Morphological and molecular characterization of two populations ex situ regenerated from a germplasm bank accession of the wild potato Solanum chacoense Bitter

Anabela Poulsen Hornum
Instituto Nacional de Tecnología Agropecuaria: Instituto Nacional de Tecnologia Agropecuaria
Elsa Lucila Camadro (ecamadro@balcarce.inta.gov.ar)
Universidad Nacional de Mar del Plata

Research Article

Keywords: germplasm conservation, ex-situ accession regeneration, allele frequencies, effective number of progenitors (Ne), wild potatoes.

Posted Date: July 10th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3111070/v1

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on November 13th, 2023. See the published version at https://doi.org/10.1007/s10722-023-01742-1.

Abstract

Crop wild relatives are conserved in germplasm banks as original accessions or their *ex situ* regenerated progenies. The commonly used sexual regeneration method in wild potatoes (FAO 2013) -cultivation of 20-25 plants/accession (N), hand-pollination with pollen mixtures when a percentage of the plants are in bloom, and composition of the regenerated accession without controlling the number of seeds provided by each female progenitor- does not consider the action of internal reproductive barriers. However, wild potatoes -mostly diploid and obligate out-crossers- may present internal pre-zygotic, post-zygotic, or both types of complete or incomplete reproductive barriers, which can reduce the effective number of progenitors (Ne) in relation to N. To evaluate possible phenotypic and genetic changes during a sexual reproduction cycle, two regeneration protocols were followed with one accession of *Solanum chacoense* Bitter: the current one and the proposed by Camadro (2012). The original accession and the two regenerated populations were morphologically and molecularly characterized. Twenty-one morphological characters (quantitative and qualitative) were considered, and the data were subjected to univariate and multivariate analyses. Allele frequencies were calculated for six microsatellite markers (SSR) and compared among populations with a chi-square test (α= 0.05). Great morphological variability for morphological phenotypes and four of the six analyzed SSR, as well as changes in allele frequencies (including allele loss) were revealed within and among the three populations. Therefore, it is advisable to consider aspects of reproductive and population genetics in developing regeneration protocols to prevent or reduce the risks of genetic erosion in sexual regeneration cycles.

Introduction

The potato, *Solanum tuberosum ssp. tuberosum* (2n = 4x = 48), is one of the most important crops worldwide, together with wheat and rice (FAO 2013). This cultivated species has many close wild relatives (CWR = Crop Wild Relatives) taxonomically classified as “species” by comparison of morphological phenotypes with holotypes or other specimen types of reference (taxonomic species = TS). The classification of plant materials as TS results from the application of the taxonomic species concept (TSC) (Grant 1981). The number of TS in potatoes varies with the taxonomist, but it has been estimated to be around 200 (Hawkes 1990); more recently, and with the aid of cytogenetics and molecular tools, this number has been reduced to approximately 100 (Spooner et al. 2014). Potato CWRs are important sources of genetic diversity for breeding purposes (Jansky et al. 2013)

Potato CWRs are widely distributed in the Americas, where they spontaneously grow in very diverse and contrasting environments from southwestern USA to southern Argentina and Chile following the Andes and, towards the east, in Paraguay, Uruguay, and Brazil (Hawkes 1990). They form a polyploid series (2n = 2x to 2n = 6x; x = 12) and can reproduce sexually (by seeds), asexually (by tubers and stolons), or both. Sexual reproduction is predominantly allogamous -with entomophilous pollination- because potatoes possess an *S*-locus gametophytic self-incompatibility system that prevents self-fertilization if there is an identity between the *S*-alleles (*S*-haplotypes) of pollen and style. Diploids are obligate out-crossers but polyploids can have a percentage of autogamy if self-pollinated by pollen heterozygous for the *S*-haplotype due to a phenomenon known as "competitive interaction" (de Nettancourt 2001).

In nature, potato CWR populations or plants can be reproductively isolated by external, internal, or both types of reproductive barriers. The external barriers can be spatial, temporal, or ecological (Hawkes and Hjerting 1969) whereas the internal ones can be pre-zygotic or post-zygotic, both of them either complete or incomplete (Camadro et al. 2004). In potato CWR, the action of internal reproductive barriers in a sexual reproduction cycle -either *in situ* or *ex situ*- can reduce the effective number of progenitors (Ne) in relation to the real number (N) of plants that composes the population; this reduction in size could alter the original allele frequencies and lead, over time, to eventual losses of genetic diversity due to allele fixation (Camadro 2012).

In germplasm banks in general, the most commonly used regeneration method (FAO 2013) consists of the cultivation of 20–25 plants/accession (N), controlled pollinations with a mixture of pollen collected when a certain percentage of the plants are in bloom, and composition of the regenerated accession with variable numbers of seeds/female progenitor. It is important to note that records on pollen sterility in the progenitors and other internal reproductive barriers that might be expressed from pollination to seed formation are not kept. To try to minimize possible changes in allele frequencies in *ex situ* regeneration cycles of potato CWR, Camadro (2012) proposed a method that takes into consideration both the action of internal reproductive barriers to estimate Ne in relation to N and the composition of the regenerated accession with the same number of seeds/female parent (given that only the female contribution is amenable of control).

In a previous work, Poulsen and Camadro (2020) analyzed the action of internal reproductive barriers to estimate Ne in relation to N in one accession of *Solanum chacoense* Bitter (Cl 898) cultivated in three growing seasons (S_1 = 2013/14, S_2 = 2014/ 15 and S_3 = 2015/16). Pollen viability and size, pollen-pistil relationships, and number and type (according to endosperm development) of seeds/fruit were studied. In the three seasons, both pre-zygotic and post-zygotic barriers were detected, which led to a reduction in the effective number of progenitors (Ne < N).

Following this line of investigation, the same accession was regenerated by the two previously mentioned methods. Thus, three populations were available for morphological and molecular studies to ascertain if one of those methods, if any, would better preserve the genetic diversity of the original accession in the progenies. The results of these studies are herein presented.

Materials and methods

Plant material

Original accession

Seeds of accession Cl 898 –classified according to the morphological phenotype as *S. chacoense* Bitter (2n = 2x = 24; 2EBN) for *ex situ* conservation, exchange, and other ends- were provided by the Potato and Forage Germplasm Bank, Agropecuarian Experimental Station “Domingo R. Pasquale”, National Institute of Agropecuarian Technology (INTA), Balcarce, Argentina. These seeds had been obtained at the mentioned germplasm bank by subjecting the seed

sample collected in the natural hábitat to one regeneration cycle following the standard FAO´s (2013) method. For the purpose o the present work, and because it was the accession we were randomly provided with, we will refer to it as the "original accession".

Derived populations

As described by Poulsen and Camadro (2020), the original accession was cultivated in a screen house in three growing seasons: S_1 = 2013/14, S_2 = 2014/ 15, and S_3 = 2015/16). In each season, when at least 75% of the plants were in bloom, pollen samples were taken from individual genotypes to estimate pollen viability, pollen size, and 2n pollen production. Then, controlled crosses were carried out with pollen samples from individual plants following an incomplete diallel mating design, and pollen-pistil relationships were determined in the generated genotypic combinations. Pollinations were also carried out with a mixture of pollen (pooled pollen) to generate a batch of seeds for applying both regeneration methods, FAO´s (2013) and Camadro´s (2012). Either no fruits or a low number of fruits (and, therefore, of seeds) were obtained from these pollinations (S_1 = 0 fruit, S_2 = S_3 = 1 fruit). Thus, the present investigation had to be carried out with seeds obtained from the previously mentioned controlled crosses, which were performed with individual pollen samples. Furthermore, seeds obtained in S_1 were used because the number of parental plants in that season (N = 25) was the highest of the three seasons (N = 16 in S_2; N = 20 in S_3) and the one recommended for regeneration of accessions.

Methods

To obtain the regenerated populations, seeds from the controlled crosses were treated with gibberellic acid (1500 ppm) for 24 h to standardize germination. Subsequently, they were placed in Petri dishes with damp filter paper (one seed = one genotype) at room temperature until germination. In the testa of seeds that did not germinate within the following 10 days, a small incision was made with a scalpel (Roa et al. 2007) to facilitate radicle emergence. When the derived seedlings reached the 2- to 4-leaf stage, they were transplanted into individual pots with a 3:1 mixture (v/v, soil: peat moss) and placed in a screen house. During plant cultivation, watering and agrochemicals (insecticide and fungicide) were applied as needed.

Each regenerated population was composed of 25 plants that reached the vegetative and reproductive stages. For the method proposed by Camadro (2012), the same number of plants was randomly taken from each genotypic combination (PR-MC). For the method of current use in germplasm banks (FAO 2013), the remaining plants of all the genotypic combinations were pooled and numbered, and then 25 of them were selected by using a table of random numbers (PR-MGB). Nonetheless, the designation "Regenerated Population by the Germplasm Bank Method (FAO 2013)" will be maintained for this population throughout this work.

Morphological characterization

Character evaluation

Twenty-one morphological characters were recorded – 11 quantitative (nine vegetative and two reproductive) and ten qualitative (eight vegetative and two reproductive) characters- in all the plants that composed the original accession (OA-S_1) and the two regenerated populations (PR-MC and PR-MGB) (Table 1). All measurements were performed in anthesis and, for leaf characters, on the fifth developed leaf from the base. For leaf and flower measurements, the free access ImageJ software (https://imagej.nih.gov/ij/; accessed: January 5, 2020) was used.

Data analysis

Univariate analysis

Qualitative characters. A visual description of qualitative (vegetative and reproductive) morphological characters was made and the proportions of plants that fell into each of the following categories was calculated: (a) terminal leaflet shape (*Tls*) and largest lateral leaflet shape (*Llls*): lanceolate, elliptical, ovate, heart-shaped, and ovate-lanceolate; (b) terminal leaflet base shape (*Tlbs*) and largest lateral leaflet base shape (*Llbs*): acute, obtuse, cordate, truncated and asymmetric; (c) terminal leaflet apex shape (*Tlas*) and largest lateral leaflet apex shape (*Llas*): acute, obtuse, acuminate and caudate; stem type (*St*): smooth, winged; (d) stem pigmentation (*Sp*): green, green with few purple spots, green and purple, green with many purple spots and purple; (e) flower color (*Fc*): white and white with cream spots; (f) corolla shape (*Cs*): very stellate, stellate, semi-stellate, pentagonal, very stellate-stellate, stellate-semi-stellate, stellate-pentagonal, semi-stellate-pentagonal, and deformed (actinomorphic). Percentages were calculated within characters.

Quantitative characters. Means, standard deviations, coefficients of variation, and ranges of vegetative and reproductive quantitative characters were calculated. Mean comparison tests were performed by Analysis of Variance (ANOVA, $p < 0.05$).

Multivariate analysis

Multivariate analyses were performed with the statistical program InfoGen (Balzarini and Di Rienzo 2014). The individual plants were the Operational Taxonomic Units (OTUs).

Quantitative characters

Principal Component Analysis (PCA). A basic data matrix (BDM) of the quantitative characters was built and standardized to unify the measuring scale of the characters. Then, the correlation matrix was estimated, and the eigenvalues and eigenvectors were calculated. The results were analyzed according to the projection of the first two principal components in bidimensional graphics. To obtain a better interpretation of the results, a Minimum Spanning Tree (MST), originating from the Euclidean distances matrix of the standardized quantitative data, was overlapped on the two-dimensional graphic for the first two Principal Components.

Cluster Analysis. Based on the standardized BDM, a similarity matrix was obtained for the OTU set by using Euclidean distances. The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was applied to obtain the phenogram. The cophenetic correlation coefficient was calculated to measure the degree of distortion of the phenogram with respect to the corresponding similarity matrix.

Qualitative characters

Principal Coordinate Analysis (ACooP). With the data of the qualitative characters, a distance matrix was calculated by using the Jaccard coefficient, and the eigenvalues were calculated. The results were analyzed according to the projection of the first two principal coordinates in bidimensional graphics. To obtain a better interpretation of the results, a Minimum Spanning Tree (MST), originating from the Euclidean distances matrix of the standardized quantitative data, was overlapped on the two-dimensional graphic for the first two Principal Coordinates.

Cluster Analysis. The analysis was performed as described for the quantitative characters.

Molecular analysis with microsatellite markers (SSR)

DNA extraction and quality analysis

DNA was extracted from young leaves of seedlings -approximately 10 cm high- according to Haymes (1996). The DNA A_{260}/A_{280} ratio and the concentration were recorded by using a spectrophotometer (Epoch™ Multi-Volume Spectrophotometer System). Samples with an A260/A280 ratio of 1.7–2.0 were considered of appropriate quality. DNA integrity was also verified by both agarose gel (0.7% p/v) electrophoresis and amplifications of the constitutive ubi3 gene. Samples were conserved at -20° C.

SSR markers and PCR amplification conditions

Eight SSR primer pairs were used (Table 2), three of which (stm1045, stm1049, stm0019) were developed by Milbourne et al. (1998), and the remaining five (stl001, stl004, stl018, stl019 and stl023) by Feingold et al. (2005). They were chosen because they are present in eight of the 12 chromosomes of the basic potato set and have shown polymorphism in various diploid wild potatoes (Raimondi et al. 2005, Bedonni and Camadro 2009, Leofanti et al. 2020).

For each primer pair, DNA samples of 75 plants were amplified and analyzed (25 plants of each population, OA- S_1, PR-MC, and PR-MGB). Each SSR stm amplification reaction was performed in a 25 µl solution containing 1.8 mM of ClMg, 1X reaction buffer, 0.2 mM of dNTPs, 0.2 µM of forward and reverse primers, 0.5 U of *GoTaq* DNA Polymerase (Promega), and 80–100 ng of DNA. The amplifications were carried out in a thermal cycler (VeritiTM 96-well, Applied Biosystems) as follows: 4 min at 94° C (initial denaturation), 35 cycles of 30 s at 94° C, 1 min at 55°C and 1 min at 72° C, with a final extension step of 5 min at 72° C. The amplified samples were kept at 4° C until removal from the thermal cycler. A similar protocol was followed for the stl SSRs, except that the amplification reactions were carried out in a 20 µl solution containing 1.5 mM of ClMg and 1 U of GoTaq DNA Polymerase.

SSR fragment amplified identification

Amplification products were separated by electrophoresis in 6% denaturing polyacrylamide gels with 1x TBE buffer. The approximate size of the amplified DNA fragments was determined with a 50bp marker (Promega). Gels were stained with an AgNO3 (0.15% p/v) solution. For each primer pair, fragment presence was coded as 1 and its absence as 0, and a binary data matrix was generated.

Allele identification

Amplified fragments in which the differences in size (bp) were proportional to the number of SSR repeats were considered alleles. Fragments that could not be related to repeat motifs were considered "polymerase stutters" and, therefore, discarded from the analysis. When only one amplification band was present for a given locus, its thickness was taken into account to decide whether the genotype had one or two copies of the same allele.

Data analysis

To facilitate the analysis, each SSR locus was identified with a capital letter and, following, a subscript number was given to each of its alleles (i.e., stl001 locus = A locus; alleles A_1 and A_2).

Allele frequencies for SSRs located at single loci and for SSRs located at duplicated loci were calculated. Allele frequency values equal to 1 indicated allele fixation in the population, whereas those equal to zero indicated allele loss.

To determine if there were significant differences in allele frequencies among OA-S_1 and the two regenerated populations, a chi-square test ($\alpha = 0.05$; with Yates correction) was performed.

Results

Morphological characterization

Univariate analysis

Qualitative characters

In the three populations, differences were observed in the proportions of plants that fell into the categories established for the ten evaluated vegetative and reproductive qualitative characters (Table 3). OA-S1 presented the greatest variability for most of the characters. In the three populations, *Tls* varied from elliptical (most frequently) to lanceolate, ovate, heart-shaped, and ovate-lanceolate in OA-S1; elliptical (most frequently) to lanceolate in PR-MC; and only

elliptical in PR-MG. *Tlbs* varied from obtuse (more frequently) to acute, truncated, and asymmetric in OA-S1; and from acute (more frequently) to obtuse and asymmetric in both regenerated populations. *Tlas* was acuminate (more frequently), acute, obtuse, or caudate in the OA-S1; and acute (more frequently) or acuminate in both regenerated populations. *Flls* varied from elliptical, lanceolate, ovate, heart-shaped to ovate-lanceolate in both OA-S1 and PR-MC, and from elliptical (more frequently) to ovate and ovate-lanceolate in PR-MGB. *Llbs* varied from asymmetric (more frequently), to cordate and truncated in OA-S1; asymmetric (more frequently) and cordate in PR-MC; and it was only asymmetric in PR-MGB. *Llas* varied from acute (more frequently) to obtuse, acuminate, and caudate in both OA-S1 and PR-MC, but it was only acute in PR-MGB. Both types of stems, smooth and winged, were observed in OA-S1 and the two regenerated populations, with green with purple spots (in different proportions) to purple pigmentation. *Fc* was white in the three populations and, in addition, white with cream spots in OA-S1. *Cs* varied from stellate (most frequently) to semi-stellate, pentagonal, stellate-semi-stellate, stellate-pentagonal to semi-stellate-pentagonal in OA-S1; very stellate to stellate (most frequently), pentagonal, stellate-pentagonal, and semi-stellate-pentagonal in PR-MC; and very stellate to stellate (most frequently), stellate-semi-stellate, stellate-pentagonal and semi-stellate-pentagonal in PR-MGB. Actinomorphous flowers were observed in both OA-S1 and PR-MGB (Table 3).

Quantitative characters

Significant differences (ANOVA, $p < 0.05$) were detected between OA-S_1 and both regenerated populations for the mean of the following characters: *5thLl*, *5thLw*, *no.Ll*, *Llll*, *Lllw*, *Tll*, and *Fd*. OA-S_1 presented the highest mean, followed by PR-MGB and PR-MC, respectively. In OA-S_1 and both regenerated populations, the most variable characters were *no.il* and Lill (Table 4). OA-S_1 presented the highest number of characters with a greater range (*5thLl*, *5thLw*, *Llll*, *Tll*, *no.P*), followed by PR-MC (*5thLl*, *5thLw*, *Tll*, *no.P*), and PR-MGB (*5thLl*, *5thLw*, *no.P*), respectively (Table 4).

Multivariate analysis

Quantitative characters

Principal Component Analysis (PCA)

In OA-S_1, the first two Principal Components (PC1 and PC2) of the Principal Component Analysis explained 61.0% of the total observed variability. PC1 explained 40.0% of the total variability, being *5thLw*, *Llll*, *Lllw*, *Tll* and *Tlw* the characters with the greatest contribution. PC2 explained 21.0% of the total variability, with *5thLl, no.Ll*, *no.il* and *Lill* being the characters with the greatest contribution (Fig. 1). By overlapping the Minimum Spanning Tree (MST) on the two-dimensional graphic for the first two PC, three groups with similar numbers of plants were observed and distributed, respectively, in the two upper and lower central quadrants (Group 1, n = 9 and Group 2, n = 7; Group 3, n = 8) (Fig. 1).

In PR-MC, the first two Principal Components (PC1 and PC2) explained 65.0% of the total observed variability. PC1 explained 47.0% of the total variability, being *5thLw*, *5thLl*, *Llll*, *Lllw*, *Tll* y *Tlw* the characters with the greatest contribution. PC2 explained 17.0% of the total variability, being *no.il* y *Lill* the characters with the greatest contribution (Fig. 1). By overlapping the MST on the two-dimensional graphic for the first two PC, four groups were observed, a central one with the highest number of plants (Group 3; n = 13) and three smaller ones distributed, respectively, in the left (Group 1, n = 3) and the right lower quadrants (Group 4, n = 2), and in the left upper quadrant (Group 2, n = 1) (Fig. 1).

In PR-MGB, the first two Principal Components (PC1 and PC2) explained 63.0% of the total observed variability. PC1 explained 43.0% of the total variability, being *5thLw*, *Llll*, *Lllw*, *Tll* and *Tlw* the characters with the greatest contribution. PC2 explained 20.0% of the total variability, with *no.il* and *Lill* as the characters with the greatest contribution (Fig. 1), and *no.P* the reproductive character that contributed negatively to this component. By overlapping MST on the two-dimensional graphic, four groups were observed, a central one with the highest number of plants (Group 3, n = 7), and three smaller ones distributed, respectively, in the left (Group 1, n = 3) and in the right lower quadrants (Group 4, n = 4), and the third in the central upper quadrant (Group 2, n = 1) (Fig. 1).

In the phenogram obtained from the Clustering Analysis, the same groups were observed in the three populations, as observed in the previously described PCA (Figure S1, Electronic Supplementary Material).

Qualitative characters

Principal Coordinate Analysis (PCooA)

In OA-S_1, the first two Principal Coordinates (PCoo1 and PCoo2) of the Principal Coordinate Analysis (PCooA) represented 31.7% of the total dispersion of the genotypes, corresponding 17.4% of this value to the first coordinate, and 14.3% to the second one. By overlapping the Minimum Spanning Tree (MST) on the two-dimensional graphic for the first two Principal Coordinates, eight groups were observed, with variable numbers of plants and widely distributed in the four quadrants (Group 1, n = 4 and Group 2, n = 5; Group 3, n = 2, Group 4, n = 1, Group 5, n = 6, Group 6, n = 1, Group 7, n = 2, Group 8, n = 4) (Fig. 2).

In PR-MC, the first two Principal Coordinates (PCoo1 and PCoo2) of the Principal Coordinate Analysis (PCooA) represented 44.7% of the total dispersion of the plants, corresponding 26.7% of this value to the first coordinate and 18.0% to the second one. By overlapping the Minimum Spanning Tree (MST) on the two-dimensional graphic for the first two Principal Coordinates, four groups differing in the number of plants were observed distributed in the four quadrants (Group 1, n = 7; Group 2, n = 3; Group 3, n = 5; Group 4, n = 3) (Fig. 2).

In PR-MGB, the first two Principal Coordinates (PCoo1 and PCoo2) of a Principal Coordinate Analysis (PCooA) represented 39.4% of the total dispersion of the plants, corresponding 22.7% of this value to the first coordinate and 16.7% to the second. By overlapping the Minimum Spanning Tree (MST) on the two-dimensional graphic for the first two Principal Coordinates, four groups (three with similar numbers of plants and a smaller one) were observed (Group 1, n = 5; Group 2, n = 1; Group 3, n = 5; Group 4, n = 4) in the four quadrants (Fig. 2).

In the phenogram obtained from the Clustering Analysis, the groups observed in the three populations were like those previously described in the PCooA (Figure S2, Electronic Supplementary Material).

Molecular analysis with microsatellite markers (SSR)

DNA extraction and quality analysis

All the samples presented an A_{260}/A_{280} relation between 1.8–2.0, thus, they were considered of good quality. These samples presented amplification products of the ubi3 constitutive gene, corroborating their integrity.

Amplified fragment identification

Amplification products were obtained with six of the eight primer pairs (stl001, stm1049, stl018, stl004, stl019 y stm1045) (Fig. 3).

Allele identification

For all analyzed SSR markers (except for stl004 and stm1045), the number of amplified fragments corresponded to the expected number of alleles. For the stl004 marker, the 89, 86, and 83 bp fragments were considered alleles, whereas the rest were considered polymerase enzyme stutters. For stm1045, five amplified fragments of 198 bp, 189 bp, 180 bp, 177 bp, and 170 bp were observed. The 177 bp band was discarded in the analysis because the staining was very light and, therefore, it was poorly defined. The remaining bands corresponded to this marker's four alleles (Table 5).

Data analysis

Allele frequencies

From the electrophoretic pattern of the SSR markers, a binary data matrix was obtained, and allele frequencies were estimated in the three populations (Tables 6, 7, and 8).

Significant statistical differences ($\alpha = 0.05$) in allele frequencies were observed between (1) OA-S_1 and PR-MC for stm1049 and stl018, and (2) OA-S_1 and both regenerated populations for stl004 and stl019. For stl004, loss of the D3 allele was observed in PR-MC (Table 9).

Discussion

Morphological characterization

The main assumption underlying *ex situ* conservation of CWR in germplasm banks is that the genetic diversity of natural populations can be captured in the sampling process and conserved over the years in the form of either the original sample (accession) or the populations that are *ex situ* derived from it. In sexually reproducing species, *ex situ* seed regeneration and/or multiplication cycles are carried out when the percentage of germination diminishes below a certain value or when propagules are required for exchanging with other germplasm banks or their use in basic or applied studies, or in breeding programs. Moreover, for taxonomic species, homogeneity in morphological phenotypes is expected in natural populations and their samples. In fact, plants that do not respond to the type are not sampled in the field or they are discarded during the regeneration or multiplication processes, with the consequent reduction in the gene pool and the eventual loss of genetic diversity (Camadro 2012, Camadro and Rimieri 2021).

In the present work, the Univariate Data Analysis revealed great morphological variability for the evaluated qualitative and quantitative characters -both vegetative and reproductive- within each of the three populations (original and the two regenerated ones), and among them, with OA-S_1 being the most variable population. The extent of this variability is not the expected in a taxonomic species according to the Taxonomic Species Concept (= TSC). In this regard, it must be considered that gene flow in natural wild potato populations can occur through sexual propagules (seeds) and/or asexual (tubers and stolons) disseminated in the landscape by abiotic and biotic agents, and by pollen transported by their pollinating insects. In the absence of internal reproductive barriers (or if present, but the barriers are incomplete) hybrids are generated whose morphological phenotypes might resemble those described for a given taxonomic species or be intermediate between taxonomic species (see examples in Hawkes and Hjerting 1969, and in Bedonni and Camadro 2009).

In Argentina, Hawkes and Hjerting (1969) described collections of the wild species *Solanum microdontum* Bitter, *Solanum venturii* Hawkes and Hjerting, and *Solanum vernei* Bitter and Wittm. growing in the same geographical area where the accession analyzed in this work -classified as *S. chacoense* for conservation purposes- was collected (Tafí del Valle; altitude: 2,200-3,000 m.a.s.l.). Likewise, in nearby sites, the same authors reported the presence of natural hybrids between *S. chacoense* and *S. microdontum*. Nowadays, populations of wild potatoes are very frequent in this area, particularly of *S. microdontum* and *S. venturii* Hawkes and Hjerting, where they grow both in the open field and as weeds of commercial potato crops (Camadro and Erazzú, personal comm.). Therefore, it can be inferred that the morphological variability observed in OA-S_1 could be the result of hybridization and introgression between sympatric populations within the limits of the sampling site, followed by gene segregation and recombination during sexual reproduction both in the first *ex situ* regeneration cycle performed at the germplasm bank and in the present work.

The extent of the variability observed in each of the three analyzed populations could be due to the unequal number of plants (genotypes) evaluated in each of them, phenotypic plasticity in the expression of quantitative characters in response to an environmental factor -which is a common phenomenon in wild potatoes (Hawkes and Hjerting 1969), with an epigenetic base (Cara et al. 2013)-, or both. In fact, a higher number of plants were morphologically

characterized in OA-S_1 because the action of a post-zygotic barrier (hybrid weakness) prevented several plants from reaching the reproductive stage in the regenerated populations. Moreover, internal barriers in the original accession led to Ne < N, thus, only 36% of the plants made a genetic contribution to the next generation. A decrease in morphological variability and genetic diversity might be, therefore, expected after *ex situ* regeneration cycles if the action of reproductive barriers in the parental populations results in the narrowing of the allele pool of the regenerated ones. This would create a conflict for *ex situ* conservation of wild population samples under specific names if the taxonomic classification is based only on morphological phenotypes. This type of classification is still the most frequently used in wild potatoes even though molecular tools have been incorporated (see Spooner et al. 2014); notwithstanding, these tools are applied to plant materials previously classified according to morphological phenotypes. Despite the taxonomic classification, in our opinion, it would be advisable to conserve individual accession as gene pools from their natural sites.

Regarding the Multivariate Data Analysis, the two-dimensional graphs of the PCA and the ACooP for the three populations revealed a tendency for the plants to concentrate in various groups. However, a more homogeneous distribution would have been expected for a taxonomic species, with plants tending to concentrate in a single group. The previous discussion on the origin and extent of the observed morphological variability holds in this case; however, it cannot be ruled out that the differences observed in both the number of groups and the number of plants included in each of them could be due to the different numbers of plants evaluated in each population.

Molecular analysis

In this work, allele frequencies were calculated for all analyzed SSR -and first reported for a wild potato from Argentina- because the number of amplified fragments per plant corresponded to the maximum number of alleles expected in diploids. Previously, Leofanti et al. (2020) worked with seven SSRs for the study of a morphologically heterogeneous wild potato natural population from Tucumán, Argentina, two of which (stm1049 and stm1045) were used in the present work. These authors observed that the maximum number of fragments obtained was higher than expected for diploid plants; therefore, they directly performed the analysis of amplified fragments, without genetic interpretation. For their part, Bedonni and Camadro (2009) also reported difficulties in calculating allele frequencies in accessions of several taxonomic species from western Argentina with 11 SSR markers, three of which (stm1045, stm1049 and stm0019) were used in the present study. The fact that these markers were designed for cultivated potatoes may explain the difficulties encountered when trying to use them in wild potatoes because rearrangements and/or duplications might have occurred in the wild genome evolution. In this sense, Westman and Kresovich (1998) reported differences in the number of copies of the repeated sequences and/or the flanking regions in other plant groups.

As previously discussed, the presence of pre-zygotic and post-zygotic internal reproductive barriers led to Ne < N in the crossing work, thus, not all cultivated plants contributed alleles to the regenerated populations. Changes in allelic frequencies were observed between OA-S_1 and the two regenerated populations, however, homozygosity for the SSR alleles in most of the parental plants involved in the compatible combinations obscured the results. In this regard, it must be considered that these changes can only be detected if the parental genotypes are segregating for the loci under consideration.

Conclusions

The presence of pre- and post-zygotic reproductive barriers in OA-S_1 reduced the effective number of parental plants (Ne) in relation to the real number (N) of plants cultivated for the *ex situ* seed regeneration study. Since OA-S_1 was generated in the germplasm bank from the original field sample, it cannot be ruled out that a previous reduction in Ne occurred in that sexual cycle. Notwithstanding, both the field sample and this population are labeled and conserved as accession Cl 898, a common practice in germplasm banks, even though there could be differences in their gene pools.

When the size of a population is reduced, the probability of inter-crossing between related individuals (inbreeding) increases. Consequently, there is an increase in homozygosity and the expression of masked potentially recessive lethal or deleterious alleles in allogamous plants such as potatoes. In this way, the genetic diversity that could have been captured at the sample site might be reduced in the seed regeneration or multiplication cycles. Therefore, it is important to register both the Ne and the presence of internal reproductive barriers in the accessions´ passports as a warning to curators and breeders (see Camadro 2012). To curators, because it would be advisable to carry out new collections in the originally sampled geographical sites when the Ne is low, to be able to *ex situ* conserve a large proportion of the available natural genetic diversity. To breeders, because it would allow them to evaluate if it is more appropriate to use either many plants of a few accessions or a few plants of many accessions for their purposes, augmenting the efficiency of the screening work. Moreover, the contribution of the female parent should always be controlled to compose the regenerated accession, to minimize changes in allelic frequencies.

Finally, and as Camadro and Rimieri (2021) have discussed, the knowledge of the genetic structure of populations is of utmost importance to developing germplasm conservation approaches. This knowledge would allow the maximum available natural genetic diversity of a species to be represented in the accessions and in their regenerated populations throughout the various germplasm banks´ operations.

Declarations

Compliance with ethical standards:

Funding: This study was partially funded by Instituto Nacional de Tecnología Agropecuaria (INTA; PNHFA-1106074), Universidad Nacional de Mar del Plata (UNMdP; AGR 344/11), and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET; first author´s scholarship).

Conflict of interest: Authors Poulsen Hornum Anabela and Elsa Lucila Camadro declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Balzarini MG, Di Rienzo JA (2004) InfoGen. [software estadistico] Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina. http://www.info-gen.com.ar
2. Bedonni MC, Camadro EL (2009) Morphological and molecular evidence of natural interspecific hybridization in the diploid potato *Solanum kurtzianum* from Argentina. Botany 87:78-87. https://doi.org/10.1139/B08-116
3. Camadro EL (2012) Relevance of the genetic structure of natural populations, and sampling and classification approaches for conservation and use of wild crop relatives: potato as an example. Botany 90:1065-72. https://doi.org/10.1139/b2012-090
4. Camadro EL, Rimieri P (2021) *Ex situ* plant germplasm conservation revised at the light of mechanisms and methods of Genetics. BAG. J Basic Appl Genet 32 (1): 11-24. https://doi.org/10.35407/bag.2020.32.01.02
5. Camadro EL, Carputo D, Peloquin SJ (2004) Substitutes for genome differentiation in tuber-bearing Solanum: interspecific pollen–pistil incompatibility, nuclear-cytoplasmic male sterility, and endosperm. Theor Appl Genet 109:1369–1376. https://doi.org/10.1007/s00122-004-1753-2
6. Cara N, Marfil CF, Masuelli RW (2013) Epigenetic patterns newly established after interspecific hybridization in natural populations of *Solanum*. Ecol Evol 3 (11):3764–3779. https://doi.org/10.1002/ece3.758
7. de Nettancourt D (2001) Incompatibility and incongruity in wild and cultivated plants (2nd ed.) Springer-Verlag, Berlin
8. FAO (2013) Genebank Standars for Plant Genetic Resources for Food and Agriculture. Commission of genetic resources for food and agricultura. Food and Agriculture Organization of the United Nations, Roma. http://www.fao.org/3/a-i7788e.pdf. Accessed 13 March 2019
9. Feingold S, Lloyd J, Norero N, Bonierbale M, Lorenzen J (2005) Mapping and characterization of new EST-derived microsatellites for potato (*Solanum tuberosum* L.). Theor Appl Genet, 111 (3): 456-466. https://doi.org/10.1007/s00122-005-2028-2
10. Grant V (1981) Plant speciation (2nd ed.) Columbia. University Press, New York.
11. Haymes KM (1996) Mini-prep method suitable for a plant breeding program. *Plant Molecular Biology Reporter*, 14, 280-284. https://doi.org/10.1007/BF02671664
12. Hawkes JG (1990) The potato: evolution, biodiversity and genetic resources. Belhaven Press, London
13. Hawkes JG, Hjerting JP (1969) The potatoes of Argentina, Brazil, Paraguay and Uruguay. A biosystematic study. Oxford University Press, Oxford
14. Jansky SH, Dempewolf H, Camadro EL, Simon R, Zimnoch-Guzowska E, Bisognin DA, Bonierbale M (2013) A case for crop wild relative preservation and use in Potato. Crop Sci 53:746-54. https://doi.org/10.2135/cropsci2012.11.0627
15. Leofanti GA, Camadro EL, Erazzú LE (2021) Over Time Variation in Microsatellite (SSR) Patterns in a Natural Wild Potato Population from Northwest Argentina. Genet Resour Crop Evol. https://doi.org/10.1002/csc2.20498
16. Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterisation and mapping of simple sequence repeat loci in potato. Molecular and General Genetics MGG, 259, 233-245. https://doi.org/10.1007/s004380050809
17. Poulsen Hornum A, Camadro EL (2020) Expression of internal reproductive barriers in a germplasm bank accession of the wild potato Solanum chacoense Bitter in three *ex situ* regeneration cycles. Genet Resour Crop Evol. https://doi.org/10.1007/s10722-020-01034-y
18. Raimondi JP, Peralta IE, Masuelli RW, Feingold, S, Camadro EL (2005) Examination of the hybrid origin of the wild potato *Solanum ruiz-lealii* Brücher. *Plant Systematics and Evolution*, 253, 33-51. https://doi.org/10.1007/s00606-005-0299-3
19. Roa NK, Hanson J, Dulloo ME, Ghosh K, Nowell D, Larinde M (2007) Manual para el manejo de semillas en bancos de germoplasma. Manual para Bancos de Germoplasma Nº8. Bioversity International, Roma, Italia
20. Spooner DM, Ghislain M, Simon R, Jansky SH, Gavrilenko T (2014) Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. Bot Rev 80:283–383. https://doi.org/10.1007/s12229-0149146-y
21. Westman AL, Kresovich S (1998) The potential for cross-taxa simple-sequence repeat (SSR) amplification between Arabidopsis thaliana L. and crop brassicas. Theor Appl Genet 96 (2): 272-281

Tables

Table 1 Quantitative and qualitative (vegetative and reproductive) morphological characters evaluated in the original accession CI 898 of *Solanum chacoense* Bitter, in the population regenerated by the method proposed by Camadro (2012) and the population regenerated by the method commonly used in germplasm banks (FAO 2013)

Character (abbreviation)	Unit/Category
Quantitative	
Vegetative	
1- 5th leaf length (5^{th}Ll)	cm
2- 5th leaf width (5^{th}Lw)	cm
3- Lateral leaflets (no.Ll)	pairs
4- Longest lateral leaflet length (Llll)	cm
5- Longest lateral leaflet width (Lllw)	cm
6- Terminal leaflet length (Tll)	cm
7- Terminal leaflet width (Tlw)	cm
8- Interjected leaflets (no.il)	pairs
9- Longest interjected leaflets length (Lill)	cm
Reproductive	
1- Flower diameter (Fd)	cm
2- Petals (no.P)	no.
Qualitative	
Vegetative	
1- Stem pigmentation (Sp)	(1) green, (2) green with few purple spots, (3) green and purple, (4) green with many purple spots, (5) purple
2- Stem type (St)	(1) smooth, (2) winged
3- Largest lateral leaflet shape (Llls)	(1) lanceolate, (2) elliptical, (3) ovate, (4) heart-shaped, (5) ovate-lanceolate
4- Largest lateral leaflet base shape (Lllbs)	(1) acute, (2) obtuse, (3) cordate, (4) truncated, (5) asymmetric
5- Largest lateral leaflet apex shape (Llas)	(1) acute, (2) obtuse, (3) acuminate, (4) caudate
6- Terminal leaflet shape (Tls)	(1) lanceolate, (2) elliptical, (3) ovate, (4) heart-shaped, (5) ovate-lanceolate
7- Terminal leaflet base shape (Tlbs)	(1) acute, (2) obtuse, (3) cordate, (4) truncated, (5) asymmetric
8- Terminal leaflet apex shape (Tlas)	(1) acute, (2) obtuse, (3) acuminate, (4) caudate
Reproductive	
1- Corolla shape (Cs)	(1) very stellate, (2) stellate, (3) semi-stellate, (4) pentagonal, (5) very stellate-stellate, (6) stellate-semi-stellate, (7) stellate-pentagonal, (8) semi-stellate- pentagonal, (9) deformed (actinomorphic)
2- Flower color (Fc)	(1) white, (2) white with cream spots

Table 2 Assayed SSR microsatellite markers

SSR	Primer sequence 5'-3'[a]	Repeat	Expected product size (bp)	Chromosome location
stm1045	F GAAGTTTTATCAGAAC R ATCACCTCATCAGCAC	(TGG)n	181	II – XII
stm1049	F CTACCAGTTTGTTGATTGTGGTG R AGGGACTTTAATTTGTTGGACG	$(ATA)_6$	189	I
stm0019	F AATAGGTGTACTGACTCTCAATG R TTGAAGTAAAAGTCCTAGTATGTG	(AT)n(GT)n(AT)n (GT)n(GC)n(GT)n	204	VI
stI001	F CAGCAAAATCAGAACCCGAT R GGATCATCAAATTCACCGCT	(AAT)n	188	IV
stI004	F GCTGCTAAACACTCAAGCAGAA R CAACTACAAGATTCCATCCACAG	(AAG)n	103	VI
stI018	F CCACTACTGCTTCCTCCACC R GCAGCAACAACAAGCTCAAC	(GTT)n	189	XI
stI 019	F TCCCTGTTGCCTTGAACAAT R TGGGAAAAGGTACAAAGACGA	(ATCT)n	126	VII
stI 023	F GCGAATGACAGGACAAGAGG R TGCCACTGCTACCATAACCA	(GGA)n	193	X

[a]F= forward, R= reverse; stm: Milbourne et al. (1998); stI: Feingold et al. (2005)

Table 3 Proportions of plants in various categories of 10 qualitative (vegetative and reproductive) morphological characters evaluated in 25 plants of the original accession CI 898 of *Solanum chacoense* Bitter (OA-S1)[a], 19 plants of the population regenerated by the method proposed by Camadro (2012) (PR-MC)[b] and in 15 plants of the population regenerated by the method commonly used in germplasm banks (FAO 2013) (PR-MGB)[c]

Qualitative character	Character/category		Population		
			OA-S_1[a]	PR-MC[b]	PR-MGB[c]
Vegetative	Tls	1: lanceolate *	0.04	0.05	0.00
		2: elliptical*	0.76	0.95	1.00
		3: ovate*	0.12	0.00	0.00
		4: heart-shaped	0.04	0.00	0.00
		5: ovate-lanceolate	0.04	0.00	0.00
	Tlbs	1: acute	0.20	0.58	0.40
		2: obtuse *	0.60	0.21	0.33
		3: cordate *	0.00	0.00	0.00
		4: truncated	0.08	0.00	0.00
		5: asymmetric	0.12	0.21	0.27
	Tlas	1: acute	0.36	0.63	0.73
		2: obtuse	0.08	0.00	0.00
		3: acuminate *	0.44	0.37	0.27
		4: caudate	0.12	0.00	0.00
	Llls	1: lanceolate *	0.12	0.11	0.00
		2: elliptical *	0.04	0.37	0.60
		3: ovate *	0.24	0.21	0.20
		4: heart-shaped	0.04	0.05	0.00
		5: ovate-lanceolate	0.56	0.26	0.20
	Llbs	1: acute	0.00	0.00	0.00
		2: obtuse *	0.00	0.00	0.00
		3: cordate *	0.12	0.11	0.00
		4: truncated	0.04	0.00	0.00
		5: asymmetric	0.84	0.89	1.00
	Llas	1: acute	0.64	0.70	1.00
		2: obtuse	0.08	0.05	0.00
		3: acuminate *	0.24	0.20	0.00
		4: caudate	0.04	0.05	0.00
	St	1: smooth *	0.64	0.63	0.33
		2: winged	0.36	0.37	0.67
	Sp	1: green *	0.00	0.21	0.27
		2: green with few purple spots	0.56	0.53	0.47
		3: green and purple	0.12	0.15	0.00
		4: green with many purple spots	0.28	0.11	0.20
		5: purple	0.04	0.00	0.06
Reproductive	Fc	1: white *	0.96	1.00	1.00
		2: white with cream spots	0.04	0.00	0.00
	Cs	1: very stellate	0.00	0.26	0.13
		2: stellate *	0.20	0.34	0.53
		3: semi-stellate	0.04	0.00	0.00
		4: pentagonal	0.16	0.06	0.00
		5: very stellate-stellate	0.00	0.00	0.00
		6: stellate-semi-stellate	0.08	0.00	0.07
		7: stellate-pentagonal	0.20	0.17	0.07
		8: semi-stellate-pentagonal	0.16	0.17	0.13
		9: deformed (actinomorphic)	0.16	0.00	0.07

Ref: Tls= Terminal leaflet shape; Tlbs = Terminal leaflet base shape; Tlas= Terminal leaflet apex shape; Llls = Largest lateral leaflet shape; Lllbs = Largest lateral leaflet base shape; Lllas= Largest lateral leaflet apex shape; St = Stem type; Sp = Stem pigmentation; Fc = Flower color; Cs = Corolla shape.

*Note: these categories match the ones given for *S. chacoense* Bitter by Hawkes y Hjerting (1969)

Table 4 Mean (), standard deviation (SD), coefficient of variation (CV), and range of 11 quantitative morphological (vegetative and reproductive) characters evaluated in 25 plants of the original accession Cl 898 of *Solanum chacoense* Bitter (OA-S_1)[a], 19 plants of the population regenerated by the method proposed by Camadro (2012) (PR-MC)[b] and 15 plants of the population regenerated by the method commonly used in germplasm banks (FAO 2013) (PR-MGB)[c]

Quantitative character	Population								
	OA-S_1[a]			PR-MC[b]			PR-MGB[c]		
	x̄ ± SD	CV	Range	x̄ ± SD	CV	Range	x̄ ± DS	CV	Range
Vegetative									
5ª Ll	17.4* ± 1.9	10.9	13.1 – 20.5 [b]	15.5* ± 2.7	17.2	11,2 - 19,1 [b]	16.6* ± 2.3	14.2	12.8 – 19.8 [b]
5ª Lw	9.9* ± 1.4	13.9	7.5 – 14.0 [b]	8.7* ± 1.3	14.8	6.5 – 11.6 [b]	9.1* ± 0.9	10.4	7.8 – 10.8 [b]
no.Ll	4.3* ± 0.7	17.3	2.0 – 5.0	3.3* ± 0.5	14.5	3.0 – 4.0	3.4* ± 0.5	14.9	3.0 – 4.,0
Llll	5.1* ± 0.8	15.7	3.9 – 7.5 [b]	4.5* ± 0.7	14.9	3.0 – 5.5	4.7* ± 0.6	11.8	3.9 – 5.6
Lllw	2.1* ± 0.5	21.9	1.2 – 3.0	1.8* ± 0.4	23.0	1.0 – 2.5	1.9* ± 0.3	15.4	1.4 – 2.3
Tll	5.9* ± 1.1	18.6	4.3 – 7.8 [b]	6.6* ± 0.9	13.7	5.0 – 8.8 [b]	6.8* ± 0.7	10.3	6.0 – 7.9
Tlw	2.5 ± 0.6	23.4	1.5 – 3.7	2.3 ± 0.5	19,.9	1.3 – 3.0	2.4 ± 0.3	13.7	1.9 – 3.2
no.il	2.0 ± 1.4	68.6 [a]	0.0 – 4.0	1.7 ± 1.3	76.8 [a]	0.0 – 4.0	1.8 ±1.5	81.9 [a]	0.0 – 5,0
Lill	0.4 ± 0.3	68.8 [a]	0.0 – 1.0	0.4 ± 0.3	81.9 [a]	0.0 – 0.9	0.4 ± 0.3	67.0 [a]	0.0 – 1.0
Reproductive									
Fd	2.4* ± 0.3	12.1	1.7 – 3.0	2.7* ± 0.3	12.3	2.0 – 3.3	2.9* ± 0.3	10.8	2.1 – 3.5
no.P	5.0 ± 0.4	7.0	4.0 – 6.0 [b]	5.1 ± 0.2	4.5	5.0 – 6.0 [b]	4.9 ± 0.3	5.2	4.0 - 5,0 [b]

Ref: 5ª Ll = 5th leaf length, 5ª Lw = 5th leaf width, no.Ll = number of pairs of lateral leaflets, Llll = longest lateral leaflet length, Lllw = Longest lateral leaflet width, Tll = Terminal leaflet length, Tlw = Terminal leaflet width, no.il = number of pairs of interjected leaflets, Lill = Longest interjected leaflets length, Fd= Flower diameter, no.P = number of petals. *significant differences (ANOVA, $p < 0.05$). [a]Most variable characters; [b]characters with a higher range

Table 5 Number (no.) and size of amplified fragments and alleles for six SSR markers

SSR	amplified fragment		alleles	
	no.	size (pb)	no.	size (pb)
stl001	2	193, 184	2	193, 184
stm1049	3	210, 180, 164	3	210, 180, 164
stl018	3	204, 201, 192	3	204, 201, 192
stl004	9	89, 88, 87,86, 85, 84, 83, 82, 81	3	89, 86, 83
stl019	3	146, 138, 130	3	146, 138, 130
stm1045	5	198, 189, 180, 177, 170	4	198, 189, 180, 170

Table 6 Binary data matrix and allele frequencies for the stl001 marker in 25 plants of the original Cl 898 accession of *Solanum chacoense* Bitter (OA-S_1)

Attribute	Plant no.																									Allele Frequency
	1 (chc2)	2 (chc4)	3 (chc5)	4 (chc10)	5 (chc11)	6 (chc12)	7 (chc13)	8 (chc15)	9 (chc16)	10 (chc17)	11 (chc18)	12 (chc19)	13 (chc20)	14 (chc21)	15 (chc22)	16 (chc23)	17 (chc24)	18 (chc26)	19 (chc27)	20 (chc29)	21 (chc30)	22 (chc31)	23 (chc33)	24 (chc35)	25 (chc37)	
Allele																										
A1	0	1	0	0	0	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	$p_{(A1)}$= 0.44[b]
A2	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	1	1	0	0	1	1	1	0	1	1	$q_{(A2)}$= 0.56[c]
Genotype	A2 A2	A1 A2	A2 A2	A2 A2	A2 A2	A1 A1	A2 A2	A1 A1	A1 A2	A1 A2	A1 A2	A1 A1	A2 A2	A1 A1	A2 A2	A2 A2	A1 A2	A1 A1	A1 A1	A1 A2	A1 A2	A1 A2	A1 A1	A2 A2	A2 A2	

[a]$p_{(A1)} = [2(_{A1A1}) + (_{A1A2})]/2n = [2(7) + (8)]/2(25) = 0.44$; [c]$q_{(A2)} = [2(_{A2A2}) + (_{A1A2})]/2n = [2(10) + (8)]/2(25) = 0.56$

Table 7 Binary data matrix and allele frequencies for the stl001 marker in 25 plants of the population regenerated by the method proposed by Camadro (2012)

Attribute	Genotypic combination																			
	2 × 26			18 × 2			18 ×10			18 × 12			18 × 19			18 × 26			31 × 4	
Plant no.	1	2	3	4	5	6	7	8	9	10	11	12	13[a]	14	15	16	17	18	19	20
Allele																				
A1	1	1	1	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1
A2	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	0	0	1
Genotype	A1 A2	A1 A2	A1 A2	A2 A2	A2 A2	A2 A2	A2 A2	A2 A2	A1 A2	A1 A2	A1 A1	A1 A1	A2 A2	A2 A1	A2 A2	A1 A2	A1 A2	A1 A1	A1 A1	A1 A2

[a]non-amplified allele; [b]$p_{(A1)}$= [2($_{A1A1}$) + ($_{A1A2}$)]/2(25)= [2(5) + (12)]/2(25)= 0.44; [c]$q_{(A2)}$= [2($_{A2A2}$) + ($_{A1A2}$)]/2(= [2(8) + (12)]/2n=0.56

Table 8 Binary data matrix and allele frequencies for the stI001 marker in 25 plants of the population regenerated by the method commonly used in germplasm banks (FAO 2013)

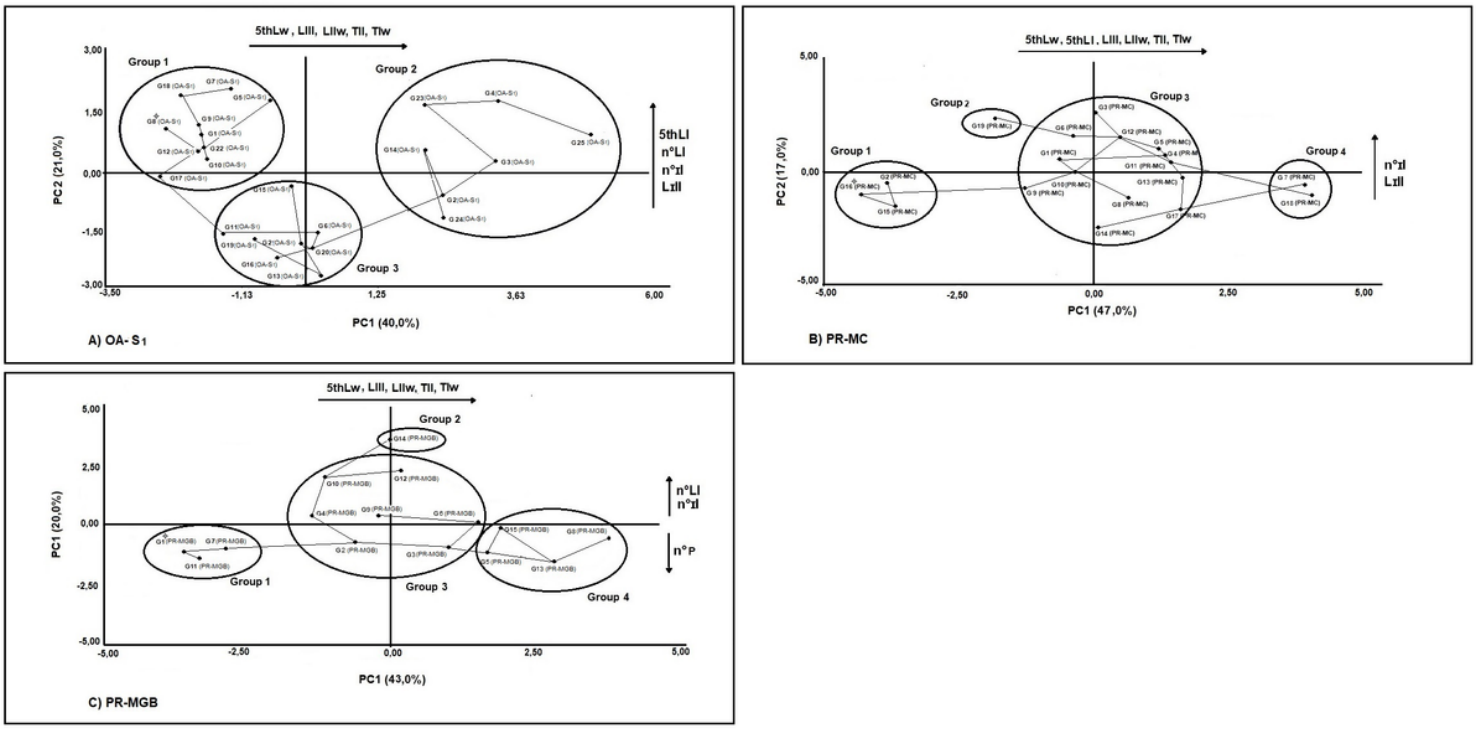
Attribute	Genotypic combination																									Allele frequency
	2×26	10 × 26							18 × 2				18×10	18×12	18 × 20					18×26		31 × 4			18×19	
Plant no.	1	2[a]	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
Allele																										
A1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	$p_{(A1)}$ =0.48[b]
A2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	$q_{(A2)}$= 0.52[c]
Genotype	A1 A2	A1 A2	A1 A2	A1 A2	A1 A2	A1 A2	A1 A2	A1 A2	A1 A2	A2 A2	A2 A2	A2 A2	A2 A2	A1 A2	A1 A2	A1 A2	A1 A2	A2 A2	A1 A2	A1 A2	A1 A1	A1 A1	A1 A1	A1 A1	A1 A2	

Ref: [a] a number assigned to each genotype of a genotypic combination [b]$p_{(A1)}$= [2($_{A1A1}$) + ($_{A1A2}$)]/2n= [2(5) + (14)]/2(25)= 0.48; [c]$q_{(A2)}$= [2($_{A2A2}$) + ($_{A1A2}$)]/2n= [2(6) + (14)]/2(25)=0.52

Table 9 **Allele frequencies for six SSR markers in 25 plants of each, the original accession CI 898 of *Solanum chacoense* Bitter (OA-S_1)[a], the population regenerated by the method proposed by Camadro (2012) (PR-MC)[b], and the population regenerated by the method commonly used in germplasm banks (FAO, 2012) (PR-MGB)**

SSR	Population		
	OA-S_1[a]	PR-MC[b]	PR-MGB[c]
stl001			
$p_{(A1)}$	0.44	0.44	0.48
$q_{(A2)}$	0.56	0.56	0.52
stm1049[a]			
$p_{(B1)}$	**0.08**	**0.18**	0.10
$q_{(B2)}$	**0.18**	**0.28**	0.26
$r_{(B3)}$	**0.74**	**0.54**	0.64
stl018[a]			
$p_{(C1)}$	**0.32**	**0.40**	0.46
$q_{(C2)}$	**0.42**	**0.56**	0.38
$r_{(C3)}$	**0.26**	**0.04**	0.16
stl004[b]			
$p_{(D1)}$	**0,52**	**0,20**	**0.14**
$q_{(D2)}$	**0,42**	**0,80**	**0.82**
$r_{(D3)}$	**0.06**	**0.00**	**0.04**
stl019[b]			
$p_{(E1)}$	**0.04**	**0.02**	**0.02**
$q_{(E2)}$	**0.70**	**0.86**	**0.86**
$r_{(E3)}$	**0.26**	**0.12**	**0.12**
Stm1045			
$p_{(F1)}$	0,25	0,25	0,26
$q_{(F2)}$	0,31	0,28	0,30
$r_{(F3)}$	0,28	0,29	0,26
$s_{(F4)}$	0,16	0,18	0,18

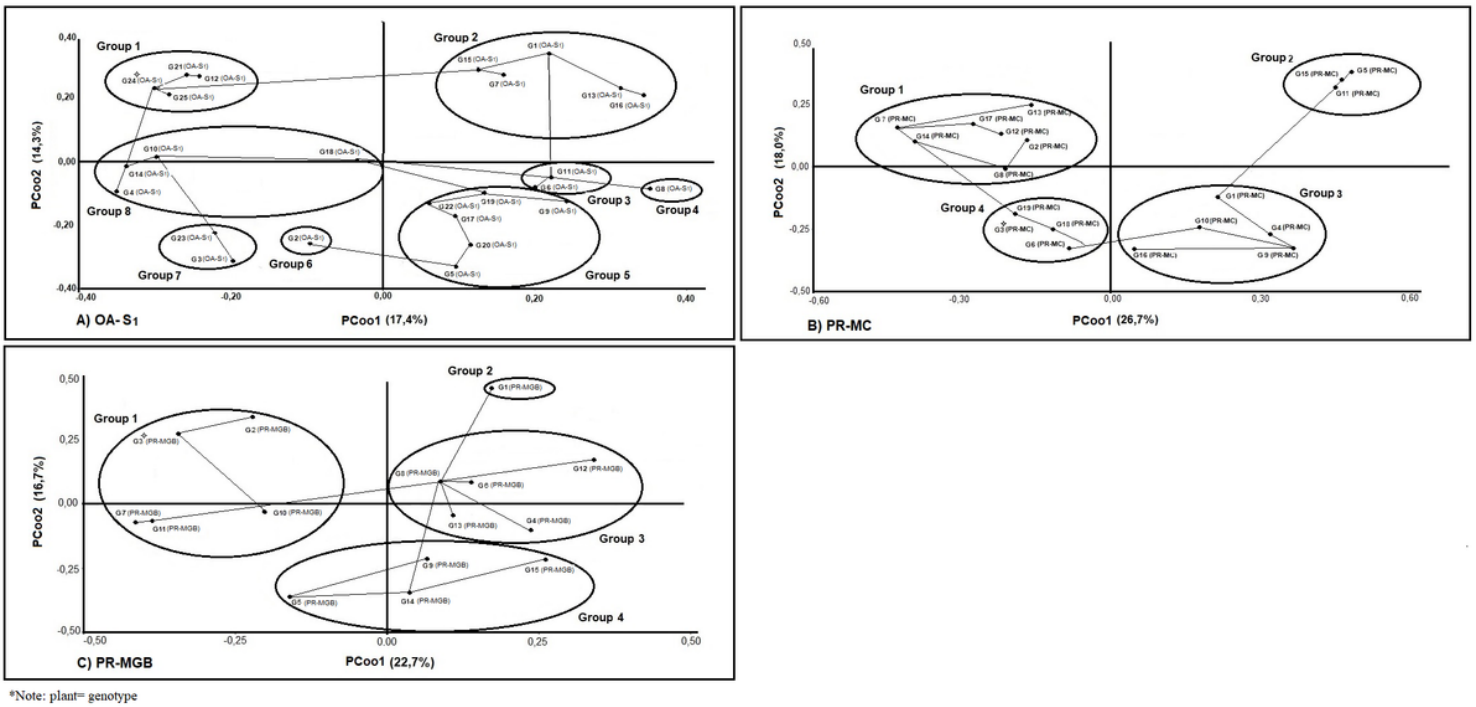
Figures



*Note: plant= genotype.

Figure 1

Two-dimensional graphic for the first two Principal Components (PC1 and PC2) of a Principal Component Analysis and Minimum Spanning Tree (MST) based on 11 quantitative (vegetative and reproductive) morphological characters evaluated in A) 25 plants of the original accession Cl 898 of *Solanum chacoense* Bitter (OA-S1), B) 19 plants of the population regenerated by the method proposed by Camadro (2012) (PR-MC), and 15 plants of the population regenerated by the method commonly used in germplasm banks (FAO 2013) (PR-MGB)



*Note: plant= genotype

Figure 2

Two-dimensional graphic for the first two Principal Coordinates (PCoo1 and PCoo2) of a Principal CoordinateAnalysis (PCooA) and Minimum Spanning Tree (MST) based on 10 qualitative (vegetative and reproductive) morphological characters evaluated in: A) 25 plants of the original accession Cl 898 of *Solanum chacoense* Bitter (OA-S1), B) 19 plants of the population regenerated by the method proposed by Camadro (2012) (PR-MC), C) 15 plants of the population regenerated by the method commonly used in germplasm banks (FAO 2013) (PR-MGB)

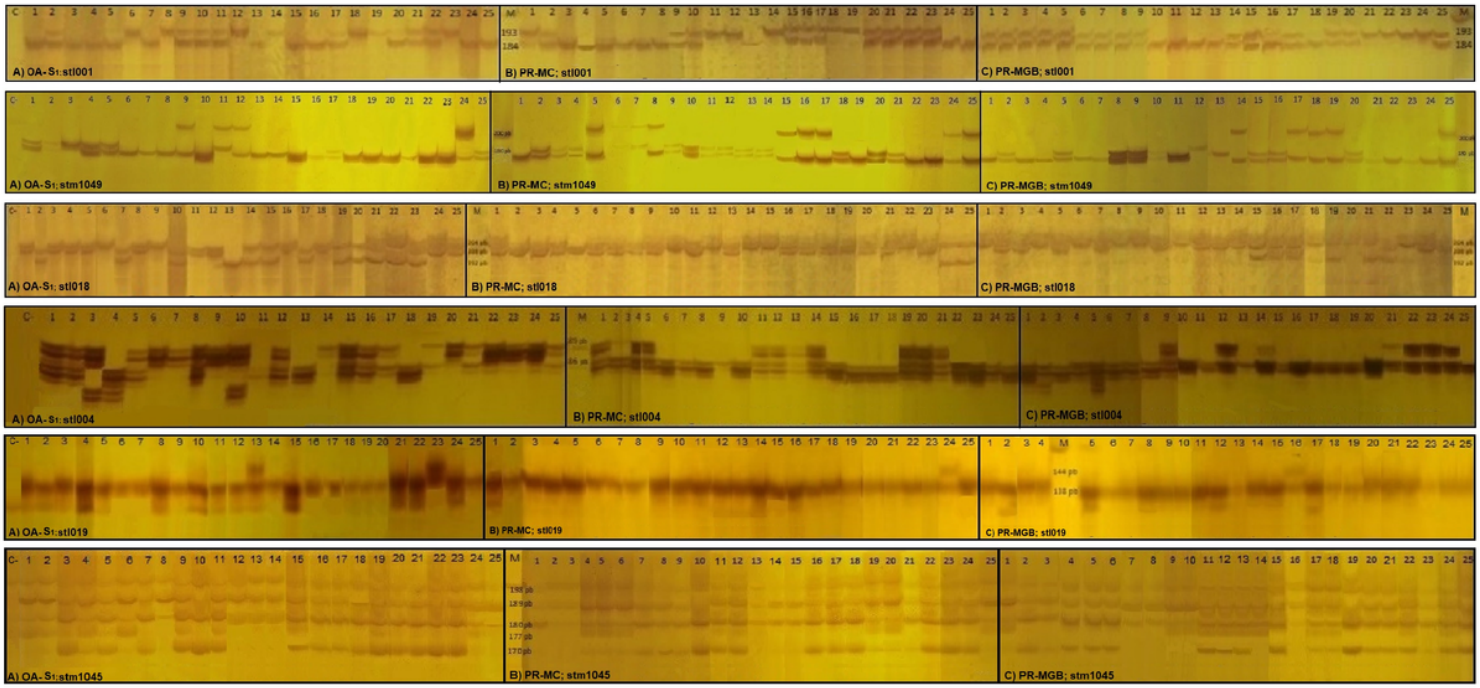


Figure 3

Electrophoretic patterns for SSR stl001, stm1049, stl018, stl004, stl019, and stm1045 in 25 plants of A) original accession Cl 898 of *Solanum chacoense* Bitter (OA-S_1), B) population regenerated by the method proposed by Camadro (2012) (PR-MC), C) population regenerated by the method commonly used in germplasm banks (FAO 2013) (PR-MGB)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf
- ESM2.pdf