horticulturae

MDPI

Article

Assessment of Genetic Diversity and Relatedness in an Andean Potato Collection from Argentina by High-Density Genotyping

Sofía Sucar [1,†], Martín Federico Carboni [2,3,†], María Florencia Rey Burusco [1,4], Martín Alfredo Castellote [1], Gabriela Alejandra Massa [1,3,4], Marcelo Nicolás Monte [1] and Sergio Enrique Feingold [1,*]

[1] Laboratorio de Agrobiotecnología, IPADS, CONICET-INTA, Balcarce B7620, Argentina; sucar.sofia@inta.gob.ar (S.S.); rey.mariaflorencia@inta.gob.ar (M.F.R.B.); castellote.martin@inta.gob.ar (M.A.C.); massa.gabriela@inta.gob.ar (G.A.M.); monte.marcelo@inta.gob.ar (M.N.M.)
[2] Museo Argentino de Ciencias Naturales, Buenos Aires C1405, Argentina; mfcarboni@macn.gov.ar
[3] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires C1425, Argentina
[4] Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, Balcarce B7620, Argentina
* Correspondence: feingold.sergio@inta.gob.ar; Tel.: +54-2266-439100
† These authors contributed equally to this work.

Abstract: Native potatoes are the most diverse among cultivated potato species and thus constitute a valuable source for identifying genes for potato improvement. Nevertheless, high-density mapping, needed to reveal allelic diversity, has not been performed for native Argentinian potatoes. We present a study of the genetic variability and population structure of 96 Andigena potatoes from Northwestern Argentina performed using a subset of 5035 SNPs with no missing data and full reproducibility. These high-density markers are distributed across the genome and present a good coverage of genomic regions. A Bayesian approach revealed the presence of: (I) a major group comprised of most of the Andean accessions; (II) a smaller group containing the out-group cv. Spunta and the sequenced genotype DM; and (III) a third group containing colored flesh potatoes. This grouping was also consistent when maximum likelihood trees were constructed and further confirmed by a principal coordinate analysis. A group of 19 accessions stored as Andean varieties clustered consistently with group Tuberosum accessions. This was in agreement with previous studies and we hypothesize that they may be reintroductions of European-bred long day-adapted potatoes. The present study constitutes a valuable source for allele mining of genes of interest and thus provides a tool for association mapping studies.

Keywords: Andean potato; germplasm; DArTseq markers; population structure; genetic diversity

Citation: Sucar, S.; Carboni, M.F.; Rey Burusco, M.F.; Castellote, M.A.; Massa, G.A.; Monte, M.N.; Feingold, S.E. Assessment of Genetic Diversity and Relatedness in an Andean Potato Collection from Argentina by High-Density Genotyping. *Horticulturae* **2022**, *8*, 54. https://doi.org/10.3390/horticulturae8010054

Academic Editors: Xin Wang and Jiantao Zhao

Received: 20 October 2021
Accepted: 29 November 2021
Published: 7 January 2022

1. Introduction

Potato (*Solanum tuberosum* L.) is the most important non-cereal food crop and the third most important food crop for direct human consumption [1]. Its haploid genome is composed of 12 chromosomes with an estimated size of 840 Mb. The reference genome DM1-3516-R44 was completely sequenced in 2011 [2] and 96% of the identified genes have been localized in a physical map [3].

Molecular and phytogeographical data indicate the high south-central Andes as the place of potato domestication [4,5]. More precisely, in the basin of Lake Titicaca, starch grains dating to the Late Archaic period (~3400 BC) were discovered [6]. Dodds classified *Solanum tuberosum* as a unique species composed of five cultivar groups: Stenotonum, Phureja, Chaucha, Tuberosum, and Andigena [7,8]. The Tuberosum group includes lowland tetraploid potatoes introduced in Europe 400 years ago and spread from there to the rest of the world as "modern potatoes" [9]. They possess a narrow genetic base, probably originating from genetic drift on their way to Europe [5,10]. Contrastingly, the Andigena group is composed of native Andean potatoes that were domesticated 7000–10,000 years

ago in the Andean uplands of South America. These are by far the most diverse among cultivated potato species, not only morphologically, but also as described by molecular markers [11]. Hence, they constitute a highly valuable source for identifying genes and allelic variants of agronomic, nutritional, and industrial quality interest [12,13]. Previous studies using SSRs have evidenced their great genetic variability [14–17]. Colman et al. [18] have reported a wide phenotypic variation in reducing sugar content and chip quality traits of Andean potatoes from Northwestern Argentina (NWA) [18].

Several Andigena landraces are still cultivated in NW Argentina, specifically in the provinces of Jujuy, Salta, and Catamarca. For more than 40 years, the INTA-Balcarce Germplasm Bank (BAL) has been collecting and preserving Andean potato materials. The characterization of genetic diversity and population structure is of great value for Germplasm bank management, and is crucial to ensure germplasm protection [19] under international agreements for the fair and equitable sharing of the benefits arising from the use of plant genetic resources [20]. Moreover, this group was traditionally selected by local farmers for several centuries and constitutes a source of allelic variants of relevant genes that can be used for improving commercial varieties [21,22]. In this sense, molecular markers constitute a very useful tool for characterizing genetic variability.

Simple sequence repeats (SSRs), also known as microsatellite markers, are short tandem repeats of di-, tri-, or tetra-nucleotides that are highly polymorphic. They have been used for varietal identification, germplasm characterization, genetic linkage maps, QTL localization, population structure, and phylogeny in potato [23–29]. They are a reliable method for genetic analysis [4] and require only a small quantity of DNA to be analyzed [30–32].

By employing a set of functional SSR markers, we have determined the genetic diversity and population structure in a collection of 88 Andigena potatoes from NWA [17], revealing the presence of two distinctive groups and the existence of group-specific alleles. Functional markers are advantageous over intergenic DNA markers owing to their linkage to characterized genes [33].

Diversity Array Technology by Sequencing (DArTseqTM) [34] is a genotyping-by-sequencing platform based on genome complexity reduction that provides single nucleotide polymorphism (SNP) marker information for high-throughput and cost-effective genotyping without the need for prior sequence data [35,36]. It has been used for diversity studies, population structure determination, and genome-wide association analyses in different species, such as rice [37], soybean [38], lesquerella [39], canola [40], wheat [41], watermelon [42], genus *Secale* [43], chickpea [44], and safflower [45]. It was also used to build one of the first maps from the genus *Solanum* [46]. DArT linkage maps have been constructed for *S. bulbocastanum* Dunal [47], *S. commersonii* Dunal, and the Mexican *Solanum pinnatisectum* Dunal [48]. In this last case, it was also useful for the mapping of a novel major late blight resistance locus.

When performing a detailed mapping analysis, such as association mapping, the use of DArTseq marks is more appropriate. This has to do with the fact that the distance between SSR markers constitutes a great portion of DNA that can potentially contain a high number of associated candidate genes [49,50]. On the other hand, the greater abundance of SNP markers and the possibility of some of them being within genetic regions makes a DArTseq screening more relevant and capable of identifying functional alleles of interest.

To our knowledge, the present study is the first genetic diversity characterization performed for Andean potatoes that makes use of DArTseq markers. The aim of this work was to serve as a first attempt to describe the genetic diversity and population structure of 96 Andigena potatoes from NWA performed with DArTseq SNPs, which are high-density markers distributed across the genome and concentrated in genomic regions.

2. Materials and Methods

2.1. Plant Material

In total, 114 accessions were used: a selection of 96 Andean potato accessions (*S. tuberosum* group Andigena) provided by the Germplasm Bank of the EEA INTA-Balcarce (BAL), 3 by Jujuy National University, 7 accessions of Imilla Negra provided by Cauqueva Cooperative, 3 accessions collected over several trials in Jujuy (*guacha* potato); 3 commercial varieties (*S. tuberosum* Group Tuberosum; cv. Spunta, cv. Pampeana INTA, and cv. Bintje) provided by Potato Group of the EEA INTA-Balcarce, and the doubled monoploid line used for the reference genome DM1-3 516 R44 (*S. tuberosum* group Phureja) (2×) provided by Potato Genome Sequencing Consortium (PGSC) [2], hereafter referred to as DM (Table 1).

Table 1. Collection used for the present study.

Group (Atencio, 2011)	Accession Code	Study Code	Landrace	Location (Province, Department, Locality)
1	CCS 1327	28	Bayista	Jujuy, Cochinoca, Rachaite
1	CL 621	23	Chorcoyeña	Salta, Santa Victoria, Nazareno
1	CL 516	44	Chorcoyeña	Salta, Santa Victoria, Chorro
1	CCS 1205	21	Churqueña	Jujuy, Humahuaca, Varas
1	CCS 1378	90	Churqueña negra	Jujuy, Tumbaya, Patacal
1	CCS 1224	60	Collareja	Jujuy, Humahuaca, Coctaca
1	CL 634	67	Collareja	Salta, Santa Victoria, Arpero
1	CL 636	68	Collareja	Salta, Santa Victoria, Abra Colorada
1	CS 1432	3	Collareja or cuarentona overa	Jujuy, General Belgrano, Cuevas
1	CCS1381	29	Runa	Jujuy, Tumbaya, Patacal
1	CL 650	45	Runa	Salta, Santa Victoria, Poscaya
1	CL 576	46	Runa	Salta, Santa Victoria, Lizoite
1	CL 641	49	Runa	Salta, Santa Victoria, Poscaya
1	CL 708	56	Runa	Salta, Iruya, Colanzulí
1	CL 489	63	Runa	Salta, Santa Victoria, Rodeopampa
1	CL 739	70	Runa	Jujuy, Humahuaca, Chaupi Rodero
1	CL 750	71	Runa	Jujuy, Humahuaca, Chaupi Rodero
1	CCS 1218	84	Runa	Jujuy, Humahuaca, Ocumazo
1	CCS 1340	116	Yaguana	Jujuy, Susques, Sala
1	CL 641	144	Runa	Salta, Santa Victoria, Poscaya
2	LC 348	62	Imilla Negra	Jujuy, Humahuaca, Huachichocana
2	CL 631	40	Allo	Salta, Iruya, Campo Carreras
2	CCS 1227	109	Azul or Sallama	Jujuy, Humahuaca, Coctaca
2	CCS 1196	80	Azul overa	Jujuy, Humahuaca, Palca de Aparzo
2	CL 815	77	Boliviana	Jujuy, Valle Grande, Santa Ana
2	CCS 1201	14	Condorilla	Jujuy, Humahuaca, Varas
2	CCS 1384	41	Corbatilla	Jujuy, Tumbaya, Patacal
2	CCS 1330	6	Moradita	Jujuy, Cochinoca, Rachaite
2	CCS 1307	8	Moradita	Jujuy, Santa Catalina, Cabreria
2	CCS 1374	39	Moradita	Jujuy, Cochinoca, Agua Castilla
2	CCS 1172	4	Moradita redonda	Jujuy, Tilcara, Casa Colorada
2	CL 783	98	Navecilla	Jujuy, Valle Grande, Santa Ana
2	CL 820	57	Negra Redonda	Jujuy, Valle Grande, Santa Ana
2	CCS 1305	47	Ojosa	Jujuy, Santa Catalina, Casir
2	CCS 1257	86	Ojosa	Jujuy, Rinconada, Rinconada
2	CCS 1366	27	Overa	Jujuy, Tumbaya, El Moreno
2	CL 748	35	Overa	Jujuy, Humahuaca, Chaupi Rodero
2	CL 790	52	Overa	Jujuy, Valle Grande, Santa Ana
2	CL 832	79	Abajeña overa	Jujuy, Valle Grande, Santa Ana
2	CL 793	74	Sallama	Jujuy, Valle Grande, Santa Ana
2	CL 804	139	Sallama	Jujuy, Valle Grande, Santa Ana
2	CL 769	32	Sallama Grande	Jujuy, Valle Grande, Santa Ana
2	CCS 1284	26	Sani	Jujuy, Yavi

Table 1. *Cont.*

Group (Atencio, 2011)	Accession Code	Study Code	Landrace	Location (Province, Department, Locality)
2	CCS 1303	37	Yuruma	Jujuy, Santa Catalina, Casira
2	CCS 1385	42	Moradita	Jujuy, Tumbaya, Patacal
3	CL 835	51	Airampía	Jujuy, Valle Grande, Santa Ana
3	CL 836	54	Airampía	Jujuy, Valle Grande, Santa Ana
3	CL 528	64	Colorada	Salta, Santa Victoria, Chorro
3	CCS1221	102	Colorada	Jujuy, Humahuaca, Coctaca
3	CL 508	123	Colorada	Salta, Santa Victoria, Chorro
3	CCS 1349	12	Coloradita	Jujuy, Tumbaya, El Angosto
3	CCS 1184	108	Colorana or Señorita	Jujuy, Humahuaca, Aparzo
3	CL 821	78	Cuarentilla Toscra	Jujuy, Valle Grande, Santa Ana
3	CS 1430	16	Cuarentona	Jujuy, General Belgrano, Cuevas
3	CL 728	31	Cuarentona	Salta, Iruya, Colanzulí
3	CCS 1166	9	Cuarentona colorada	Jujuy, Tilcara, Casa Colorada
3	CS 1414	94	Cuarentona morada	Jujuy, General Belgrano, Papachacra
3	CS 1425	122	Cuarentona oquecha	Jujuy, General Belgrano, Papachacra
3	CCS 1353	117	Cuarentona Redonda	Jujuy, Tumbaya, El Angosto
3	CS 1416	120	Cuella	Jujuy, General Belgrano, Papachacra
3	CCS 1288	15	Desiree	Jujuy, Santa Catalina, Cieneguillas
3	CL 712	69	Huareña	Salta, Iruya, Colanzulí
3	CCS 1170	53	Ojos colorados	Jujuy, Tilcara, Casa Colorada
3	CS 1402	91	Ojos colorados	Jujuy, Tumbaya, Carcel
3	CCS 1383	34	Pera or señorita	Jujuy, Tumbaya, Patacal
3	CCS 1321	89	Rosada	Jujuy, Cochinoca, Agua Caliente
3	CL 658	17	Santa María	Jujuy, Yavi, Yavi
3	LC 335	97	Tonca	Jujuy, Humahuaca, Patacal
3	CL 482	33	Rosada	Salta, Santa Victoria, Rodeopampa
3	CCS 1255	7	Desiree	Jujuy, Rinconada, Rinconada
3	CCS 1323	36	Colorada	Jujuy, Cochinoca, Agua Caliente
4	CL 849	99	Balcacha	Salta, Rosario de Lerma, El Gólgota
4	CCS 1350	1	Blanca	Jujuy, Tumbaya, El Angosto
4	CS 1419	10	Blanca	Jujuy, General Belgrano, Papachacra
4	CCS 1310	115	Blanca alargada	Jujuy, Santa Catalina, Cabreria
4	CCS 1309	58	Blanca redonda	Jujuy, Santa Catalina, Cabreria
4	CCS 1251	5	Chacarera	Jujuy, Cochinoca, Cochinoca
4	CCS 1371	20	Chacarera	Jujuy, Cochinoca, Quebraleña
4	CS 1418	2	Chaqueña	Jujuy, General Belgrano, Papachacra
4	CS 1408	92	Chaqueña overa	Jujuy, General Belgrano, Papachacra
4	CCS 1300	88	Holandesa	Jujuy, Santa Catalina, Casira
4	CCS 1209	83	Luqui	Jujuy, Humahuaca, Chorcan
4	CCS 1299	100	Malgacha	Jujuy, Santa Catalina, Casira
4	CCS 1200	81	Papa oca	Jujuy, Humahuaca, Varas
4	CCS 1206	82	Papa oca	Jujuy, Humahuaca, Varas
4	CL 548	48	Papa palta	Salta, Santa Victoria, Trigohuaico
4	CS 1413	93	Papa vallista	Jujuy, General Belgrano, Papachacra
4	CCS 1185	13	Tuni	Jujuy, Humahuaca, Aparzo
4	CCS 1199	11	Tuni blanca	Jujuy, Humahuaca, Palca de Aparzo
4	CCS 1247	22	Tuni blanca	Jujuy, Cochinoca, Ojo de Agua
4	CCS 1375	101	Tuni Blanca	Jujuy, Tumbaya, Tumbaya
4	CCS 1393	119	Tuni morada	Jujuy, Tumbaya, Cieneguillas
4	CL 782	73	Tuni rosilla	Jujuy, Valle Grande, Santa Ana
4	CCS 1271	25	Blanca	Jujuy, Santa Catalina, Morco Esquina
4	Cl 752B	72	Blanca	Jujuy, Valle Grande, Santa Ana
4	CL 814A	75	Holandesa colorada	Jujuy, Valle Grande, Santa Ana
	Berta	104	Azul	Provided by Jujuy National University
	Berta	105	Santa María	Provided by Jujuy National University

Table 1. *Cont.*

Group (Atencio, 2011)	Accession Code	Study Code	Landrace	Location (Province, Department, Locality)
	Berta	106	Santa María pulpa blanca	Provided by Jujuy National University
		70B		Guacha potato
		89A		Guacha potato
		89B		Guacha potato
		DM		Potato reference genome
		IN1	Imilla Negra	Provided by Cauqueva
		IN2	Imilla Negra	Provided by Cauqueva
		IN3	Imilla Negra	Provided by Cauqueva
		IN4	Imilla Negra	Provided by Cauqueva
		IN5	Imilla Negra	Provided by Cauqueva
		IN6	Imilla Negra	Provided by Cauqueva
		IN7	Imilla Negra	Provided by Cauqueva
		BINTJE	Bintje	Commercial variety
		PAMPEANA-INTA	Pampeana-Inta	Commercial variety
		SPUNTA	Spunta	Commercial variety
		SPUNTA2	Spunta	Commercial variety

Group assignment was made by Atencio (2011) [51] in accordance with tuber morphologic traits (tuber shape, skin color, eyes description and distribution, and flesh color). Accession code was given by the BAL. The Study Code column refers to denominations used in this study and was used for all the figures in the manuscript. The landrace column corresponds to the traditional name given by the local farmer who provided the material. Location of collection is described in the last column, where province, department, and locality are detailed.

The *Solanum tuberosum* Andigena group in vitro collection used for this study comprises a selection of accessions collected and propagated in vitro by the BAL [51,52].

Selection was based on tuber morphologic traits, such as shape, flesh and skin color, and distribution and depth of eyes, in order to encompass variability [51]. The BAL material was collected between 1976 and 2001 from farmers in the provinces of Salta and Jujuy [53–57] and stored in vitro under a 16/8 photoperiod, at a controlled temperature of 21 °C.

Tubers were obtained from plants grown in greenhouses and chambers in counterseason and planted in fields in Jujuy, Argentina (23°45′ S 65°30′ O, 3600 masl) from November to March/April, and were used to obtain tissue for DNA extraction.

2.2. DNA Extraction and Preparation

DNA was extracted from young leaf tissue according to [58], as described elsewhere [17]. DNA was examined in 1% agarose gels run in 1× TBE (89 mM Tris–borate, 20 mM ethylenediaminetetraacetic acid) and stained with GelRed (GenBiotech, Argentina) for quality and concentration. DNA quantification was performed in a SmartSpecTM 3000 Spectrophotometer (BIORAD, Hercules, CA, USA).

In order to ensure DNAse-free samples, the samples were incubated at 37 °C with 10× Restriction Enzyme buffer (Promega, Madison, WI, USA) for 2 h. Samples were also tested for the presence of DNAses by Diversity Arrays Pty Ltd. Canberra, Australia, before the DArTseq assay was conducted.

2.3. Genotyping

Genotyping was performed by Diversity Arrays Pty Ltd. (Canberra, Australia), using a combined technology called DarTseq™. This technique makes use of the traditional Diversity Arrays Technology (DArT) combined with next-generation sequencing. It involves a two-step procedure that includes genome complexity reduction by selection of fractions of the genome that correspond predominantly to active genes and the removal of large repetitive sequences. This is achieved by digestion with a combination of restriction

enzymes. Samples were digested using a rare cutter, Pst I, and subsequently incubated with a frequent-cutter restriction enzyme, such as Taq I. Adaptors were ligated afterwards to the ends of the Pst I fragments and amplified by PCR using specific primers complementary to the adaptor sequences. This step was followed by Illumina short-read sequencing.

2.4. Diversity Analysis

Percentage heterozygosity was calculated for the potato panel using a subset of 5035 SNPs with no missing data and full reproducibility. It was calculated for each accession as:

$$(\text{number of heterozygous loci}/\text{number of total loci}) \times 100.$$

A locus was considered homozygous when the same SNP was detected for both copies and heterozygous if one copy contained the reference allele while the other copy contained a SNP.

The amount of variation among clusters was assessed by partitioning genetic diversity using analysis of molecular variance (AMOVA).

2.5. Population Structure Analysis

In order to identify the number of populations (K) capturing the major structure in the data, STRUCTURE v.2.3.4 software was used [59], choosing an admixture model, with independent allele frequencies, a burn-in period of 50,000 MCMC iterations, and 100,000 run length. Ten independent iterations were performed for each simulated value of K ranging from 1 to 10. The most likely number of K was then resolved by the DeltaK method [60] with the Structure Harvester software [61]. STRUCTURE analysis output permutations were performed employing CLUMPP software [62], using independent runs to obtain a consensus matrix.

To further assess the relationship between the accessions, the SNPhylo pipeline [63] was employed to construct a maximum likelihood tree, with the following steps: (1) filtering of the SNP datasets using minor allele frequency of 10% and 10% missing data; (2) remotion of redundant SNPs based on linkage disequilibrium information (cutoff threshold 1) using SNPRelate [64]; (3) construction of multiple sequence alignment of the SNP dataset using MUSCLE [65]; (4) construction of the maximum likelihood tree using DNAML from Phylip [66]; (5) performing 1000 bootstraps using Phangorn [67]. Before running the pipeline, the script file was edited to set the transition/transversion ratio equal to the previously calculated value of 1.5. The phylogenetic tree was drawn and visualized with the online software Interactive Tree Of Life (iTOL) v3 [68]. The bar graph of the population structure results was added in the same way as a multi-value bar chart dataset in iTOL.

A principal coordinate analysis (PCoA) was carried out to explore group conformation within the collection using the GenAlEx software [69] and accessions were plotted based on the first two principal coordinates. To this end, 5035 SNPs were selected under the two main criteria of (i) high reproducibility and (ii) no missing data for the collection studied.

3. Results

3.1. Genetic Diversity Analysis

In total, 56,163 SNPs were generated (Figure 1). The depth of reads ranged from 5× to 1600× with a mean value of 18×. They were mapped altogether with genomic regions (Figure 1) using an in-house developed tool, Agrobiotechtools (Agrobiotechnology Laboratory, IPADS, CONICET-INTA) to visualize their distribution. This tool was particularly useful for plotting density in a 1 Mbp bin and exact positions of DArTseq markers throughout the 12 potato chromosomes along with their colocalization with genomic regions.

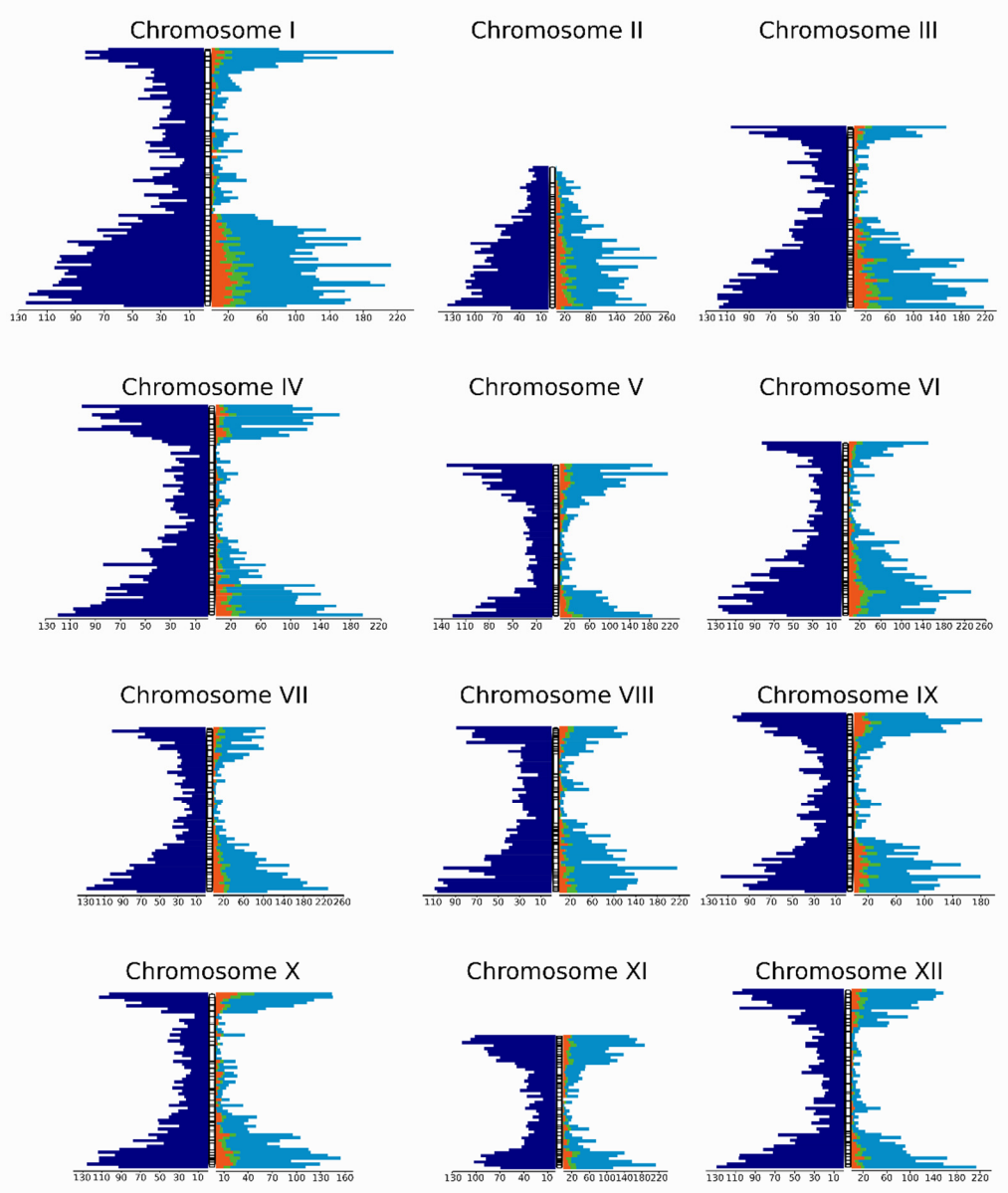

Figure 1. Density of genes and SNP markers per chromosome. Each bar indicates density in a 1 Mbp bin. Blue bars show gene density; light blue, the total SNPs mapped; green, SNPs without missing data; and orange, SNPs without missing data and with best quality. Number of genes or SNPs per Mbp are indicated at the bottom of each sub-figure.

The majority of SNPs detected were transitions (A/G or C/T). In total, 33,806 were found, whereas 22,357 transversions were found, giving a transition/transversion ratio of about 1.5.

It was possible to map the 45,159 SNPs found in the PGSC Version 4.03 Pseudomolecules of the reference potato (*S. tuberosum* group Phureja DM1-3 516 R44) [3]. On average, 3670 SNPs were detected per chromosome, ranging from 5643 SNPs in chromosome 1 to 2699 in chromosome 10. Only 1123 SNPs were mapped to the unanchored chromosome 00. This chromosome includes superscaffolds that remain unanchored in the published reference genome map [3]. The distribution of SNPs was coincident with regions of high density of coding sequences of the PGSC_DM_v4.03_gene database (Figure 1).

Using a subset of 5035 SNPs with no missing data and full reproducibility, we found that the average percentage heterozygosity observed within the potato panel was 32%. The

minimum heterozygosity was found, as expected, in the doubled monoploid accession DM, whereas the maximum heterozygosity was 43.8%, as found in accession 31 (Figure 2).

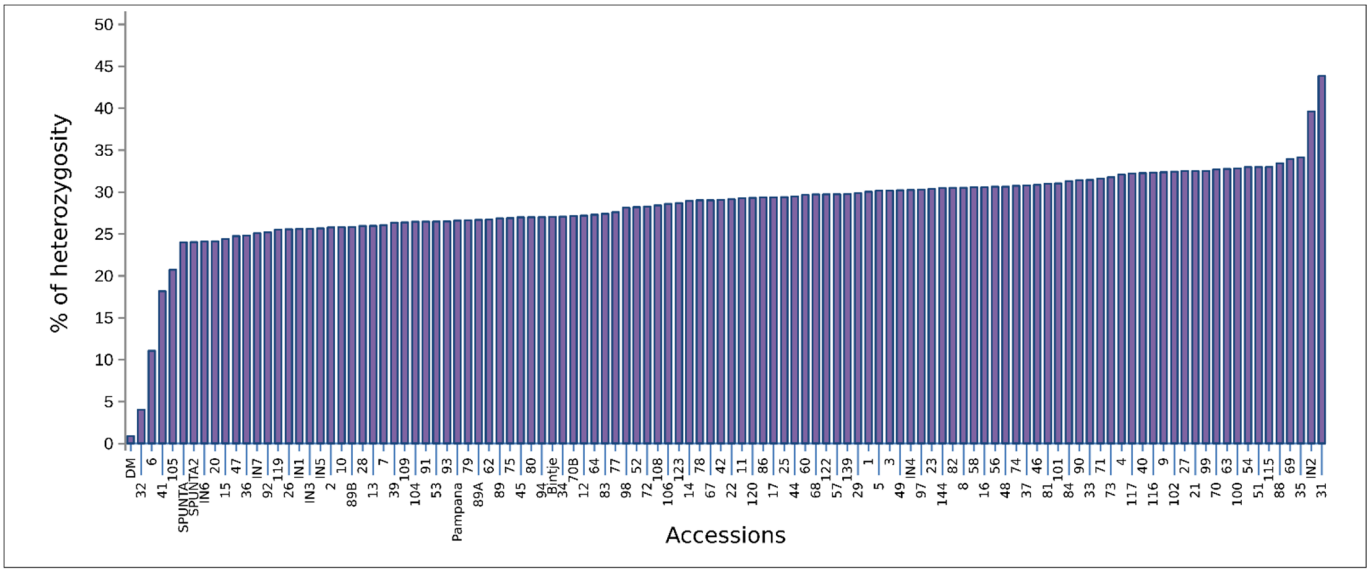

Figure 2. Percentage of heterozygous loci for each of the 114 accessions was determined with 5035 SNPs and a diploid genotyping model.

3.2. Population Structure Analysis

To explore the population structure, a Bayesian approach was conducted using the program STRUCTURE [59] version 2.3.4 and the aforementioned subset of 5035 SNPs scattered across the genome. Using the method of Evanno et al. [60], it was determined that the 114 accessions were partitioned into three clusters (K = 3) (Figure 3, bottom left box). The green group includes cv. Spunta, Bintje, and Pampeana-INTA (*S. tuberosum* Tuberosum group), DM (*S. tuberosum* Phureja group), and 19 accessions from the BAL. The smaller group, shown in cyan, comprises 15 accessions and includes colored flesh potatoes, such as Santa María and Azul. The major group is shown in red and contains accessions with different tuber shapes belonging to the *S. tuberosum* Andigena group.

3.3. Clustering Analysis

To visualize detailed information about the genetic diversity among different groups of potato accessions, two maximum likelihood trees were constructed. The grouping of accessions in the first tree was in general agreement with the Bayesian population structure analysis with only three mislocated accessions (Figure 3). When the same analysis was performed without the ambiguous accessions, the clustering obtained was completely coincident with structure results (Figure S1).

A principal coordinate analysis (PCoA) was conducted to complement the clustering analysis and further visualize the pattern of genetic relationships. The first two principal coordinates collectively explained 40.45% of the total genetic variance (Coord.1 = 22.7% and Coord.2 = 17.8%). The PCoA results showed clear separation into distinct groups, which agreed with the STRUCTURE results (Figure 4).

The analysis of molecular variance (AMOVA) was used for hierarchical partitioning of the genetic variation among the three clusters, as revealed by the STRUCTURE analysis. There was a higher proportion of genetic variation within clusters (68%) than among clusters (32%) in the potato collection.

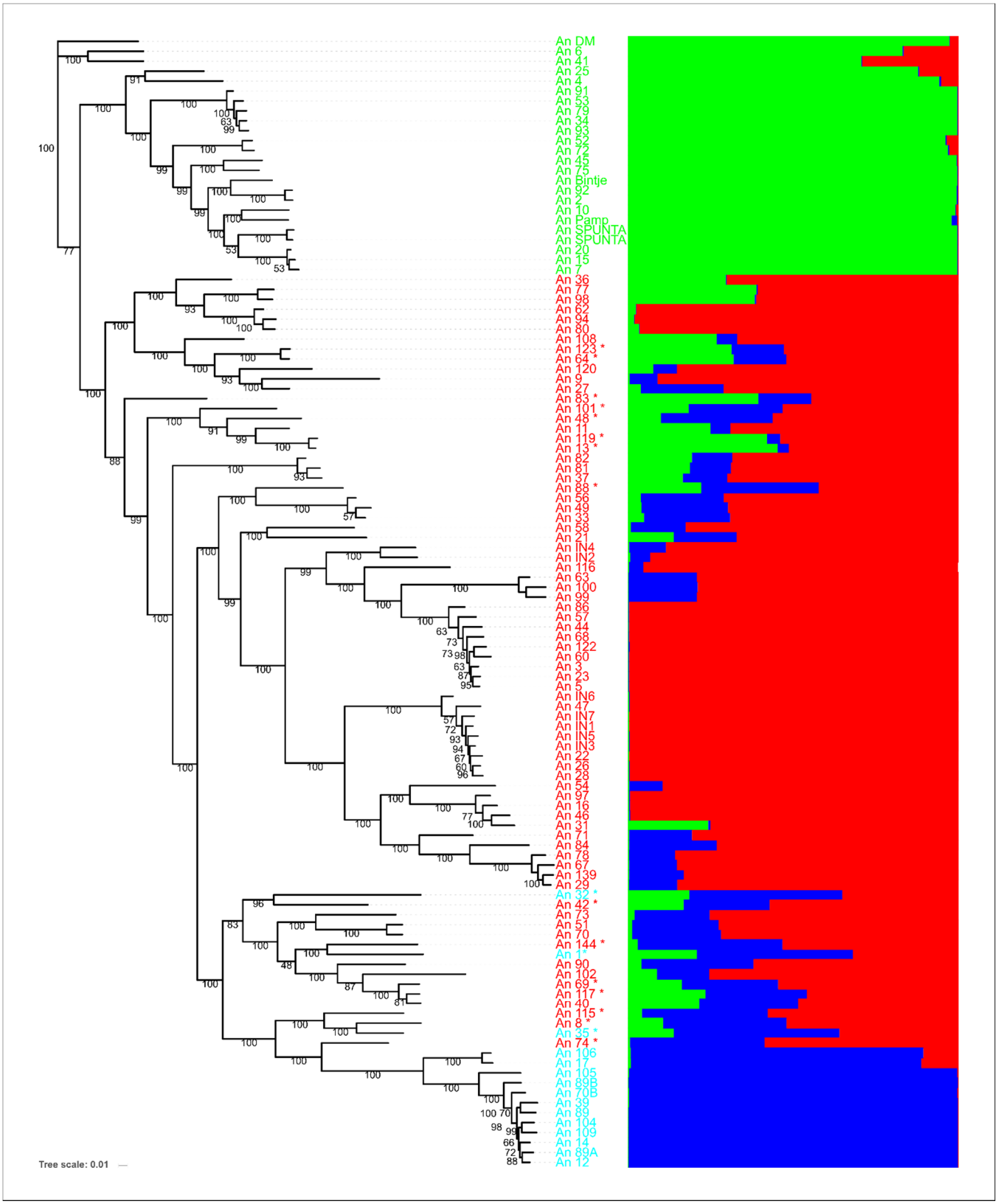

Figure 3. STRUCTURE plot of the 114 potato accessions based on 5035 SNP markers. Each column corresponds to one accession and is partitioned into colored segments that are in proportion to the estimated membership of the three subpopulations. Maximum likelihood trees are shown on the left side. Accessions were labeled by color according to STRUCTURE results. Accessions with a membership proportion lower than 0.6 were labeled with an asterisk.

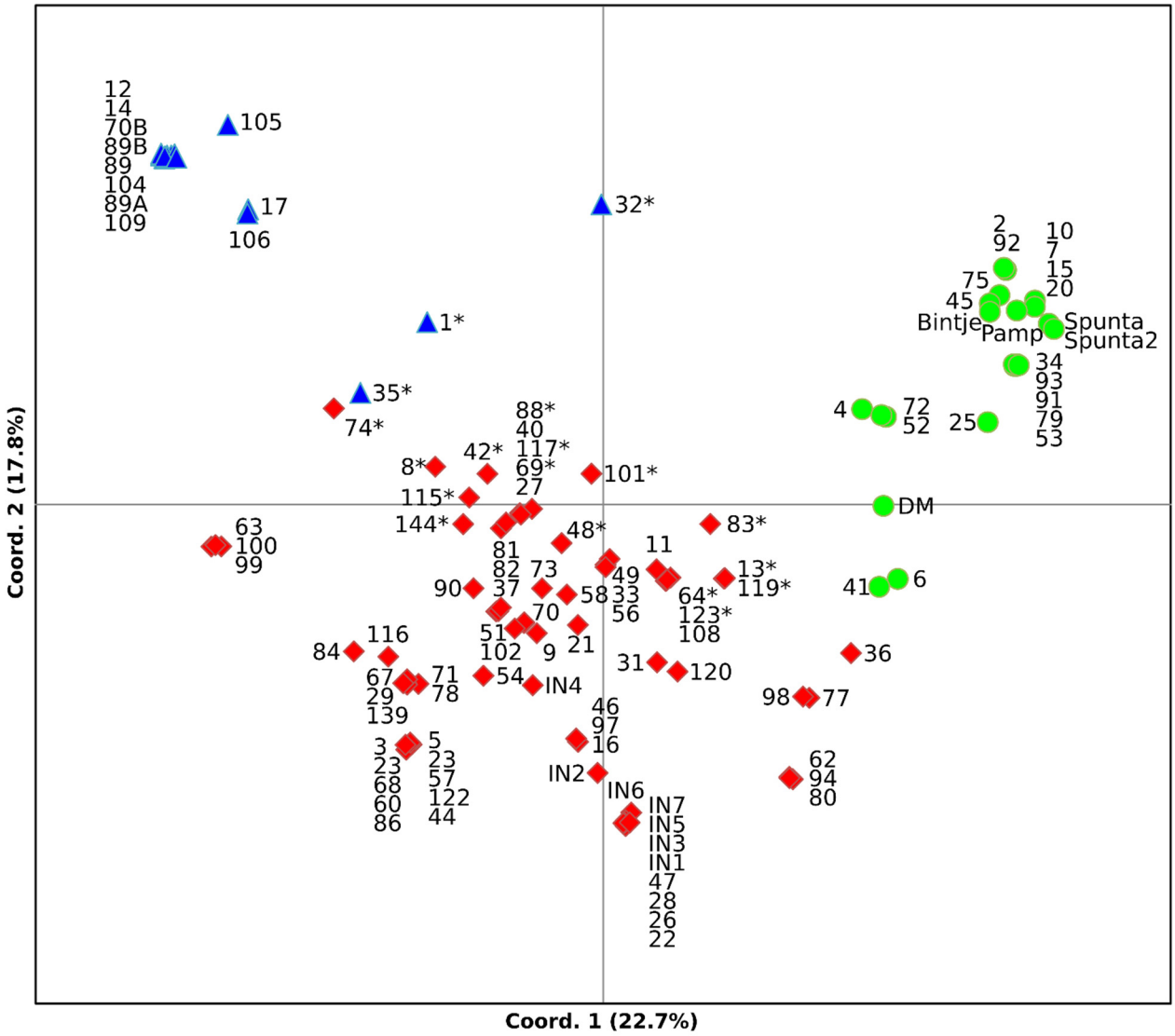

Figure 4. Principal coordinate analysis (PCoA) of the 114 potato accessions in the panel based on 5035 DArTseq markers. Color coding corresponds to colors used for the STRUCTURE grouping assignment. Accessions with a membership proportion lower than 0.6 are labeled with an asterisk. Numbers in parentheses in the axes depict the percentage of genetic variance.

4. Discussion and Conclusions

The assessment of genetic relatedness, performed by three different analyses, revealed that the collection was composed of three main subpopulations. The fact that this grouping was supported by a Bayesian approach, a genetic distance-based method, and a principal coordinate analysis, provides strong evidence for the resulting population structure.

In addition, a very high marker density was used. A high-density map results is particularly useful when performing a detailed mapping analysis, such as association mapping. Since future association mapping studies will be performed for this collection, DArTseq markers were chosen, given the fact that they present high gene coverage when compared to other types of markers used for genotyping studies, such as DNA microarrays for example [70]. This advantage over other methods is related to the usage of a combination of restriction enzymes that separates low-copy sequences from the highly repetitive fraction of the genome. This was clearly visualized when plotting genomic regions and DArTseq markers together by chromosome, using *Agrobiotechtools*, an in-house developed

tool (Agrobiotechnology Laboratory, IPADS, CONICET-INTA) (Figure 1). It is worth noticing that, even after applying filters to choose markers with high reproducibility and no missing data, the percentage of genomic regions covered was still very high. The majority of SNPs detected were transitions (A/G or C/T) and the ratio obtained was consistent with other studies performed in various species, such as maize [71], oil palm [72], rubber tree [73], or Iranian cannabis germplasm [74]. Previous reports on SNP discovery programs have also detected a high frequency of transitions [75–77] showing the high frequency of the C/T mutation that occurs after methylation [71]. Population structure analysis revealed the existence of three defined groups. One of them was composed by modern potatoes and accessions that possibly resulted from reintroductions of commercial potatoes, and were mislabeled as Andean.

Local farmers tend to prefer modern potatoes because—unlike native genotypes—they lack photoperiod sensitivity; thus, in lower latitudes they present a shorter crop cycle that represent an advantage for soil use efficiency where multiple crops are grown within a year.

The mislabeling of accessions in germplasm banks constitutes a problem; the curation of entries becomes necessary. In this study, it is possible that these mislabeled Andigena entries have been erroneously named as such during collection. It is usual that Jujuy farmers treat tuberosum landraces that do not present photoperiodic requirements (such as "Holandesa" or "Desireé", for example) as short-day potatoes. This can lead to the collection of tuberosum landraces as if they were, in fact, Andigena. Another possibility is that mislabeling occurred during in vitro conservation, when morphologic characteristics used for identification are sometimes not evident. All the above indicate the need for using markers when managing accessions in a germplasm bank.

A study performed by Monte and Rey Burusco et al. [17] with the same potato collection using 26 SSR markers, revealed the presence of only two groups (with the colored flesh accessions as a subgroup of the major one). Given the fact that these DArT markers are homogeneously distributed across the genome and have a high coverage of genomic regions, they allow the discovery of relationships among accessions that cannot be found using other types of markers.

It is worth noticing that most of the accessions were assigned to any of the three clusters with probabilities higher than 0.6. Only 18 out of 114 accessions (15.78%) presented probabilities lower than 0.6. One was accession 32, which displays a rare position in the PCoA analysis, distantly located from all of the three groups. Interestingly, this accession showed a maximum of two alleles for each of the 26 SSRs reported by Monte and Rey Burusco et al. [17], which might indicate that it is a diploid genotype. Moreover, it is highly homozygous (see Figure 2).

Seven 'Imilla Negra' (IN) accessions provided by the Cauqueva Cooperative were included in this study. These were sub-selected within the same landrace seed pool due to the morphological variability observed in preliminary analyses (data not shown). IN is a largely diverse landrace that is very common in Bolivia and South Peru [78]. Five of the Cauqueva accessions were grouped together in the dendrogram and apart from the remaining two, which appeared in a separate branch (Figure 3). This grouping was consistent with distribution seen in the PCoA analysis (Figure 4). It is worth pointing out that accession 62, the only IN accession provided by BAL, appears in a distant branch, separated from any of the other INs under study. The results from Monte and Rey Burusco et al. [17], using SSRs, are in agreement with this.

Atencio et al. [11] previously noted the occurrence of genetic diversity within landraces. The study showed that 24 individuals belonging to the "Collareja" landrace, collected from the same parcel in a Jujuy local producer field, presented a high molecular diversity, despite having similar tubers and crop cycle. Moreover, these accessions did not show genetic identity when compared with other "Collareja" individuals collected in different geographic sites in Jujuy. This was indicative of the high level of polymorphism existing within this landrace and agreed with the results from the present study. In addition, this

IN grouping can be explained given the fact that landraces names do not necessarily reflect genetic relationships, but respond to the tuber morphological characteristics that describe them.

When analyzing homozygosity, DM was the accession that presented the higher homozygosity level (0.6% heterozygosity) due to its doubled monoploid nature. This was followed by 32 (2.5%), a possible diploid genotype [17]. Heterozygosity percentages were in agreement with those previously described [79]. Accessions 6 and 41, which do not show a consistent grouping throughout the studies performed, showed heterozygosity percentages of 8 and 13%, respectively.

As regards to heterozygosity of potato commercial varieties, Spunta presents 24%, Pampeana 26%, and Bintje 27%. Hirsh et al. [79] reported between 53 and 59 percent heterozygosity in market classes of cultivated potato. The only Andean potato analyzed in their study presented 29% heterozygosity, which is mostly in agreement with values obtained for our Andean cluster. It is worth noting that overall heterozygosity levels could have been underestimated due to the simplification of tetraploid genotypes to two state markers derived from DArTseq data. On the other hand, even though a given SNP could present four different allelic states (A, T, C or G), most SNPs often present two [80,81].

It was not possible to find a consistent pattern that relates genetic composition with collection site. This is due to high tuber interchange between local farmers during fairs and potato producers that are then used as a seed potato. As of today, there is no commercial source of propagules, since they have been traditionally interchanged informally for many years.

Our collection is composed of material collected mostly in areas of great interchange. In addition, collection campaigns were performed in multiple years and different departments were covered in each year of collection [53–57]. Thus, population clustering was neither related to year nor location of collection. Colman [82] proved that there was no statistically significant association between population structure and year or site of collection for a sample of 50 Andean accessions that were included in the present study [82].

For accessions with group assignment higher than 0.6 in the Bayesian analysis, clustering was the same when analyzed by a maximum likelihood tree. On the other hand, for accessions with probabilities lower than 0.6, the results were not consistent, giving rise to branches forming new groups in the tree. When these accessions are removed from the study, the clustering obtained matches the Bayesian analysis with only three mislocated accessions (Supplemental Figure S1). In addition, three accessions, DM, 6, and 41 formed a group with Tuberosum group accessions when analyzed with STRUCTURE, but clustered either in a new group (Figure 3) or with the Andean group (Supplemental Figure S1) in the tree. The fact that some accessions group together with Andean accessions or in a different group depending on the type of analysis shows that these studies might not be sufficient to characterize them. Additional studies including other landraces and wild species could be conducted to determine the appropriate grouping of these genotypes.

Surprisingly, a group of 19 accessions provided by the BAL, classified as Andean varieties, consistently clustered with Tuberosum group accessions. Given the fact that previous studies [83] on Andean potatoes have shown similar results, it could be hypothesized that these 19 accessions were, in fact, reintroductions of lowland commercial potatoes. It should be noted that the 26 SSRs used in the study mentioned above were able to group all these accessions, except for 41, 6, and 4, together with the Tuberosum group. Nonetheless, the grouping of accessions 4 and 6 with Andean potatoes was not strong in the PCoA analyses. On the other hand, 41 is the only accession for which DArTseq markers were able to reveal grouping with commercial genotypes, but SSRs were not.

In vitro studies showed that tuberization behavior for accession 2 (one of the aforementioned accessions) was photoperiod independent, a characteristic shown by commercial potatoes [17]. This was also supported by the *cdf1* allele characterization for this genotype. This gene is responsible for triggering cellular signaling cascades for tuberization in response to the photoperiod. Commercial accessions, such as Bintje and Spunta, were found

to have an allelic variant with a 7 bp insertion in the *cdf1* gene, which makes tuberization independent from photoperiod. The same insertion was found for accession 2, as well as for accessions 25, 45, and 92, Andean accessions that were also grouped together in the commercial potato cluster [17].

The detailed characterization of genetic diversity with a high number of molecular markers is of great importance when performing association mapping studies [22]. A genetically diverse collection, such as the one employed for the present study, provides a good starting point for future association analyses. Unique alleles were identified in all three clusters obtained. It is important to point out that the group with the highest number of unique alleles, 417, was the Andean group, when compared to 225 alleles for the commercial potatoes group and only 29 unique alleles for the colored flesh accessions group. This was in agreement with previous studies revealing a great genetic diversity for Andean potato landraces from Northwestern Argentina [14]. Such a characteristic makes them a valuable source of genes of interest for potato breeding programs.

In conclusion, the present study has revealed genetic diversity of native potatoes from NWA using high-density markers.

Diversity studies that include high marker densities will be useful for future association mapping studies to identify candidate genes and alleles of characteristics of interest.

Supplementary Materials: The following data are available online at https://www.mdpi.com/article/10.3390/horticulturae8010054/s1, Figure S1: STRUCTURE plot of a subset of 96 potato accessions with memberships higher than 0.6 to any of the 3 subpopulations based on 5035 SNP markers.

Author Contributions: Conceptualization, S.S., M.F.C., M.F.R.B. and M.N.M.; methodology, S.S., M.F.C., M.F.R.B. and M.N.M.; software, M.A.C.; validation, G.A.M. and S.E.F.; investigation, S.S., M.F.C., M.F.R.B. and M.N.M.; data curation, M.A.C.; writing—original draft preparation, S.S., M.F.C., M.F.R.B., M.A.C. and M.N.M.; writing—review and editing, S.S., M.F.C., M.F.R.B., M.A.C.; supervision, G.A.M. and S.E.F.; project administration, G.A.M. and S.E.F.; funding acquisition, G.A.M. and S.E.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by INTA projects PNAIyAV 1130043 "Estrategias para la Diferenciación de Alimentos y el Desarrollo de Nuevos Productos Alimentarios" [Strategies for food differentiation and development of new food products] INTA 2013-2019, PNBIO 1131024 ¨Desarrollo de sistemas alternativos de generación y utilización de variabilidad genética y su aplicación al mejoramiento de los cultivos¨ [Development of Alternative systems of generation and utilization of genetic variation, and its application on cultivars breeding] INTA 2013-2019; PNBIO 1131023 ¨Prospección y caracterización funcional de genes de interés biotecnológico' [Prospection and functional characterization of genes with biotechnological interest] INTA 2013-2019, PNBIO 1131042 "Genómica aplicada a la caracterización de la diversidad genética" [Genomics applied to genetic diversity characterization] INTA 2013-2019, AETA 282811 INTA 2009-2012 and AERG- 231221 INTA 2009-2012. In addition, this work was possible thanks to funding from FONCYT, MINCyT ARGENTINA: PICT- 2010-2037 and PE/09/02 and CONCyTEC 2011-2012. MM and MFC hold CONICET doctoral fellowships and MFRB by a CONICET post-doctoral fellowship. 1. PICT2017-1070 "Bases genéticas y moleculares de la tuberización en Solanum tuberosum: Mo-dulación del ciclo de cultivo." 2. "Proyecto INTA PE-I114: Caracterización de la diversidad genética de plantas, animales y mi-croorganismos mediante herramientas de genómica aplicada."

Acknowledgments: We would like to thank the BAL for providing most of the germplasm material used in the present study. Also, Javier Rodriguez from CAUQUEVA Cooperative and Berta Andrade from Jujuy National University for providing the remaining accessions for the analyses. This material is maintained in vitro by Silvina Divito at INTA EEA Balcarce. Tubers are annually planted in Jujuy, Argentina (23°45′ S 65°30′ O, 3600 masl) in collaboration with Dario Castro, Daniel Martinez, Pablo Garcia, Irma Cusi and Kevin Dante Tolab from the Agroexperimental Station INTA Abrapampa, and Guillermo Aracena and Damián Alcoba from INTA IPAF NOA.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. FAOSTAT. FAO Production Statistics. Food and Agriculture Organization: Crops Statistics Database. Available online: www.faostat.fao.org (accessed on 7 March 2021).
2. Potato Genome Sequencing Consortium. Genome sequence and analysis of the tuber crop potato. *Nature* **2011**, *475*, 189–195. [CrossRef] [PubMed]
3. Sharma, S.K.; Bolser, D.; de Boer, J.; Sønderkær, M.; Amoros, W.; Carboni, M.; D'Ambrosio, J.M.; de la Cruz, G.; Di Genova, A.; Douches, D.S.; et al. Construction of Reference Chromosome-Scale Pseudomolecules for Potato: Integrating the Potato Genome with Genetic and Physical Maps. *G3 Genes Genomes Genet.* **2013**, *3*, 2031–2047. [CrossRef]
4. Spooner, D.M.; Núñez, J.; Rodríguez, F.P.; Naik, S.; Ghislain, M. Nuclear and chloroplast DNA reassessment of the origin of Indian potato varieties and its implications for the origin of the early European potato. *Theor. Appl. Genet.* **2005**, *110*, 1020–1026. [CrossRef]
5. Milbourne, D.; Pande, B.; Bryan, G.J. *Genome Mapping and Molecular Breeding in Plants Pulses, Sugar and Tuber Crops*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 12, pp. 205–236.
6. Rumold, C.U.; Aldenferder, M.S. Late Archaic–Early Formative period microbotanical evidence for potato at Jiskairumoko in the Titicaca Basin of southern Peru. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13672–13677. [CrossRef]
7. Dodds, K.S. Classification of cultivated potatoes. In *the Potato and Its Wild Relatives*; Series of Botanical Studies 4; Correll, D.S., Ed.; Texas Research Foundation: Renner, TX, USA, 1962; pp. 517–539.
8. Gavrilenko, T.; Antonova, O.; Shuvalova, A.; Krylova, E.; Alpatyeva, N.; Spooner, D.M.; Novikova, L. Genetic diversity and origin of cultivated potatoes based on plastid microsatellite polymorphism. *Genet. Resour. Crop. Evol.* **2013**, *60*, 1997–2015. [CrossRef]
9. Vos, P.G.; Uitdewilligen, J.G.; Voorrips, R.E.; Visser, R.G.; van Eck, H.J. Development and analysis of a 20K SNP array for potato (Solanum tuberosum): An insight into the breeding history. *Theor. Appl. Genet.* **2015**, *128*, 2387–2401. [CrossRef]
10. Ames, M.; Spooner, D.M. DNA from herbarium specimens settles a controversy about origins of the European potato. *Am. J. Bot.* **2008**, *95*, 252–257. [CrossRef]
11. Atencio, H.M.; Ispizúa, N.V.; Feingold, S.; Clausen, A.M. Conservación ex situ de variedades de papas nativas. Caso de estudio de la variedad 'Collareja' del noroeste de la Argentina. *Rev. Investig. Agropecu. (RIA)* **2019**, *45*, 242–251.
12. Andre, C.M.; Ghislain, M.; Bertin, P.; Oufir, M.; del Rosario Herrera, M.; Hoffmann, L.; Hausman, J.F.; Larondelle, Y.; Evers, D. Andean potato cultivars Solanum tuberosum L. as a source of antioxidant and mineral micronutrients. *J. Agric. Food Chem.* **2007**, *55*, 366–378. [CrossRef]
13. Clausen, A.M. Tesoros de la Biodiversidad Andina: Las Papas Nativas y su Valor Para la Humanidad. Recuperado el 18 de Marzo de 2019, del Sitio Web del Instituto Nacional de Tecnología Agropecuaria. Available online: http://inta.gob.ar/noticias/tesoros-de-la-biodiversidad-andina-las-papas-nativas-y-su-valor-para-la-humanidad (accessed on 18 March 2021).
14. Ispizúa, V.N.; Guma, I.R.; Feingold, S.; Clausen, A.M. Genetic diversity of potato landraces from northwestern Argentina assessed with simple sequence repeats (SSRs). *Genet. Resour. Crop. Evol.* **2007**, *54*, 1833–1848. [CrossRef]
15. Ortiz, R. The state of the use of potato genetic diversity. In *Broadening the Genetic Base of Crop Production*; CABI Publishing: Wallingford, UK, 2011; pp. 181–200.
16. Hardigan, M.A.; Bamberg, J.; Buell, C.R.; Douches, D.S. Taxonomy and genetic differentiation among wild and cultivated germplasm of Solanum sect. Petota. *Plant Genome* **2015**, *8*, 1–16. [CrossRef]
17. Monte, M.N.; Burusco, M.F.R.; Carboni, M.F.; Castellote, M.A.; Sucar, S.; Norero, N.S.; Colman, S.L.; Massa, G.A.; Colavita, M.L.; Feingold, S.E. Genetic Diversity in Argentine Andean Potatoes by Means of Functional Markers. *Am. J. Potato Res.* **2018**, *95*, 286–300. [CrossRef]
18. Colman, S.L.; Massa, G.A.; Carboni, M.F.; Feingold, S.E. Cold sweetening diversity in Andean potato germplasm from Argentina. *J. Sci. Food Agric.* **2019**, *678*, 499–524. [CrossRef]
19. Esquinas-Alcázar, J. Protecting crop genetic diversity for food security: Political, ethical and technical challenges. *Nat. Rev. Genet.* **2005**, *6*, 946–953. [CrossRef] [PubMed]
20. FAO. Tratado Internacional Sobre los Recursos Fitogenéticos Para la Alimentación y la Agricultura. 2001. Available online: http://www.fao.org/plant-treaty/es/ (accessed on 8 March 2021).
21. Brush, S.B.; Taylor, J.; Bellon, M.R. Technology adoption and biological diversity in Andean potato agriculture. *J. Dev. Biol.* **1992**, *39*, 365–387. [CrossRef]
22. Ortiz, R. Genomic-Led Potato Breeding for Increasing Genetic Gains: Achievements and Outlook. *Crop. Breed. Genet. Genom.* **2020**, *2*, e200010. [CrossRef]
23. Bradshaw, J.E.; Hackett, C.A.; Meyer, R.C.; Milbourne, D.; McNicol, J.W.; Phillips, M.S.; Waugh, R. Identification of AFLP and SSR markers associated with quantitative resistance to *Globodera pallida* (Stone) in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*) with a view to marker-assisted selection. *Theor. Appl. Genet.* **1998**, *97*, 202–210.
24. Ghislain, M.; Trognitz, B.; Herrera, M.D.R.; Solis, J.; Casallo, G.; Vasquez-Robinet, C.; Hurtado, O.; Castillo, R.; Portal, L.; Orrillo, M. Genetic loci associated with field resistance to late blight in offspring of *Solanum phureja* and *S. tuberosum* grown under short-day conditions. *Theor. Appl. Genet.* **2001**, *103*, 433–442. [CrossRef]
25. Norero, N.; Malleville, J.; Huarte, M.; Feingold, S.E. Cost efficient potato *Solanum tuberosum* L. cultivar identification by microsatellite amplification. *Potato Res.* **2004**, *45*, 131–138. [CrossRef]

26. Feingold, S.; Lloyd, J.; Norero, N.; Bonierbale, M.; Lorenzen, J. Mapping and characterization of new EST-derived microsatellites for potato (*Solanum tuberosum* L.). *Theor. Appl. Genet.* **2005**, *111*, 456–466. [CrossRef]
27. Fu, Y.B.; Peterson, G.W.; Richards, K.W.; Tarn, T.R.; Percy, J.E. Genetic diversity of Canadian and exotic potato germplasm revealed by simple sequence repeat markers. *Am. J. Potato Res.* **2009**, *86*, 38–48. [CrossRef]
28. Yıldırım, A.; Kandemir, N.; Sönmezoğlu, Ö.A.; Güleç, T.E. Transferability of microsatellite markers among cool season cereals. *Biotechnol. Biotechnol. Equip.* **2009**, *23*, 1299–1302. [CrossRef]
29. Favoretto, P.; Veasey, E.A.; Melo, P.C.T. Molecular characterization of potato cultivars using SSR markers. *Hortic. Bras.* **2011**, *29*, 542–547. [CrossRef]
30. Provan, J.; Powell, W.; Hollingsworth, P.M. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* **2001**, *16*, 142–147. [CrossRef]
31. Zane, L.; Bargelloni, L.; Patarnello, T. Strategies for microsatellite isolation: A review. *Mol. Ecol.* **2002**, *11*, 1–16. [CrossRef]
32. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2011**, *177*, 309–334. [CrossRef]
33. Andersen, J.R.; Lübberstedt, T. Functional markers in plants. *Trends Plant Sci.* **2003**, *8*, 554–560. [CrossRef]
34. Yadav, P.; Vaidya, E.; Rani, R.; Yadav, N.K.; Singh, B.K.; Rai, P.K.; Singh, D. Recent Perspective of Next Generation Sequencing: Applications in Molecular Plant Biology and Crop Improvement. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* **2018**, *88*, 435–449. [CrossRef]
35. Jaccoud, D.; Peng, K.; Feinstein, D.; Kilian, A. Diversity arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* **2001**, *29*, e25. [CrossRef]
36. Wenzl, P.; Carling, J.; Kudrna, D.; Jaccoud, D.; Huttner, E.; Kleinhofs, A.; Kilian, A. Diversity arrays technology (DArT) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9915–9920. [CrossRef]
37. Courtois, B.; Audebert, A.; Dardou, A.; Roques, S.; Herrera, T.G.; Droc, G.; Frouin, J.; Rouan, L.; Gozé, E.; Kilian, A.; et al. Genome-Wide Association Mapping of Root Traits in a Japonica Rice Panel. *PLoS ONE* **2013**, *8*, e78037. [CrossRef]
38. Hahn, V.; Würschum, T. Molecular genetic characterization of Central European soybean breeding germplasm. *Plant Breed.* **2014**, *133*, 748–755. [CrossRef]
39. Von Mark, V.C.; Kilian, A.; Dierig, D.A. Development of DArT marker platforms and genetic diversity assessment of the US collection of the new oilseed crop lesquerella and related species. *PLoS ONE* **2013**, *8*, e64062.
40. Raman, H.; Raman, R.; Coombes, N.; Song, J.; Prangnell, R.; Bandaranayake, C.; Tahira, R.; Sundaramoorthi, V.; Killian, A.; Meng, J.; et al. Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant Cell Environ.* **2016**, *39*, 1228–1239. [CrossRef] [PubMed]
41. Dracatos, P.M.; Haghdoust, R.; Singh, R.P.; Huerta Espino, J.; Barnes, C.W.; Forrest, K.; Hayden, M.; Niks, R.E.; Park, R.F.; Singh, D. High-Density Mapping of Triple Rust Resistance in Barley Using DArT-Seq Markers. *Front. Plant Sci.* **2019**, *10*, 467. [CrossRef] [PubMed]
42. Yang, X.; Ren, R.; Ray, R.; Xu, J.; Li, P.; Zhang, M.; Liu, G.; Yao, X.; Kilian, A. Genetic diversity and population structure of core watermelon (Citrullus lanatus) genotypes using DArTseq-based SNPs. *Plant Genet. Resour.* **2016**, *14*, 226–233. [CrossRef]
43. Al-Beyroutiová, M.; Sabo, M.; Sleziak, P.; Dušinský, R.; Birčák, E.; Hauptvogel, P.; Kilian, A.; Švec, M. Evolutionary relationships in the genus Secale. *Plant Syst. Evol.* **2016**, *302*, 1083–1091. [CrossRef]
44. Farahani, S.; Maleki, M.; Mehrabi, R.; Kanouni, H.; Scheben, A.; Batley, J.; Talebi, R. Whole Genome Diversity, Population Structure, and Linkage Disequilibrium Analysis of Chickpea (*Cicer arietinum* L.) Genotypes Using Genome-Wide DArTseq-Based SNP Markers. *Genes* **2019**, *10*, 676. [CrossRef]
45. Hassani, S.M.R.; Talebi, R.; Pourdad, S.S.; Naji, A.M.; Fayaz, F. In-depth genome diversity, population structure and linkage disequilibrium analysis of worldwide diverse safflower (*Carthamus tinctorius* L.) accessions using NGS data generated by DArTseq technology. *Mol. Biol. Rep.* **2020**, *47*, 2123–2135. [CrossRef]
46. Śliwka, J.; Jakuczun, H.; Chmielarz, M.; Hara-Skrzypiec, A.; Tomczyńska, I.; Kilian, A.; Zimnoch-Guzowska, E. A resistance gene against potato late blight originating from *Solanum* × *michoacanum* maps to potato chromosome VII. TAG. *Theor. Appl. Genet.* **2011**, *124*, 397–406. [CrossRef] [PubMed]
47. Iorizzo, M.; Gao, L.; Mann, H.; Traini, A.; Chiusano, M.L.; Kilian, A.; Aversano, R.; Carputo, D.; Bradeen, J.M. A DArT marker-based linkage map for wild potato *Solanum bulbocastanum* facilitates structural comparisons between Solanum A and B genomes. *BMC Genet.* **2014**, *15*, 123. [CrossRef]
48. Nachtigall, M.; Konig, J.; Thieme, R. Mapping of a novel, major late blight resistance locus in the diploid (1EBN) Mexican *Solanum pinnatisectum* Dunal on chromosome VII. *Plant Breed.* **2018**, *137*, 433–442. [CrossRef]
49. Dracatos, P.M.; Zhang, P.; Park, R.F.; McIntosh, R.A.; Wellings, C.R. Complementary resistance genes in wheat selection 'Avocet R'confer resistance to stripe rust. *Theor. Appl. Genet.* **2016**, *129*, 65–76. [CrossRef] [PubMed]
50. Mogga, M.; Sibiya, J.; Shimelis, H.; Lamo, J.; Yao, N. Diversity analysis and genome-wide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. *PLoS ONE* **2018**, *13*, e0198012. [CrossRef] [PubMed]
51. Atencio, M. Diversidad en Variedades Andinas de Papa Solanum Tuberosum ssp. Andigena Evaluada con Microsatélites [Diversity on Andean Potato Varieties *Solanum tuberosum* ssp. Andigena Evaluated with Microsatellites]. Master's Thesis, Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina, 2011.

52. GCDT. Background on the development of the 'Global Strategy for the Ex situ Conservation of Potato'. 2006. Available online: https://www.croptrust.org/wp-content/uploads/2014/12/Potato-Strategy-FINAL-30Jan07.pdf (accessed on 5 April 2019).
53. Okada, K.A. Exploration, conservation and evaluation of potato germplasm in Argentina. *Potato Res.* **1976**, *19*, 263–269. [CrossRef]
54. Okada, K.A. Collection and taxonomy of Argentine wild species (tuber-bearing Solanum). In *Report of the Planning Conference on the Exploration, Taxonomy and Maintenance of Potato Germplasm, 15–19 October 1979*; CIP: Lima, Peru, 1979; pp. 98–113.
55. Okada, K.A.; Clausen, A.M. Collecting potatoes in northwestern Argentina. *Am. Potato J.* **1983**, *64*, 301–305.
56. Clausen, A.M. Collecting indigenous potato varieties in Northwest Argentina. *Plant Genet. Resour. Newsl.* **1989**, *80*, 38–39.
57. Clausen, A.M.; Colavita, M.; Butzonitch, I.; Carranza, A.V. A potato collecting expedition in the province of Jujuy, Argentina and disease indexing of virus and fungus pathogens in Andean cultivars. *Genet. Resour. Crop. Evol.* **2005**, *52*, 1099–1109. [CrossRef]
58. Haymes, H. Mini-prep method suitable for plant breeding program. *Plant Mol. Biol. Rep.* **1996**, *14*, 280–284. [CrossRef]
59. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef] [PubMed]
60. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef]
61. Earl, D.A.; von Holdt, B.M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2011**, *4*, 359–361. [CrossRef]
62. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [CrossRef] [PubMed]
63. Lee, T.H.; Guo, H.; Wang, X.; Kim, C.H.; Paterson, A.H. SNPhylo: A pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genom.* **2014**, *15*, 162. [CrossRef]
64. Zheng, X.; Levine, D.; Shen, J.; Gogarten, S.M.; Laurie, C.; Weir, B.S. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **2012**, *28*, 3326–3328. [CrossRef]
65. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef] [PubMed]
66. Felsenstein, J. PHYLIP-phylogeny inference package (Version 3.2). *Cladistics* **1989**, *5*, 164–166.
67. Schliep, K.P. Phangorn: Phylogenetic analysis in R. *Bioinformatics* **2011**, *27*, 592–593. [CrossRef] [PubMed]
68. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3, an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, 242–245. [CrossRef]
69. Peakall, R.; Smouse, P.E. GENALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
70. Nadeem, M.A.; Nawaz, M.A.; Shahid, M.Q.; Doğan, Y.; Comertpay, G.; Yıldız, M.; Hatipoğlu, R.; Ahmad, F.; Alsaleh, A.; Labhane, N.; et al. DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* **2018**, *32*, 261–285. [CrossRef]
71. Batley, J.; Barker, G.; O'Sullivan, H.; Edwards, K.J.; Edwards, D. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiol.* **2003**, *132*, 84–91. [CrossRef]
72. Pootakham, W.; Jomchai, N.; Ruang-Areerate, P.; Shearman, J.R.; Sonthirod, C.; Sangsrakru, D.; Tragoonrung, S.; Tangphatsornruang, S. Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics* **2015**, *105*, 288–295. [CrossRef]
73. Shearman, J.R.; Sangsrakru, D.; Jomchai, N.; Ruang-Areerate, P.; Sonthirod, C.; Naktang, C.; Theerawattanasuk, K.; Tragoonrung, S.; Tangphatsornruang, S. SNP identification from RNA sequencing and linkage map construction of rubber tree for anchoring the draft genome. *PLoS ONE* **2015**, *10*, e0121961. [CrossRef]
74. Soorni, A.; Fatahi, R.; Haak, D.C.; Salami, S.A.; Bombarely, A. Assessment of Genetic Diversity and Population Structure in Iranian Cannabis Germplasm. *Nat. Sci. Rep.* **2017**, *7*, 15668. [CrossRef]
75. Hayashi, K.; Hashimoto, N.; Daigen, M.; Ashikawa, I. Development of PCR-based SNP markers for rice blast resistance genes at the Piz locus. *Theor. Appl. Genet.* **2004**, *108*, 1212–1220. [CrossRef]
76. Picoult-Newberg, L.; Ideker, T.E.; Pohl, M.G.; Taylor, S.L.; Donaldson, M.A.; Nickerson, D.A.; Boyce-Jacino, M. Mining SNPs from EST databases. *Genome Res.* **1999**, *9*, 167–174. [CrossRef] [PubMed]
77. Deutsch, S.; Iseli, I.; Bucher, P.; Antonarakis, S.E.; Scott, H.S. A cSNP map and database for human chromosome 21. *Genome Res.* **2011**, *11*, 300. [CrossRef]
78. Ochoa, C.M. *The Potatoes of South America: Bolivia*; Cambridge University Press: Cambridge, UK, 1990.
79. Hirsch, C.N.; Hirsch, C.; Felcher, K.; Coombs, J.; Zarka, D.; Van Deynze, A.; De Jong, W.; Veilleux, R.E.; Jansky, S.; Bethke, P.; et al. Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3 Genes Genomes Genet.* **2013**, *3*, 1003–1013. [CrossRef] [PubMed]
80. Bali, S.; Robinson, B.R.; Sathuvalli, V.; Bamberg, J.; Goyer, A. Single Nucleotide Polymorphism (SNP) markers associated with high folate content in wild potato species. *PLoS ONE* **2018**, *13*, e0193415. [CrossRef] [PubMed]
81. Mammadov, J.; Aggarwal, R.; Buyyarapu, R.; Kumpatla, S. SNP Markers and Their Impact on Plant Breeding. *Int. J. Plant Genom.* **2012**, *2012*, 728398. [CrossRef] [PubMed]

82. Colman, S. Contribución Alélica de Genes Candidatos al Endulzamiento Inducido por Frío en Papa *Solanum tuberosum* L. [Candidate Genes Allelic Contribution to Cold Sweetening Induce on Potato *Solanum tuberosum* L.]. Ph.D. Thesis, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina, 2014.
83. Wang, Y.; Rashid, M.A.R.; Li, X.; Yao, C.; Lu, L.; Bai, J.; Li, Y.; Xu, N.; Yang, Q.; Zhang, L.; et al. Collection and Evaluation of Genetic Diversity and Population Structure of Potato Landraces and Varieties in China. *Front. Plant Sci.* **2019**, *10*, 139. [CrossRef] [PubMed]