 doi:10.20944/preprints202106.0162.v1

Article

Genetic Diversity and Population Structure of Capirona (*Calycophyllum spruceanum* Benth.) from the Peruvian Amazon Assessed by RAPD Markers

Carla L. Saldaña [1], Johan D. Cancan [1], Wilbert Cruz [1], Mirian Y. Correa [2], Miriam Ramos [3], Eloy Cuellar [4], and Carlos I. Arbizu [1,*]

[1] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru; carla18317@gmail.com (C.L.S); clg.johan@gmail.com (J.D.C); wilbertcruz2002@gmail.com (W.C)

[2] Universidad Nacional Pedro Ruíz Gallo, Facultad de Ciencias Biológicas, Juan XXIII No 391, Lambayeque, Perú; miryacormel15@gmail.com.

[3] Universidad Nacional Agraria La Molina, Facultad de Ciencias Forestales, Av. La Molina s/n, Lima, Lima 15024, Peru; miriam.piaforestal@gmail.com.

[4] Universidad Nacional Agraria La Molina, Av. La Molina s/n, La Molina, Lima, Lima, 15024, Peru; eloycuellar@lamolina.edu.pe.

* Correspondence: carbizu@inia.gob.pe

Abstract: Capirona (*Calycophyllum spruceanum* Benth) is a tree species of commercial importance widely distributed in South American forests and is traditionally used for its medicinal properties and wood quality. Studies on this tree species have been focused mainly on wood properties, propagation and growth. Genetic studies on capirona are very limited to date. Today it is possible to explore genetic diversity and population structure in a fast and reliable manner by using molecular markers. We here used 10 Random Amplified Polymorphic DNAs (RAPDs) markers to analyze genetic diversity and population structure of 59 samples of capirona that were sampled from four provinces located in the eastern region of the Peruvian amazon. A total of 186 bands were manually scored, generating a 59 x 186 presence/absence matrix. We used R software to calculate genetic distances based on provesti coefficient. A dendrogram was generated using the UPGMA clustering algorithm and showed four groups that correspond to the geographic origin of the capirona samples. Similarly, a discriminant analysis of principal components (DAPC) confirmed that capirona is grouped into four clusters. However, we also noticed few accessions are intermingled. Genetic diversity estimation was conducted considering the four groups (populations) identified by *adegenet* package in R. Nei's genetic diversity estimate varied from 0.26 to 0.39 and Shannon index ranged from 2.48 to 2.83. AMOVA analysis revealed the greatest variation exist within populations (69.7%) and indicated that variability among populations is 31.5%. To our best knowledge, this is the first investigation employing molecular markers in capirona in Peru considering their natural distribution, and sheds light towards its modern genetic improvement and for the sustainable management of forests in Peru.

Keywords: Amazon forest; capirona; molecular markers; genetic diversity; population structure.

1. Introduction

Calycophyllum spruceanum (Benth) (Rubiaceae) is a fast-growing forest specie [1] with a natural origin in the Amazon basin, covering the territories of Bolivia, Peru, Colombia, and Brazil [2]. Capirona is a valuable timber forest species in the Peruvian Amazon, widely used by local communities and is exported around the world. It has a high-density, durable wood, therefore, it is used for the construction of economically valuable products [3]. Capirona has a high potential for cultivation in agroforestry systems, which has countered unsustainable land use due to activities such as slash-and-burn agriculture, and this

 doi:10.20944/preprints202106.0162.v1

specie represents an alternative to increase rural incomes and development [4]. Thus, capirona is currently the object of research in the Peruvian Amazon basin in order to establish participatory research with local communities, optimizing genetic and silvicultural management strategies including community conservation methods [1] since it is known that the indiscriminate use of forest species is generating deforestation in various regions of the planet [5-7]. Genetic diversity studies are indispensable for conducting conservation programs and sustainable management. Moreover, these studies based on molecular markers provide important information on the genetic makeup of the population because they are independent of the flexing effects of the environmental factors [8]. A demographic study would not provide this kind of information, influencing decision making. In recent years, molecular markers such as DNA-based markers like random-amplified polymorphic DNA (RAPD), microsatellite or simple sequence repeat (SSR) and ISSR have been used in studies for the analysis of phylogeny, inter-species relationships, and genetic diversity for forest species like *Pinus leucodermis*, *Cedrela Odorata Eucalyptus globulus*, *Swietenia macrophylla*, *Bertholletia excels* and *Populus deltoides* [9-14].

Previous research determined the genetic variation of capirona by using amplified fragment length polymorphisms (AFLP) among nine populations from the Peruvian Amazon [1]. They demonstrated that variation among individuals within populations is predominant. However, they did not employed samples from Madre de Dios region, which is considered a place where capirona exists naturally [15]. On the other hand, Tauchen et al., [3] identified the genetic variability of capirona at the level of DNA polymorphism evaluated by non-specific ITS primers of plant materials from Pucallpa, department of Ucayali in the Peruvian Amazon. Their results showed that morphological diversity is greater than genetic diversity of capirona. They also resolved that the environmental factor had a greater impact on the phenotype in those accessions. Finally, their results are in line with the claims of previous studies on *C. spruceanum*, suggesting greater variation within accessions than between accessions. RAPDs is an especially useful DNA-based method for initial research of genetic diversity in plant species [16, 17] as capirona. Also, RAPDs have demonstrated great potential in the analysis of polymorphism and genetic diversity as they are fast, easy to test and require low concentrations of DNA [17].

The objective of this study was to determine the genetic diversity and population structure of capirona from primary forests located in the regions of Ucayali, San Martín and Madre de Dios located in the Peruvian Amazon to establish in the near future a modern breeding program, and a sustainable conservation of this tree species.

2. Materials and Methods

2.1. Plant material

We collected 59 samples of capirona from Ucayali, San Martín and Madre de Dios regions in the Peruvian Amazon considering their natural range of distribution. Leaf samples were collected in paper envelopes, stored in airtight container with silicone gel, and transported to the National Institute of Agricultural Innovation (INIA for its acronym in Spanish) for genomic DNA extraction. Further details of the samples examined in this study are available in Table 1.

2.2. DNA amplification

We extracted genomic DNA by using the CTAB method with minor modifications [18,19]. All procedures were performed in 1.5 mL plastic microfuge tubes. About 0.1 g dry leaves were grounded in liquid nitrogen and suspended in 600 mL of 2x CTAB buffer containing 0.2% b-mercaptoethanol, followed by incubating at 65 °C for 60 min, then equal volume of chloroform: isoamylalcohol (24:1, v/v) was added, and the sample was shaken gently and then centrifuged. The supernatant was extracted, then 10X CTAB buffer and chloroform: isoamylalcohol (24:1, v/v) was added again. The supernatant was extracted and then

 doi:10.20944/preprints202106.0162.v1

mixed with ice cold isopropanol. DNA was recovered as a pellet by centrifugation, washed with ice-cold ethanol twice (70 y 90%), and then DNA was dried in the air. Finally, DNA was resuspended in nuclease-free water. The RNA contamination in all the samples was removed by digesting the extract with RNase-A (100 µg ml-1) for 30 min at 37ºC. DNA quality and quantity were checked in 1% agarose gels using Gelred (Biotium®, USA) and by standard spectrophotometry.

Table 1. The 59 samples of capirona characterized in this study, code, and geographic information.

Number	Code	Location	Region	Country	Number	Code	Location	Region	Country
1	Cap 018	Irazola	Ucayali	Peru	31	Cap 057	LBS	San Martín	Peru
2	Cap 019	Irazola	Ucayali	Peru	32	Cap 058	LBS	San Martín	Peru
3	Cap 020	Irazola	Ucayali	Peru	33	Cap 059	LBS	San Martín	Peru
4	Cap 021	Irazola	Ucayali	Peru	34	Cap 060	LBS	San Martín	Peru
5	Cap 022	Irazola	Ucayali	Peru	35	Cap 061	LBS	San Martín	Peru
6	Cap 023	Irazola	Ucayali	Peru	36	Cap 062	LBS	San Martín	Peru
7	Cap 024	Irazola	Ucayali	Peru	37	Cap 063	LBS	San Martín	Peru
8	Cap 030	Masisea	Ucayali	Peru	38	Cap 064	LBS	San Martín	Peru
9	Cap 031	Masisea	Ucayali	Peru	39	Cap 065	LBS	San Martín	Peru
10	Cap 032	Masisea	Ucayali	Peru	40	Cap 066	LBS	San Martín	Peru
11	Cap 033	Masisea	Ucayali	Peru	41	Cap 067	LBS	San Martín	Peru
12	Cap 034	Masisea	Ucayali	Peru	42	Cap 068	LBS	San Martín	Peru
13	Cap 035	Masisea	Ucayali	Peru	43	Cap 069	LBS	San Martín	Peru
14	Cap 037	Masisea	Ucayali	Peru	44	Cap 070	Iñapari	Madre de Dios	Peru
15	Cap 038	Masisea	Ucayali	Peru	45	Cap 071	Iñapari	Madre de Dios	Peru
16	Cap 039	Masisea	Ucayali	Peru	46	Cap 072	Iñapari	Madre de Dios	Peru
17	Cap 042	Masisea	Ucayali	Peru	47	Cap 073	Iñapari	Madre de Dios	Peru
18	Cap 043	Masisea	Ucayali	Peru	48	Cap 074	Iñapari	Madre de Dios	Peru
19	Cap 044	Masisea	Ucayali	Peru	49	Cap 075	Iñapari	Madre de Dios	Peru
20	Cap 045	Masisea	Ucayali	Peru	50	Cap 076	Iñapari	Madre de Dios	Peru
21	Cap 046	Masisea	Ucayali	Peru	51	Cap 078	Iñapari	Madre de Dios	Peru
22	Cap 047	Masisea	Ucayali	Peru	52	Cap 079	Iñapari	Madre de Dios	Peru
23	Cap 048	Masisea	Ucayali	Peru	53	Cap 080	Iñapari	Madre de Dios	Peru
24	Cap 049	Masisea	Ucayali	Peru	54	Cap 081	Iñapari	Madre de Dios	Peru
25	Cap 050	LBS	San Martín	Peru	55	Cap 082	Iñapari	Madre de Dios	Peru
26	Cap 051	LBS	San Martín	Peru	56	Cap 084	Iñapari	Madre de Dios	Peru
27	Cap 052	LBS	San Martín	Peru	57	Cap085	Iñapari	Madre de Dios	Peru
28	Cap 053	LBS	San Martín	Peru	58	Cap 086	Iñapari	Madre de Dios	Peru
29	Cap 054	LBS	San Martín	Peru	59	Cap 087	Iñapari	Madre de Dios	Peru
30	Cap 055	LBS	San Martín	Peru					

A total of 10 RAPD markers (Operon Technologies Inc., USA) were employed to assess genetic diversity among 59 accessions of *C. spruceanum*. Amplification was achieved with 10 µl reaction volume containing 2ng DNA, kit Kapa HiFi Hotstart ReadyMix, 1 µM of primers. The PCR Amplification was performed at an initial denaturation temperature

of 94 °C for 4 min followed by 40 cycles of 1 min denaturation at 94 °C, 45 s annealing at 37 °C and 2 min extension at 72 °C with a final extension of 10 min at 72 °C [20]. in a Thermal Cycler Simplyone (Applied Biosystems™, USA). Amplified products were separated on 1% (w/v) agarose gel in TBE buffer by electrophoresis, and were then visualized with Gelred staining and photographed using Gel Documentation System. The size of the amplification products was estimated by comparing the amplicons with a 100 bp ladder (New England Biolabs, USA).

2.3. Data analysis

The RAPD bands patterns were scored visually for the presence (1) or absence (0) of various molecular weight sizes. Only polymorphic and reproducible bands were considered for the analysis [20]. Similar to the procedure employed by [21], loci with more than 10% missing data were excluded from the analysis. A total of 186 bands were manually scored, generating a 59 x 186 presence/absence data matrix. We then used RStudio software v1.2.5033 to calculate genetic distances based on provesti coefficient, then a dendrogram was generated using the UPGMA clustering algorithm with 1000 bootstrap replicates from poppr package v2.9.2. We also used *ade4* v1.7-16 and *adegenet* v.2.1.3 packages in Rstudio to conduct a principal coordinate analysis (PCoA) and a discriminant analysis of principal component analysis (DAPC), respectively, in order to infer capirona population structure. Number of populations (K) were set from 1 to 10 by k-means clustering with 100,000 iterations. Selection of the most likely number of clusters was based on the lowest Bayesian Information Criterion (BIC) value.

An analysis of molecular variation (AMOVA) was conducted to examine significant differences between the following grouped variance components: within populations and between populations within geographical regions. In addition, genetic diversity parameters were estimated using the *poppr* package in RStudio.

3. Results

3.1 RAPD analysis

The 10 RAPD primers employed in this work revealed 16 to 23 fragments in 59 samples of capirona, with 18.6 fragments as average. Of the total 186 fragments, 100% were polymorphic. Polymorphic information content (PIC) ranged from 0.19 to 0.34.

3.2 Genetic diversity and population structure

Our presence/absence data was used to construct a dendrogram based on Provesti ge-netic distances and shows that all members of capirona were placed in four clades according to their geographic locality: (1) LBS, (2) Irazola, (3) Masisea, and (4) Iñapari. Clade 1 was made up of individuals from the locality of La Banda de Shilcayo (LBS), San Martín region. Clade 2 included individuals from Irazola and the third clade from Masisea; these two clades belong to Ucayali region. Finally, the fourth clade is made up of individuals belonging to the locality of Iñapari in Madre de Dios region (Figure 1). In addition, our dendrogram shows that cluster Masisea and Iñapari are the only clusters supported with a bootstrap value of more than 70% (Figure 1).

Principal coordinate analysis (PCoA) shows that the first two axis explained 22.8% of the variation, and reveals that individuals, except for two samples from Masisea are separated into one group. Two samples of capirona from Iñapari are intermingled within the LBS group. In addition, there are two samples from Irazola that are mixed with the LBS group (Figure 2a). To explore the genetic structure of capirona from the Peruvian Amazon basin, we used the *find.clusters* function to determine the best K value to our capirona samples, obtaining that K = 4 is the most likely number of groups, according to the BIC criteria (Figure S1). In accordance with the dendrogram result, discriminant analysis of principal components (DAPC) also evidence that all samples of capirona are separated

 doi:10.20944/preprints202106.0162.v1

into the four clusters (Figure 2b). Moreover, Figure 3 demonstrates that capirona samples cluster according to their geographic origin (Irazola, Masisea, LBS, Iñapari). However, four samples from cluster 1 (Cap 030, Cap 033, Cap 035, Cap 065) grouped within the Masisea and LBS population. Similarly, sample Cap 069 from cluster 4 were placed within the LBS population. Sample Cap 085 belongs to cluster 3, but it was intermingled within the Iñapari population.

Genetic diversity estimation was conducted considering the four clusters (populations) identified by *adegenet* package, taking into account that RAPD are dominant markers. The Nei´s genetic diversity [22] estimates ranged from 0.26 to 0.39, showing a considerable similarity between the samples studied here. The Nei´s genetic diversity is thought to be a more comprehensible index to evaluate the similarity between species in an ecological population [23]. Shannon-Wiener index ranged from 2.48 to 2.83, and Simpson´s index from 0.92 – 0.94, indicating high genetic diversity.

The analysis of molecular variance (AMOVA) allows the study of genetic variation within and between populations. AMOVA revealed that 31.47% of the total variation was found among populations of capirona while 69.69% was within populations (Table 3). Variation among regions was 0%.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0162.v1

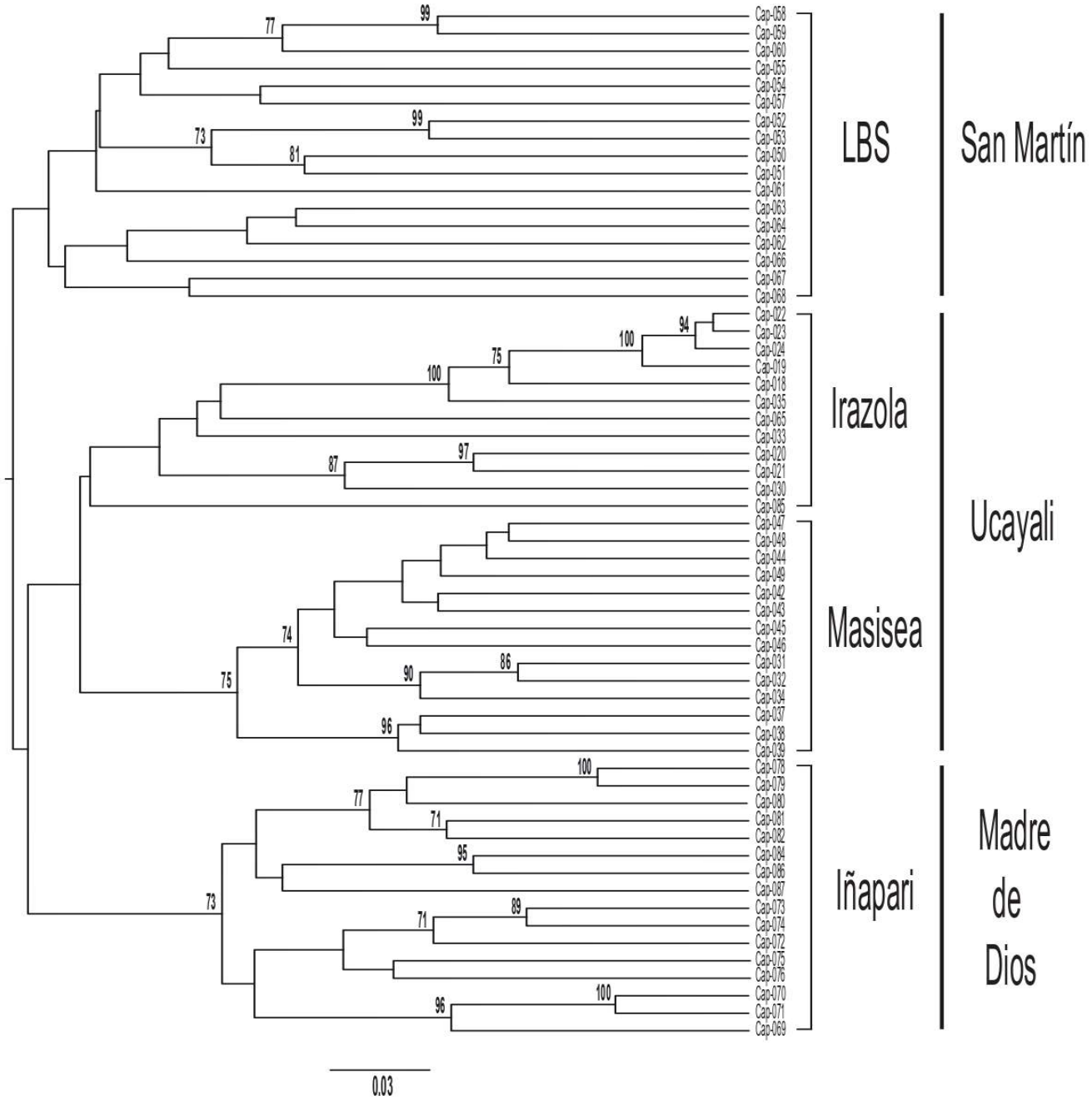

Figure 1. Dendrogram based on Provesti genetic distance and the UPGMA clustering method of 59 samples of capirona using 10 RAPD markers. Numbers above the branches represent bootstrap values, with only values higher than 70% shown. Names given to clades refer to the geographic origin.

 doi:10.20944/preprints202106.0162.v1

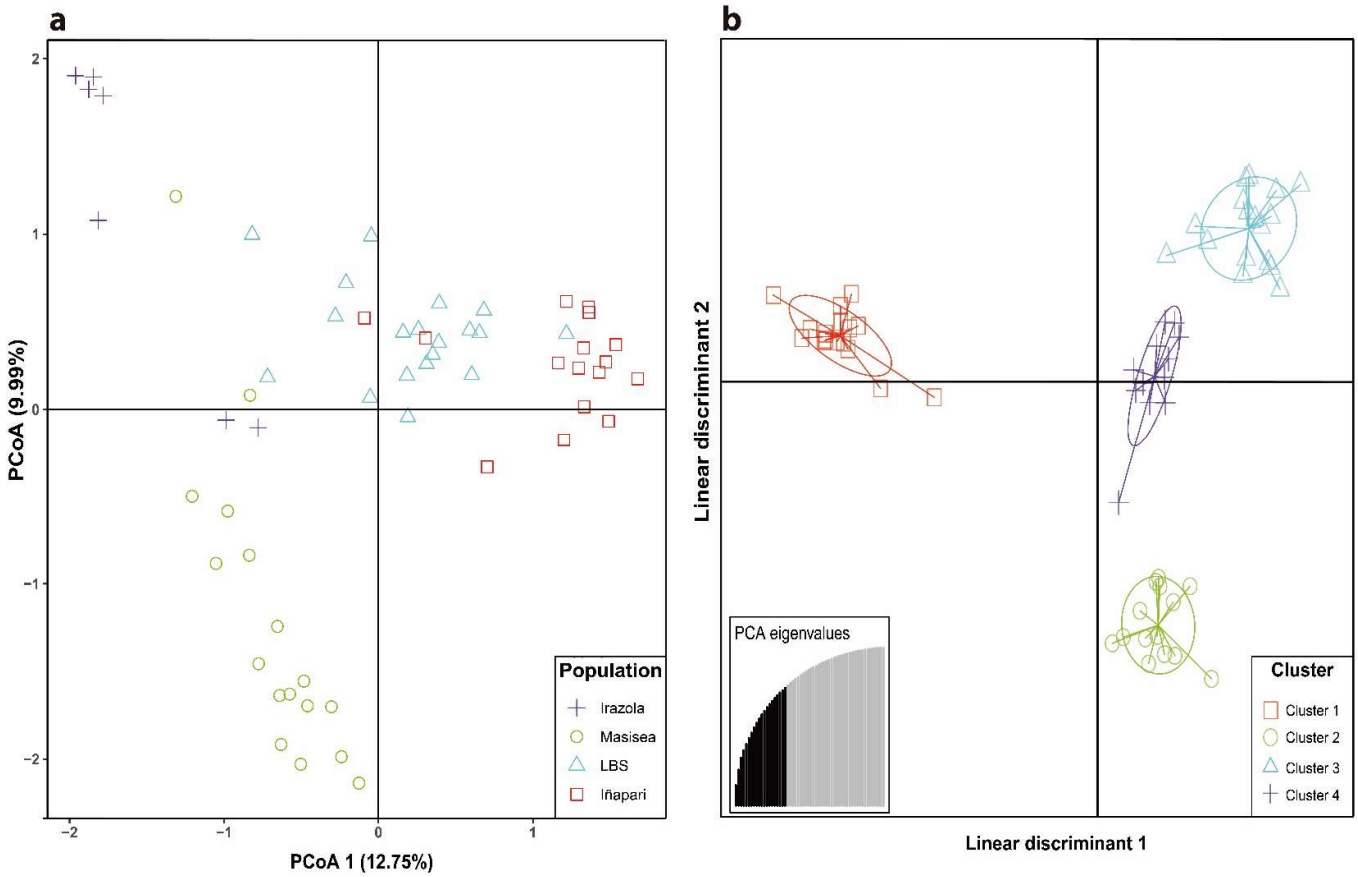

Figure 2. (a) Principal coordinates analysis (PCoA) of 59 samples of capirona from the eastern region of the Peruvian Amazon based on RAPD data. Percentages of total variance explained by each coordinate are noted in parentheses. Population symbols and colors indicate geographic affiliation. b) Discriminant analysis of principal components (DAPC) of 59 samples of capirona. The axes represent the first two Linear Discriminants (LD). Each circle represents a cluster and each symbol represents an individual. Numbers represent the different clusters identified by DAPC analysis.

Table 2. Analysis of molecular variance (AMOVA) using 10 RAPD markers of the genetic variation of 59 samples of capirona.

Source	df	SS	MS	Est.Var.	%
Among Regions	2	452.77	226.39	0	0.00%
Among Pops	1	201.06	201.06	13.28	31.47%
Within Pops	55	1617.72	29.41	29.41	69.69%
Total	58	2271.56		42.2	100.00%

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0162.v1

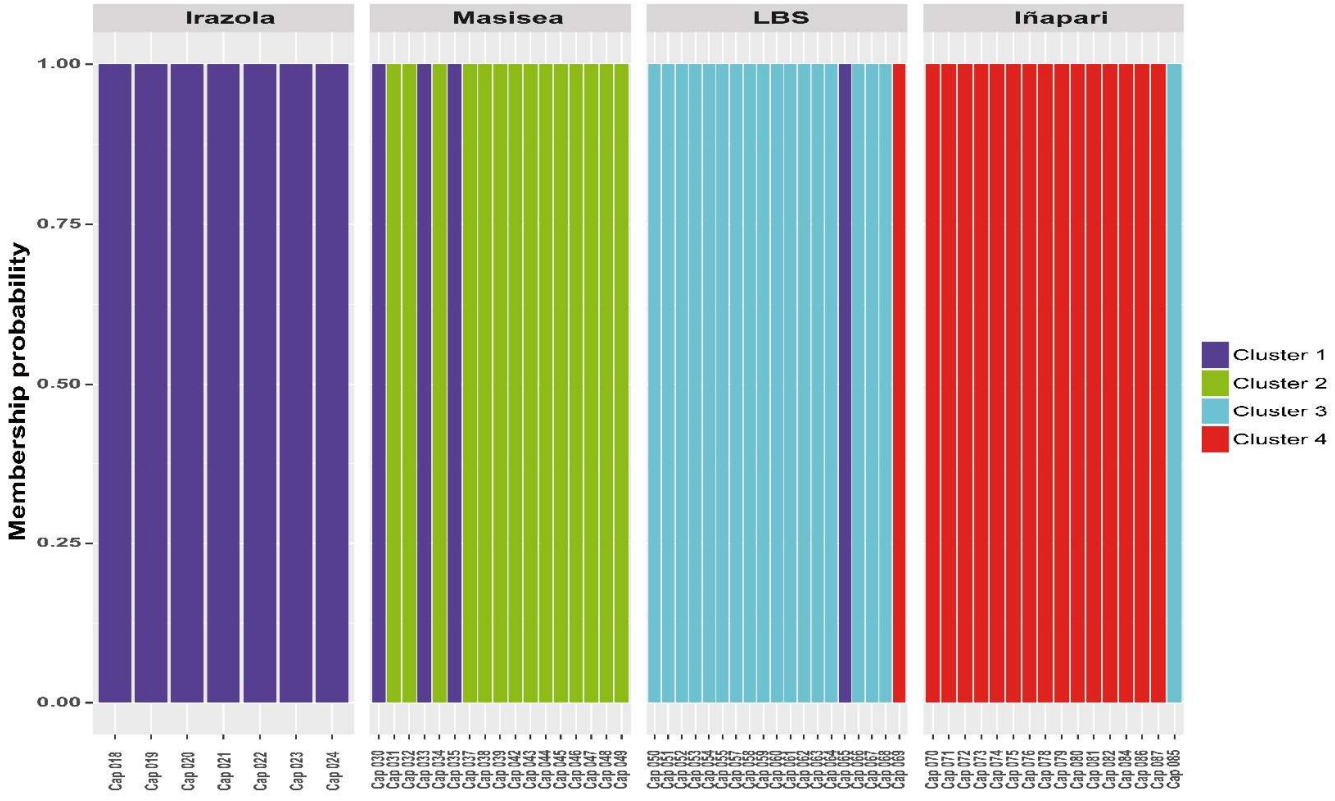

Figure 3. Population structure of 59 samples of capirona given by DAPC using 10 RAPD markers. Locality names refer to the geographic origin of the samples. Colors represent different assigned clusters. The x-axis provides samples names; the y-axis provides the probability of each accession belonging to the assigned cluster.

Table 2.- Genetic diversity based on RAPD markers among four clusters.

Cluster	N	H	lambda	He
1	16	2.77	0.94	0.29
2	14	2.64	0.93	0.26
3	17	2.83	0.94	0.39
4	12	2.48	0.92	0.30
Total	59	4.08	0.98	0.41

Pop: Population name, N: population size, H: Shannon-Wiener index of diversity, lambda: Simpson's index, He: Nei's 1978 expected heterozygosity.

4. Discussion

Genetic diversity sheds the light on the evolutionary pressure on the alleles and the mutation rate a locus might have undergone over a time period [26]. To our best knowledge, this is the first study that employs molecular markers to estimate population structure and genetic diversity of capirona from a wide range of natural distribution in Peru. Molecular markers based on DNA (RAPD, AFLP, SSRs) are used to study genetic variation in forest species [1,8,19,24]. Advantages of these methods over others include

 doi:10.20944/preprints202106.0162.v1

their better representation of the variation present within species [12]. RAPD markers is applied for initial studies of genetic diversity in forest species, like big-leaf mahogany (*Swietenia macrophylla*), pine (*Pinus leucodermis*), tornillo (*Cedrelinga cateniformis*) [9,12,16]. To date, only a small number of studies have used molecular markers to examine genetic diversity within and between populations of capirona, and this study represents the first population structure assessment in capirona from the Peruvian Amazon. We here report that genetic diversity of capirona based on RAPD markers are slightly high (0.26-0.39) among four populations sampled across the Peruvian Amazon. These estimates are in agreement with the results obtained by Russell et al [1]. On the contrary, Dávila-Lara et al [26] reported lower genetic diversity parameters across 13 populations of capirona in Nicaragua (Central America). This discrepancy may be explained due to sampling origin. The Amazon basin is known as the center of origin of this forest species [2], therefore, this geographic region might hide plenty of beneficial genes as germplasm around this center had the opportunity over a long period of time to evolve adaptation capacity to multiple environmental challenges [27]. Therefore, it is expected the capirona samples from the Amazon basin possess high genetic diversity. On the other hand, low genetic diversity of samples from Nicaragua are explained by biogeographic history, Pleistocene range dynamics and recent anthropogenic deforestation [26]. Moreover, Tauchen et al. [3] used Internal Transcribed Spacer (ITS) to molecularly characterize populations of capirona, and failed to separate individuals by multivariate analysis. They concluded that morphological diversity is higher than the genetic one in this tree species. Similarly, Montes et al. [25] analyzed the morphological diversity of *Callycophyllum spruceanum* in eleven natural populations of the Peruvian Amazon. They grew seeds of capirona and analyzed stem-growth and branch-wood traits and found no significant structuring populations.

The different values obtained for the three diversity indices (Shannon-Wiener, Simpson and Nei´s indices) may be explained due to the fact that whereas Nei´s index measures the level of heterozygosity, Shannon-Wiener and Simpson assumes that any difference between the analyzed individuals means a different species. Thus, although genetic relationship between species is relatively low using those markers (low Nei´s index), the diversity of the sequences and therefore of the population is relatively high (high Simpson and Shannon indices) [22, 23]. Similarly, a recent study on *Guazuma crinita* reported low Nei´s gene diversity and high Shannon information index [24]. In addition, our expected heterozygosity is similar to the populations of tornillo (*Cedrelinga cateniformis*) [16]. On the other hand, in this study 59 samples of capirona were characterized by an extreme dearth of genetic diversity as revealed by an overall Simpon´s index of 0.98.

Population structure analysis is informative to understand genetic diversity and facilitates subsequent association mapping studies [27]; its presence in those studies can lead to false positive associations between markers and traits [27]. Consequently, assessing population structure is a crucial step to conduct association between molecular markers and traits. DAPC analysis divided the population of capirona into well-defined clusters associated with their genetic structure. According to Rosyara et al. [28], DAPC exhibits the ability to control population structure in association with mapping studies, and it is slightly better than STRUCTURE analysis for discriminating among populations as revealed in *Prunus avium* [29] and *Camelia sinensis* [30]. In our study, DAPC and dendrogram (UPGMA tree) results revealed that 59 samples of capirona from the Peruvian Amazon clustered into four groups. Moreover, PCoA gave similar results. The presence of structure in this 59 samples of capirona meet our expectations since according to the pedigree of the genotypes, all the individuals can be divided into four geographical locations, and intentional selection of traits by farmers might lead to population structure. However, as seen in Figure 3, there are few accessions intermingled between the four clusters. This might be due to the common process of exchanging genetic material by farmers living in close geographic areas (Irazola, Masisea, Iñapari), or migration as it may have occurred for sample Cap 069 (LBS).

AMOVA demonstrated that there was 0% of variation among regions (San Martín, Ucayali and Madre de Dios), suggesting absence of genetic structure across these locations. A plausible explanation for this result is the geographic distance that exist between San Martín (northern Peru) vs Ucayali and Madre de Dios (southern Peru), which limits seed exchange. On the contrary, the greatest variation exist within populations of capirona. This may be explained due to the sexual propagation method of capirona. In addition, this is expected in this species, because it produces seeds that are dispersed over long distances by wind and water. In general, this is the case of tropical trees, and especially those that are outcrossing, and pollinated and/or dispersed by wind, so they tend to harbor high levels of genetic diversity at local and regional spatial scales [24]. Moreover, our results are in agreement with similar studies by Russell et al [1] and Tauchen et al. [3].

Even though, there are some efforts to study capirona at the molecular and morphological level, further research is needed aiming to identify putative genes in this tree species. Next generation sequencing (NGS) is a modern tool for the study of diversity and population structure. However, very few investigations were conducted using this tool for forest species [31, 32]. Our next step is to develop extra molecular tools for capirona and other tree species using NGS techniques and collaborating with other researchers worldwide. Future work will decipher the evolution history of genus *Calycophyllum*.

5. Conclusions

We here demonstrated that RAPDs markers were successful and effective for the assessment of the genetic diversity and structure of *C. spruceanum* populations from the Peruvian Amazon. High levels of genetic diversity were registered using different indices, reflecting probable extensive gene flow due to long-distance seed dispersal. Moreover, capirona samples were grouped into four clusters, according to their geographic affiliation. However, six samples were intermingled probably due to the common activity of seed exchange followed by farmers. More work is still needed using additional collections of capirona from a wider geographic origin. In addition, extra molecular tools should be developed for this tree species using NGS techniques in order to promote the establishment of a modern breeding program of forest species in Peru.

Supplementary Materials: Figure S1: Number of populations. Plot of K ranging from 1 to 10. All values were obtained from *adegenet* analysis. Four populations were considered in a data set of 10 RAPD markers and 59 samples.

Author Contributions: Conceptualization, Carla Saldaña, Wilbert Cruz and Eloy Cuellar; Data curation, and Carlos Arbizu; Formal analysis, Carla Saldaña and and Carlos Arbizu; Funding acquisition, Miriam Ramos and Eloy Cuellar; Methodology, Carla Saldaña, Johan Cancan and Mirian Correa; Project administration, Wilbert Cruz; Resources, Miriam Ramos and Eloy Cuellar; Supervision, Wilbert Cruz and and Carlos Arbizu; Writing – original draft, Carla Saldaña and and Carlos Arbizu; Writing – review & editing, Carla Saldaña, Johan Cancan, Wilbert Cruz, Mirian Correa, Miriam Ramos, Eloy Cuellar and and Carlos Arbizu.

Funding: This research was funded by National Agrarian Innovation Program (PNIA for its acronym in Spanish) - Project No. 120-PI.

Data Availability Statement: Not applicable

Acknowledgments: To Joel Reyes for supporting technical activities in the laboratory. We also thank Haru Garcia, Yanett Chumbimune and Lhia Aguirre for their support in the sample collections.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Russell, J. R.; Weber, J. C.; Booth, A.; Powell, W, and Dawson, C. S. Genetic variation *of Calycophyllum spruceanum* in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. Molecular Ecology. **1999**, 8, 199-204.
2. Sears, R. New forestry on the floodplain: the ecology and management of *Calycophyllum spruceanum* (Rubiaceae) on the Amazon landscape. Doctoral dissertation, Columbia University, EEUU. **2003**
3. Tauchen, J.; Lojka, B.; Hlasna-Cepkova, P.; Svobodova, E.; Dvorakova, Z.; and Rollo, A. Morphological and genetic diversity of *Calycophyllum spruceanum* (Benth) k. Schum (Rubiaceae) in Peruvian amazon. Agric. Trop. Subtrop. **2011**, 44, 212-218.
4. Weber, J.; Sotelo-Montes, C.; Labarta-Chávarri, R. Tree domestication in the Peruvian Amazon Basin – working with farmers for community development. Agroforestry Today. **1997**, 9, 4–8.
5. Busch, J. and Amarjargal, O. Authority of Second-Tier governments to reduce deforestation in 30 tropical countries. Frontiers in Forests and Global Change. **2020**, 3, 1.
6. Estoque, R. C.; Ooba, M.; Avitabile, V.; Hijioka, Y.; Dasgupta, R.; Togawa, T.; and Muruyama, Y. The future of Southeast Asia's forests. Nature Communications. **2019**, 10, 1–12.
7. Rasolofoson, R. A.; Ferraro, P. J.; Jenkins, C. N.; Jones J.P. Effectiveness of Community Forest Management at reducing deforestation in Madagascar. Biological Conservation. **2015,** 184, 271–277.
8. Singh, P., Singh, S. P., Tiwari, A. K., & Sharma, B. L. Genetic diversity of sugarcane hybrid cultivars by RAPD markers. Biotech. **2017**, 7, 1-5.
9. Bucci, G.; Vendramin, G. G.; Lelli, L.; Vicario, L. Assessing the genetic divergence of *Pinus leucodermis* Ant. endangered populations: Use of molecular markers for conservation purposes. Theoretical and Applied Genetics. **1997**, 95, 1138–1146.
10. Gillies, A. C. M.; Cornelius, J. P.; Newton, A. C.; Navarro, C.; Hernández, M.; and Wilson. Genetic variation in Costa Rican populations of the tropical timber species *Cedrela odorat*a L., assessed using RAPDs. Molecular Ecology. 1997, 6, 1133-1145.
11. Nesbitt, K. A.; Potts, B. M.; Vaillancourt, R. E.; Reid J.B. Fingerprinting and pedigree analysis in *Eucalyptus globulus* using RAPDs. Silvae Genetica. **1997**, 46,6-11.
12. Gillies, A. C. M.; Navarro, C.; Lowe, A. J.; Newtons, A.C.; Hernández, M.; Wilson, J. Genetic diversity in Mesoamerican populations of Mahogany (*Swietenia macrophylla*), assessed using RAPDs. Heredity. **1999**, 83,722-732.
13. Baldoni, A. B.; Teodoro, L. P. R.; Teodoro, P. E.; Tonini, H.; Tardin, F. D.; Botin, A. A.; Hoogerheide, E. S.; Botelho, S; Lulu, J.; Neto, A.; Azevedo, V. C. R. Genetic diversity of Brazil nut tree (*Bertholletia excelsa* Bonpl.) in southern Brazilian Amazon. Forest Ecology and Management. **2020**, 458, 117795.
14. Chen, C.; Chu, Y.; Ding, C.; Su, X., Huang, Q. Genetic diversity and population structure of black cottonwood *(Populus deltoides*) revealed using simple sequence repeat markers. BMC Genetics. **2020**, 21,1–12.
15. Kember, M. Plantas medicinales con potencial para eco negocios. Revista forestal del Perú. **2001,** 25.
16. Cruz, W.; Saldaña, C.; Ramos, H.; Baselly, R.; Loli, J. C.; Cuellar, E. Estructura y diversidad genética de poblaciones naturales de *Cedrelinga Cateniformis* "tornillo" en la región oriental del Perú. Scientia Agropecuaria. **2020**, 11, 521-528.
17. Tijerino, A.; Callejas, L. y Cerda-Granados, D. A. Assessment of Genetic Diversity in Five Nicaraguan Populations of *Cedrela odorata* L. (Meliaceae) using RAPD Markers. Encuentro. **2016**,103, 28–39.
18. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin**.** **1987**, 19, 11-15.
19. Cruz, W.; Ramos, H.; y Cuellar, J. Manual de protocolos para el estudio de diversidad genética en especies forestales nativas: Tornillo (*Cedrelinga cateniformis* Ducke), Capirona (*Calycophyllum spruceanum* Benth.), Shihuahuaco (*Dipteryx* sp.), Ishpingo (*Amburana* sp.) y Castaña (*Bertholletia excelsa*). Instituto Nacional de Innovación Agraria. **2019**. Lima-Perú.
20. Goyal, P.; Jain, R.; Kachhwaha, S.; and Kothari, S. L. Assessment of genetic diversity in *Pithecellobium dulce* (Roxb.) Benth. germplasm using RAPD and ISSR markers. Trees. **2014**, 29, 637-653.
21. Chia-Wong, J. A.; Garcia, C. L.; Suni, N. M.; and Eskes, B. Characterization of *Theobroma cacao* L. collection at Tingo Maria using ISSR molecular markers. Revista Aporte Santiaguino. **2011**, 4, 195-202.
22. Nei, M. Genetic distance between populations. In Molecular Evolutionary Genetics. **1987**, 208-253. Columbia University Press.
23. Hennink, S.; Zeven, A. C. The interpretation of Nei and Shannon-Weaver within population variation indices. Euphytica. **1990**, 51, 235-240.
24. Tuisima-Coral, L. L.; Hlásná Čepková, P.; Weber, J. C.; Lojka, B. Preliminary evidence for domestication effects on the genetic diversity of *Guazuma crinita* in the Peruvian Amazon. Forests. **2020**, 11, 795.
25. Montes, C. S.; Vidaurre, H.; and Weber, J. Variation in stem-growth and branch- wood traits among provenances *of Calycophyllum spruceanum* Benth. from the Peruvian Amazon. New Forests. **2003**, 26, 1-16.
26. Dávila-Lara, A.; Affenzeller, M.; Tribsch, A.; Díaz and Comes, H. P. AFLP diversity and spatial structure of *Calycophyllum candidissimum* (Rubiaceae), a dominant tree species of Nicaragua's critically endangered seasonally dry forest. Heredity. **2017**, 4, 275-286.
27. Eltaher, S.; Sallam, A.; Belamkar, V.; Emara, H. A.; Nower, A. A.; Salem, K. F. M.; Poland, J.; Baenziger, P.S. Genetic diversity and population structure of F3:6 nebraska winter wheat genotypes using genotyping-by-sequencing. Front. Genet. **2018,** 9 - 76.

 doi:10.20944/preprints202106.0162.v1

28. Rosyara, U.R.; De Jong, W.S.; Douches, D.S.; Endelman, J.B. Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome. **2016,** 9.
29. Campoy, J.A.; Lerigoleur-Balsemin, E.; Christmann, H.; Beauvieux, R.; Girollet, N.; Quero-García, J.; Dirlewanger, E.; Barreneche, T. Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of Prunus avium L. landraces and bred cultivars. BMC Plant Biol. **2016**, 16, 49.
30. Lee, K.J.; Lee, J-R.; Sebastin, R.; Shin M-J, Kim, S-H.; ho, G-T.; Hyun, D.Y.; Assessment of genetic diversity of tea germplasm for its management and sustainable use in Korea genebank. Forests. **2019**, 10, 780.
31. González, N. A. Identificación y validación de single nucleotide polymorphism (SNPs) distribuidos en el genoma de *Eucalyptus globulus*. Tesis de maestría, Universidad de Concepción. Chile. **2015**.
32. Durán Reyes, R. F. Modelos de predicción genómicos para la selección de genotipos de *Eucalyptus globulus* en base a densidad de la madera y volumen. Tesis de doctorado, Universidad de Concepción. Chile. **2017**.