biology and life sciences forum

MDPI

Proceeding Paper

Genetic Diversity and Population Structure Assessed by SSR in a Peruvian Germplasm Collection of Loche Squash (*Cucurbita moschata*, Cucurbitaceae) †

Carlos I. Arbizu [1,*], Raúl H. Blas [2] and Roberto Ugás [2]

[1] Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru
[2] Facultad de Agronomía, Universidad Nacional Agraria la Molina, Av. La Molina s/n, Lima 15024, Peru; rblas@lamolina.edu.pe (R.H.B.); rugas@lamolina.edu.pe (R.U.)
* Correspondence: carbizu@inia.gob.pe; Tel.: +51-9-86288181
† Presented at the 2nd International Electronic Conference on Diversity (IECD 2022)—New Insights into the Biodiversity of Plants, Animals and Microbes, 15–31 March 2022; Available online: https://sciforum.net/event/IECD2022.

Abstract: Loche is an ancient landrace of squash from Northern Peru, notable for its vegetative reproduction and lack of seeds in fruits. To date, very little is known about its genetics. Here, we used 21 simple sequence repeats to assess the genetic diversity and population structure of a collection of 100 samples of loche from three localities in Peru, and 10 samples of related species, *C. pepo* and *C. maxima* (110 accessions in total). A total 85 bands were manually scored, obtaining an average of 4.05 alleles per locus. The UPGMA clustering method and principal coordinate analysis showed a clear identification between the three species of *Cucurbita*. Population structure analysis clustered the 110 accessions into the following five populations: (i) three of loche, (ii) one of *C. pepo*, and (iii) one of *C. maxima*. Genetic diversity estimation was conducted considering only the three groups (populations) of loche identified, which was 0.024 as an average. AMOVA revealed the greatest variation between populations (79.66%) and indicated that variability within populations is 20.33%. Vegetative propagation by means of stem cuttings and cultivation in a very restricted geographical area would explain the rather low diversity of loche. This in turn would suggest that the apparent variation observed in fruit shape may be explained by somatic mutation and/or environmental factors.

Keywords: germplasm; microsatellites; genetic resources

check for updates

Citation: Arbizu, C.I.; Blas, R.H.; Ugás, R. Genetic Diversity and Population Structure Assessed by SSR in a Peruvian Germplasm Collection of Loche Squash (*Cucurbita moschata*, Cucurbitaceae). *Biol. Life Sci. Forum* **2022**, *15*, 6. https://doi.org/10.3390/IECD2022-12420

Academic Editor: Ben-Erik Van Wyk

Published: 15 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The loche squash (*Cucurbita moschata*) is a vegetable that is grown exclusively and traditionally in the north coast of Peru, mainly in the department from Lambayeque and is practically unknown elsewhere [1]. However, it is necessary to promote this landrace since it constitutes one of the crops of the Moche culture and can be used as a new product for international markets, recognizing the gastronomy of Lambayeque in the world. Currently, agro-export crops threaten to displace this exceptional pumpkin from its production area [1]. There are no studies on the genetic diversity of the loche squash in Peru. This study represents a fundamental step to understanding the way loche is structured. In addition, the estimation of genetic diversity is crucial for genetic improvement, conservation of species and gain in selection. To date, genetic diversity of loche cultivated in northern Peru is unknown. Furthermore, we also do not know its phenotypic diversity, that is, we do not we know how many cultivars are planted by farmers in northern Peru.

Here, we successfully employed 21 SSRs for the genetic characterization of loche from three departments of Peru, Amazonas, Lambayeque and Pasco, to initiate a modern genetic improvement program and also to promote better conservation strategies for this species.

2. Materials and Methods

2.1. Samples Examined and DNA Amplification

Young fresh leaves from 100 individuals of loche from Amazonas, Lambayeque and Pasco departments in Peru were collected, considering their natural distribution range, and 10 samples of related species, *C. pepo* and *C. maxima* (110 accessions in total) as control. Samples were handled according to the protocol followed by Saldaña et al. [2,3]. Loche DNA extraction was performed by the method of Doyle and Doyle [4] with minor modifications. We used a lower concentration of ethanol in washes (both at 70%) and shorter centrifugation time. To determine the genetic diversity, we tested 21 SSRs. The amplification procedure was conducted in a final volume of 10 µL containing 5 ng of DNA. For the PCR reaction, we followed the steps of Saldaña et al. [2]. Amplified products size was tested with 100 bp marker (New England Biolabs, MA, USA).

2.2. Data Analysis

SSR band patterns were inspected to score the presence (1) or absence (0) of these. To construct a dendrogram considering the UPGMA clustering algorithm, we employed R to calculate the Nei's genetic distances by using the *poppr* package v2.9.2 [5]. Additionally, 1000 bootstrap replicates were conducted. A principal coordinate analysis (PCoA) and a Bayesian approach to infer the genetic population structure was performed in R and STRUCTURE software [6], respectively. We used the R package *poppr* considering the number of clusters inferred by STRUCTURE to conduct an analysis of molecular variance (AMOVA).

3. Results and Discussion

3.1. Data Analysis

The 21 SSR primers utilized for the molecular analysis revealed 85 fragments in 110 samples of *Cucurbita* spp., with 18.6 fragments as average with an average of 4.05 alleles per locus.

3.2. Genetic Diversity Estimates and Population Structure Analysis

With the 85 scored fragments, we constructed a 110 × 95 presence–absence data set. The Provesti's genetic distances based UPGMA tree did not clearly discriminate loche samples according to their geographic locality (Amazonas, Lambayeque, Pasco). A total of two clusters of loche are supported by our dendrogram, but they present a bootstrap support of lower than 70% (Figure 1). The first two axes of the PCoA explained that 60.81% of the variation is in agreement with our dendrogram, and showed that some samples of loche are intermingled. As expected, a clear separation among *Cucurbita* species was observed (Figure 2).

The Evanno method [7] depicted that the best K value (number of populations) is five for our data set: (i) cluster 1 includes loche from Pasco, (ii) cluster 2 is composed of loche from Lambayeque and Amazonas, (iii) cluster 3 contains samples of *C. maxima*, (iv) cluster 4 corresponds to samples of *C. pepo*, and (v) cluster 5 is composed of loche from Lambayeque and Amazonas. The STRUCTURE analysis demonstrated admixture for few samples. This analysis also confirmed that loche samples are not clustered based on the geographic origin (Figure 3).

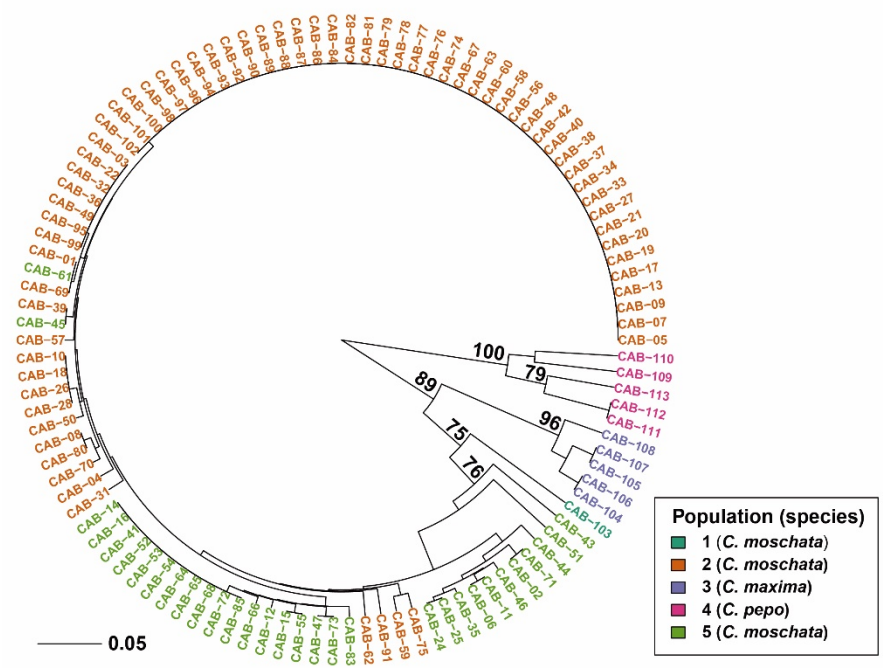

Figure 1. Dendrogram of 110 samples of *Cucurbita* using 21 SSR markers based on UPGMA clustering method and Provesti's genetic distance. Numbers above the branches represent bootstrap values, with only values higher than 70% shown.

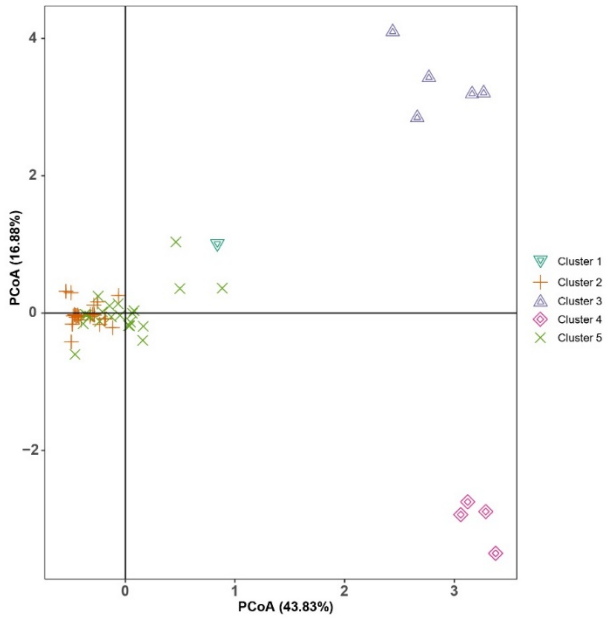

Figure 2. Principal coordinates analysis (PCoA) of 110 *Cucurbita*. Symbols and colors refer to clusters assigned by STRUCTURE software.

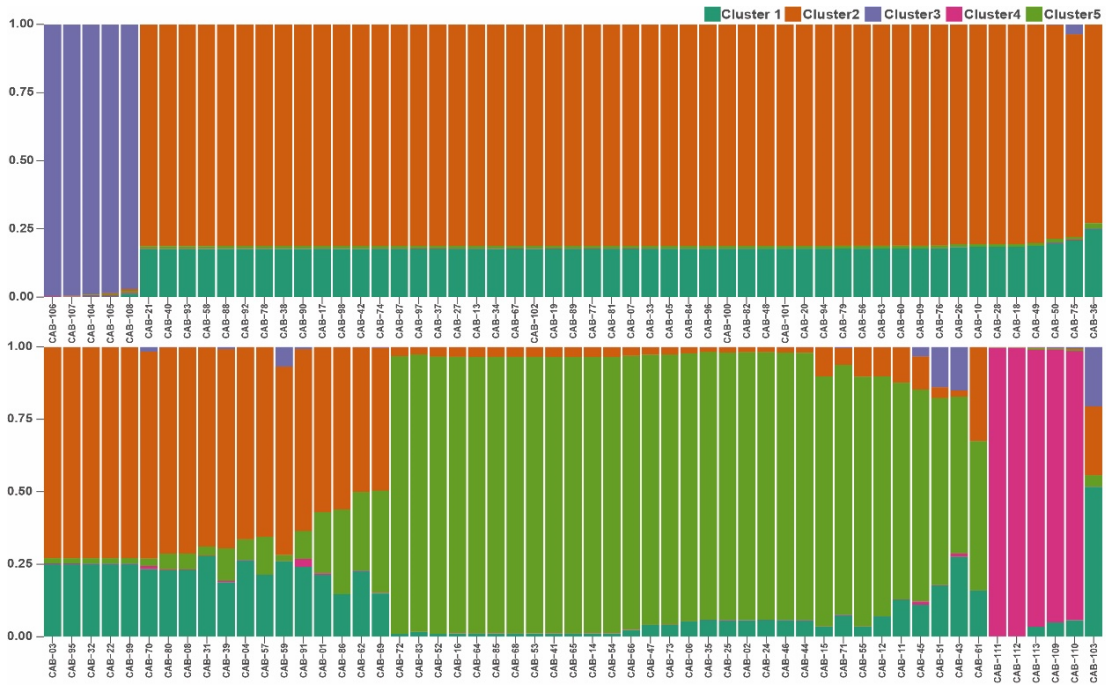

Figure 3. Population structure analysis of 110 samples of a Peruvian collection of loche and related species estimated by the software STRUCTURE employing 21 SSRs.

The genetic diversity indices were estimated by considering the five clusters (K = 5) inferred by STRUCTURE program. The Shannon.Wiener index ranged from 1.33 to 2.91, and Simpon´s index varied from 0.72 to 0.92. The expected heterozygosity (Nei´s genetic diversity index) was 0.012 for cluster 2 and 0.144 for cluster 4 (Table 1). The AMOVA revealed the greatest variation between populations (79.66%) and indicated that variability within populations is 20.33% (Table 2).

Table 1. The genetic diversity parameters based on 21 SSR markers in five clusters.

Cluster	N	H	Lambda	He
1	1	0	0	0
2	68	2.16	0.721	0.012
3	5	1.61	0.8	0.089
4	5	1.33	0.72	0.144
5	31	2.91	0.92	0.035
Total	110	3.26	0.886	0.0973

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

Table 2. Analysis of molecular variance (AMOVA) for the five clusters identified in STRUCTURE analysis.

Source	df	SS	MS	Est.Var.	%
Between clusters	4	365.25	91.31	6.11	79.66%
Within clusters	105	163.75	1.56	1.56	20.34%
Total	109	528.99	4.85	7.67	100.00%

Loche is an orphan landrace as it has been under-researched during recent years, even though it possesses traits of high agronomic and economic potential. Genetic studies of this landrace are missing. In Peru, most research is focused on its botany [8]. Ferriol et al. [9]

used AFLP and SRAP to study the genetic diversity of a collection of pumpkin germplasm (*C. moschata*) maintained in Spain, finding that the accessions studied were grouped according to their geographical origin. This suggests the existence of two independent centers of domestication in the American continent, and/or introgression of species related to *C. moschata*. Wu et al. [10] studied the genetic diversity of this species using AFLP markers and 74 accessions from China, and 15 accessions from other countries. Accessions from China were classified into two groups, differing from the accessions from Mexico, Guatemala, Honduras, and Ecuador. The authors concluded that the American accessions present a higher number of loci than those from China. Genetic diversity of accessions of *C. moschata* from Mesoamerica were studied by chloroplast DNA sequences, demonstrating a high level of genetic diversity especially in the Mexican germplasm [11]. Most of the genetic variation among squash germplasm collections was attributed to variations between individuals within of the respective localities. In the present research study, we observed that the genetic diversity of the loche in Peru is quite low. These results appear to be consistent with populations of almost homogeneous plants observed in commercial fields of the provinces of Lambayeque. Since loche is vegetatively propagated, very few closely related clones are being cultivated in the communities of Pacora and Illimo, as well as in other localities of Peru. Therefore, there is no opportunity for genetic recombination. It is most likely that the morphological variation arose from somatic mutation, which is maintained and propagated by cuttings. On the other hand, the "loche de montaña" that is cultivated in Bagua, Amazonas does not have genotypic variation of the loche grown in Lambayeque. The main reason for this is that the cuttings used by the farmers in Bagua would be the same as those employed in Lambayeque because farmers who decided to live in Bagua took their vegetative seeds from loche from there in order to continue cultivating loche. Therefore, the little morphological difference that exists between these fruits of loche cultivated in these two localities can be explained by environmental factors. It was also possible to confirm the transferability of the microsatellite markers within the genus *Cucurbita*, as described for other genera including *Phaseolus* and Vigna [12], among others.

Author Contributions: Conceptualization, R.U. and R.H.B.; methodology, C.I.A., R.H.B.; formal analysis, C.I.A.; investigation, C.I.A., R.H.B.; resources, R.H.B., R.U.; data curation, C.I.A.; writing—original draft preparation, C.I.A.; writing—review and editing, C.I.A., R.U., R.H.B.; supervision, R.H.B.; project administration, R.U., R.H.B.; funding acquisition, R.U. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by VLIR/UNALM.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Luis Gutierrez and Joel Flores for supporting the technical activities in the laboratory.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Andres, T.; Ugás, R.; Bustamente, F. *Loche: A Unique Pre-Columbian Squash Locally Grown in North Coastal Peru*; Holmes, G.J., Ed.; Proc. Cucurbitaceae; Universal Press: Raleigh, NC, USA, 2006; pp. 333–340.
2. Saldaña, C.L.; Cancan, J.D.; Cruz, W.; Correa, M.Y.; Ramos, M.; Cuellar, E.; Arbizu, C.I. Genetic diversity and population structure of capirona (*Calycophyllum spruceanum* benth.) from the peruvian amazon revealed by rapd markers. *Forests* **2021**, *12*, 1125. [CrossRef]
3. Calycophyllum, C.; Saldaña, C.L.; Rodriguez-grados, P.; Ch, J.C.; Feijoo, S.; Guerrero-abad, J.C.; Vásquez, H.V.; Maicelo, J.L.; Jhoncon, J.H.; Arbizu, C.I. Unlocking the Complete Chloroplast Genome of a Native Tree Spruceanum, Rubiaceae), and Its Comparative Analysis with Other Ixoroideae Species. *Genes* **2022**, *13*, 113.
4. Doyle, J.J.; Doyle, J.L. A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.

5. 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**, *2014*, e281. [CrossRef] [PubMed]
6. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef] [PubMed]
7. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef] [PubMed]
8. Bazo, S.I.; Espejo, J.R.; Palomino, A.C.; Flores, P.M.; Chang, L.M.; López, B.C.; Mansilla, S.R. Estudios de biología floral, reproductiva y visitantes florales en el "Loche" de Lambayeque (*Cucurbita moschata* DUCHESNE). *Ecol. Apl.* **2018**, *17*, 191–205. [CrossRef]
9. Ferriol, M.; Picó, B.; Nuez, F. Morphological and Molecular Diversity of a Collection of *Cucurbita maxima* Landraces. *J. Am. Soc. Hortic. Sci.* **2004**, *129*, 60–69. [CrossRef]
10. Wu, J.; Chang, Z.; Wu, Q.; Zhan, H.; Xie, S. Molecular diversity of Chinese *Cucurbita moschata* germplasm collections detected by AFLP markers. *Sci. Hortic.* **2011**, *128*, 7–13. [CrossRef]
11. Barboza, N.; Albertazzi, F.J.; Sibaja-Cordero, J.A.; Mora-Umaña, F.; Astorga, C.; Ramírez, P. Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data. *Sci. Hortic.* **2012**, *134*, 60–71. [CrossRef]
12. Yu, K.; Park, S.J.; Poysa, V. Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome* **1999**, *42*, 27–34. [CrossRef]