Small Ruminant Research 193 (2020) 106246

Contents lists available at ScienceDirect

Small Ruminant Research

journal homepage: www.elsevier.com/locate/smallrumres

Genetic diversity and validation of a microsatellite panel for parentage testing for alpacas (*Vicugna pacos*) on three Peruvian farms

J.A. Morón [a,*], E.A. Veli [b], A. Membrillo [c], M.M. Paredes [a], G.A Gutiérrez [a]

[a] Laboratorio de Biologia Molecular y Genómica, Instituto de Investigación de Bioquímica y Biología Molecular, Universidad Nacional Agraria la Molina, Av. La Molina s/n, La Molina, Lima, Peru
[b] Laboratorio de Biología Molecular y Genómica, Instituto Nacional de Innovación Agraria, Av. La Molina 1981, La Molina, Lima, Peru
[c] Department of Genetic, University of Córdoba, Campus Rabanales, 14071, Córdoba, Spain

ARTICLE INFO

Keywords:
Microsatellite
Alpaca
Genetic diversity
Parentage testing

ABSTRACT

The alpaca is of greatest economic importance in the Peruvian High Andes. This study aimed to determine the genetic diversity of three Peruvian alpaca farms, as well as, to validate a microsatellite markers panel for paternity testing. In this study, 247 samples of Huacaya alpacas were taken from three different localities (Sanjo, San Pedro de Raco and Cachipampa) from Pasco Region in Peru. DNA was obtained from hair follicles and genotyped for 15 microsatellites markers in multiplex electrophoresis runs.

A total of 225 alleles were detected across the 15 loci investigated. The polymorphism information content considering all loci was 0.82, which indicated that the microsatellite panel was very polymorphic and highly informative. The estimated diversity parameter showed that farms have high levels of genetic diversity ($H_E = 0.826$), and revealed the existence of genetic differentiation among the farms ($F_{ST} = 2.8$ %). The highest inbreeding coefficient was in the Sanjo farm ($F_{IS} = 0.303$). The results of the parentage testing indicated that all loci showed values greater than 70 % probability of discrimination. However, the highest values found were 94 % (YWLL08) and 90 % (YWLL36). The average of the probability of exclusion obtained was 0.999994 if the genotype for one alleged parents is known, and 0.99999 if the genotypes for both alleged parents are known.

The results obtained show that there is a high genetic diversity and validate the panel of microsatellite markers, that would help to improve the identification system and genealogical data collection.

1. Introduction

Peru hosts with approximately 3 685 516 animals the world's largest alpaca population. Alpacas are well adapted to the harsh conditions of the High Andes and are of great economic importance as their fiber is in high demand on the national and international markets (FAO, 2005; Gutierrez et al., 2018).

At present, the largest alpaca populations in Peru are found in the regions Puno, Cusco, Arequipa, Huancavelica, Apurímac, Ayacucho and Pasco (CENAGRO, 2012). Alpacas are usually kept by smallholder farmers (85–90%) and medium size farmers (about 10 %).

Correct pedigree information is important for performing genetic evaluations as errors lead to incorrect estimates and low accuracies of estimated breeding values (Maichomo et al., 2008). Biotechnology such as DNA markers is a viable option for several domestic animal breeding programs to detect/validate correct pedigree assignments (Souza et al., 2012).

DNA genotyping using the genetic markers has become the most common procedure for paternity test and pedigree inferences, not only in human but also in livestock species (Al-Atiyat, 2015). The microsatellite markers are considered the best tool for genetic identification, genetic relationships within and among populations, as well as, for parentage testing in breeds (Schlötterer and Harr, 2001; Yilmaz, 2016).

In this regard, some studies of genetic diversity and parentage testing have been carried out for different Peruvian alpaca populations. Rodriguez et al. (2009) report results from the central Peruvian Andes (Junín, Huancavelica) and other authors (Lang et al., 1996; Agapito et al., 2008; Rodriguez et al., 2009; La Manna et al., 2011; Paredes et al., 2013; Yalta et al., 2014; Paredes et al., 2014) cover the Southern region (Puno, Cusco, Arequipa).

* Corresponding author at: Laboratorio de Biología Molecular y Genómica, Instituto de Investigación de Bioquímica y Biología Molecular, Universidad Nacional Agraria la Molina, Lima, Peru.
E-mail address: jmoron@lamolina.edu.pe (J.A. Morón).

https://doi.org/10.1016/j.smallrumres.2020.106246
Received 21 December 2019; Received in revised form 31 July 2020; Accepted 20 September 2020
Available online 22 September 2020

The genetic diversity studies of Peruvian alpaca populations would provide important information for the High-Andean communities and alpaca breeding centers, to undertake breeding and conservation strategies. Also, the parentage testing studies would help these communities, where the information on parent-offspring is incomplete or unavailable.

The aim of this research was to determinate the genetic diversity of three Peruvian alpaca farms, as well as, to validate a microsatellite markers panel for paternity testing.

2. Material and methods

2.1. Collection of biological samples and microsatellite analysis

Hair follicle samples from 247 Huacaya alpacas were collected on three different farms of the Central Andes in Pasco Region in Peru. These farms were chosen as they had pedigree information available. Young animals ($n = 115$), their mothers ($n = 103$) and their assigned fathers and other males, which were possible candidates ($n = 29$) were sampled. These localities are: Cachipampa ($n = 50$), Sanjo ($n = 83$) and San Pedro de Raco ($n = 114$) (Fig. 1). All samples were used in both studies, paternity test and genetic diversity. DNA extraction was carried out with Sambrook and Russell (2001) method, modified by the Laboratorio de Biologia Molecular y Genomica of the Instituto de Investigación en Bioquimica y Biología Molecular, Universidad Nacional Agraria la Molina, Lima, Peru.

A selection of fifteen microsatellite loci (Table 1) was used in the present work, some of these were used in previous studies for parentage verification in Peruvian alpacas. VOLP32, YWLL08 and YWLL44 are recommended by the International Society of Animal Genetic (ISAG--FAO, 2011).

The microsatellites were amplified in three multiplex PCRs, according to fragment length and fluorescence labeling (Table 2). Each PCR (Polymerase chain reaction) amplification was performed in a total volume of 20 µL containing: 10X reaction buffer; $MgCL_2$ 3 mM; dNTPs 0.4 mM, Taq DNA polymerase 0.5U, 0.07−0,4 mM of each primer (forward primers labeled with fluorochromes at its 5′); and 3 µL of approximately 20 ng/µl genomic DNA.

The PCR reactions were optimized to amplify all microsatellite in an Mastercycler® (Eppendorf), and the PCR program was: first and second PCR multiplex, initial denaturation at 95 °C for 5 min, 28 cycles of 95 °C for 30 s, 61 °C and 57 °C for 90 s,72 °C for 60 s; and a final extension at 60 °C for 30 min; and the third PCR multiplex (Table 2), initial denaturation at 95 °C for 5 min, 25 cycles of 95 °C for 30 s, 56 °C for 90 s, 72 °C for 60 s; and a final extension 72 °C for 30 min.

The PCR products were attached in the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems ABI 3130), and GeneScan 500 LIZ size standard (Thermo Fisher Scientific), GeneMapper software (version 4.01) (Applied Biosystems) was used to determinate fragment size.

2.2. Statistical analyses

The Cervus version 3.0.3 software (Marshall et al., 1998) was used to determine the following parameters: The number of alleles (Na), observed heterozygosity (Ho) expected heterozygosity (He), polymorphism information content (PIC) values and Hardy-Weinberg

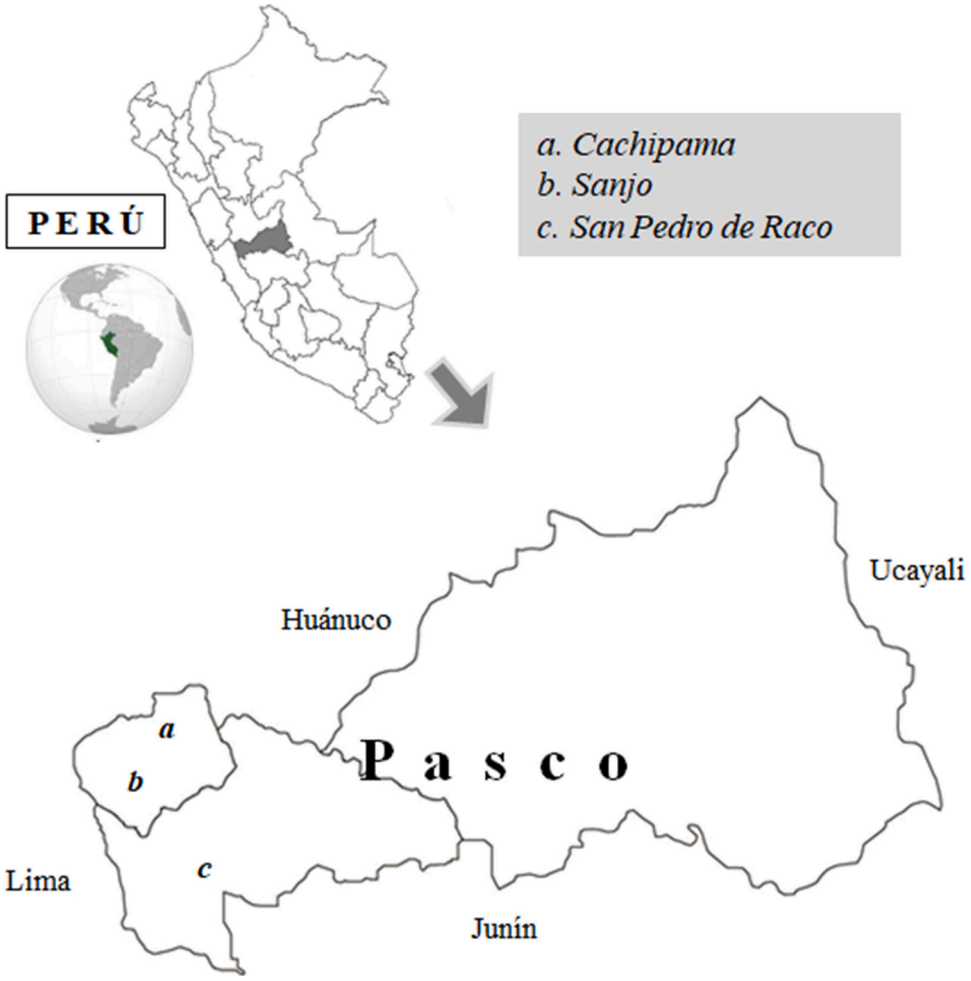

Fig. 1. Peru map showing the geographical distribution of the three alpaca farms from central Andean zone, Pasco region.

Table 1
Summary data of the 15 microsatellite markers used on three Peruvian alpaca farms.

Locus	Primer (5′–3′) Forward and Reverse	Size (bp)	Alleles (*n*)	Reference
LCA05	F:GTGGTTTTTGCCCAAGCTC R:ACCTCCAGTCTGGGGATTTC	180–204	11	Penedo et al. (1998a)
LCA08	F:GCTGAACCACAATGCAAAGA R:AATGCAGATGTGCCTCAGTT	228–262	14	Penedo et al. (1998a)
LCA37	F:AAACCTAATTACCTCCCCCA R:CCATGTAGTTGCAGGACACG	128–180	18	Penedo et al. (1998a)
LCA66	F:GTGCAGCGTCCAAATAGTCA R:CCAGCATCGTCCAGTATTCA	220–262	24	Penedo et al. (1998b)
YWLL44	F:CTCAACAATGCTAGACCTTGG R:GAGAACACAGGCTGGTGAATA	86–120	11	Lang et al. (1996)
YWLL08	F:ATCAAGTTTGAGGTGCTTTCC R:CCATGGCATTGTGTTGAAGAC	135–177	13	Lang et al. (1996)
YWLL36	F:AGTCTTGGTGTGGTGGTAGAA R:TGCCAGGATACTGACAGTGAT	148–166	7	Lang et al. (1996)
LCA90	F:TATAACCCTGGTCTCGCCAA R:CCAAGTAGTATTCCATTATGCG	229–263	14	Penedo et al. (1999)
LCA94	F:GTCCATTCATCCAGCACAGG R:ACATTTGGCAATCTCTGGAGAA	189–213	9	Penedo et al. (1999)
VOLP92	F:AGTTATCTTACTTCCAATTAAAAT R:AACATAGAAACAGCATTGAG	194–218	8	Obreque et al. (1999)
VOLP32	F:GTGATCGGAATGGCTTGAAA R:CAGCGAGCACCTGAAAGAA	192–247	12	Obreque et al. (1998)
VOLP55	F:AGTTACCGGTTTTTAACCTAT R:GACTTACTATGTGCCAATC	159–189	9	Obreque et al. (1999)
VOLP72	F:ACCAGGAAACCCAACTACTCTT R:GTCAAGGGGCAGGATGT	150–190	11	Obreque et al. (1999)
VOLP04	F:GCATTTCTCCGTAATCATTG R:TGACACCTTTTGTTTCCATT	226–258	13	Obreque et al. (1999)
VOLP77	F:TATTTGGTGGTGACATT R:CATCACTGTACATATGAAGG	144–168	11	Obreque et al. (1999)

equilibrium (HWE) test for each locus.

The statistical package GENETIX 4.05 (Belkhir et al., 2004) was used to estimate the fixation indices (F_{IS}, F_{ST}, and F_{IT}) per locus and the genetic structure of the alpaca population related to the differentiation within and between the studied alpaca population, *F* statistics (fixation indices) were calculated according to Weir and Cockerham (1984) with the Jackknife procedure applied over the loci and a confidence interval of 95 % computed with 1000 bootstraps. The genetic structure of the alpaca populations was analyzed with Wright's *F* statistics (Wright, 1965) using the pairwise distance (F_{ST}), inbreeding coefficients (F_{IS}) and gene flow (number of migrations in each generation, Nm).

The Factorial Correspondence analysis (FCA) of the individual multilocus genotype was performed to investigate the population differentiation pattern using GENETIX 4.05 (Belkhir et al., 2004) software.

Population structure was determined using Bayesian based approach implemented in the software STRUCTURE 2.3.4 (Pritchard et al., 2000), which inferred fractions on the admixture ancestry model for individuals and populations assuming a given number (K) of cluster. A Monte Carlo Markov Chain was tested for $K = 1$ to K = 4 and runs were repeated 10 times to calculate the mean L (*K*), with a burn-in period of 100,000 and a run length of 300,000 iterations.

The optimal number of genetic clusters (*K*) was determined based by ΔK statistic, the second order rate of change in L (*K*) following the procedure of Evanno et al. (2005). Graphical representations of these statistics and the best *K* value were estimated using STRUCTURE HARVESTER v0.68 (Earl and Von Holdt, 2012). Population structure was also estimated by molecular variance analysis (AMOVA) using the GenAlEx 6.4 software (Peakall and Smouse, 2012).

The Cervus program was used to calculate, probability of exclusion (PE), probability of identity (PI) and probability of discrimination (PD). The parentage of PE, was calculated in different scenarios: considering the situations where both parents are genotyped but only one parent is evaluated for exclusion (PE-1), or both parents are evaluated for exclusion (PE-2).

The paternity simulation was run to estimate the resolving power of a series of loci given their allele frequencies, and to estimate critical values of the log-likelihood statistically logarithm of the odds (LOD) or Delta. Therefore, the confidence of made parentage assignments using the parentage analysis module could be evaluated statistically, the simulation was performed for either paternal or maternal side.

3. Results

3.1. Genetic diversity

The number of alleles per locus ranged from 2 (LCA5) to 32 (YWLL08) with an average of 15.66. The number of private alleles ranged from 3.91 (LCA5) to 17.54 (YWLL08). The expected heterozygosity (He) varied from 0.744 at LCA05 to 0.943 at YWLL08 with an average of 0.840 (He) and 0.652 (Ho). Most of the loci showed significant deviation from Hardy-Weinberg test. The PIC values considering all markers were highly informative (0.82), whereas the LCA05 (0.710) and LCA90 (0.733), showed to be the least informative markers (Table 3).

Table 2
Summary data for PCR conditions and fragment size ranges for the 15 amplified microsatellite markers.

Locus	Size (bp)	Fluorescent dye	Anneling temperature (°C)	Multiplex PCR
YWLL36	135–177	FAM	61.0	1
LCA66	217–261	FAM		
LCA05	180–208	FAM		
LCA08	224–260	HEX		
LCA94	185–217	NED		
LCA37	124–182	NED		
LCA90	225–265	NED		
YWLL44	74–128	FAM	57.0	2
YWLL08	121–189	HEX		
VOLP92	191–217	HEX		
VOLP32	191–275	HEX	56.0	3
VOLP55	152–190	HEX		
VOLP72	150–190	NED		
VOLP04	220–256	FAM		
VOLP77	115–177	FAM		

Table 3
Genetic variability parameters at 15 microsatellites loci analyzed in three Peruvian alpaca farms. Number alleles per locus (Na), total number of private alleles (PA), observed heterozygosity (H_O), expected heterozygosity (H_E), significance of Hardy–Weinberg equilibrium (HW) test, mean polymorphic information content (PIC) and *F*-statistic (F_{IS}, F_{ST}, F_{IT}) according to Weir and Cockerham (1984). Probability of Exclusion using 15 loci in 247 Peruvian alpacas, probability of exclusion knowing one parent (PE-1), probability of exclusion knowing both parents (PE-2), Total exclusion probability (Total PE), and Candidate parent (CP).

Locus	Na	PA	H_O	H_E	HW	PIC	PE-1	PE-2	CP	$F_{IS}(f)$[b]	F_{ST} (θ)[b]	F_{IT} (F)[b]
YWLL36	18	10.64	0.775	0.906	NS	0.897	0.675	0.806	0.940	0.21599	0.02926	0.23893
LCA66	15	6.45	0.644	0.845	***[a]	0.829	0.539	0.703	0.879	0.20874	0.02932	0.23194
LCA5	2	3.91	0.626	0.744	NS	0.710	0.356	0.537	0.734	0.21429	0.02857	0.23673
LCA08	11	6.41	0.638	0.844	***	0.824	0.523	0.690	0.863	0.21004	0.02592	0.23052
LCA94	8	5.35	0.650	0.813	*	0.786	0.453	0.629	0.810	0.21169	0.02875	0.23435
LCA37	21	6.67	0.589	0.850	***	0.834	0.549	0.710	0.883	0.20333	0.02986	0.22712
LCA90	12	4.17	0.427	0.760	***	0.733	0.390	0.572	0.774	0.19513	0.03023	0.21946
YWLL44	19	10.20	0.824	0.902	NS	0.892	0.663	0.798	0.935	0.22107	0.02820	0.24303
YWLL08	32	17.54	0.836	0.943	***	0.938	0.787	0.881	0.976	0.21892	0.02925	0.24177
VOLP92	11	4.76	0.635	0.790	NS	0.771	0.446	0.627	0.827	0.21162	0.02939	0.23479
VOLP04	22	6.90	0.781	0.855	NS	0.841	0.565	0.724	0.895	0.22067	0.02790	0.24241
VOLP32	19	9.90	0.583	0.899	***	0.889	0.657	0.793	0.933	0.20129	0.02621	0.22222
VOLP55	12	5.78	0.826	0.827	NS	0.802	0.481	0.654	0.831	0.22587	0.02902	0.24834
VOLP72	10	4.76	0.202	0.790	***	0.761	0.420	0.599	0.787	0.17564	0.02667	0.19762
VOLP77	13	6.10	0.744	0.836	NS	0.813	0.505	0.673	0.849	0.21824	0.02887	0.24081
Total	225	109.54							99.9999 [c]			
Average	15.667	7.30	0.652	0.8402		0.8213				0.20982	0.02851	0.23241

[a] $*P \leq 0.05$, $** P \leq 0.01$, $*** P \leq 0.001$ and NS, $P \geq 0.05$ non-significant deviation.
[b] Jackknifing estimates over the loci.
[c] Percentage of candidate parents typed.

In addition, Table 3 shows that the F_{IS} values were similar with a range from 0.175 (VOL72) to 0.225 (VOLP55). The overall deficit of heterozygotes (F_{IS}) of 20.9 % and the global deficit of heterozygotes from individuals within the total farms (F_{IT}) was 0.232, all the 15 loci showed a deficit of heterozygotes.

The mean number of alleles ranged from 10.53 for Cachipampa, 11.86 for Sanjo, and 14.13 for San Pedro de Raco. The average number of alleles within each farm was lower (12.17) than the overall mean (15.66), suggesting the occurrence of lower within-farm variation. In addition, the values of He are greater than the Ho in all three farms. The highest inbreeding level was observed in San Pedro de Raco farm (0.165) and lowest in Sanjo farm (0.303) (Table 4).

Pairwise Nei's genetic distance and F_{ST} values between all the alpaca farms are shown in Table 5. Pair-wise genetic differentiation (F_{ST}) ranged from 0.027 (between Sanjo and San Pedro de Raco) to 0.032 (between Cachipampa and San Pedro de Raco). According to the global F_{ST} values 3% of the total genetic variation corresponded to the differences between alpaca farms, whereas 97 % was explained by differences among individuals. These results indicate a large variation within each alpaca farm.

Gene Flow (Nm) between the three alpaca farms ranged from 7.67 (between Cachipampa and San Pedro de Raco) to 9.05 (between Sanjo and San Pedro de Raco) as presented in Table 5. These results showed that Cachipampa farm was differentiated from other alpaca farms (Sanjo and San Pedro de Raco). The largest Nm was found between Sanjo and San Pedro de Raco (9.05), implying a relatively high genetic relationship and low genetic differentiation, which is consistent with low F_{ST} of 0.027.

According to the correspondence analysis the first two components explain 100 % of the total variation (Fig. 2). The first axis explains 55.27 % of the total variation and separates the Sanjo farm from the other two farms (San Pedro de Raco y Cachipampa). The second axis represents 44.73 % of the total variation. This finding showed the isolation of the Cachipama alpaca farm.

Fig. 3 shows the genetic structure; when $K = 2$ was assumed, Sanjo (79.3 %) and Cachipampa (79.6 %) alpaca farms were assigned to cluster 1, while San Pedro de Raco farm (75.9 %) was assigned to cluster 2. When $K = 3$ was assumed, Sanjo farm (64.4 %) was assigned to a distinct cluster 3; similarly, Cachipampa (73.4 %) and San Pedro de Raco (60.6 %) farms, were assigned to clusters 1 and 2 respectively. Considerable admixture was noticed in these three alpaca farms (Table 6). The results revelated $K = 3$ as the optimal clustering solution for the given dataset, consistent with the factorial correspondence analysis.

Table 4
Total of Number alleles per locus (Na), mean number of alleles (MNA), average expected (H_E) observed (H_O) heterozygosity and inbreeding coefficient per farm (F_{IS}) for 15 microsatellite markers analyzed in three alpaca farms.

Farm	n	Total Na	MNA	Mean heterozygosity		F_{IS}
				H_O	H_E	
Sanjo	83	178	11.867	0.5775	0.8268	0.30334
Cachipampa	50	158	10.533	0.6736	0.8179	0.17816
San Pedro de Raco	114	212	14.133	0.6887	0.8243	0.16519

Table 5
Genetic distance (F_{ST}) estimates (above the diagonal) and Gene flow (Nm) (below the diagonal) between the three alpaca farms.

Farms	Sanjo	Cachipampa	San Pedro de Raco
Sanjo	–	0.028	0.027
Cachipampa	8.72	–	0.032
San Pedro de Raco	9.05	7.67	–

3.2. Parentage validation

Table 7 shows the probability of identity for the 15 loci with values greater than 70 %. The loci with the lowest probability of discrimination of parentage were LCA90 (74 %) and LCA05 (75 %). On the other hand, loci with the highest probability of identity were YWLL08 (0.059) and YWLL36 (0.095), exhibiting these microsatellites the highest probability of discrimination being 94 % and 90 %, respectively.

The lowest and the highest values were observed in LCA05 (PE-2 = 0.537) and YWLL08 (PE-2 = 0.881) loci respectively. The values obtained in the three situations of exclusion probabilities were high (greater than 0.999), with low value for LCA5 and higher for YWLL08. In addition, the minimum number of loci to obtain a probability greater of 0.999 was 8 with highest values PE-2 (Table 3).

Information of the correct and incorrect paternity assignment, according to the results of the paternity tests performed by the fifteen

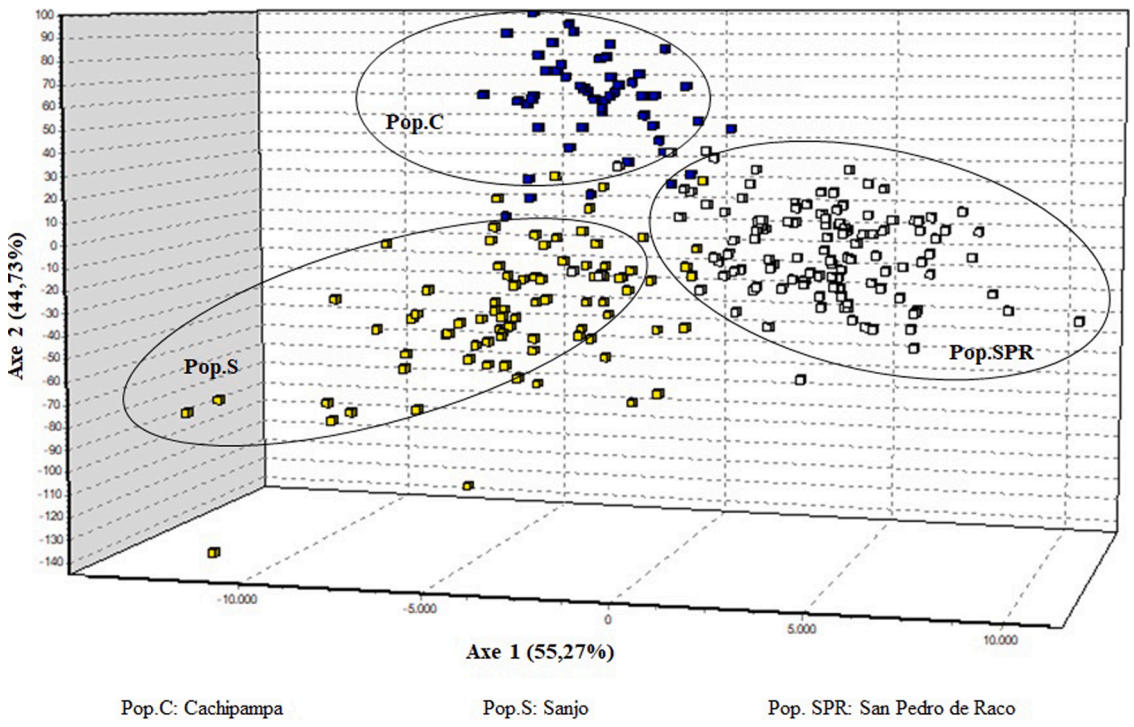

Fig. 2. Factorial Correspondence Analysis (FCA) of three alpaca farms studied on bases of 15 microsatellite loci.

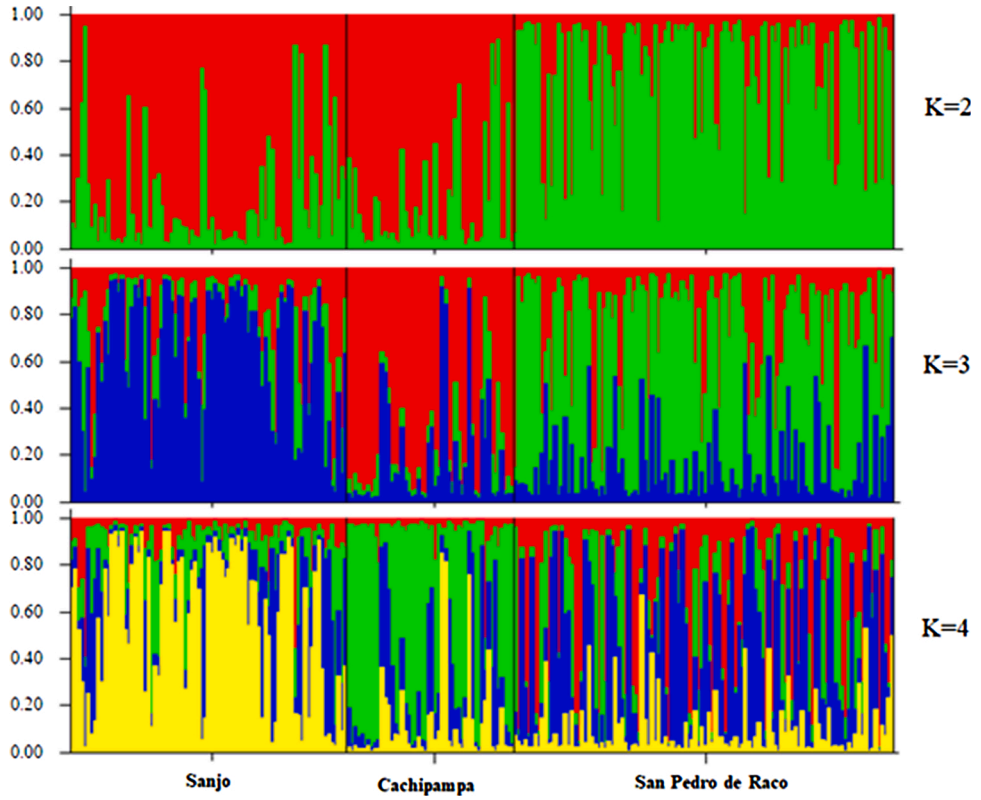

Fig. 3. Bayesian clustering analysis for the three alpaca farms, applying an admixture model with independent allele frequencies. A vertical bar that can be partitioned into colored fragments, with its length proportional to cluster contribution, represents each farm. Black lines separate the alpaca populations. *K* values represent number of ancestral farms assumed. The analyses were conducted with a burn-in period of 100,000 and 300,000 iterations.

Table 6
Proportion of membership coefficient of individual from three alpaca farms in different inferred cluster after STRUCTURE analysis.

Farms	N	Inferred clusters								
		$K=2$		$K=3$			$K=4$			
		1	2	1	2	3	1	2	3	4
Sanjo	83	0.793	0.207	0.213	0.143	0.644	0.084	0.142	0.201	0.574
Cachipampa	50	0.796	0.204	0.734	0.085	0.182	0.046	0.628	0.198	0.129
San Pedro de Raco	114	0.241	0.759	0.234	0.606	0.159	0.362	0.158	0.372	0.107

Table 7
Probability of 247 identity (PI) and discrimination (PD) by locus.

Locus	Alpaca farms				PD (%)
	Sanjo	Cachipampa	San Pedro de Raco	PI total farm	
LCA66	0.1568	0.2090	0.1651	0.1571	0.8429
LCA05	0.3348	0.2571	0.2373	0.2581	0.7419
LCA08	0.1585	0.1918	0.2254	0.1584	0.8416
LCA94	0.2073	0.2514	0.1838	0.1887	0.8113
LCA37	0.1216	0.1822	0.1797	0.1520	0.8480
LCA90	0.3262	0.2655	0.2119	0.2424	0.7576
YWLL36	0.1220	0.0928	0.1101	0.0956	0.9044
YWLL44	0.1291	0.1616	0.0997	0.0996	0.9004
YWLL08	0.0860	0.0804	0.0655	0.0593	0.9407
VOLP92	0.2053	0.2637	0.2180	0.2123	0.7877
VOLP04	0.1483	0.1308	0.2032	0.1466	0.8534
VOLP32	0.1316	0.2044	0.1201	0.1028	0.8972
VOLP55	0.1987	0.1691	0.1940	0.1753	0.8247
VOLP72	0.1909	0.2112	0.3008	0.2114	0.7886
VOLP77	0.1775	0.1936	0.1827	0.1665	0.8335

Table 8
Number and percentage of the correct and incorrect parental assignment in three alpaca farms.

Farms	Number offspring	Mating records		Error rate (%)
		Correct number	Incorrect number	
Sanjo	40	13	27	67.50
Cachipampa	20	20	0	0.00
San Pedro de Raco	56	44	12	21.42
Average	116	77	39	33.62

microsatellites are reported in Table 8. The average records error rate for three populations was calculated as 33.62 %.

4. Discussion

4.1. Genetic diversity

In terms of genetic variability, considering the dataset as a single farm, the majority number of alleles per locus ranged from 10 to 32, similar values for these loci were found in other alpaca populations (Penedo et al., 1998a; La Manna et al., 2011; Yalta et al., 2014; Paredes et al., 2014). The three alpaca farms showed in most of their loci significant deviations from the Hardy-Weinberg equilibrium. The results indicated high levels of genetic variability between farms.

The results of this study are in accordance with previous studies (La Manna et al., 2011; Yalta et al., 2014; Paredes et al., 2014), which also found a large number of polymorphic loci and the high genetic variability with other studies of alpaca populations in the southern (Rodriguez et al., 2009; La Manna et al., 2011; Paredes et al., 2013; Yalta et al., 2014; Paredes et al., 2014) and central Peruvian Andes (Rodriguez et al., 2009).

Previous studies showed relatively similar values of the heterozygote deficit (Yalta et al., 2014), and lower values (La Manna et al., 2011; Paredes et al., 2013, 2014). The results indicated values of observed heterozygosity lower than the expected in the three alpaca farms. A possible explanation for the heterozygosity deficiency might be related to inbreeding. Aranguren and Jordana (2001) explained that the loss of genetic diversity in a livestock population diminishes the ability to improve their performance.

The microsatellite loci from HWE and heterozygotes deficit indicated a departure from random mating which may be due to small population size, random genetic drift, inbreeding, selection or existence of null alleles (Young and Clarke, 2000; Dakin and Avise, 2004).

The small genetic differentiation between the three alpaca farms shows that the San Pedro de Racco farm has a slightly higher value or pair-wise FST (0.027) and the lowest value of Nm (7.67) compared with the two other alpaca farms. This result suggests that San Pedro de Raco farm has maintained genetic isolations from the Sanjo and Cachipampa farms. The AMOVA analysis showed only 3.0 % of the total genetic variation among alpaca farms, and similar results between other alpaca populations were reported (Yalta et al., 2014; Paredes et al., 2013).

The results of Bayesian analysis also revealed migration and admixture specially between Sanjo and San Pedro de Raco. This result is consistent with the PCA analysis, showed Cachipampa in a defined cluster.

Further analysis using more individuals covering the whole Peruvian Andean range would be required to get the full panorama of the genetic diversity of alpacas. This information would be a starting point for the proper design of conservation programs.

4.2. Parentage validation

The probability of identity (PI) estimates were less than 70 % in the total alpaca farms, for the 15 microsatellites loci evaluated. However, Yalta et al. (2014), showed PI values greater than found in this study. The PE values obtained were also higher than reported by Agapito et al. (2008) and Yalta et al. (2014).

The mean value of records error detected (33.62 %) of parental assignment was higher than in other reports and showed large differences between farms Yalta et al. (2014) and Agapito et al. (2008) obtained lower values (9.67 % and 4.44 %, respectively). These results indicate problems of correct data recording.

5. Conclusion

The present study reports the successful amplification of 15 microsatellite markers. The study reveals the richness of genetic resources raised in the central Peruvian Andes and can serve as a basis for decision on the development of breeding programs for these small and medium alpaca producers.

The probabilities of exclusion obtained and paternity test results indicated that this microsatellite panel is suitable for parentage testing

In conclusion, these markers can contribute to improve the identification system and support the genealogical data collection and management system.

Funding

The authors acknowledge the financial support from VLIR-UOS funding of the UNALM (IUC) programme. Opinion of the authors do not automatically reflect those of either the Belgian government or VLIR-UOS, and can bind neither the Belgian Government nor VLIR-UOS.

Declaration of Competing Interest

The authors reported no declarations of interest.

Acknowledgements

The authors would like to express their special thanks to the alpaca farmers who provided data for this study.

References

Agapito, J., Rodriguez, J., Herrera, P., Timoteo, O., Rojas, P., Boettcher, P.J., García, F., Espinoza, J.R., 2008. Parentage testing in alpacas (*Vicugna pacos*) using semi-automated fluorescent multiplex PCRs with 10 microsatellite markers. Anim. Genet. 39, 201–203.

Al-Atiyat, R.M., 2015. The power of 28 microsatellite markers for parentage testing in sheep. J. Biotechnol. 18, 116–121.

Aranguren, J., Jordana, J., 2001. Utilización de Marcadores de ADN (microsatélites) En Poblaciones de Animales Domésticos En Peligro De Extinción. Memorias Asociación Venezolana de Producción Animal (AVPA), Venezuela.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 2004. GENETIX 4.05, Logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171. Université de Montpellier II, Montpellier, France.

CENAGRO, 2012. Censo Nacional Agropecuario. http://siea.minagri.gob.pe/siea/?q=iv-censo-nacional-agropecuario-2012/iv-cenagro-2012.

Dakin, E.E., Avise, J.C., 2004. Microsatellite null alleles in parentage analysis. Heredity 93, 504–509.

Earl, Dent A., Von Holdt, Bridgett M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of cluster of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.

FAO (Food and Agriculture Organization), 2005. Situación Actual de Los Camélidos Sudamericanos En Bolivia, Proyecto de Cooperación Técnica En Apoyo a la Crianza Y Aprovechamiento de Os Camélidos Sudamericanos En la Región Andina TCP/RLA/2914. http://www.fao.org/tempref/GI/Reserved/FTP_FaoRlc/old/prior/segalim/animal/paises/pdf/2914bol.pdf.

Gutierrez, G., Gutierrez, J.P., Huanca, T., Wurzinger, M., 2018. Challenges and Opportunities of Genetic Improvement in Alpacas and Llamas in Peru. World Congress on Genetics Applied to Livestock Production, February 11-16, 2018, Auckland, New Zealand.

ISAG- F.A.O. (International Society of Animal Genetics – Food and Agriculture Organization), 2011. Molecular Genetic Characterization of Animal Genetic Resources, Commission on Genetic Resources for Food and Agriculture. ISBN 978-92-5-107032-1. http://www.fao.org/3/i2413e/i2413e00.pdf.

La Manna, V., La Terza, A., Ghezzi, S., Saravanaperumal, S., Apaza, N., Huanca, T., Bozzi, R., Renieri, C., 2011. Analysis of genetic distance between Peruvian Alpaca (*Vicugna pacos*) showing two distinct fleece phenotypes, Suri and Huacaya, by means of microsatellite Markers. Ital. J. Anim. Sci. 10, 60.

Lang, K.D.M., Wang, Y., Plante, Y., 1996. Fifteen polymorphic dinucleotide microsatellites in llamas and alpacas. Anim. Genet. 27, 285–294.

Maichomo, K.M., Gitau, W.G.K., Gathuma, J.M., Ndung'u, J.M., Kosura, O., Olivier, H.O., 2008. Extent and implications of incorrect offspring-sire relationships in pastoral production system in Kajiado District, Kenya. Livest. Res. R. Dev. 20, 80.

Marshall, T.C., Slate, J., Kruuk, L., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7, 639–655.

Obreque, V., Coogle, L., Henney, P.J., Bailey, E., Mancilla, R., García-Hiudobro, J., Hinrichsen, P., Cothran, E.G., 1998. Characterization of 10 polymorphic alpaca dinucleotide microsatellites. Anim. Genet. 29, 460–477.

Obreque, V., Mancilla, R., García, J., Cothran, E.G., Hinrichsen, P., 1999. Thirteen new dinucleotide microsatellites in alpaca. Anim. Genet. 30, 382–405.

Paredes, M.M., Membrillo, A., Azor, P.J., Machaca, J., Torres, D., Muñoz Serrano, A., 2013. Genetic and phenotypic variation in five populations of Huacaya Alpacas (Vicugna pacos) from Peru. Small Rumin. Res. 111, 31–40.

Paredes, M.M., Membrillo, A., Gutiérrez, J.P., Cervantes, I., Azor, P.J., Morante, R., Alonso-Moraga, A., Molina, A., Muñoz-Serrano, A., 2014. Association of microsatellite markers with fiber diameter trait in Peruvian alpacas (*Vicugna pacos*). Livest. Sci. 161, 6–16.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 1–3. https://doi.org/10.1093/bioinformatics/bts460.

Penedo, M.C.T., Caetano, A.R., Cordova, K.I., 1998a. Microsatellite markers for South American camelids. Anim. Genet. 29, 398–413.

Penedo, M.C.T., Caetano, A.R., Cordova, K., 1998b. Eight microsatellite markers for South American camelids. Anim. Genet. 30, 161–168.

Penedo, M.C.T., Caetano, A.R., Cordova, K.I., 1999. Six microsatellite markers for south american camelids. Anim. Genet. 30, 382–405.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Rodriguez, J., Agapito, J., De la Flor, I.D., Barreto, T., Timoteo, O., Herrera-Velit, P., Espinoza, J.R., 2009. Genetic diversity in Alpaca (*Vicugna pacos*) populations using 10 microsatellite markers. FAO/IAEA International Symposium on Sustainable Improvement of Animal Production and Health. https://inis.iaea.org/collection/NCLCollectionStore/_Public/41/033/41033544.pdf?r=1&r=1.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, New York.

Schlötterer, C., Harr, B., 2001. Microsatellite instability. Encyclopedia of Life Sciences (eLS). J. Wiley & Sons, pp. 1–4. https://doi.org/10.1038/npg.els.0000840.

Souza, C.A., Paiva, S.R., McManus, C.M., Azevedo, H.C., Mariante, A.S., Grattapaglia, D., 2012. Genetic diversity and assessment of 23 microsatellite markers for parentage testing of Santa Inês hair sheep in Brazil. Genet. Mol. Res. 11, 1217–1229.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19, 395–420.

Yalta, C., Sotil, G., Eudosio, V., 2014. Variabilidad genética y detección de error en filiación utilizando microsatélites en dos rebaños de alpacas Huacaya (*Vicugna pacos*). Salud tecnol. vet. 2, 134–145.

Yilmaz, O., 2016. Power of different microsatellite panels for paternity analysis in sheep. Anim. Sci. Pap. Rep. 34, 155–164.

Young, A.G., Clarke, G.M., 2000. Genetics, Demography and Viability of Fragmented Populations. Cambridge University Press, Cambridge.