RESEARCH PAPER

Genetic identity based on simple sequence repeat (SSR) markers for Quinoa (*Chenopodium quinoa* Willd.)

Maria Romero, Angel Mujica, Edgardo Pineda, Yesenia Ccamapaza, and Nohely Zavalla

Universidad Nacional del Altiplano, Escuela Profesional de Biología, Puno, Perú.

Abstract

M. Romero, A. Mujica, E. Pineda, Y. Ccamapaza, and N. Zavalla. 2019. Genetic identity based on simple sequence repeat (SSR) markers for Quinoa (*Chenopodium quinoa* Willd.). Cien. Inv. Agr. 46(2): 166-178. Molecular markers based on simple sequence repeats (SSRs) constitute a highly effective instrument in the identification of quinoa genotypes (*Chenopodium quinoa*), and they are very useful in the management and conservation of germplasm banks. The present study was carried out in the Molecular Biology Laboratories of the Megalaboratory of the National University of the Altiplano and the National Agrarian University la Molina. With the objective of determining a minimum group of highly informative initiators for the cultivation of quinoa to study and identify the obtained alleles and implementing and incorporating this technology into research of genetic identity, the molecular analysis of nine loci located by microsatellite markers (SSRs) was performed on a sample of 26 varieties of quinoa: Ayrampo, Amarilla de Marangani, Choclito, Chullpi, Huariponcho, Pandela, Sajama, Witulla, Kcancolla, Negra Collana, Salcedo, Pasankalla, Blanca de Juli, and *Chenopodium petiolare* from the CIP-Camacani and Blanca de Juli, Kcancolla, Negra Collana, Pasankalla, Altiplano, Illpa INIA, Salcedo, Ayara Blanca de Juli, Ayara Blanca de Arequipa, Ayara Cancolla, Ayara Pasankalla, and Ayara Salcedo from the INIA. Genomic DNA was extracted by PCR (GeneJET Plant Genomic DNA Purification), and 20 microsatellite regions were amplified. The amplified fragments were loaded on polyacrylamide gels to determine their size in base pairs, of which only nine showed products with reading quality (QCA012, QCA015, QCA021, QCA029, QCA034, QCA040, QCA053, QCA055 and QCA067). The fragments were evaluated for their allelic richness, heterozygosity (H) and polymorphic information content (PIC). The data were processed with Gen Alex software ver. 3.5 A total of 67 alleles were detected among the different regions analyzed, with an average of 7 alleles for loci ranging from 142 to 240 bp and an effective number of alleles (ENA) of 5.36. The mean heterozygosity was 0.80, and the mean Polymorphic Information Content (PIC) was 0.81. The markers were highly polymorphic; therefore, the most informative SSR primers in the present study would be made up of three markers with PIC, QCA053 (0.87), QCA015 (0.86) and QCA034 (0.86), for determining the genetic identity of *Chenopodium quinoa* Willd. These markers can be easily interpreted and are useful for the molecular characterization of quinoa varieties. Analysis of the hierarchical clusters using UPGMA (Unweighted Pair Group Method) clustering identified 5 groups at a similarity coefficient of 0.77 among the quinoa varieties studied in this research.

Keywords: *Chenopodium quinoa,* marker, molecular genetics, polymorphism.

Introduction

Currently, *Chenopodium quinoa* Willd. genetic improvement programs need to preserve and manage germplasm collections, which are made up of numerous genotypes carrying valuable genes and used as progenitors to improve varieties and transfer to new cultivars (Biasutti *et al*., 2000). Most of these genotypes are carriers of valuable genes. The genetic diversity in *C. quinoa* represents a great reserve whose genetic variability in grain quality, panicle colors, leaves, plant height, saponin content and proteins needs to be studied (Mujica, 1988). ISSR molecular markers are distributed in the coding and noncoding regions of the genome and allow differentiation of genotypes. Polymorphic microsatellite loci are widely used for the analysis of population genetic structure (Botstein *et al*., 1980).

The maintenance of germplasm banks under i*n vitro* cultivation conditions requires that each genotype be regenerated at least two to three times a year and have an efficient logging system. Molecular identification based on DNA fingerprinting using microsatellite markers (SSRs) has been described as a powerful tool for identifying clones and quinoa cultivars that could greatly facilitate the conservation and management of the *Chenopodium* germplasm (Biasutti *et al*., 2000).

In the last twenty years, great advances in molecular biology have been made; molecular techniques were developed that allow the genome of living beings to be analyzed quickly and accurately. Through these techniques, a large number of molecular markers were obtained scattered throughout the entire genome, genetic maps were constructed of various horticultural species, and genes with resistance to diseases, pests, etc. were identified. Markers linked to agronomically interesting genes can be used for early genotypic selection in seedlings and thus avoid handling hundreds or thousands of plants (Maughan *et al*., 2009).

Microsatellites are DNA sequences consisting of repetitions of nucleotide motifs from 1 to 6 base pairs (Hancock, 1999). This type of sequence has been found in both eukaryotic and prokaryotic genomes and has even been identified in the genomes of mitochondria and chloroplasts. Microsatellites are distributed in coding and noncoding regions and are characterized by being highly polymorphic in terms of their length, so they are suitable regions for use as molecular markers at the population level (Zane *et al.,* 2002). This high degree of polymorphism is the result of a high mutation rate (from $10{-}6$ to $10{-}2$ mutations per site per generation (Schlotterer, 2000)) and is attributed to insertion and deletion events during DNA replication. Due to the high degree of polymorphism in marker size, they are codominant.

SSR microsatellites are important in the definition of individual genotypes, and in studies of gene flow in forest species, due to their great variability and codominance, they manage to distinguish heterozygous from homozygous individuals. In domestication programs, microsatellites allow the identification of germplasm and the construction of linkage maps, reflecting the mapping of traits of interest through the localization of QTLs (quantitiave trait loci) useful in assisted selection (Butcher *et al*., 2007).

SSRs present greater technological simplicity in relation to RFLP and AFLP and do not require high concentration or DNA quality. In contrast to RAPD markers, SSRs have high reproducibility between laboratories and are genetically codominant (Ghislain *et al*., 1999). Considering the relevance of these markers, they are considered ideal for obtaining information on polymorphisms to discriminate between genotypes and to identify original and representative materials of the genetic variability of a collection. They may also be used in genetic improvement and germplasm management programs and have the potential to maximize the conservation of genetic diversity.

For this reason, it was considered important to evaluate with the use of simple sequence repeat (SSR) markers with the aim of characterizing the quinoa germplasm, determining a minimum group of highly informative initiators for the cultivation of quinoa and identifying the obtained alleles. Additionally, we established polymerase chain reaction (PCR) amplification conditions to obtain highly informative SSRs and identify markers of known molecular weight for analyzing the genetic identity of quinoa varieties.

Materials and Methods

The vegetal material (quinoa varieties) was obtained from the Germplasm Bank of the CIP Camacani, the National University of the Altiplano, and the National Institute of Agrarian Innovation. Both centers are located in the Platería district, Puno Province and region. For sampling, we proceeded with the greenhouse germination of 26 varieties of quinoa, sown and properly codified; the varieties were allowed to grow to obtaining young leaves. These were transferred to the laboratory to begin the research. The research was carried out in the Molecular Biology Laboratory of the Megalaboratory of the National University of the Altiplano and the National Agrarian University La Molina in Puno and Lima, Peru, respectively.

DNA extraction.

For easier and faster DNA extraction than with the CTAB (Bromuro Trimetil Amonio of Cetilo) method, reported protocols (Doyle and Doyle, 1987) with modifications were applied by using the extraction kit (Gen Jet plant Genomic DNA) for purification from young leaves. The quality assessment and quantitation of DNA were performed by UV spectrophotometry at 260 nm. This first phase was carried out in the laboratory of Molecular Biology-Megalaboratory of the National University of the Altiplano, Puno.

Molecular characterization by PCR (polymerase chain reaction) was performed in the Molecular Biology Laboratory of the National Agrarian University La Molina Lima by using a group of 9 pairs of microsatellite markers described by (Maughan *et al.*, 2004). From the preselection results from the DNA amplification of the 26 quinoa varieties, each microsatellite represents a locus that contains simple repetitions of one, two, three or four nucleotides. Through the PCR, the amplification of the white sequence was performed. The primers applied recognize the flanking sequences of the SSR and start the amplification of the repetitive sequence.

PCR amplification products were resolved by vertical electrophoresis with 6% polyacrylamide gels. The polyacrylamide gel electrophoresis was performed as by (Laemmli, 1970) with minimal modifications. Subsequently, the gels were developed by staining with silver nitrate according to a standardized protocol. The SSR fragments were separated according to their size in base pairs.

Each SSR allele corresponding to an amplified band was recorded with 1 or 0, according to presence or absence, respectively, of each genotype; subsequently, the values were entered into a data matrix. Each SSR was characterized according to the sharpness of bands, width or amplitude of its alleles and presence of shadows, defining a reading quality index for each marker. The size of the alleles was determined by comparison of mobility in the gel with the marker Molecular Gene Ruler™ 50 pb DNA standard Low Mass Ladder (MBI Fermentas, Ontario, Canada). Polymorphism was sought according to the presence or absence of bands in the entire population analyzed. The pattern of bands obtained for each initiator was recorded in a binary matrix. Each of the gels dyed with silver salts was scanned and stored in a database. The interpretation and analysis of the data obtained were carried out according to the research objectives.

Results and Discussion

To determine a minimum group of highly informative initiators and the identification of alleles, the results of 20 SSR markers from a group of 26 varieties of *C. quinoa*, were reported. The SSR markers used correspond to part of the group described by Mason *et al.* (2005), Christensen *et al.* (2007) and Jarvis *et al.* (2008), products of the evaluations developed; 11 microsatellites did not amplify despite the testing of several PCR conditions. Based on the resolved amplification product, nine initiators were selected, with sharpness, a high number of polymorphic bands and genetic identity in the molecular characterization in all cases. The dinucleotide microsatellites were treated for homogeneous repeats (CA, TG), as described in Table 1. To identify the highly informative initiators, an evaluation was made of the polymorphism, diversity and genetic structure of the population under study.

To achieve the amplification conditions, hybridization temperatures were applied using different gradient tests, varying between 52–65 °C. (Table 2).

Microsatellite type sequences (SSRs or STRs, simple sequence repeats or short tandem repeats) are abundant in the genomes of eukaryotes and some prokaryotes and are composed of short units (basic motifs) of 1 to 6 base pairs that are repeated in tandem many times. Each SSR sequence is defined by the type of repeated unit (the most frequent are one, one, three or four nucleotides,

although pantafive and six nucleotides also exist), and the frequency and type of repetition varies in the genomes of different species, as does the site they occupy in the genome (locus) (Jarvis *et al.*, 2008). These are highly variable sequences between and within individuals. Variation is usually manifested as differences in length between different alleles. The amplification patterns of nine microsatellite initiators identified a total of 67 alleles with an average of 7 per locus. Table 3, shows the number, size and frequency of each allele per microsatellite. In general, the amplification of the nine initiators proved to be of good quality. Likewise, the presence of a maximum of 3 to 4 bands was identified by locus (QCA034, QCA015, QCA053, QCA055 and QCA012), which corresponded to the expected tetraploid disomic inheritance of quinoa (originating from two diploid species) (Gandarillas, 1984).

Table 2. Amplification range, hybridization temperature used for each marker.

Marker	Amplification range (pb)	Hybridization temperature (°C)
QCA012	172 to 196	57°
QCA015	190 to 222	52°
QCA021	192 to 216	55°
QCA029	150 to 184	57°
QCA034	142 to 188	59°
QCA040	208 to 230	62°
QCA053	178 to 198	56°
QCA055	196 to 240	67°
QCA067	198 to 210	62°

Table 1. Microsatellites (SSRs) selected for the molecular study of *Chenopodium quinoa* (Mason *et al.,* 2005).

N°	Locus	Sequence SSR	Forward primer (5_-3_)	Reverse primer (5_-3_)
1	QCA012_	(TG)9	Tcccatatgcctacgtaccaa	Tggtcatcaacatccaaagg
2	QCA015_	(AC)17	Tgggaccctgatagcttgac	Tgtcctttgcatgtgctatga
3	QCA021_	(CA)16	Cagggtatcagaatactgggaaa	Ccaagattggaggacaggaa
4	QCA029	CA)10	Tctacttgcaacccgaatgtc	Cgcaaagcaaatcaggtaca
5	QCA034	(CA)16	Agggagaatgcggagaaga	Tcaacaaacaagcacgaagg
6	QCA040	(CA)13	Tgtggtgacaagcaactttga	Aacctacttcaattagaccaacttcc
7	QCA053	(TG)25	Agatgtggtgcgttggatct	Aaggagagctctaaccgcttg
8	QCA055	(TG)14	Gggcatatctgaagagaatcca	Acgcaggtagcacttccagt
9	9QCA067	(CA)12	Gcaagacctgctcacaacaa	Tatcaacagcaacggaagca

Table 3. Size in (bp) and frequency of amplified alleles per microsatellite.

SSR	Size in base pairs (bp) and frequency of nine microsatellite alleles Evaluated in 26 varieties of *Chenopodium quinoa* Willd. 2017									
QCA012	172	174	182	186	188	192	196			
	0.063	0.413	0.016	0.111	0.190	0.143	0.063			
QCA015	190	198	204	206	208	210	214	218	220	222
	0.090	0.060	0.106	0.212	0.015	0.030	0.151	0.166	0.121	0.045
QCA021	188	192	194	198	210	216				
	0.265	0.294	0.176	0.059	0.147	0.059				
QCA029	150	158	162	166	170	176	180	182	184	
	0.09	0.033	0.082	0.295	0.033	0.016	0.033	0.082	0.328	
QCA034	142	150	152	160	166	172	174	178	180	188
	0.104	0.075	0.090	0.224	0.015	0.030	0.149	0.060	0.119	0.134
QCA040	208	212	214	216	224	226	230			
	0.113	0.189	0.094	0.057	0.038	0.075	0.433			
QCA053	178	180	184	190	192	194	196	198		
	0.186	0.084	0.135	0.220	0.084	0.084	0.067	0.135		
QCA055	196	212	220	234	240					
	0.262	0.147	0.213	0.262	0.115					
QCA067	198	200	204	208	210					
	0.092	0.129	0.277	0.351	0.148					

Genetic diversity is of fundamental importance in the continuity of a species, providing the means to adapt to the prevailing in environmental conditions, both biotic and abiotic, and allowing changes in genetic composition to cope with habitat changes.

The analysis of genetic diversity was assessed on the basis of the following parameters: Number of Alleles (NA), Average Number of Alleles per locus (NPA), Effective Number of Alleles (ENA), expected (He) and observed (Ho) Heterozygosity and Polymorphic Information Content (PIC), applying the software Gen Alex version 6.5 (Peakall and Smouse, 2012).

The first item necessary for studying the survival of a species is knowledge of the level of genetic diversity. This refers to the determination of the number of alleles, effective alleles, polymorphic loci, observed and expected heterozygosity, genetic structure and spatial distribution of genetic variants (Van Delden, 1992).

The average number of effective alleles (ENA) was 5.36 according to these indexes. All markers were identified as optimal; however, the best were QCA053, QCA034 and QCA015. The number of effective alleles refers to alleles with the ability to

move to the next generation (Kimura, 1965) and is a good indicator of markers that make important contributions to diversity studies, as the value of ENA approaches the number of alleles found. Of the nine microsatellites assessed, 78% reflected values close to those found in each of the evaluated populations. QCA053 presented the highest value (8.31), and the fewest number of identified alleles was detected by QCA040 (3.92) (Table 4).

Table 4. Number of alleles (NA), effective alleles (ENA) and average number of alleles per locus (NPA) using microsatellite markers in *C. quinoa*.

Markers	NA	ENA
QCA012	7	4.15
QCA015	10	7.00
QCA021	6	4.60
QCA029	9	4,58
QCA034	10	7.51
QCA040	7	3.92
QCA053	8	8.31
QCA055	5	4.60
QCA067	5	4.03
Total	67	
NPA	7.44	5.36

We identified 35% of exclusive alleles and 65% of alleles shared with 160 bp-QCA034, 174 bp-QCA012, 184 bp-QCA029, 190 bp-QCA053, 204 bp-QCA067 206 bp-QCA015, and 234 bp-QCA055 between 20 commercial varieties; 5 Ayaras; and the only wild relative considered *C. petiolare* making up the study population (Table 5).

Table 5. Allelic richness detected for each locus in 26 varieties of quinoa.

Locus SSR	Range (pb)	Exclusive alleles	Shared alleles	Total Alleles
QCA012	172 to 196	3	4	7
QCA015	190 to 222	3	7	10
QCA021	192 to 216	0	6	6
QCA029	150 to 184	4	5	9
QCA034	142 to 188	4	6	10
QCA040	208 to 230	4	3	7
QCA053	178 to 198	1	7	8
QCA055	196 to 240	3	2	5
QCA067	198 to 210	1	4	5
Total		23	44	67
%		35	65	100

The values obtained from the allelic richness indicate that 8 microsatellite loci were polymorphic, QCA034, QCA015, QCA053, QCA055, QCA021 QCA029 QCA040 and QCA067, chosen based on the variation in the length reflected in the differences in the size of band products of PCR and electrophoresis with silver nitrate-stained polyacrylamide.

Table 6, shows the Heterozygosity (H) and the PIC. The average values of Ho and were 0.79 and 0.80, demonstrating a high polymorphism. A molecular marker is considered polymorphic if H is ≥0.1 and highly polymorphic if H is ≥0.7; high values of Ho and He in all loci indicate a high rate of heterozygotes and an estimation of the degree of genetic variability in the population Ott (1992).

The PIC, presented an average of 0.81, identifying eight of the markers as highly polymorphic (Table 6). This parameter is often used to measure the discriminatory capacity of SSR markers; however, its value may vary for the same SSR, depending on the characteristics of the germplasm studied. The PIC calculated for a marker may vary between 0 and 1, indicating a higher level of polymorphism or variation when the value is closer to 1. The high polymorphism observed in the study population is similar to that reported by Maughan *et al.* (2004), who developed the first large-scale research of quinoa SSR markers, which consisted of evaluating 208 markers,

which were validated and characterized by 31 cultivated quinoa accessions representative of the main areas of cultivation of South America. The genetic analysis performed revealed a series of observed alleles ranging from 2 to 13, with an average of four alleles detected by locus. The values of heterozygosity oscillated between 0.2 and 0.9, with an average value of 0.57. In addition, 67 SSR markers were highly polymorphic, with values of heterozygosity equal to or greater than 0.70.

Likewise, Veramendi (2006), characterizing the genetic diversity of 90% of Bolivian quinoa crops with eight microsatellite markers, reported a high level of polymorphism, with a PIC greater than 0.73; the marker QAAT-022 was the most polymorphic, with a value of 0.95. This marker found 129 alleles, ranging from 5 to 30 alleles per locus and from 111 to 239 bp in size.

Christensen *et al.* (2007), analyzed 143 accessions from South America with 36 SSRs that detected 420 alleles with an average of 11 alleles per locus, while Fuentes *et al.* (2012) analyzed 59 entries of quinoa from Chile with 20 SSRs that detected 150 alleles with an average of 7.5 alleles per locus.

The polymorphic content of the loci was similar to that in the present study, indicating a high level of polymorphism. Polymorphism, defined as changes of a single base, is rapidly becoming

the marker system of choice in breeding programs according to Fuentes *et al.,* (2009).

This property has been used to grant each individual a particular and unique genetic footprint due to the finding of specific allele patterns for each SSR locus. A single microsatellite locus already allows for evidence of differences between individuals. The latter is of great use to establish a panel of SSR markers that allow the discrimination of closely related quinoa individuals.

This variation within quinoa cultivation has important agronomic implications. Fuentes *et al.* (2009) indicate that ecotypes in Andean and coastal areas present different responses, such as to altitude, drought, salinity and long-days. These adaptations are associated with the genetic variability of the crop *C. quinoa*; despite it being considered a species mostly pollinated, quinoa varieties show a wide range of variation in genetic diversity. Intrapopulation genetic diversity depends on the number and frequency of alleles among all loci and on the genetic structure of the population (Crossa *et al*., 1993). The native quinoa accessions characterized are from local populations that have survived in the region for a long time and may be descendants of genotypes that have been the subject of selection for many generations. As a unit, they integrate several

Table 6. Observed (Ho) and expected (He) heterozygosity and polymorphic information content (PIC) using microsatellite markers in *Chenopodium quinoa* Willd.

Microsatellite	Ho	He	PIC
QCA012	0.75	0.76	0.76
QCA015	0.86	0.87	0.86
QCA021	0.75	0.78	0.79
QCA029	0.78	0.78	0.77
QCA034	0.86	0.87	0.86
QCA040	0.74	0.75	0.73
QCA053	0.85	0.86	0.87
QCA055	0.78	0.78	0.78
QCA067	0.75	0.75	0.75
Average	0.79	0.80	0.81

Higher PIC values were obtained with the primers QCA053, QCA034 and QCA015, which would be very informative and useful in subsequent studies for determining genetic footprints.

components (agronomic and cultural) that have adjusted to each other through generations and reflect decisions made by farmers (Table 7).

Molecular identification based on DNA fingerprinting using microsatellite molecular markers (SSRs), described as a powerful tool for identifying clones and cultivars that could greatly facilitate conservation, genetic improvement and management of germplasm in plant species (Ashkenazi *et al*., 2001).

The registration and characterization of the nine microsatellite initiators QCA012, QCA015, QCA021, QCA029, QCA034, QCA040, QCA053, QCA055

Table 7. Analysis of microsatellites used in the genotyping of 26 varieties of quinoa (*Chenopodium quinoa* Willd.).

Microsatellites	BP	BM	BT	PMF (95%)	PIC	E.E.	AMP	PDICMA
QCA012	6	1	7	0.86	0.76	0.03	35.71	9.5E-09
QCA015	10	0	**10**	0.90	**0.86**	0.03	**25.77**	1.8E-13
QCA021	6	0	6	1.00	0.79	0.02	23.08	1.2E-13
QCA029	9	0	**9**	0.89	0.77	0.02	25.64	1.3E-06
QCA034	10	0	**10**	1.00	**0.86**	0.03	**26.54**	1.5E-14
QCA040	7	0	7	1.00	0.73	0.02	29.12	2.2E-08
QCA053	8	0	8	1.00	**0.87**	0.02	28.85	**2.4E-16**
QCA055	5	0	5	1.00	0.78	3.1E-03	46.92	**1.8E-16**
QCA067	5	0	5	1.00	0.75	0.01	41.54	4.3E-14
Total	66	1	67	-	-	-	30.14	1.5E-10

BP: Polymorphic Bands, BM: Bands Monomorphic, BT: Bands Total, PMF: Proportion of loci polymorphisms, he: Expected heterozygosity, AMP: Percentage of amplification, PDICMA: Probability that two individuals share the same allele.

Table 8. Range in size of base pairs, number of alleles and polymorphic information content (PIC).

N°	Molecular Marker	Range In Size Of Base Pairs (PM)	N° Of Alleles	Polymorphic Content Index (PIC)
1	QCA053	178 to 198	8	0.97
2	QCA015	190 to 222	10	0.98
2	QCA034	142 to 188	10	0.98
3	QCA021	100 to 216	5	0.93
4	QCA055	196 to 240	5	0.84
5	QCA029	150 to 184	9	0.94
6	QCA012	172 to 196	7	0.93
7	QCA067	198 to 210	5	0.79
8	QCA040	208 to 230	7	0.92

and QCA067 indicate that they are polymorphic, and according to the index of polymorphic content and genetic diversity index, the most informative SSRs loci were QCA053 (0.97), QCA015 (0.98) and QCA034 (0.98) (Table 8).

The markers were characterized by recording the molecular weight of alleles produced in base pairs by comparing the bands contained in each locus. These are highly variable sequences (Table 3). Variation is usually manifested as differences in length between different alleles of the same locus. These differences in length arise from the existence of a different number of repetitions of the basic motif in each case. It has been estimated that the rate of mutation in microsatellites varies between 10^2 and 10^5 per generation, and the mechanism that best explains the high degree of polymorphism in size is the accumulation of errors caused by the slippage of the polymerase during the replication of DNA (Ellegren, 2004).

The molecular genetic variability of quinoa analyzed with microsatellite markers showed a structure and differentiation between varieties according to the observed grouping depending on the genetic distances.

In the quinoa varieties analyzed, genetic distances are evident in Blanca de Juli, Altiplano, Amarilla de Marangani and Ayrampo, Huariponcho, Negra Collana, Ayara Kcancolla, Pasankalla, Padela, Negra collana, Illpa INIA, Witulla, Ayara Blanca de Juli, and Blanca de Arequipa, clearly visible in the PCoA (Analysis of Cordenates Major) and coincident with the analysis of conglomerates.

The analysis of conglomerates (Figure 1) allowed us to visualize 3 clusters at a distance of 0.77, with 5 groups clearly defined within the population, the first formed only by the wild variety *C. petiolare*. This species is characterized by being tolerant to environmental and soil stresses because they are resistant to diseases and present other adaptive characters useful for crop breeding. Mujica and Jacobsen (2006) cultivated varieties that make morphological identification difficult. SSR markers are ideal for selecting between these types of varieties.

A first group composed of a single variety of *C. petiolare*, resistant to drought, is present inside the cultivated fields of quinoa and possibly accompanies the locations of quinoa distribution (Mujica and Jacobsen, 2001). The tender plant is used for feeding cattle, and its tender leaves are similar to leaf vegetables in human food. The grains are used in cooking to elaborate the quispiño (elaborated dark muffin of that flour), and it also has medicinal use, mainly for bone fractures.

Without continuous genetic improvement, a variety of germplasm is used from wild sources; the increase in crop yields obtained over the last seven decades would not be sustainable, and yields would grow more slowly. Agricultural production
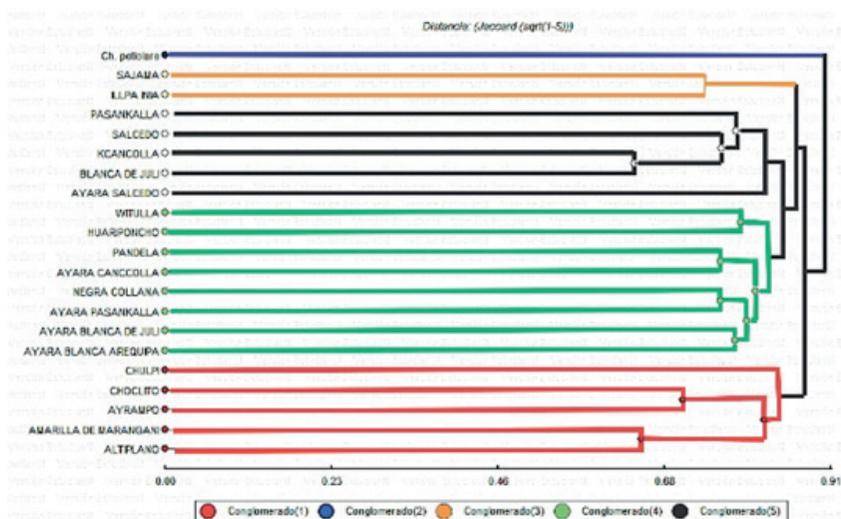
Figure 1. Cluster analysis of 26 varieties of quinoa according to molecular markers used in their genotyping.

is increasingly dependent on 'temporal diversity', which means replacing varieties frequently to maintain resistance to pests and diseases (Rubenstein *et al*., 2005).

Modern biotechnological tools offer new opportunities for the greater and more effective use of wild species in crop improvement as research tools. A second group is composed of Sajama and Illpa INIA; this variety is generated from the cross Sajama x Blanca of Juli carried out in the experimental fields of Salcedo-Puno in 1984. This variety is generated from the cross of two lines, Real 547 × Dulce 559, of Bolivian origin; has precocious high yield and white and large grains 2 to 2.2 mm in diameter; is a sweet variety free of saponin; has a glomerular panicle and a 170 day vegetative period; reaches a height of 1.10 m; is susceptible to ornithological attack and mildew due to its sweet character; has a yield of 3000 kg ha^{-1}; and adapts well in Azangaro, Ayaviri and Lampa. It is a contribution to our culture for everyone. According to scholars, this crop has become increasingly important for its diversity and usefulness in countries with fragile ecosystems, adding to its nutritional benefits that satisfy the necessities of basic food (food safety) of the producer, in addition to generating economic income by the sale of its surplus production.

The third group was formed by Pasankalla, Salcedo, Kcancolla, Blanca de Juli and Ayara Salcedo; the fourth group by Witulla, Huariponcho, Pandela, Ayara Canccolla, Negra Collana, Ayara Pasankalla, Ayara Blanca de Juli and Ayara Blanca Arequipa; and the fifth group by Chulpi, Choclito, Ayrampo, Amarilla de Marangani and Altiplano.

According to our dendrological tree, the accessions Chullpi, Choclito, Ayrampo, Amarilla de Marangani and Altiplano are separated from all other groups; *C. petiolare* is commonly observed in fields of the Andean highlands, among 3,830 to 3,900 M.A.S.L. (Mujica and Jacobsen, 2006). Sajama, Illpa INIA are completely separate from all the aforementioned groups.

Main Coordinates Analysis (PCoA)

The main coordinate analysis indicates that component 1 contributes to the explanation of total variance by only 11%, component 2 9.7% and component 3 8.2%; these contributions are insufficient to describe the ordering of 26 varieties. However, the proximity between Ayrampo and Choclito, as well as between Altiplano and Amarilla de Marangani is shown in Figure 8. In Figure 8 is shown the proximity between Salcedo,

Sajama and Pasankalla, among other commuters, which coincides with the results of the analysis of conglomerates.

C. petiolare was used as a source node, and the first phylogenetic path indicates Witulla, giving rise to three different lines. In the second level, Kcancolla, Negra collana and Huariponcho are located. The third level includes Ayara, Blanca de Juli, Pandela, and Illpa INIA, whereas in the fourth level includes Ayrampo and Choclito, among others, and the fifth level Salcedo, Altiplano and Pasankalla. Finally, the variety of greatest distance from the ancestor is Amarilla Marangani.

According to Mujica (1992), cultivated quinoas have great genetic diversity, showing variability in the coloration of the plant and inflorescence, which is the only way that we can guarantee food sovereignty. The competitiveness of agricultural products is by consolidating continuous improvement programs that use all the tools available today.

The genetic variation in plants is found in the set of genes that they possess, and the spectrum of this variability within cultivated species and their wild relatives is commonly maintained in germplasm banks. The importance of the maintenance of these resources is in the measurement and characterization of this diversity. In fact, conserving genetic information is very important for germplasm banks. In this research, we can also assume that within the same varieties, there is still intrapopulation variability, so the conservation of accessions in germplasm banks should apply to molecular markers.

The genetic diversity of quinoa (*C. quinoa*) has been maintained over time thanks to the traditional and ancestral knowledge of peasant populations, which is considered a culture of family inheritance. However, in recent years, with increasing demand in both domestic and international markets, it has led to the substitution of native varieties by commercial ones, which generates loss of genetic diversity and increased homozygosity in the population, making the plants more vulnerable to phytosanitary problems and decreasing their ability to adapt to the changing environment (Fuentes *et al*., 2012). On the other hand, industrial development has also caused the abandonment of rural areas by mass migration to cities, and this social and economic situation is changing the use of land and affecting the genetic variability of Andean crops. Therefore,

Figure 2. Minimum travel network (phylogeny) of 26 varieties of quinoa according to molecular markers used in their gene typing.

it is essential to preserve, on a small scale, the genetic diversity of the local quinoa materials, which are strongly associated principally with culture. Preserving agrobiodiversity also means preserving the culture associated with farmers and the peasants who live in the Andean region (Bazile, 2014).

Conclusions

The minimum group of highly informative initiators for quinoa crops under study and the identification of alleles obtained is made up of 9 microsatellite markers (SSRs): QCA012, QCA015, QCA021, QCA029, QCA034, QCA040, QCA053, QCA055 and QCA067. These markers identified 67 alleles with a range of 5 (QCA055 and QCA067) to 10 (QCA015 and QCA034) alleles per locus ranging from 142 to 240 bp with an effective number of alleles (ENA) of 5.36 with PIC 0.81 as the average value and therefore proved to be highly polymorphic and useful for the molecular characterization of quinoa varieties.

The established and standardized PCR parameters for highly informative SSR identification were hybridization temperature (Tm, annealing temperature), DNA concentration, primers and components such as free water of nucleases or NFW (nuclease free water) and MgCl2 (magnesium chloride).

According to the genetic diversity index, genetic structure and molecular behavior, the most informative initiators in the genetic identity of 26 varieties of quinoa were QCA053, QCA015 and QCA034, which were highly polymorphic, discriminant, easily read and easy for interpreting genetic diversity, which corroborates the results of the analysis of hierarchical conglomerates and genetic distances.

Molecular markers are essential for the investigation of genetic variability and understanding of genome dynamics.

Resumen

M. Romero, A. Mujica, E. Pineda, Y. Ccamapaza, y N. Zavalla. 2019. Identidad genética basada en marcadores de repetición de secuencia simple (SSR) para Quinua (*Chenopodium quinoa* Willd.). Cien. Agr. 46(2): 166-178. Los marcadores moleculares basados en repeticiones de secuencia simple (SSRs) constituyen un instrumento altamente efectivo en la identificación de genotipos de quinua (*Chenopodium quinoa*), y son muy útiles en el manejo y conservación de bancos de germoplasma. El presente estudio se realizó en los Laboratorios de Biología Molecular del Megalaboratorio de la Universidad Nacional del Altiplano y de la Universidad Nacional Agraria de la Molina. Con el objetivo de determinar un grupo mínimo de iniciadores altamente informativos para el cultivo de la quinua para estudiar e identificar los alelos obtenidos e implementar e incorporar esta tecnología en la investigación de la identidad genética, se realizó el análisis molecular de nueve loci localizados por marcadores microsatélites (SSRs) en una muestra de 26 variedades de quinua: Ayrampo, Amarilla de Marangani, Choclito, Chullpi, Huariponcho, Pandela, Sajama, Witulla, Kcancolla, Negra Collana, Salcedo, Pasankalla, Blanca de Juli, y Chenopodium petiolare del CIP-Camacani y Blanca de Juli, Kcancolla, Negra Collana, Pasankalla, Altiplano, Illpa INIA, Salcedo, Ayara Blanca de Juli, Ayara Blanca de Arequipa, Ayara Cancolla, Ayara Pasankalla y Ayara Salcedo del INIA. El ADN genómico fue extraído por PCR (GeneJET Plant Genomic DNA Purification), y se amplificaron 20 regiones microsatélites. Los fragmentos amplificados se cargaron en geles de poliacrilamida para determinar su tamaño en pares de bases, de los cuales sólo nueve mostraron productos con calidad de lectura (QCA012, QCA015, QCA021,

QCA029, QCA034, QCA040, QCA053, QCA055 y QCA067). Los fragmentos fueron evaluados por su riqueza alélica, heterocigosidad (H) y contenido de información polimórfica (CFP). Los datos fueron procesados con el software Gen Alex ver. 3.5 Se detectaron un total de 67 alelos entre las diferentes regiones analizadas, con un promedio de 7 alelos para los loci que oscilan entre 142 y 240 pb y un número efectivo de alelos (ENA) de 5.36. La heterocigosidad media fue de 0,80 y la media del Contenido de Información Polimórfica (PIC) fue de 0,81. Los marcadores eran altamente polimórficos, por lo que los primers SSR más informativos del presente estudio estarían compuestos por tres marcadores con PIC, QCA053 (0,87), QCA015 (0,86) y QCA034 (0,86), para determinar la identidad genética de Chenopodium quinoa Willd, los cuales pueden ser fácilmente interpretados y son útiles para la caracterización molecular de variedades de quinua. El análisis de las agrupaciones jerárquicas utilizando el método UPGMA (Unweighted Pair Group Method) identificó 5 grupos con un coeficiente de similitud de 0,77 entre las variedades de quinua estudiadas en esta investigación.

Palabras clave: *Chenopodium quinoa*, genética molecular, marcador, polimorfismo.

References

Ashkenazi, V., E. Chani, U. Lavi, D. Levy, J. Hillel, and E. Veilleux. (2001). Development of microsatellite markers in potato and their use in phylogenetic and fingerprinting analyses. Genome 44:50–62.

Bazile D. (2014). Quinua, a model crop to examine the dynamics of biodiversity within agricultural systems. Biodiversity, Issue 15.1:3–4.

Biasutti C.A, F. Casanoves, and D.A. Peiretti. 2000. Response to different adaptive mass selection criteria in a maize exotic population. Maydica 45:89–90.

Botstein, D., R.White, H. Sholmick, and S.R. Davu. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. American Journal of Human Genetics 32:314–331.

Butcher P, and S. Southerton. 2007. Marker-assisted selection in forestry species. In Marker-assisted selection – Current status and future perspectives in crops, livestock, forestry and fish. Roma: FAO; pp. 284–296.

Christensen, S., D. Pratt, and C. Pratt. 2007. Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers. Plant Genet Resources: Charact Util 5:82–95.

Crossa, J., C.M. Hernandez, P. Bretting, S.A. Eberhart, and S. Taba. 1993. Statistical genetic considerations formaintaining germplasm collections. Theoretical and Applied Genetics 86:673–678.

Doyle, J.J., and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15.

Ellegren, H. 2004. Microsatellites: Simple Sequences with Complex Evolution. Genetics. Nature Reviews 5:435–445.

Fuentes, F., D. Bazile, A. Bhargava, and E. Martínez. 2012. Implications of farmers' seed exchanges for on-farm conservation of quinoa, as revealed by its genetic diversity in Chile. Journal of Agricultural Science 150:702–716.

Fuentes, F.F., P.J. Maughan, and E.N. Jellen. 2009. Diversidad genética y recursos genéticos para el mejoramiento de la quinua (*Chenopodium quinoa* Willd.). Revista geográfica de Valparaíso 42:20–33.

Gandarillas, H. 1984. Obtención experimental de *Chenopodium quinoa* Willd. MACA IBTA. La Paz, Bolivia. 21 p.

Ghislain, M., D. Zhang, Z. Fajardo. 1999. Marker assisted sampling of the cultivated Andean potato *Solanum phureja*. Colección using RAPD marker. Genetic resources and crop evolution.

Hancock, J.M. 1999. Microsatellites and other simple sequences: Genomic context and mutational

mechanisms. *In:* D. B. Goldstein and C. Schlötterer [eds.], Microsatellites: Evolution and applications, 1–9. Oxford University Press, Oxford, United Kingdom.

Jarvis, D., O. Kopp, E. Jellen. 2008. Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). J Genet 87:39–51.

Kimura, M. 1965. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics 52:875–890

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259):680–685.

Mason, S.L., M.R. Stevens, E.N. Jellen, A. Bonifacio, D.J. Fairbanks, C.E. Coleman, R.R. McCarty, A.G. Rasmussen, and P.J. Maughan. 2005. Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). Crop Sci 45:1618–163.

Maughan, P., A. Bonifacio, and E. Jellen. 2004. A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. Theor Appl Genet 109:1188–1195.

Maughan, P.J., T.B. Turner, C.E. Coleman, D.B. Elzinga, E.N. Jellen, J.A. Morales, J.A. Udall, D.J. Fairbanks, and A. Bonofacio. 2009. Characterization of salt overly sensitive (SOS1) gene homoeologs in quinoa (*Chenopodium quinoa* Willd.). Genome 52:647–650

Mujica, A. 1988. Parámetros genéticos e índices de selección en quinua (*Chenopodium quinoa* Willd.). Tesis de Doctor en Ciencias. Colegio de Postgraduados, Montecillo, México.113 p.

Mujica, A. 1992. Granos andinos: Quinua, Kiwicha y Kañiwa. En: *Cultivos marginados: otra perspectiva de 1492. 1992. Organización de la Naciones Unidas para la Agricultura y la Alimentación FAO.* Lima: Roma. pp. 129–146.

Mujica, A., and E. Jacobsen. 2001. Biodiversidad un desafío en la región centro oeste de Sudamérica. En: Agricultura andina., Puno, Perú. 6 p.

Mujica, A., and S.E. Jacobsen. 2006. La quinua (*Chenopodium quinoa* Willd.) y sus parientes silvestres. *In*: Moraes, M., Ollgaard, B., Kvist, L.P., Borchsenius, F. and Balslev, H. (eds.). *Botánica Económica de los Andes.* Universidad Mayor de San Andrés, La Paz, Bolivia, pp. 449–457.

Ott, J. 1992. Strategies for characteriziting highly polymorphic markers in human gene mapping. Am. J. Hum. Genet. 51:283–290.

Peakal, R., and P. Smouse. 2012 GenAlex 6.5: Genetic analysis in Excel. Population genetic software for teaching and research. Bioinformatics, 28(19):2537–2539.

Rubenstein, K.D., J.P. Heisey, R. Shoemaker, J. Sullivan, and G. Frisvold. 2005. Crop Genetic Resources: An Economic Appraisal, Economic Information Bulletin Number 2. In United States Department of Agriculture (USDA). Washington, DC.

Schlotterer, C. 2000 Evolutionary dynamics of microsatellite DNA. Chromosoma 109(6):365–371.

Van Delden, W. 1992. Genetic diversity and its role in survival of species. Biodiversity and Global Change. *In*.: CAB international and IUBS. pp. 41–56.

Veramendi, S. 2006. Estudio de la Diversidad Genética de la Colección Boliviana de Quinua (*Chenopodium quinoa* Willd.) correspondiente al Altiplano sur utilizando Tres Marcadores Microsatélite QAAT-074, QAAT-076, QAAT- 022. Tesis Magister Sciantiae. Universidad Mayor de San Simón. Cochabamba, Bolivia.

Zane, L., L. Bargelloni, and T. Patarnello. 2002. Strategies for microsatellite isolation: a review. Molecular Ecology 11:1–16.