [CrossRef]
30. Steane, D.A.; Conod, N.; Jones, R.C.; Vaillancourt, R.E.; Potts, B.M. A Comparative Analysis of Population Structure of a Forest Tree, *Eucalyptus globulus* (Myrtaceae), Using Microsatellite Markers and Quantitative Traits. *Tree Genet. Genomes* **2006**, *2*, 30–38. [CrossRef]
31. Rossetto, M.; Slade, R.W.; Baverstock, P.R.; Henry, R.J.; Lee, L.S. Microsatellite Variation and Assessment of Genetic Structure in Tea Tree (*Melaleuca alternifolia*–Myrtaceae). *Mol. Ecol.* **1999**, *8*, 633–643. [CrossRef]
32. Doyle, J.J.; Doyle, J.L. A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phyt. Bull.* **1987**, *19*, 11–15.
33. Sambrook, J.; Frisch, E.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: New York, NY, USA, 1989; Volume 2.
34. Glenn, T.C.; Schable, N.A. Isolating Microsatellite DNA Loci. In *Methods in Enzymology*; Molecular Evolution: Producing the Biochemical Data; Academic Press: Cambridge, MA, USA, 2005; Volume 395, pp. 202–222.

35. Schuelke, M. An Economic Method for the Fluorescent Labeling of PCR Fragments. *Nat. Biotechnol.* **2000**, *18*, 233–234. [CrossRef]
36. Maddox, J.D.; Feldheim, K.A. A Cost-Effective Size Standard for Fragment Analysis That Maximizes Throughput on Five Dye Set Platforms. *Conserv. Genet. Resour.* **2014**, *6*, 5–7. [CrossRef]
37. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]
38. Peakall, R.; Smouse, P. GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—An Update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef] [PubMed]
39. Excoffier, L.; Smouse, P.; Quattro, J. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* **1992**, *131*, 479–491. [CrossRef] [PubMed]
40. Slatkin, M. Gene Flow in Natural Populations. *Annu. Rev. Ecol. Syst.* **1985**, *16*, 393–430. [CrossRef]
41. Reynolds, J.; Weir, B.S.; Cockerham, C.C. Estimation of the Coancestry Coefficient: Basis for a Short-Term Genetic Distance. *Genetics* **1983**, *105*, 767–779. [CrossRef] [PubMed]
42. Jombart, T. Adegenet: A R Package for the Multivariate Analysis of Genetic Markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef]
43. Adegenet: Exploratory Analysis of Genetic and Genomic Data. Available online: https://rdrr.io/cran/adegenet/ (accessed on 20 May 2024).
44. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
45. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **2003**, *164*, 1567–1587. [CrossRef]
46. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of Population Structure Using Multilocus Genotype Data: Dominant Markers and Null Alleles. *Mol. Ecol. Notes* **2007**, *7*, 574–578. [CrossRef]
47. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring Weak Population Structure with the Assistance of Sample Group Information. *Mol. Ecol. Resour.* **2009**, *9*, 1322–1332. [CrossRef]
48. Miller, J.M.; Cullingham, C.I.; Peery, R.M. The Influence of a Priori Grouping on Inference of Genetic Clusters: Simulation Study and Literature Review of the DAPC Method. *Heredity* **2020**, *125*, 269–280. [CrossRef] [PubMed]
49. Saitou, N.; Nei, M. The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef] [PubMed]
50. Nei, M. Genetic Distance between Populations. *Am. Nat.* **1972**, *106*, 283–292. [CrossRef]
51. Pembleton, L.W.; Cogan, N.O.I.; Forster, J.W. StAMPP: An R Package for Calculation of Genetic Differentiation and Structure of Mixed-Ploidy Level Populations. *Mol. Ecol. Resour.* **2013**, *13*, 946–952. [CrossRef] [PubMed]
52. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction. *PeerJ* **2014**, *2*, e281. [CrossRef]
53. Poppr: Genetic Analysis of Populations with Mixed Reproduction. Available online: https://rdrr.io/cran/poppr/ (accessed on 2 June 2024).
54. De Beukelaer, H.; Davenport, G.F.; Fack, V. Core Hunter 3: Flexible Core Subset Selection. *BMC Bioinform.* **2018**, *19*, 203. [CrossRef]
55. Wright, S. *Evolution and the Genetics of Populations: Variability Within and Among Natural Populations*; University of Chicago Press: Chicago, IL, USA, 1984; Volume 4, ISBN 978-0-226-91041-3.
56. Goudet, J. Hierfstat, a Package for r to Compute and Test Hierarchical F-Statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186. [CrossRef]
57. R: The R Project for Statistical Computing. Available online: https://www.r-project.org/ (accessed on 13 May 2024).
58. Wickham, H.; Navarro, D.; Lin, T. *Ggplot2: Elegant Graphics for Data Analysis*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2009; ISBN 0387981403.
59. Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Available online: https://ggplot2.tidyverse.org/ (accessed on 15 May 2024).
60. Pheatmap: Pretty Heatmaps. Available online: https://cran.r-project.org/web/packages/pheatmap/index.html (accessed on 18 May 2024).
61. Tang, D.; Chen, M.; Huang, X.; Zhang, G.; Zeng, L.; Zhang, G.; Wu, S.; Wang, Y. SRplot: A Free Online Platform for Data Visualization and Graphing. *PLoS ONE* **2023**, *18*, e0294236. [CrossRef]
62. de Campos Telles, M.P.; Coelho, A.S.G.; Chaves, L.J.; Diniz-Filho, J.A.F.; Valva, F.D. Genetic Diversity and Population Structure of *Eugenia dysenterica* DC. ("cagaiteira"—Myrtaceae) in Central Brazil: Spatial Analysis and Implications for Conservation and Management. *Conserv. Genet.* **2003**, *4*, 685–695. [CrossRef]
63. Alves, R.M.; Silva, C.R.D.S.; Silva, M.S.D.C.; Silva, D.C.D.S.; Sebbenn, A.M. Diversidade Genética Em Coleções Amazônicas de Germoplasma de Cupuaçuzeiro [*Theobroma grandiflorum* (Willd. Ex Spreng.) Schum.]. *Rev. Bras. Frutic.* **2013**, *35*, 818–828. [CrossRef]

64. Baldoni, A.B.; Ribeiro Teodoro, L.P.; Eduardo Teodoro, P.; Tonini, H.; Dessaune Tardin, F.; Alves Botin, A.; Hoogerheide, E.S.S.; de Carvalho Campos Botelho, S.; Lulu, J.; de Farias Neto, A.L.; et al. Genetic Diversity of Brazil Nut Tree (*Bertholletia excelsa* Bonpl.) in Southern Brazilian Amazon. *For. Ecol. Mang.* **2020**, *458*, 117795. [CrossRef]
65. Zhang, D.; Boccara, M.; Motilal, L.; Mischke, S.; Johnson, E.S.; Butler, D.R.; Bailey, B.; Meinhardt, L. Molecular Characterization of an Earliest Cacao (*Theobroma cacao* L.) Collection from Upper Amazon Using Microsatellite DNA Markers. *Tree Genet. Genomes* **2009**, *5*, 595–607. [CrossRef]
66. Gaiotto, F.A.; Grattapaglia, D.; Vencovsky, R. Genetic Structure, Mating System, and Long-Distance Gene Flow in Heart of Palm (*Euterpe edulis* Mart.). *J. Hered.* **2003**, *94*, 399–406. [CrossRef] [PubMed]
67. Sytsma, K.J.; Litt, A.; Zjhra, M.L.; Chris Pires, J.; Nepokroeff, M.; Conti, E.; Walker, J.; Wilson, P.G. Clades, Clocks, and Continents: Historical and Biogeographical Analysis of Myrtaceae, Vochysiaceae, and Relatives in the Southern Hemisphere. *Int. J. Plant Sci.* **2004**, *165*, S85–S105. [CrossRef]
68. Migicovsky, Z.; Warschefsky, E.; Klein, L.L.; Miller, A.J. Using Living Germplasm Collections to Characterize, Improve, and Conserve Woody Perennials. *Crop Sci.* **2019**, *59*, 2365–2380. [CrossRef]
69. Cardoso, R.; Ruas, C.F.; Giacomin, R.M.; Ruas, P.M.; Ruas, E.A.; Barbieri, R.L.; Rodrigues, R.; Gonçalves, L.S.A. Genetic Variability in Brazilian *Capsicum baccatum* Germplasm Collection Assessed by Morphological Fruit Traits and AFLP Markers. *PLoS ONE* **2018**, *13*, e0196468. [CrossRef]
70. Manel, S.; Schwartz, M.K.; Luikart, G.; Taberlet, P. Landscape Genetics: Combining Landscape Ecology and Population Genetics. *Trends Ecol. Evol.* **2003**, *18*, 189–197. [CrossRef]
71. Oakley, C.G.; Lundemo, S.; Ågren, J.; Schemske, D.W. Heterosis Is Common and Inbreeding Depression Absent in Natural Populations of *Arabidopsis thaliana*. *J. Evol. Biol.* **2019**, *32*, 592–603. [CrossRef] [PubMed]
72. Rawat, S.; Annapurna, D.; Arunkumar, A.N.; Geeta, J. Genetic Diversity and Population Structure in Fragmented Natural Populations of *Melia dubia* Cav. Revealed by SSR Markers—Its Implications on Conservation. *Plant Mol. Biol. Rep.* **2021**, *40*, 247–255. [CrossRef]
73. Yu, K.; Wang, H.; Liu, X.; Xu, C.; Li, Z.; Xu, X.; Liu, J.; Wang, Z.; Xu, Y. Large-Scale Analysis of Combining Ability and Heterosis for Development of Hybrid Maize Breeding Strategies Using Diverse Germplasm Resources. *Front. Plant Sci.* **2020**, *11*, 660. [CrossRef]
74. Mr, L.; Aj, S.; Je, R. Heterosis and Hybrid Crop Breeding: A Multidisciplinary Review. *Front. Genet.* **2021**, *12*, 643761. [CrossRef]
75. Ramírez, F. Breeding Programs. In *Latin American Blackberries Biology*; Springer: Cham, Switzerland, 2023; pp. 157–162.
76. León, N.; Murillo, O.; Badilla, Y.; Ávila, C.; Murillo, R. Expected Genetic Gain and Genotype by Environment Interaction in Almond (*Dipteryx panamensis* (Pittier) Rec. and Mell) in Costa Rica. *Silvae Genet.* **2017**, *66*, 9–17. [CrossRef]
77. Bost, J. *Persea schiedeana*: A High Oil "Cinderella Species" Fruit with Potential for Tropical Agroforestry Systems. *Sustainability* **2013**, *6*, 99–111. [CrossRef]
78. Lemes, M.R.; Gribel, R.; Proctor, J.; Grattapaglia, D. Population Genetic Structure of Mahogany (*Swietenia macrophylla* King, Meliaceae) across the Brazilian Amazon, Based on Variation at Microsatellite Loci: Implications for Conservation. *Mol. Ecol.* **2003**, *12*, 2875–2883. [CrossRef]
79. Gamba, D.; Muchhala, N. Global Patterns of Population Genetic Differentiation in Seed Plants. *Mol. Ecol.* **2020**, *29*, 3413–3428. [CrossRef]
80. Slatkin, M. Gene Flow and the Geographic Structure of Natural Populations. *Science* **1987**, *236*, 787–792. [CrossRef]
81. Dick, C.W.; Hardy, O.J.; Jones, F.A.; Petit, R.J. Spatial Scales of Pollen and Seed-Mediated Gene Flow in Tropical Rain Forest Trees. *Trop. Plant Biol.* **2008**, *1*, 20–33. [CrossRef]
82. Wright, S. Isolation by Distance. *Genetics* **1943**, *28*, 114–138. [CrossRef]
83. Sujii, P.S.; Martins, K.; Wadt, L.H.D.O.; Azevedo, V.C.R.; Solferini, V.N. Genetic Structure of *Bertholletia excelsa* Populations from the Amazon at Different Spatial Scales. *Conserv. Genet.* **2015**, *16*, 955–964. [CrossRef]
84. Dick, C.W.; Bermingham, E.; Lemes, M.R.; Gribel, R. Extreme Long-Distance Dispersal of the Lowland Tropical Rainforest Tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Mol. Ecol.* **2007**, *16*, 3039–3049. [CrossRef]
85. Scotti-Saintagne, C.; Dick, C.W.; Caron, H.; Vendramin, G.G.; Troispoux, V.; Sire, P.; Casalis, M.; Buonamici, A.; Valencia, R.; Lemes, M.R.; et al. Amazon Diversification and Cross-Andean Dispersal of the Widespread Neotropical Tree Species *Jacaranda copaia* (Bignoniaceae). *J. Biogeogr.* **2013**, *40*, 707–719. [CrossRef]
86. Motamayor, J.C.; Lachenaud, P.; Mota, J.W.; Loor, R.; Kuhn, D.N.; Brown, J.S.; Schnell, R.J. Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). *PLoS ONE* **2008**, *3*, e3311. [CrossRef] [PubMed]
87. Wu, H.; Duan, A.; Wang, X.; Chen, Z.; Zhang, X.; He, G.; Zhang, J. Construction of a Core Collection of Germplasms from Chinese Fir Seed Orchards. *Forests* **2023**, *14*, 305. [CrossRef]
88. Yao, Z.; Feng, Z.; Wu, C.; Tang, L.; Wu, X.; Chen, D.; Wang, Q.; Fan, K.; Wang, Y.; Li, M. Analysis of Genetic Diversity and Construction of a Core Collection of *Ginkgo Biloba* Germplasm Using EST-SSR Markers. *Forests* **2023**, *14*, 2155. [CrossRef]
89. Teodosiu, M.; Mihai, G.; Ciocîrlan, E.; Curtu, A.L. Genetic Characterisation and Core Collection Construction of European Larch (*Larix decidua* Mill.) from Seed Orchards in Romania. *Forests* **2023**, *14*, 1575. [CrossRef]
90. Zhu, Y.; An, W.; Peng, J.; Li, J.; Gu, Y.; Jiang, B.; Chen, L.; Zhu, P.; Yang, H. Genetic Diversity of Nanmu (*Phoebe zhennan* S. Lee. et F. N. Wei) Breeding Population and Extraction of Core Collection Using nSSR, cpSSR and Phenotypic Markers. *Forests* **2022**, *13*, 1320. [CrossRef]

91. Yan, P.; Zhang, L.; Hao, J.; Sun, G.; Hu, Z.; Wang, J.; Wang, R.; Li, Z.; Zhang, H. Construction of a Core Collection of Korean Pine (*Pinus koraiensis*) Clones Based on Morphological and Physiological Traits and Genetic Analysis. *Forests* **2024**, *15*, 534. [CrossRef]
92. Dos Santos, C.; de Andrade, L.; do Carmo, C.; de Oliveira, E. Development of Cassava Core Collections Based on Morphological and Agronomic Traits and SNPS Markers. *Front. Plant Sci.* **2023**, *14*, 1250205. [CrossRef]
93. Koorevaar, T.; Willemsen, J.H.; Visser, R.G.F.; Arens, P.; Maliepaard, C. Construction of a Strawberry Breeding Core Collection to Capture and Exploit Genetic Variation. *BMC Genom.* **2023**, *24*, 740. [CrossRef]
94. Bidot Martínez, I.; Valdés de la Cruz, M.; Riera Nelson, M.; Bertin, P. Establishment of a Core Collection of Traditional Cuban *Theobroma Cacao* Plants for Conservation and Utilization Purposes. *Plant Mol. Biol. Rep.* **2016**, *35*, 47–60. [CrossRef]