Genetic diversity of arracacha (*Arracacia xanthorrhiza*) in Peru

Jorge Biondi[1], Cinthya Zorrilla[*], Iván Manrique[2], Carlos Arbizu[2], William Roca[2], Tulio Medina[3], Juan Seminario[4], José Quispe[5], David Tay[2], Raúl Blas[1**]

1 Universidad Nacional Agraria La Molina (UNALM). Av. La Molina s/n., Lima, Perú
2 Centro Internacional de la Papa (CIP). Av. La Molina 1895, La Molina, Lima, Perú
3 Instituto Nacional de Innovación Agraria (INIA). Av. La Molina 1981, Lima, Perú
4 Universidad Nacional de Cajamarca (UNC). Av. Atahualpa 1050, Cajamarca, Perú
5 Universidad Nacional de San Cristóbal de Huamanga (UNSCH). Portal Independencia 57, Ayacucho, Perú
* Current address: University of Wisconsin-Madison, Department of Horticulture. 1575 Linden Drive Madison, WI 53706
** Corresponding author: biondi_th_j@hotmai.com

Abstract

Arracacha (Umbelliferae family) is a traditional Andean root crop that has been used as a food for thousands of years. It is an important source of calcium and contains a special starch easily digestible. It has become a significant crop in Brazil, where 12,000 ha are grown annually. Peru is the country with the largest gene bank collections, but the diversity is poorly understood. The aim of this work was to study the genetic diversity of 334 accessions from five gene banks collections in Peru using 148 polymorphic AFLP markers, obtained from five primer combinations. The molecular variance analysis (AMOVA) and the genetic dissimilarity (Nei and Shannon-Weaver index) indicated that arracacha in Peru is diverse. It was found that all five collections were significantly different among them and no duplicates were identified among 334 accessions studied. The UPGMA dendogram showed three clusters related to geographical distribution patterns of arracacha diversity: northern, central and southern of Peru. The northern cluster contains a much higher diversity than other clusters.

Keywords: *Arracacia xanthorrhiza*, AFLP, genetic diversity, gene banks.

Introduction

Arracacha is a starchy food and an important source of β-carotene, ascorbic acid and calcium. Several studies point to Andean South America as the place of domestication of arracacha. Although the genus *Arracacia* is particularly diverse in Mexico, the wild species most closely resembling arracacha are known from Peru and especially Ecuador. Arracacha is produced mainly in four countries: Brazil, Colombia, Ecuador and Venezuela (Hermann, 1997). Preliminary work using PCR amplified DNA from random sequence decamer primers has yielded promising results in terms of molecular polymorphism for application in fingerprinting of arracacha cultivars. Thus Blas *et* al. (1997) and Castillo (1997) report occurrence of DNA polymorphism in 48% and 85% of primer assayed, respectively. Castillo, however, concludes from his work that overall molecular diversity in his Ecuadorean material is low. Another work is the study of Blas (2005), where compared cultivated with wild species of arracacha, but there are lot of publications about morphological characterization. The AFLP analysis is considered by now as the most efficient technique to study the DNA in the case of species in which the genome is little known. In Peru there are arracacha's gene banks in different zones, each gene bank collected arracacha mostly from their zones; in the last years some accessions have been lost and then they agreed to make one gene bank with all the accessions. The purpose of the present study is to develop the AFLP technique on the arracacha genome to assess its genetic diversity of this set of gene bank and give information of what is the genetic diversity of arracacha in Peru.

Materials

A total of 334 accessions of arracacha of five institutions of Peru were tested (Table 1).

Table 1. Number of accessions per arracacha gene banks

Name of Institutions that maintain arracacha accessions	Number of accessions
Centro Internacional de la papa (**CIP)**	48
Instituto Nacional de Innovación Agraria (**INIA)**	144
Universidad Nacional de Cajamarca (**UNC)**	121
Universidad Nacional San Cristóbal de Huamanga (**UNSCH)**	20
Universidad Nacional Agraria La Molina (**UNALM)**	1
Total	**334**

Methods

DNA extraction

DNA extraction was made from young leaves using CTAB technique (Doyle and Doyle, 1990) medium scale protocol standardized and reported by CIP (2000). The DNA concentration was calculated in comparison with known concentrations of Phage Lambda DNA digested with *Pst I* restriction enzyme and visualized in agarose gel 1%.

AFLP analysis

AFLP protocols are based on Vos et al. (1995), adapted for use with silver staining of polyacrylamide gels 6%. The protocols of DNA restriction and ligation, PCR amplifications and electrophoresis were used as reported in the manual of the CIP (2000).

Data analysis

The AFLP bands pattern for each accessions were registered in a binary matrix in an excel sheet, where the present bands registered by 1 and the absent bands by 0. The unique genotype recognizing and the identification of unique genotypes and genetic redundancy is achieved through a cluster analysis based in the Jaccard index and UPGMA algorithm. The genetic heterozygosity was estimated using the diversity index of Nei and Shannon, and the genetic structure was estimated with the factorial analysis.

Results

AFLP primers combinations

A screening with 120 primers combination was made for pick up the AFLP primers combination used; these primers were selected due to its high polymorphisms index and good resolution of the bands (Table 2).

Table 2. Primers used for AFLP markers

Laboratory Code	Primer	Total Bands	Polymorphic Bands	Percentage of Polymorphism
EGA-M36	E-GA	36	27	75
	M-ACC			
E32-M52	E-AAC	50	30	60
	M-CCC			
E35-M48	E-ACA	45	38	85
	M-CAC			
E37-M55	E-ACG	40	28	70
	M-CGA			
E38-M60	E-ACT	35	25	72
	M-CTC			
Total		**206**	**148**	**71**

AFLP ,arkers evaluation

AFLP pattern bands were different in all 334 accessions tested. From a total of 148 polymorphic markers obtained by 5 combinations of AFLP primers, 77 are share bands for all collections and only 11 exclusive markers were identified (Table 3).

Table 3. Markers present in each collection

Genebank	No. of present bands	No. of exclusive band	No. of share band
CIP	129	1	77
INIA	139	2	
UNC	140	5	
UNSCH	127	3	
UNALM	78	0	

Cluster analysis

The cluster analysis identified three main molecular groups of diversity. The largest group is formed only for accessions collected from the north of Peru; the other two groups are formed by arracachas from the center and the south, respectively (Figure 1) but some accessions grouped with arracachas from other collections (Figure 2). Also there are 33 accessions that not grouped in any cluster. The factorial analysis showed that these three groups are not very separated between them because there are relations between accessions from north south and center of Peru (Figure 3).

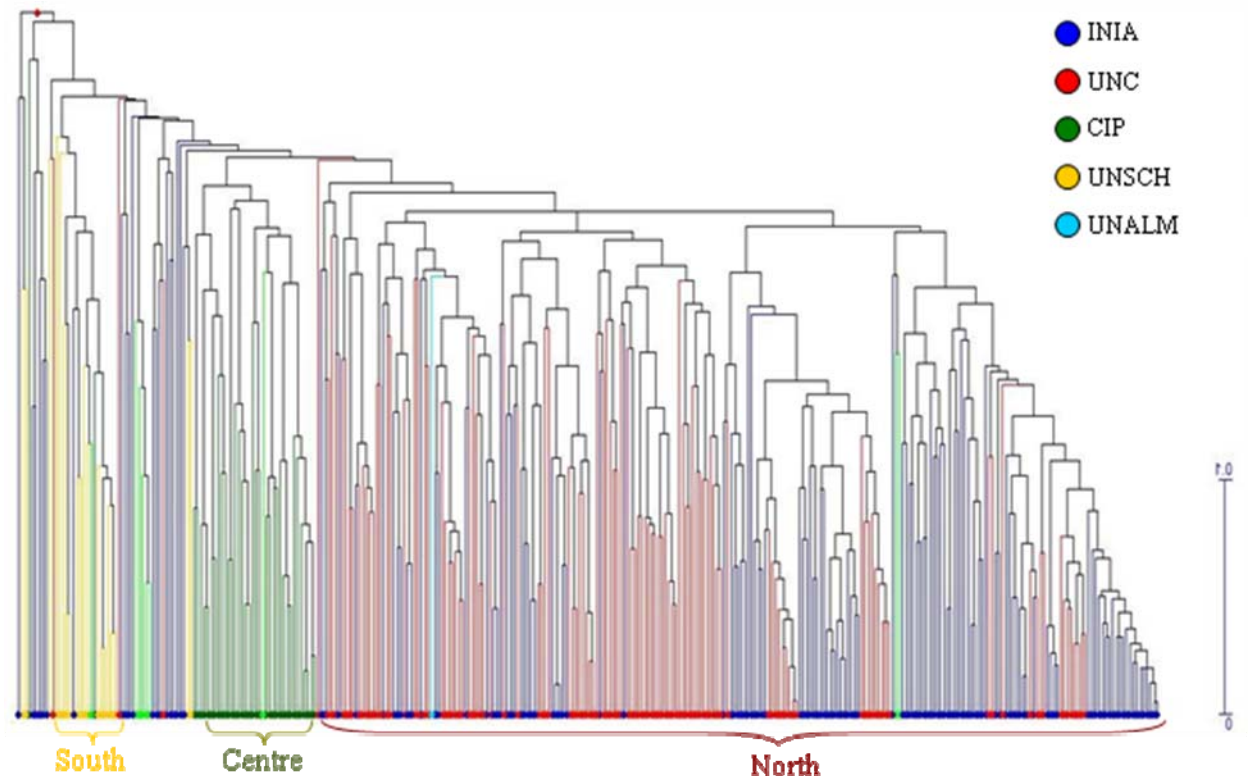

Figure 1. Clusters of the arracacha gene banks, based in 148 AFLP markers using Jaccard index and UPGMA algorithm

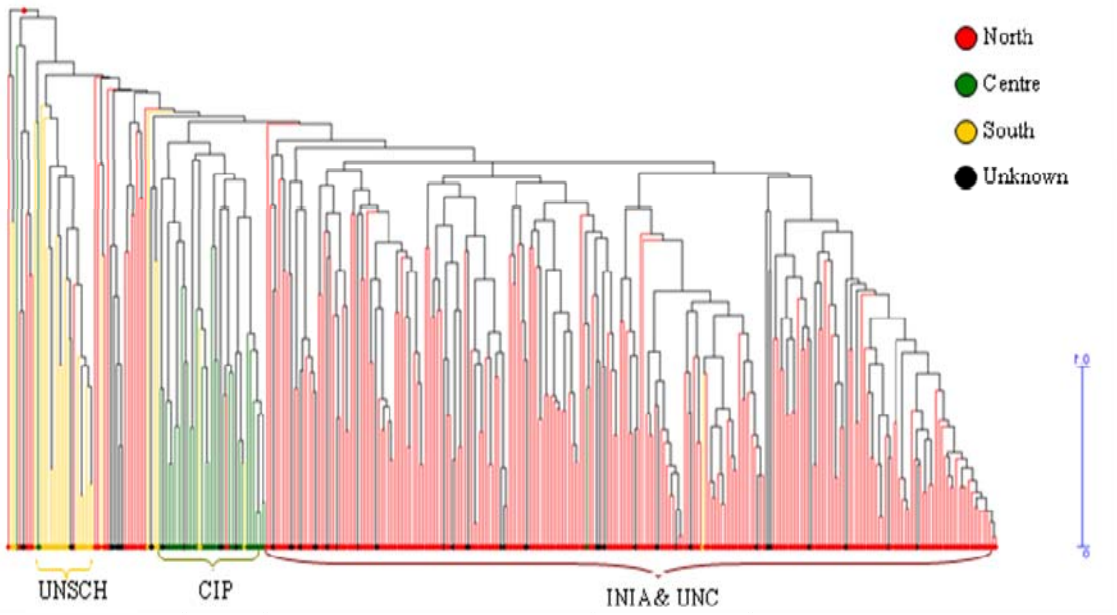

Figure 2. Clusters of the arracacha gene banks per region of origin, based in 148 AFLP markers using Jaccard index and UPGMA algorithm

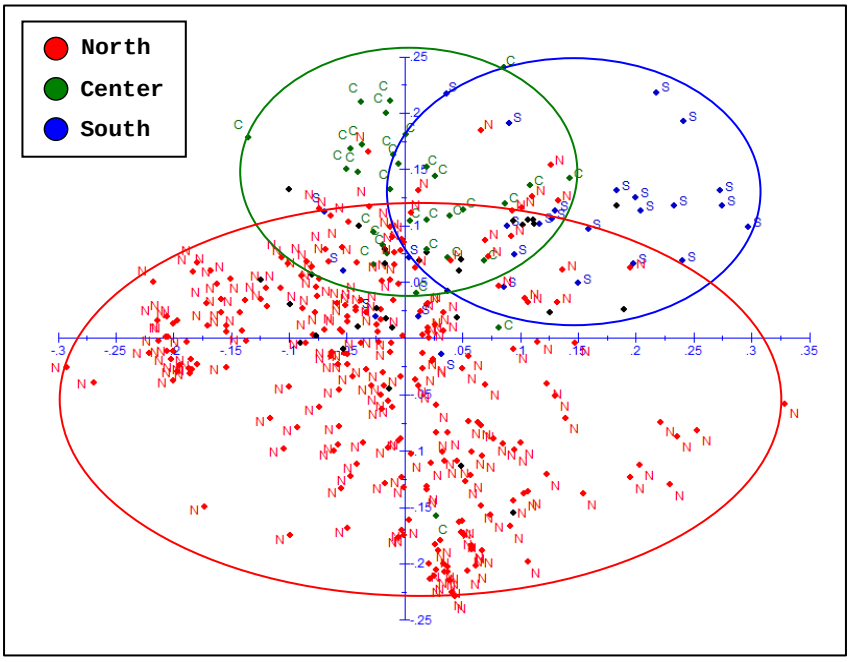

Figure 3. Factorial Analysis of principal coordinates per region of origin

Genetic diversity index

The Shannon-Weaver index (5.59) and the Nei index (0.28) indicate that the genetic diversity is high for all collections (Table 4). The highest values of Nei index were found in INIA and UNC (0.29), which are the largest collections. Although CIP collection is small its genetic diversity resulted high (0.28).

Table 4. Nei and Shannon-Weaver Index for genebank

Genebank	No. of accessions	Nei index	Shannon-Weaver index
CIP	46	0.28	0.849
INIA	145	0.29	2.305
UNC	122	0.29	2.060
UNSCH	20	0.26	0.382
UNALM	1	-	-
Total	**334**	**0.28**	**5.596**

Analysis of Molecular Variance (AMOVA)

The AMOVA shows a significant difference (Table 5) and the genetic distance between collections (Table 6) which shows a significant genetic distance between collections.

Table 5. Percentage of variation per collections

Source of variation	g.l	Percentage of variation
Among collections	4	12.68
Within collections	329	84.32
Fst.		0.12684

Table 6. Genetic distances per collection

	UNSCH	CIP	INIA	UNC
UNSCH	-			
CIP	0.178 (+)	-		
INIA	0.230 (+)	0.154 (+)	-	
UNC	0.239 (+)	0.133 (+)	0.070 (+)	-
Distance average	**0.215**	**0.155**	**0.151**	**0.147**

Also the Table 7 shows a significant difference between regions from Peru and Table 8 shows that there is significant genetic distance between regions.

Table 7. Percentage of variation per regions

Source of variation	g.l	Percentage of variation
Among regions	4	14.93
Within regions	329	85.07
Fst.		0.14928

Table 8. Genetic distances per region

	Centre	North	South
Centre	-		
North	0.139 (+)	-	
South	0.165 (+)	0.159 (+)	-
Distance average	**0.152**	**0.149**	**0.162**

Discussion

Blas (2005) obtained 76 markers with 3 AFLP primer combinations in cultivated arracacha. These primers were not use in the present work because their lack of resolution and high percentage of monomorphic bands. We obtained 148 markers with 5 AFLP primer combinations, which have high resolution and high polymorphism. Although the arracacha has few AFLP markers in comparison with others crops (Kim *et al.,* 1998).

Diversity of arracacha is specially organized by a geographical distribution, which is higher in the north of Peru (Nei index=0.29), where most accessions are from INIA and UNC collection. Besides CIP collection (Nei index=0.28) was more representative from center of Peru because with a low number of accessions has a high index of diversity, probably because the accessions were strategically better collected.

The Analysis of Molecular Variance shows differences among collections, this is possibly because there are genetic differences between the areas of collections, and also, each collection represents nearly one region. Also the analysis shows a high significant genetic distance between collection, being the largest distance between UNSCH (south accession) with INIA and UNC (north accessions) but is closer genetically with CIP collection (principally accessions from center). The INIA and UNC collections are genetically close; this agrees with the shown in the cluster analysis. According with the Fst. Analysis the differences between collections (0.12684) and between regions (0.14928) are moderate.

Conclusion

According to these results, the genetic diversity from Peru seems to be high and the higher genetic diversity is from Northern region. All the 334 accessions analyzed are unique and there are no duplicates in the gene banks. Besides, the genetic differentiation among Northern, Centre and South region was moderate, and accessions from south are more different in relation with the other regions.

References

Arbizu, C. y Robles, E. (1986). *La colección de los cultivos de raíces y tubérculos andinos de la Universidad de Huamanga.* In Anales del V congreso Internacional de Sistemas Agropecuarios Andinos. Puno, Peru.

Blas, R., Arbizu, C., Ghislain, M. and Herrera, R. (1997). *Caracterización de arracachas (*Arracacha xanthorrhiza *Bancroft) peruanas.* Pp. 46-47 in libro de resúmenes IX Congreso Internacional de Cultivos Andinos, Universidad Nacional San Antonio Abad del Cusco, Cusco, 22-25 Abril 1997 (Abstract)

Blas, R. (2005). *Diversity of Arracacia species in Peru.* Dissertation originale presenteé en vue de l´obtention du grade de docteur en Sciences Agronomiques et Ingénierie Biologique. Faculte Universitaire des Sciences Agronomiques de Gembloux.

Castillo, R. (1997). *Caracterizacion molecular de 29 morfotipos de arracacha* (Arracacha xanthorrhiza *Bancroft) de la colección ecuatoriana.* IX Congreso Internacional de Cultivos Andinos, Universidad Nacional San Antonio Abad del Cusco, Cusco, 22-25 Abril 1997 (Abstract).

Centro Internacional de la Papa, CIP (2000). *Molecular Biology Laboratory Protocols: Plant Genotyping.* Ma. Rosario Herrera, M. Ghislain, D. Zhang (eds.) Crop Improvement and Genetic Resources Department Training Manual, tercera edición.

Doyle, J.J. and Doyle, J.L. (1990). *Isolation of plant DNA from fresh tissue*. Focus (Gibco BRL) 12: 13-15.

Excoffier, L., Smouse, P., Quattro, J. (1992). *Analysis of Molecular Variace inferred from metric distance among haplotypes.* Genetics 131: 479-491.

Hermann, M. (1997). Arracacha (*Arracacia xanthorrhiza* Bancroft). In: Hermann M & J Heller (eds.): Andean roots and tubers: Ahipa, arracacha, maca yacon. Promoting the conservation and use of underutilized and neglected crops. 21. Institute of Plant Genetics and Crop Plant Research. Gatersieben/International Plant Genetic Resources Institute. Rome, Italy, p. 199-242.

Kim, J., Joung, H., Kim, H., Lim, Y. (1998) *Estimation of genetic variation and relationship in Potato (*Solanum tuberosum *L.) Cultivars using AFLP markers.* American journal of Potato Research 75(2): 107-112.

Perrier, X., Flori, A. y Bonnot, F. (2003). *Data analysis methods. In* : *Genetic diversity of cultivated tropical plants.* Enfield, Science Publishers. Montpellier. Pp 43 – 76.

Rohlf, F. (2000). *NTSYS-pc: Numerical taxonomy and multivariate analysis system*. Version 2.1. Publicaciones Exeter, New York, USA.

Saitou, N. y Nei, M. (1987). *The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees.* Mol. Biol. Evol. 4(4): 406-425.

Vos P.; Hogers, R.; Bleeker, M.; Reijans, M.; Theo van de Lee; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M. y Zabeau, M. (1995). AFLP *A new technique for DNA fingerprinting*. Nucleic Acid Research 23: 4407-4414.