 Chilean J. Agric. Anim. Sci., ex Agro-Ciencia (2022) 38(3):374-389.
https://doi.org/10.29393/CHJAA38-35CEBS10035

ISSN 0719-3890 online

GENETIC STRUCTURE AND DIVERSITY OF A PERUVIAN COLLECTION OF A HIGH-QUALITY WOOD TREE SPECIES, ULCUMANO (*Retrophyllum rospigliosii*, PODOCARPACEAE)

Carla L. Saldaña[1a], Johan D. Cancan[2], Evelyn J. Salazar[1b], Sheyla Y. Chumbimune[1c], Jorge H. Jhoncon[3], and Carlos I. Arbizu[1d]*

[1a] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru
https://orcid.org/0000-0002-9725-3578
[1b] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru
https://orcid.org/0000-0002-8878-430X
[1c] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru
https://orcid.org/0000-0002-2485-0988
[1d] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru
https://orcid.org/0000-0002-0769-5672
[2] Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Av. Carlos Germán Amezaga 375, Lima 15081, Peru
https://orcid.org/0000-0002-6395-2047
[3] Centro de Investigación de Plantas Andinas y Nativas, Facultad de Ciencias, Universidad Nacional de Educación Enrique Guzmán y Valle, Av. Enrique Guzmán y Valle s/n, Lima 15472, Peru
https://orcid.org/0000-0002-0119-6223
* Corresponding author: carbizu@inia.gob.pe

ABSTRACT

Ulcumano, which is native to South America, is an important conifer in Peru. Molecular studies are scarce, limiting modern breeding and appropriate conservation activities. Currently, molecular markers are widely employed to explore genetic structure and diversity parameters of plant species in a fast and precise manner. The objective of this study was to analyze the genetic diversity and population structure of ulcumano in Peru by using DNA-based molecular markers. Nine Randomly Amplified Polymorphic DNA (RAPD) markers were used, while 95 individuals of ulcumano were sampled from three departments of Peru. A total of 265 DNA fragments were manually scored, but 247 of them were kept after removing the non-polymorphic markers. Genetic distances were calculated using R software based on Provesti´s coefficient. A dendrogram was obtained using the UPGMA clustering algorithm, showing no clear clustering. The principal coordinate analysis agreed with two population structure analyses, demonstrating that ulcumano is contained within two clusters, (i) Junín + Pasco, and (ii) Cajamarca, while very few individuals are intermixed. Genetic diversity parameters were estimated considering the two groups (populations) identified by STRUCTURE software. Nei's genetic diversity estimate varied between 0.22 and 0.28, while Shannon index ranged from 3.43 to 4.16. Population divergence (F_{st}) between the two clusters revealed low

Received: June 8, 2022 Accepted: November 28, 2022

genetic differentiation (0.064). AMOVA analysis revealed that 87.31 and 12.69% of the total genetic variation were found within populations and between individuals, respectively. To the best of our knowledge, this is the first molecular study in ulcumano in Peru, and provides valuable information for the genetic improvement and sustainable management of this conifer in the country.

Key words: ulcumano, molecular markers, genetic diversity, population structure, germplasm.

INTRODUCTION

Retrophyllum rospigliosii "ulcumano" is a conifer that belongs to the Podocarpaceae family that grows in rainforests. This family is distributed in the Southern Hemisphere, with populations also in China, Japan, Mexico, and the Caribbean in the neotropics (Cernusak et al., 2011; Pujana et al., 2020). Ulcumano is native to Ecuador, Peru, Colombia, Venezuela (Farjon, 2010) and Bolivia (Zenteno-Ruiz, 2007). It is distributed from 1,470 to 3,300 m.a.s.l. in Peru and Colombia, being reported up to 3,750 m.a.s.l. (Farjon, 2010). This species forms extensive masses in exposed sites, but it is frequently found as dispersed individuals due to deforestation (Zenteno-Ruiz, 2007; Reynel and Marcelo, 2009). Its regeneration is scarce because the seeds are attacked by insects (beetles), and thus there are few seed trees (Arteaga et al., 2020). Furthermore, reforestation of the species is difficult due to its deficient sexual reproduction (Arteaga et al., 2020; Gómez et al., 2013). In fact, *R. rospigliosii* is a dioecious species (Cueva et al., 2016), with a germination power that decreases significantly in a short time (recalcitrant seeds), and a very long germination process (Cueva et al., 2013; Cueva et al., 2016; Mill, 2016).

Increase in productivity and adaptability based on forest genetic improvement, as well as the proper management of genetic resources for conservation, require prior knowledge of the magnitude and distribution of the genetic variation of the species (Sotolongo et al., 2013). For the development of genetic diversity studies, molecular markers have a series of advantages over morphological ones (Gómez et al., 1997). The current trend in forest biotechnology is genetic improvement by molecular markers (Zapata and Hasbun, 2011), which are obtained by different techniques for diversity studies. Molecular markers have many advantages as they are not influenced by environmental factors (Valadez and Kahl, 2000). Herbert et al. (2002) indicated that the starting point for conservation and management of genetic resources is the characterization of the species. Likewise, Cheliak (1993) mentioned that estimation of genetic diversity facilitates the identification of different genotypes. The use of molecular markers in forest species has increased efficiency of genetic improvement programs (Araya et al., 2005). Specifically, molecular markers are used to (i) estimate the rate of gene migration and (ii) characterize and analyze systems of mating, as well as for (iii) paternity or kinship analysis (Valadez and Kahl, 2000), and (iv) phylogenetic studies and genetic diversity estimation (Saldaña et al., 2021). Unfortunately, the use of molecular markers in the forest genetic resources of Peru has not emerged yet.

Randomly Amplified Polymorphic DNA (RAPD) markers, which were initially described by Williams et al. (1990), are still used for initial research in genetic diversity and population structure of plant species. These baseline studies are crucial for the development of conservation programs, being RAPD markers a useful tool to obtain key information in a short time (Renau-Morata et al., 2005; Tijerino et al., 2016; Saldaña et al., 2021). In fact, the efficacy of the RAPD technique in revealing DNA-level genetic variation has been demonstrated for different species such as capirona (*Calycophyllum spruceanum* Benth.) (Saldaña et al., 2021). In addition, RAPD markers are inexpensive, easy to use, and require low DNA concentrations to generate genetic profiles in a short time (Saldaña et al., 2021). In this sense, Gómez et al. (1997) have indicated that RAPD markers can be employed for genetic mapping in conifers, and to estimate their genetic diversity. Stark (2005) carried out taxonomic and phylogenetic studies in the Podocarpaceae family through internal transcribed spacer 2 (ITS2) marker and concluded that the use of markers such as RAPD, amplified fragment length polymorphism (AFLPs) or microsatellites allowed resolving phylogenetic relationships within the species of this family. Quiroga (2008) evaluated the genetic characteristics, phylogeography and phylogeny of some podocarps by isoenzymes and chloroplast markers and showed similar rates of gene flow and differences in the geographic structuring pattern. Ulcumano is an orphan species as no genetic diversity studies have been reported. Even genetic studies within the *Retrophyllum* genus are still limited. Herbert et al. (2002) sequenced the trnL-F region of chloroplast DNA to determine the phylogeny *Retrophyllum, Podocarpus, Nageia*¸ and *Afrocarpu*s species and demonstrated that the Melanesian species (*R. minus, R. comptonii, R.*

vitiense) formed a monophyletic clade, sister to the South American species *R. rospigliosii.*

To date, there is no genetic information of ulcumano in Peru. Therefore, the objective of this study was to analyze the genetic diversity and population structure of ulcumano in Peru by using DNA-based molecular markers. Nine RAPD markers were used, while ulcumano samples were obtained from primary forests located in the Peruvian departments of Cajamarca, Junín and Pasco.

MATERIALS AND METHODS

Plant material

Ninety-five ulcumano trees, with a diameter at breast height (DBH) of ≥ 30 cm and spaced at least 40–80 m from each other, were sampled from the departments of Cajamarca, Pasco, and Junín (Peru), considering their natural range of distribution. Young fresh leaves were collected in paper envelopes, stored in an airtight container with silicone gel, and then transported to the National Institute of Agrarian Innovation (INIA for its acronym in Spanish) for genomic DNA extraction. Further details of the ulcumano samples examined in this study are available in Table 1.

DNA Amplification

The CTAB method with minor modifications (Doyle et al., 1987; Saldaña et al., 2021) was used to extract genomic DNA. About 0.1 g dry leaves were ground in liquid nitrogen, suspended in 1ml of sorbitol buffer (0.1 M Tris-HCl (pH 6.4), 5 mM EDTA (pH 8.0), 2.5% PVP-40, 0.35 M sorbitol, and 1% β-mercaptoetanol), and then centrifuged. This step was repeated three times. Then, a volume of 1 mL of 2x CTAB buffer containing 0.2% b-mercaptoethanol and PVP (polyvinylpyrrolidone) 1% was added, and incubated at 65 °C for 60 min. Subsequently, an equal volume of chloroform: isoamyl-alcohol (24:1, v/v) was added, and the sample was shaken gently and then centrifuged. For residue removal, the supernatant was extracted by adding 10X CTAB buffer and chloroform: isoamyl alcohol (24:1, v/v), and then mixed with ice-cold isopropanol. DNA was recovered as a pellet by centrifugation, washed with ice-cold ethanol twice (70 and 90%), and then air-dried. Finally, DNA was resuspended in nuclease-free water. RNA contamination was removed from all the samples of ulcumano by digesting the extract with RNase-A (100 µg ml-1) at 37ºC for 30 min. DNA quality was determined by 1% agarose gel electrophoresis using Gelred (Biotium®, USA) and a Implen NanoPhotometer.

Nine RAPD markers (Operon Technologies Inc., USA) were used to assess the genetic diversity among 95 individuals of ulcumano: OPA-04, OPA-10, OPA-12, OPF-05, OPF-06, OPF-07, OPF-12, OPT-05, and OPT-08 (Table 2). The amplification procedure was carried out according to Saldaña et al. (2021), in a final volume of 10 µl with Kapa HiFi HotStart ReadyMix PCR Kit (Roche) containing 5 ng of DNA, 0.2 µM primers. PCR amplification was performed using the following cycle profile: 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 (Goyal et al., 2014) in a SimpliAmp TM Thermal Cycler (Applied Biosystems™, USA). The PCR products were separated on 1.2% (w/v) agarose gel in TBE buffer by electrophoresis, and then visualized with Gelred® staining and photographed using Gel Documentation System. Size of the amplification products was estimated by comparing the amplicons with a 100 bp ladder (New England Biolabs, MA, USA) and 1 µl of DNA + 9µl of dye buffer 1x and 0.035 µl of Gelred®.

Data analysis

RAPD band patterns were visually inspected, recoding the presence (1) or absence (0) of them. Polymorphic information content (PIC) was calculated from dominant markers by the following equation:

$$PIC = 1 - [fi^2 + (1 - fi^2)^2]$$

Where, fi frequency of amplified band (1) and (1 - fi) is frequency of absence of band (0) (Chesnokov and Artemyeva, 2015).

Only polymorphic DNA fragments were considered for further analysis. R software v4.0.2 was used to calculate genetic distances based on Provesti´s distance. Subsequently, a dendrogram was constructed using the unweighted pair group method with arithmetic mean (UPGMA) clustering algorithm, with 1000 bootstrap replicates from *poppr* package v2.9.2 (Kamvar et al., 2014). To determine ulcumano population structure, *ade4* v1.7-16 and *adegenet* v.2.1.3 packages were used in R to carry out a principal coordinate analysis (PCoA) and a discriminant analysis of the principal component analysis (DAPC), respectively. The number of populations (K) was set from 1 to 10 by k-means clustering with 100,000 iterations. The most likely number of clusters was determined based on the lowest Bayesian Information Criterion (BIC) value.

Population structure of ulcumano was inferred using the STRUCTURE program v.2.3.4. (Pritchard et al., 2000) with ten runs for

Table 1. Samples of ulcumano from the Peruvian Amazon by code and geographic origin.

Code*	Location	Region	Latitude	Longitude
Ulc01	Oxapampa	Pasco	10° 28' S	75° 25' W
Ulc02	Oxapampa	Pasco	10° 28' S	75° 25' W
Ulc03	Oxapampa	Pasco	10° 28' S	75° 25' W
Ulc04	Oxapampa	Pasco	10° 25' S	75° 31' W
Ulc05	Oxapampa	Pasco	10° 25' S	75° 31' W
Ulc06	Oxapampa	Pasco	10° 25' S	75° 31' W
Ulc07	Oxapampa	Pasco	10° 45' S	75° 17' W
Ulc08	Oxapampa	Pasco	10° 45' S	75° 16' W
Ulc09	Oxapampa	Pasco	10° 46' S	75° 16' W
Ulc10	Oxapampa	Pasco	10° 42' S	75° 15' W
Ulc11	Oxapampa	Pasco	10° 42' S	75° 15' W
Ulc12	Oxapampa	Pasco	10° 43' S	75° 15' W
Ulc13	Oxapampa	Pasco	10° 43' S	75° 15' W
Ulc14	Oxapampa	Pasco	10° 43' S	75° 16' W
Ulc15	Oxapampa	Pasco	10° 42' S	75° 16' W
Ulc16	Oxapampa	Pasco	10° 44' S	75° 13' W
Ulc17	Oxapampa	Pasco	10° 44' S	75° 12' W
Ulc18	Oxapampa	Pasco	10° 45' S	75° 17' W
Ulc19	Oxapampa	Pasco	10° 46' S	75° 19' W
Ulc20	Oxapampa	Pasco	10° 42' S	75° 15' W
Ulc21	Oxapampa	Pasco	10° 42' S	75° 15' W
Ulc22	Oxapampa	Pasco	10° 42' S	75° 15' W
Ulc23	Oxapampa	Pasco	10° 42' S	75° 14' W
Ulc24	Oxapampa	Pasco	10° 42' S	75° 14' W
Ulc25	Oxapampa	Pasco	10° 36' S	75° 25' W
Ulc26	Oxapampa	Pasco	10° 39' S	75° 23' W
Ulc27	Oxapampa	Pasco	10° 34' S	75° 24' W
Ulc28	Oxapampa	Pasco	10° 25' S	75° 31' W
Ulc29	Oxapampa	Pasco	10° 29' S	75° 27' W
Ulc30	Oxapampa	Pasco	10° 28' S	75° 28' W
Ulc31	Oxapampa	Pasco	10° 31' S	75° 26' W
Ulc32	Chanchamayo	Junín	10° 04' S	75° 23' W
Ulc33	Oxapampa	Pasco	10° 26' S	75° 30' W
Ulc34	Satipo	Junín	11º 15' S	74º 48' W
Ulc35	Satipo	Junín	11º 16' S	74º 48' W
Ulc36	Satipo	Junín	11º 16' S	74º 48' W
Ulc36	Satipo	Junín	11º 16' S	74º 48' W
Ulc37	Satipo	Junín	11º 16' S	74º 48' W
Ulc38	Satipo	Junín	11º 16' S	74º 48' W
Ulc38	Satipo	Junín	11º 16' S	74º 48' W
Ulc39	Satipo	Junín	11º 16' S	74º 48' W
Ulc40	Satipo	Junín	11º 16' S	74º 48' W
Ulc41	Satipo	Junín	11º 16' S	74º 48' W
Ulc42	Satipo	Junín	11º 15' S	74º 48' W
Ulc43	Satipo	Junín	11º 15' S	74º 48' W
Ulc44	Satipo	Junín	11º 15' S	74º 48' W
Ulc45	Satipo	Junín	11º 15' S	74º 48' W
Ulc46	Satipo	Junín	11º 15' S	74º 48' W
Ulc47	Satipo	Junín	11º 15' S	74º 48' W
Ulc48	Oxapampa	Pasco	10° 42' S	75° 15' W

Ulc61	San Ignacio	Cajamarca	5º 13' S	78º 58' W
Ulc62	San Ignacio	Cajamarca	5º 13' S	78º 58' W
Ulc63	San Ignacio	Cajamarca	5º 15' S	79º 6' W
Ulc64	San Ignacio	Cajamarca	5º 15' S	79º 6' W
Ulc66	San Ignacio	Cajamarca	5º 14' S	79º 5' W
Ulc67	San Ignacio	Cajamarca	5º 14' S	79º 5' W
Ulc68	San Ignacio	Cajamarca	5º 14' S	79º 5' W
Ulc69	San Ignacio	Cajamarca	5º 13' S	79º 5' W
Ulc70	San Ignacio	Cajamarca	5º 13' S	79º 5' W
Ulc71	San Ignacio	Cajamarca	5º 13' S	79º 5' W
Ulc72	San Ignacio	Cajamarca	5º 13' S	79º 5' W
Ulc73	Chanchamayo	Junín	11° 02' S	75° 22' W
Ulc74	Chanchamayo	Junín	11° 02' S	75° 22' W
Ulc75	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc76	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc77	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc78	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc79	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc80	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc81	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc82	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc83	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc84	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc85	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc86	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc87	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc88	San Ignacio	Cajamarca	5º 13' S	79º 6' W
Ulc89	San Ignacio	Cajamarca	5º 14' S	79º 6' W
Ulc90	San Ignacio	Cajamarca	5º 14' S	79º 6' W
Ulc91	Oxapampa	Pasco	11º 7' S	75º 21' W
Ulc92	Oxapampa	Pasco	10º 43' S	75º 16' W
Ulc93	Oxapampa	Pasco	10º 43' S	75º 16' W
Ulc94	Oxapampa	Pasco	10º 48' S	75º 22' W
Ulc95	Oxapampa	Pasco	10º 48' S	75º 22' W
Ulc96	Oxapampa	Pasco	10º 48' S	75º 22' W
Ulc97	Oxapampa	Pasco	10º 48' S	75º 22' W
Ulc98	Oxapampa	Pasco	10º 25' S	75º 29' W
Ulc99	Oxapampa	Pasco	10º 25' S	75º 29' W
Ulc100	Oxapampa	Pasco	10º 28' S	75º 28' W
Ulc101	Oxapampa	Pasco	10º 31' S	75º 26' W
Ulc102	Oxapampa	Pasco	10º 31' S	75º 26' W
Ulc103	Chanchamayo	Junín	11°13' S	75° 21' W
Ulc104	Chanchamayo	Junín	11°13' S	75° 21' W
Ulc105	Chanchamayo	Junín	11°13' S	75° 21' W
Ulc106	Chanchamayo	Junín	11°15' S	75° 20' W
Ulc107	Chanchamayo	Junín	11°15' S	75° 20' W
Ulc108	Chanchamayo	Junín	11°15' S	75° 20' W

*Ulc49-Ulc60, Ulc65 correspond to other forest tree species that were also collected during the expedition.

Table 2. RAPD primer sequences and their polymorphic information content (PIC).

Primer	Sequence*	Total bands	% Polymorphism	PIC
OP-04	AATCAGCCAC	22	95.5	0.236
OPA-10	GTGATCGCAG	22	100	0.223
OPA-12	TCGGCGATAG	34	100	0.413
OPF-05	CCGAATTCGG	39	100	0.26
OPF-06	GGGAATTCGG	43	100	0.179
OPF-07	CCGATATCCC	35	100	0.181
OPF-12	ACGGTACCAG	27	100	0.188
OPT-05	GGGTTTGGCA	19	100	0.151
OPT-08	AACGGCGACA	24	100	0.277
	Total	265	99,5	0.234

*Goyal et al. (2015).

each number of populations (K value), ranging from 1 to 15 with a burn-in length of 100,000 Monte Carlos iterations, which was followed by 200,000 iterations. An admixture model with no previous population information was included; all other parameters were set to default values. Estimation of the most likely number of clusters was calculated by the Evanno method (Evanno et al., 2005). Membership probabilities ≥ 0.8 or the maximum membership probability was adopted to divide the ulcumano samples into different clusters. Population structure plots were generated with R package *pophelper* v.2.3.1 (Francis, 2017). In addition, the *tess3r* package v1.1.0 (Caye et al., 2016) was used to visualize the population structure of ulcumano on a geographic map of Peru, considering the STRUCTURE membership coefficient matrix of the estimated K.

An analysis of molecular variance (AMOVA) was conducted using the R package *poppr*, considering the number of populations determined by STRUCTURE. In addition, three genetic diversity parameters were calculated using the same package: (i) Shannon-Wiener index, (ii) Simpson's index, and (iii) Nei's gene diversity (expected heterozygosity). Moreover, the degree of gene differentiation among clusters in terms of allele frequencies (F_{st}) was estimated using the following formula:

$$F_{st}=1 - (H_s/H_t)$$

Where H_S is the average expected heterozygosity estimated from each cluster and H_t is total gene diversity or expected heterozygosity in the total cluster as estimated from the pooled allele frequencies.

RESULTS

RAPD Analysis

The nine primers used for molecular analysis revealed 265 fragments in 95 samples of ulcumano, with an average of 18.6 fragments. RAPD band patterns were scored visually for presence (1) or absence (0) in a 1% agarose gel image with various molecular weights. Of the total 265 bands, 99.5% were polymorphic (Fig. 1). Polymorphic information content (PIC) per primer ranged from 0.151 to 0.413, while the mean PIC value was 0.234 (Table 2).

Genetic diversity estimates and population structure

Eighteen DNA fragments were removed by the *poppr* package since they were monomorphic. Therefore, the final data set consisted of a 95 x 247 presence/absence binary matrix. A phylogenetic tree based on Provesti´s genetic distances did not clearly discriminate the ulcumano samples according to their geographic locality (Cajamarca, Pasco and Junín). However, there are some "admixture" clusters with samples from different origin, mainly Junín + Pasco, which present a bootstrap support lower than 70% (Fig. 2).

Principal coordinate analysis (PCoA) showed that the first and second axis explained 8.47% and 4.87% of the variation, respectively, and revealed that individuals from Junín and Pasco are completely mixed and clustered into one group. However, two samples from Junín (ulc73 and ulc74) and one sample of Pasco (ulc48) are intermingled within the Cajamarca group (Fig. 3). To explore the genetic structure of ulcumano from Peru, the *find.clusters* function was used to determine the best K value for our ulcumano samples, obtaining that K = 2 is the most likely

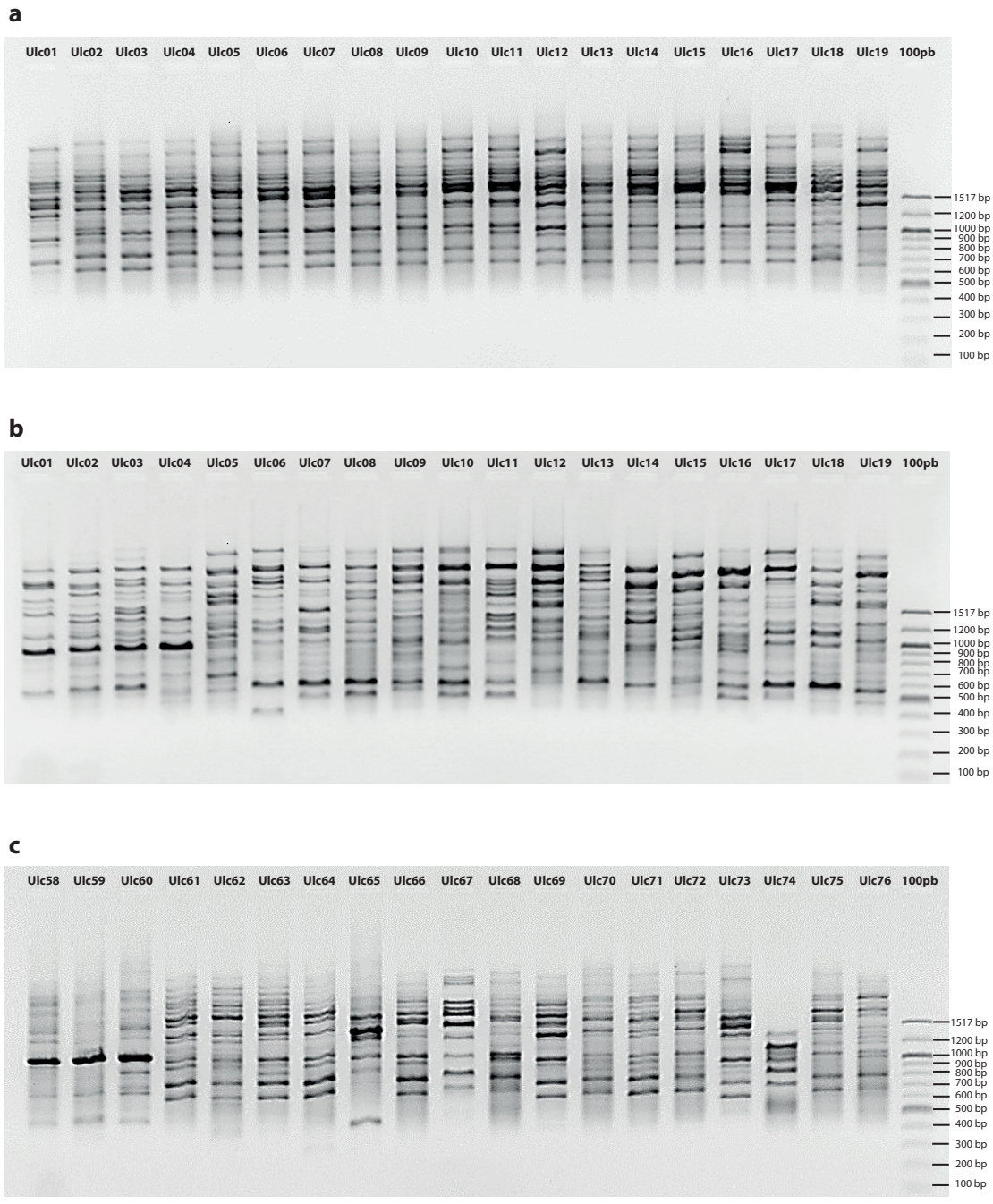

Fig. 1. RAPD banding pattern using primers OPA-04: a, OPA-10: b, OPA-12: c. Ladder 100 pb NEB. 1% agarose. 1 ul of DNA + 9 uL of dye buffer and 0.035 uL of Gelred®.

number of groups, according to the BIC criteria (Fig. 4). Unlike the dendrogram result, our discriminant analysis of principal components (DAPC) needed only one discriminant function to determine that all samples of ulcumano are separated into two clusters only (Fig. 5).

According to Evanno method (Evanno et al., 2005), the best K value (number of populations) is two for our data set. Similarly, DAPC indicated that K = 2 is the most likely number of populations, which agrees with the PCoA (Fig. 3). Moreover, Fig. 6 demonstrated that ulcumano samples do not cluster according to geographic origin (Cajamarca, Junín and Pasco)

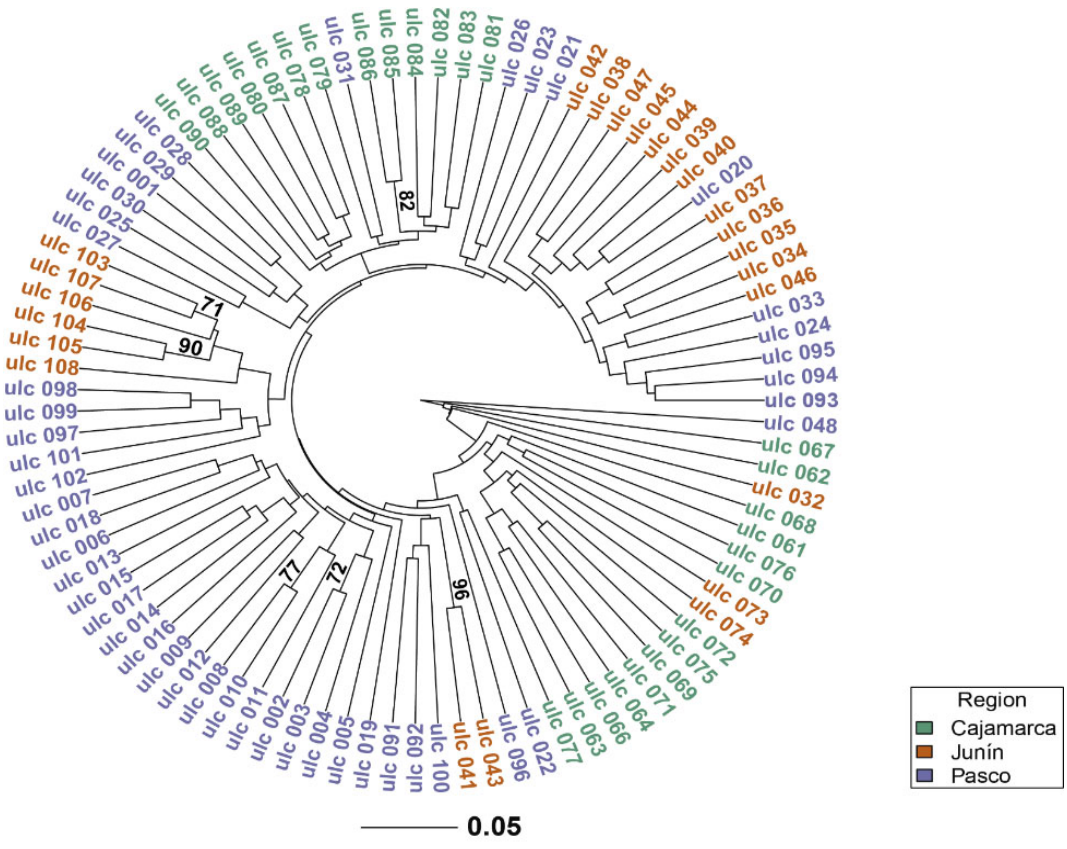

Fig. 2. Dendrogram of 95 samples of ulcumano using nine RAPD markers based on Provesti´s genetic distance and the UPGMA clustering method. Numbers above the branches represent bootstrap values, with only values higher than 70% shown.

but in two main clusters: (1) cluster 1, which includes ulcumano samples from Junín and Pasco; and (2) cluster two, which consists of ulcumano individuals from Cajamarca (Table 3). STRUCTURE analysis exhibited admixture for few samples. Samples ulc78 and ulc88 may belong to cluster 1. Similarly, samples ulc073 and ulc074 from Junín, and ulc031, ulc048, ulc091 and ulc092 from Pasco are more likely to belong to cluster 2. Fig. 7 shows STRUCTURE membership proportions to clusters spatially interpolated into a map of Peru, resulting from the analysis in TESS3 (Caye et al., 2016). Spatial interpolation of membership matrix assigned ulcumano samples from Junín and Pasco to cluster 1 mainly, while cluster 2 included samples from Cajamarca.

Genetic diversity indices and F_{st} estimate were determined considering the two clusters (populations) identified by STRUCTURE software, considering that RAPD are dominant markers. The Nei´s genetic diversity index was 0.222 for cluster 1 and 0.281 for cluster 2. Shannon-Wiener index ranged from 3.43 to 4.16, while Simpson´s index varied between 0.984 and 0.968 for cluster 1 and 2, respectively, indicating high genetic diversity. Moreover, the percentage of polymorphic loci was higher for cluster 2 (91.9%) than for cluster 1 (87.04%) (Table 4). Population divergence (F_{st}) between clusters 1 and 2 was 0.064, implying low genetic differentiation between these two populations.

The analysis of molecular variance (AMOVA) revealed genetic variation within populations and between individuals of ulcumano, reaching 12.69% and 87.31%, respectively (Table 5).

DISCUSSION

Ulcumano is an orphan crop, which possesses important traits and has the potential to generate profit. However, it has been under-researched during the past years, and thus genetic and genomic resources for this species are limited. Accordingly, there is little information about the genetic diversity of ulcumano and other members of the genus *Retrophyllum*. In fact, studies of the Podocarpaceae family using molecular markers are scarce. Stark (2005) employed ITS2 and trnL-F

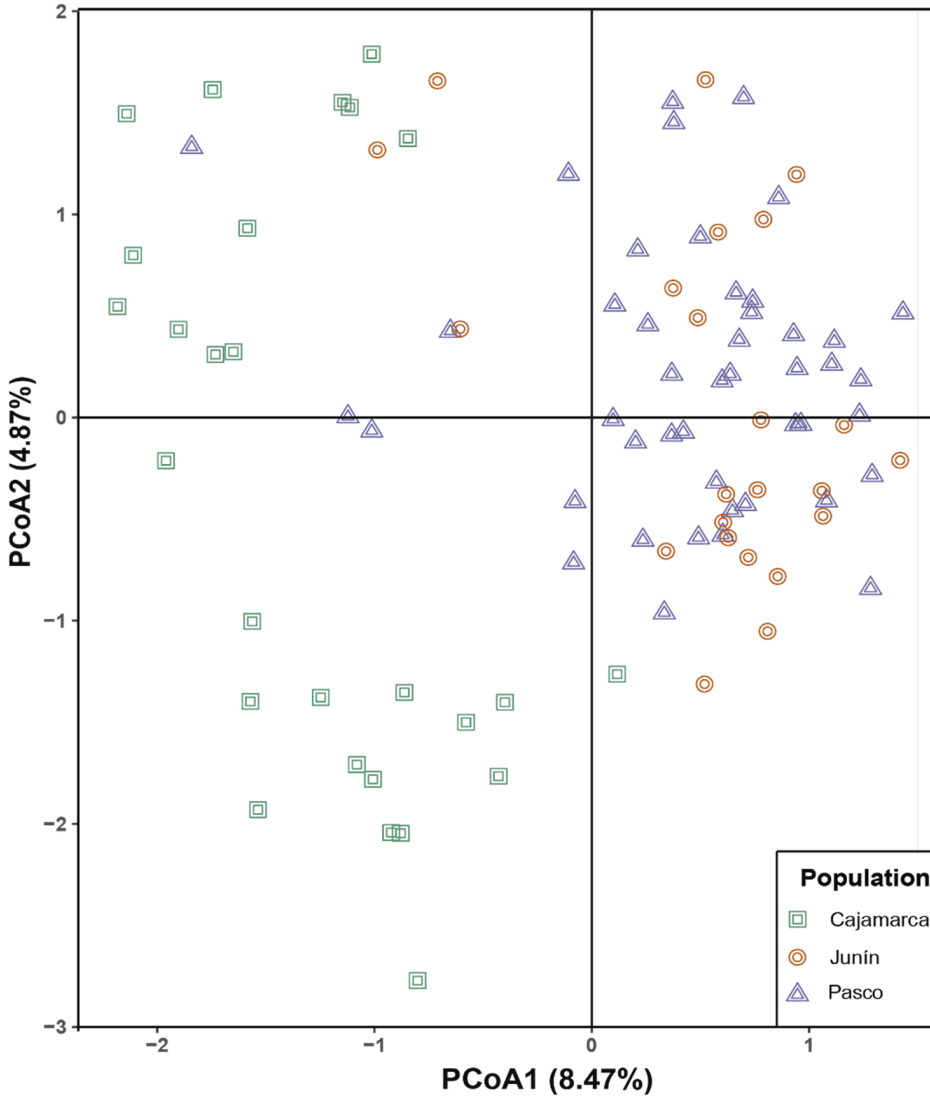

Fig. 3. Principal coordinates analysis (PCoA) of 95 samples of ulcumano from a Peruvian collection based on RAPD markers. Percentages of total variance explained by each coordinate are noted in parentheses. Population symbols and colors indicate geographic origin.

Table 3. Origin of the 95 ulcumano samples for the two clusters inferred by structure analysis.

Region	Cluster 1	Cluster 2
Cajamarca	3	24
Junín	20	3
Pasco	41	4
Total	64	31

markers to infer the phylogenetic relationship of 15 taxa of *Podocarpus* and *Foliolatus*, concluding that there is a need for further research using other molecular markers like RAPD, AFLP or microsatellites. To the best of our knowledge, this is the first study on population structure and genetic diversity of *R. rospigliosii*, a species with economic potential.

As observed in a recent study in other forest tree species (Saldaña et al., 2021), all three

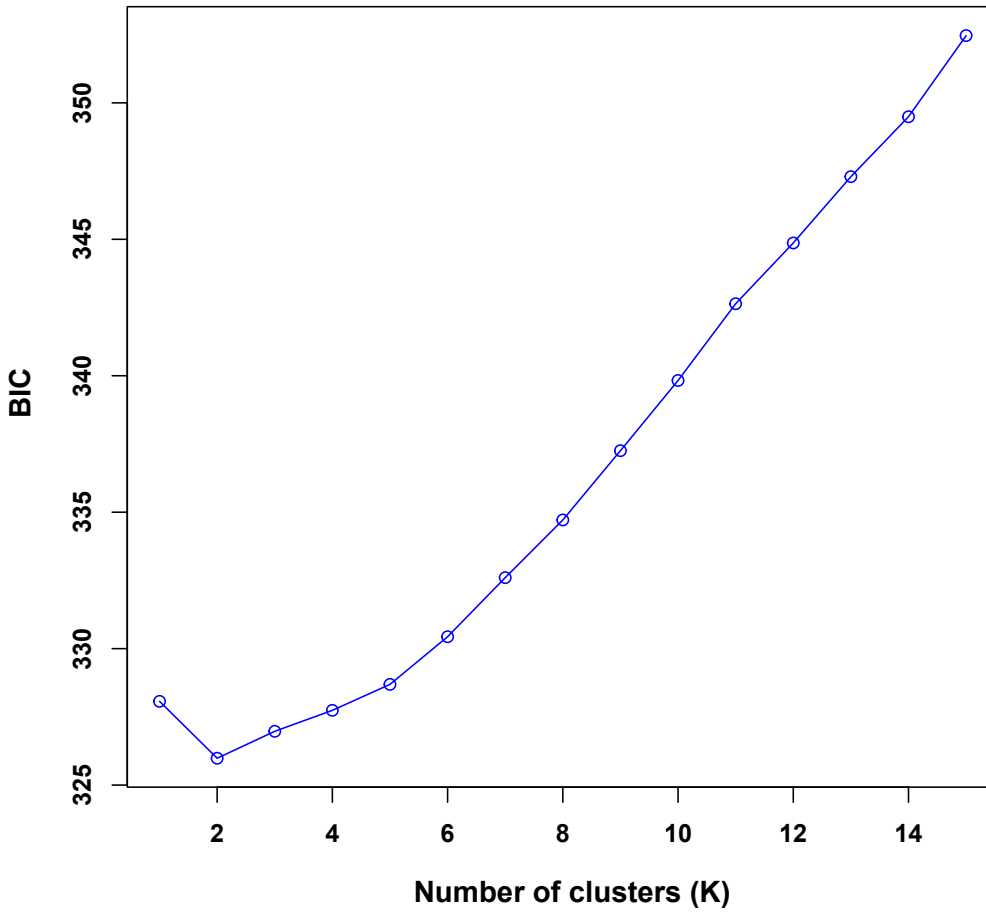

Fig. 4. Number of populations (K) inferred by discriminant analysis of principal components DAPC. K ranges from 1 to 15.

Table 4. Genetic diversity based on RAPD markers for the two clusters.

Cluster	N	H	lambda	He	PIC	PPL (%)
1	64	4.16	0.984	0.222	0.226	87.04
2	31	3.43	0.968	0.281	0.277	91.9
Total	95	4.55	0.989	0.256		

N: population size, H: Shannon-Wiener index of diversity, lambda: Simpson's index, He: Nei's 1978 expected heterozygosity, PIC: polymorphic information content, PPL: percentage of polymorphic loci.

Table 5. Analysis of molecular variance (AMOVA) using nine RAPD markers of the genetic variation of 95 samples of ulcumano between and within the two clusters inferred by STRUCTURE analysis.

Source	df	SS	MS	Est. Var.	%
Between clusters	1	211.88	211.88	4.36	120.69
Within clusters	93	2786.68	29.96	29.96	87.31
Total	94	2998.55	31.9	34.32	

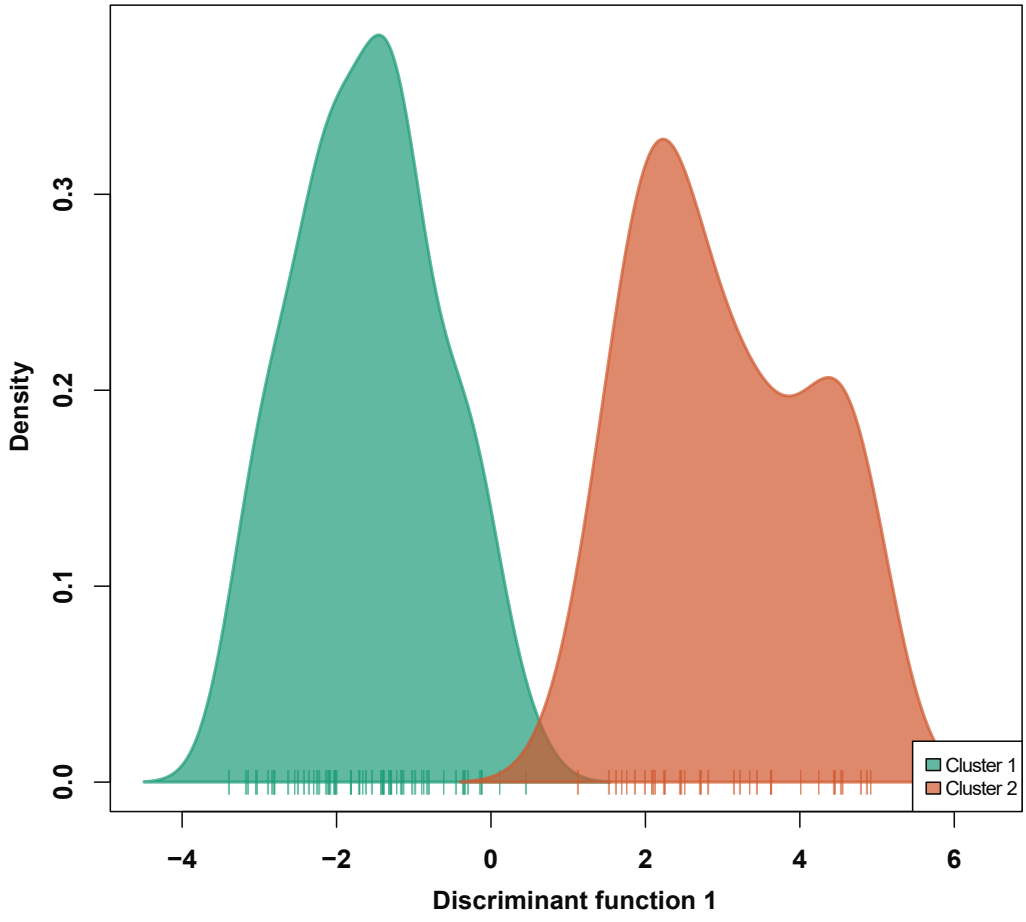

Fig. 5. Discriminant analysis of principal components (DAPC) of 95 samples of ulcumano.

genetic diversity indices used (Shannon-Wiener, Simpson and Nei´s indices) for ulcumano samples revealed different values because Nei´s index determines the level of heterozygosity, while Shannon-Wiener and Simpson's indices consider differences between individuals as different species. Therefore, these two indices showed very high values in the two clusters, which is explained by high variations between individuals. The average Nei´s index is 0.256, and thus ulcumano presents relatively low genetic diversity. In this sense, low levels of genetic diversity have been reported for other gymnosperms, including three species of Araucariaceae (Peakall et al., 2003). However, low genetic diversity does not seem to be characteristic of this species since other studies have reported high genetic diversity in wood species such as *Podocarpus sellowii* (Goncalvez, 2008) and *Dacrycarpus imbricatus* (Su et al., 2010). This low genetic diversity was also demonstrated by Dantas et al. (2015) in their analysis with ISSR and SSR on *Podocarpus sellowii*. Low genetic diversity could lead to vulnerability to external factors, anthropogenic pressure, or deleterious alleles (Dantas et al., 2015). Indeed, ulcumano is classified as VULNERABLE according to the Red List of Threatened Species, with a high risk of extinction in the wild (Gardner and Thomas, 2013); and ALMOST THREATENED in Peru, according to Supreme Decree No. 043-2006-AG (described as *Nageia rospigliosii*) (Supreme Decree, 2006). Exploitation and loss/decrease of their habitat have caused recent and constant declines in the size of ulcumano populations, having an impact on the genetic diversity of the species. Logging of this valuable timber tree has reduced or fragmented most of the formerly rather extensive stands of this species. Whole mountainsides of ulcumano in Peru were clear-felled in the 1980s for timber, and the species is now reduced to scattered individuals (Mill, 2016). Thus, policies to increase diversity of *R. rospigliosii* such as expansion of population sizes, reduction of threats, and/or introduction of genetically

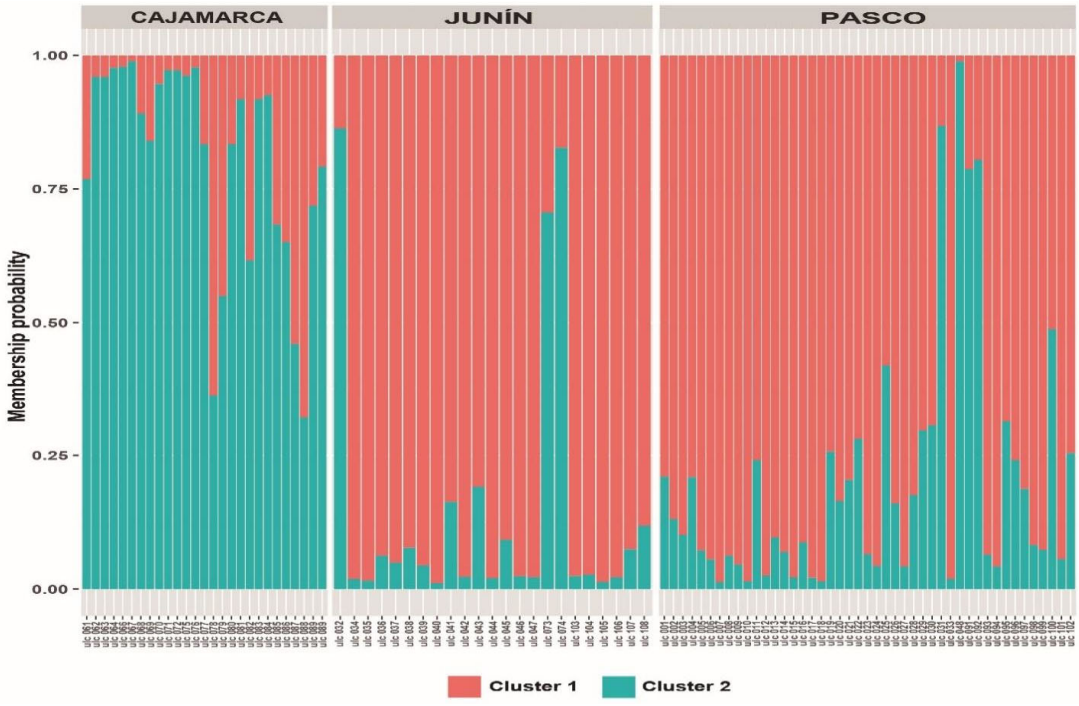

Fig. 6. Structure of 95 samples of a Peruvian collection of ulcumano inferred by STRUCTURE analysis using nine RAPD markers. Cajamarca, Junín and Pasco refer to the geographic origin of the individuals.

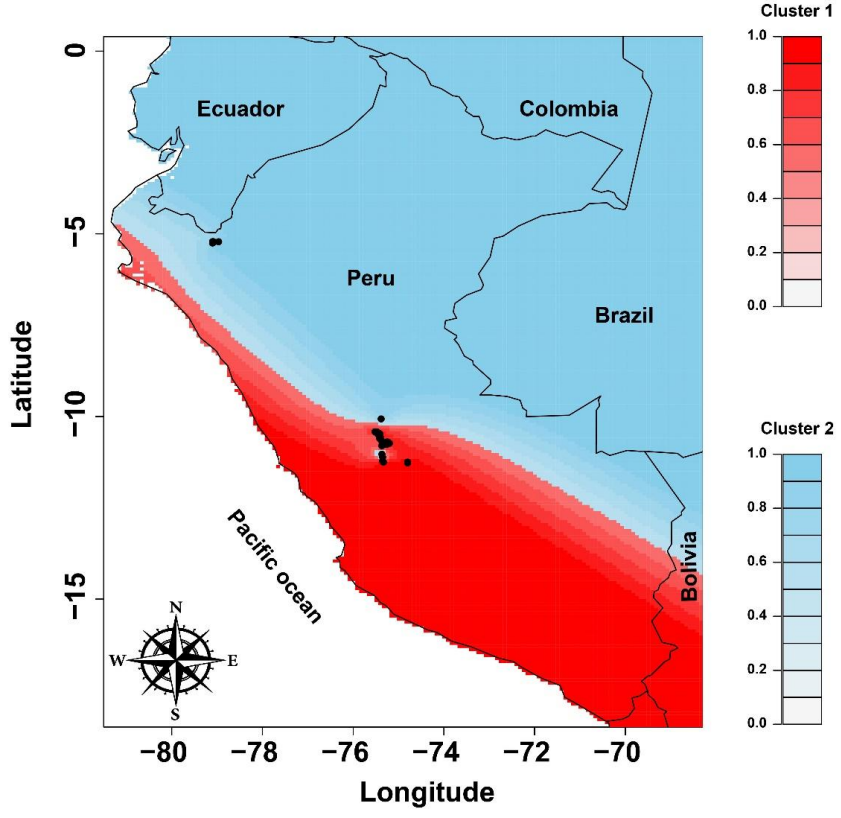

Fig. 7. Geographic map of membership matrix of STRUCTURE analysis employing 95 samples of ulcumano based on nine RAPD markers. Black dots represent ulcumano genotypes sampled.

different individuals could help mitigate the negative effects of low diversity.

In agreement with the results obtained for population structure, principal coordinate analysis also showed two different clusters associated with a geographic component (central and northern Peru) of ulcumano individuals. However, only 13.34% of the variance was captured by the first two axes (Figs. 3 and 6). One of the clusters mostly contains samples from one geographic region (Cajamarca), while the second cluster was mainly composed of samples from two geographical regions (Junín and Pasco) (Fig. 6). Since Junín and Pasco are contiguous departments, it is very likely that farmers living in those locations exchange genetic materials. Moreover, this kind of population structure was also observed for the *Retrophyllum* population in New Caledonian, where some individuals were more genetically related to others in a different geographical region rather than to individuals of the same locality (Herbert et al., 2002). Furthermore, population structure of three *Podocarpus* species was found to be mixed between different temperatures and elevations (Ornelas et al., 2019). Probably, mixed populations of related organisms allow better use of environmental resources as shown by other organisms (Frenkel et al., 2015). Including individuals of ulcumano from other localities is required to conduct extra spatial population structure analysis and allow concluding about the whole area of Peru.

AMOVA showed that the greatest variation exists (87.31%) within populations of ulcumano. This may indicate that ulcumano follows a sexual propagation method. However, pollination and seed dispersal have been little studied on podocarps (Cernusak et al., 2011). Moreover, this species is likely to produce seeds that are dispersed over long distances by wind or water. There is a total of 12.69% of genetic variation between the two clusters of ulcumano. This agrees with the analysis of Dantas et al. (2015), who also mentioned concerns about low variability among populations of *P. sellowii*. On the other hand, a study on three *Podocarpus* species in Mesoamerica showed that most of the diversity is explained by variation between populations rather than within populations. Measures to revert those results in *Retrophyllum* populations in Peru are urgently needed.

Next generation sequencing (NGS) is a modern and reliable tool, which is widely employed on many plant species. do Nascimento Vieira (2016) used NGS to characterize the plastome sequence of the endemic Amazonian conifer, *R. piresii*, and identified 120 genes. This genome will facilitate phylogenetic and germplasm characterization within the Podocarpaceae family. Regarding other forest species, NGS application in not widely used yet (González, 2015; Durán, 2017). Therefore, the next step is to develop additional molecular tools for ulcumano using NGS approaches, same as the process currently conducted for capirona (Saldaña et al., 2022). In addition, we plan to conduct a *de novo* transcriptome study to identify EST-SSR markers to benefit the establishment of a modern breeding program of ulcumano.

CONCLUSIONS

This study showed that RAPD markers were effective for the analysis of the population structure and genetic diversity of *R. rospigliosii* from a Peruvian collection. Three different indices were used, while low levels of genetic diversity were evidenced. In addition, ulcumano samples were grouped into two clusters according to their geographic origin. However, a few samples were intermingled, probably because farmers living in the area exchange seeds. Extra molecular tools should be developed for this tree species using NGS techniques in order to implement a modern breeding program of forest species in Peru.

ACKNOWLEDGEMENTS

This research was funded by the National Agrarian Innovation Program (PNIA)—Project No. 122-PI. C.I.A. and C.L.S. were supported by PP0068 "Reducción de la vulnerabilidad y atención de emergencias por desastres".

We would like to thank Miriam Correa for supporting the technical activities in the laboratory.

LITERATURE CITED

Araya, E., O. Murillo, G. Aguilar, y O. Rocha. 2005. Uso de marcadores genéticos en silvicultura clonal. Revista Forestal Kurú 2:1-14.

Arteaga, M.N., S.M. Tafur, G.P. Pérez, S.A. Pastor, and A. Batista, 2020. Characterization of colonization by micorhizae in *Retrophyllum rospigliossi* Pilger in the Huamantanga forest, Peru. Revista Cubana de Ciencias Forestales 8:535-549.

Caye, K., T.M. Deist, H. Martins, O. Michel, and O. François. 2016. TESS3: Fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources 16:540–548. doi:10.1111/1755-0998.12471.

Cernusak, L.A., H. Adie, P.J. Bellingham, E. Biffin, and T.J. Brodribb. 2011. Podocarpaceae in Tropical Forests: A Synthesis. Smithsonian Contributions to Botany 95:189–195. doi:10.5479/si.0081024X.95.189.

Cheliak, W.M. 1993. Clone identification. In: Clonal Forestry I, Ahuja, M.R., Libby, W.J., Eds.; Springer-Verlag, Berlin/Heidelberg, Germany. 107-164.

Chesnokov, Y.V., and A.M Artemyeva. 2015. Evaluation of the measure of polymorphism information of genetic diversity. Agricultural Biology 50:571–578. doi:10.15389/agrobiology.2015.5.571eng.

Cueva, N.C., D. Vélez, A. Barrios y V. Nieto, 2013. Pino romerón [*Retrophyllum rospiglio*sii (Pilger) C.N. Page], especie nativa potencial para la reforestación en zonas altoandinas de Colombia. Bogotá, Colombia: Corporación Nacional de Investigación y Fomento Forestal (CONIF®), Ministerio de Agricultura y Desarrollo Rural (MADR), Colegio Integrado Nacional Oriente de Caldas (CINOC). chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://iescinoc.edu.co/wp-content/uploads/Silvicultura-del-pino-romeron-CINOC.pdf

Cueva, N.C. y E. Trujillo. 2016. Biología reproductiva del pino romerón-*Retrophyllum rospigliosii* (Pilg.) C.N. Page. Pensilvania, Colombia, Institución de Educación Superior Colegio Integrado Nacional Oriente de Caldas (IES CINOC). http://iescinoc.edu.co/wp-content/uploads/Biologia-reproductiva-del-pino-romeron.pdf

Dantas, L., T. Esposito, A. Barbosa, L. Feliz, and L. Amorim. 2015. Low genetic diversity and high differentiation among relict populations of the neotropical gymnosperm *Podocarpus sellowii* (Klotz.) in the Atlantic Forest. Genetica 143:21-30. doi:10.1007/s10709-014-9809.

do Nascimento Vieira, L., M. Rogalski, H. Faoro, H. Pacheco de Freitas Fraga, K. Goulart dos Anjos, G.F. Assine Picchi, R. Onofre Nodari, F. de Oliveira Pedrosa, E. Maltempi de Souza, and M.P. Guerra. 2016. The plastome sequence of the endemic Amazonian conifer, *Retrophyllum piresii* (Silba) C.N. Page, reveals different recombination events and plastome isoforms. Tree Genetics & Genomes 12:1-11. doi: 10.1007/s11295-016-0968-0.

Doyle, J.J., and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15.

Durán, R.F. 2017. Modelos de predicción genómicos para la selección de genotipos de *Eucalyptus globulus* en base a densidad de la madera y volumen. 170 p. Ph.D. Thesis, Universidad de Concepción, Concepción, Chile. http://repositorio.udec.cl/jspui/bitstream/11594/2750/3/Tesis_Modelos_de_prediccion_genomicos.pdf

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620. doi:10.1111/j.1365-294X.2005.02553.x.

Farjon, A.A. 2010. Handbook of the World's Conifers. Volume I, II Brill: Leiden, Boston, USA.

Francis, R.M. 2017. pophelper: an R package and web app to analyse and visualize population structure. Molecular Ecology Resources 17:27-32. doi:10.1111/1755-0998.12509.

Frenkel, E., M. McDonal, D. Van Dyken, K. Kosheleva, and G. Lang. 2015. Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations. Proceedings of the National Academy of Sciences of the United States of America 112:11306-11311. doi:10.1073/pnas.1506184112.

Gardner, M., and P. Thomas. 2013. *Retrophyllum rospigliosii*. The IUCN Red List of Threatened Species 2013: e.T34110A2846471. IUCN.

Gómez, M.L., J.L. Toro, and E. Piedrahita. 2013. Propagación y conservación de especies arbóreas nativas. Corporación Autónoma Regional del Centro de Antioquia, Corantioquia. Medellín: Corantioquia, Colombia. https://www.corantioquia.gov.co/wp-content/uploads/2022/01/Arboreas-Nativas.pdf

Gómez, A., M. Bueno, R. Alía y G. Vendramin. 1997. Uso de RAPD y microsatélites como marcadores en *Pinus halepensis*: Estudios Preliminares. Cuadernos de la S.E.C.F. 5: 237-241. http://secforestales.org/publicaciones/index.php/cuadernos_secf/article/view/9137

Goncalvez, F.R. 2008. Estructura genética em populações naturais de *Podocarpus sel-lowii* Klotzch (Podocarpaceae) na região do Alto Rio Grande, Sul de Minas Gerais. Dissertação de mestrado. Lavras: UFLA. Brasil. http://repositorio.ufla.br/bitstream/1/2721/1/DISSERTA%C3%87%C3%83O_Estrutura%20gen%C3%A9tica%20em%20popula%C3%A7%C3%B5es%20naturais%20de%20Podocarpus%20sellowii%20Klotzsch%20(Podocarpaceae)%20na%20regi%C3%A3o%20do%20Alto%20Rio%20Grande,%20Sul%20de%20Minas%20Gerais.pdf

González, N.A. 2015. Identificación y validación de single nucleotide polymorphism (SNPs) distribuidos en el genoma de *Eucalyptus globulus*. 62 p. Master's Thesis, Universidad de Concepción, Concepción, Chile. http://repositorio.udec.cl/xmlui/handle/11594/1829

Goyal, P., R. Jain, S., Kachhwaha, and S.L. Kothari. 2014. Assessment of genetic diversity in *Pithecellobium dulce* (Roxb.) Benth. germplasm using RAPD and ISSR markers. Trees 29: 637-653. doi:10.1007/s00468-014-1141-8.

Herbert, J.M., P. Hollingsworth, M. Gardner, R. Mill, and P. Thomas. 2002. Conservation genetics and phylogenetics of New Caledonian *Retrophyllum* (Podocarpaceae) species. New Zealand Journal of Botany 40:175-188. doi:10.1080/0028825X.2002.9512781.

Kamvar, Z.N., J.F. Tabima, and N.J. Grünwald. 2014. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:281. doi:10.7717/peerj.281.

Mill, R. 2016. A Monographic Revision of *Retrophyllum* (Podocarpaceae). Edinburgh Journal of Botany 73:171–261. doi:10.1017/S0960428617000178.

Ornelas, J., A. Ortiz-Rodriguez, E. Ruiz-Sanchez, V. Sosa, and M. Perez-Farrera. 2019. Ups and downs: Genetic differentiation among populations of the *Podocarpus* (Podocarpaceae) species in Mesoamerica. Molecular Phylogenetics and Evolution 138: 17-30. doi:10.1016/j.ympev.2019.05.025.

Peakall, R., D. Ebert, L.J. Scott, P.F. Meagher, and C.A. Offord. 2003. Comparative genetic study confirms exceptionally low genetic variation in the ancient and endangered relictual conifer, *Wollemia nobilis* (Araucariaceae). Molecular Ecology 12:2331-2343. doi:10.1046/j.1365-294x.2003.01926.x.

Pritchard, J.K., M. Stephens, N.A. Rosenberg, and P. Donnelly. 2000. Association mapping in structured populations. American Journal of Human Genetics 67:170-181. doi:10.1086/302959.

Pujana, R., P. Wilf, and M. Gandolfo. 2020. Conifer wood assemblage dominated by Podocarpaceae, early Eocene of Laguna del Hunco, central Argentinean Patagonia. PhytoKeys 156:81-102. doi:10.3897/phytokeys.156.54175.

Quiroga, M. 2008. Contribución para la conservación de Podocarpaceae del sur de Sudamérica a partir de patrones genéticos y biogeográficos. 197 p. Ph.D. Thesis, Universidad Nacional del Comahue, Argentina.

Renau-Morata, B., G. Nebauer, E. Sales, J. Allainguillaume, and P. Caligari. 2005. Genetic diversity and structure of natural and managed populations of *Cedrus atlantica* (Pinaceae) assessed using random amplified polymorphic DNA. American Journal of Botany 92:875-884. doi:10.3732/ajb.92.5.875.

Reynel, C., and J. Marcelo. 2009. Árboles de los ecosistemas forestales andinos. Manual de identificación de especies. Ecosistemas Forestales 118-122.

Saldaña, C.L., J.D. Cancan, W. Cruz, M.Y. Correa, M. Ramos, E. Cuellar, and C.I. Arbizu. 2021. Genetic diversity and population structure of capirona (*Calycophyllum spruceanum* Benth.) from the Peruvian Amazon revealed by RAPD markers. Forests 12:1125. doi:10.3390/f12081125.

Saldaña, C.L., P. Rodriguez-Grados, J.C. Chávez-Galarza, S. Feijoo, J.C. Guerrero-Abad, H.V. Vásquez, J.L. Maicelo, J.H. Jhoncon, and C.I. Arbizu. 2022. Unlocking the complete chloroplast genome of a native tree species from the Amazon Basin, capirona (*Calycophyllum spruceanum*, Rubiaceae), and its comparative analysis with other Ixoroideae species. Genes 13:113. doi:10.20944/preprints202111.0533.v1.

Sotolongo, R., G. Geada, y M. Cobas. 2013. Mejoramiento Genético Forestal, Texto para estudiantes de Ingeniería Forestal. FAO. https://www.fao.org/fileadmin/user_upload/training_material/docs/Mejoramiento%20Genetico%20Forestal.pdf

Stark, D.M. 2005. Estudios Taxonómicos y Filogenéticos de las especies de *Podocarpus* del Caribe y América Central. 76 p. Bachelor's thesis, Austral University of Chile, Valdivia, Chile.

Su, Y., T. Wang, and F. Deng. 2010. Contrasting genetic variation and differentiation on Hainan Island and the Chinese mainland population of *Dacrycarpus imbricatus* (Podocarpaceae). Biochemical Systematics and Ecology 38:576-584. doi:10.1016/j.bse.2010.07.003.

Supreme Decree No 043-2006-AG. 2006. El Peruano newspaper, Lima, Perú, July 13th.

Tijerino, A., L. Callejas, and D.A. Cerda-Granados. 2016. Assessment of genetic diversity in five Nicaraguan populations of *Cedrela odorata* (Meliaceae) using RAPD markers. Encuentro 103: 28–39. doi:10.5377/encuentro.v0i103.2690.

Valadez, E., y G. Kahl. 2000. Huellas de ADN en genomas de plantas: teoría y protocolos de laboratorio. Mundi-Prensa, S.A de C.V. México D.F.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535. doi:10.1093/nar/18.22.6531.

Zapata, J., and R. Hasbun. 2011. Mejoramiento genético forestal acelerado mediante selección genómica. Bosque 32:209-213. doi:10.4067/S0717-92002011000300001.

Zenteno-Ruíz, F.S. 2007. *Retrophyllum rospigliosii* (Podocarpaceae), un nuevo registro de Pino de Monte, en el noroeste de Bolivia. Kempffiana 3:3-5.