

## Uso de la micromatriz de alta densidad de bovino para la construcción de un mapa físico de polimorfismos de nucleótido simple en alpacas (*Vicugna pacos*)

### Use of the high-density bovine microarray for the generation of an alpaca (*Vicugna pacos*) single nucleotide polymorphism physical map

Camilo Mamani<sup>1,5</sup>, Gustavo Augusto Gutiérrez Reynoso<sup>1</sup>,  
Polina Perelman<sup>2</sup>, Warren E. Johnson<sup>3</sup>, Federico Abel Ponce de León Bravo<sup>4</sup>

#### RESUMEN

El objetivo del estudio fue desarrollar un mapa físico preliminar de polimorfismos de nucleótido simple (PNSs) en alpaca usando un panel celular híbrido irradiado alpaca/hámster y una micromatriz de genotipado de alta densidad de bovino (BovineHD BeadChip-Illumina). Se realizó el genotipado de 92 clones celulares híbridos irradiados y cuatro muestras control (alpaca macho, alpaca hembra, hámster y 1:10 alpaca/hámster mezcla de ADN) con la micromatriz. Luego del genotipaje de las muestras control del ADN de alpaca y hámster solo se retuvieron PNSs de bovinos que mostraron una frecuencia de señal positiva de 1. Los PNSs identificados en el ADN de alpaca fueron filtrados para sustraer los presentes en el hámster. De estos últimos, solo se retuvieron para el análisis final aquellos que tuvieron una frecuencia de señal positiva de 0.2 a 0.8 para los 92 clones y se eliminaron los PNSs que presentaron falsos positivos. Los restantes PNSs específicos de alpaca fueron tabulados en el formato MapMaker y fueron analizados con el programa Carthagene para identificar grupos de ligamiento. Los grupos ligados fueron ubicados en el genoma referencial *Vicugna\_pacos-2.0.2*, aplicando el programa BLAST y el comando SHORTBLAST, en la plataforma Galaxy. Las dos muestras control de alpaca registraron 294 165 PNSs, con señal positiva en la micromatriz de bovino. La cantidad de PNSs de alpaca, después de eliminar los PNSs positivos comunes con hámster, fue 50 686

<sup>1</sup> Programa de Mejoramiento Animal, Universidad Nacional Agraria La Molina (UNALM), Lima, Perú

<sup>2</sup> Institute of Molecular and Cellular Biology, Novosibirsk, Russia

<sup>3</sup> Smithsonian Conservation Biology Institute, Smithsonian Institution, USA

<sup>4</sup> Department of Animal Science, University of Minnesota, USA

<sup>5</sup> E-mail: 20181395@lamolina.edu.pe

Recibido: 5 de noviembre de 2019

Aceptado para publicación: 26 de junio de 2020

Publicado: 29 de septiembre de 2020

y luego de eliminar los posibles falsos positivos quedaron 2924 PNSs. Esta cantidad se analizó con el programa Carthagene y resultó en 33 grupos de ligamiento, con un total de 216 PNSs. Las distancias calculadas entre estos PNSs oscilan entre 65.3 y 671.9 cR, con un logaritmo de probabilidades (LOD)  $\geq 6.0$ . Finalmente, se encontraron un total de 31 PNSs en el genoma referencial (E-value < 0.05).

**Palabras clave:** alpaca, genómica, PNSs, micromatriz

## ABSTRACT

The aim of this study was to develop a preliminary physical map of single nucleotide polymorphisms (SNPs) in alpaca using an alpaca/hamster radiation hybrid panel and a bovine high-density SNP genotyping microarray (BovineHD BeadChip-Illumina). The methodology included genotyping 92 alpaca/hamster hybrid cell clones, and four control samples (male alpaca, female alpaca, hamster and 1:10 DNA mixture) with the microarray. After genotyping the alpaca and hamster DNA control samples, only bovine SNPs with a call frequency of 1 were retained. The SNPs identified in the alpaca DNA samples were then filtered to remove those also found in the hamster. From the remaining alpaca SNPs, to decrease the probability of false positives, only those with a call frequency from 0.2 to 0.8 in the 92 hybrid clone samples were retained for the final analysis. The remaining alpaca specific SNPs were tabulated in MapMaker format and were analysed with the Carthagene software to identify linkage groups. The linkage groups were mapped to the reference genome *Vicugna pacos-2.0.2*, using the BLAST software and the SHORTBLAST command on the Galaxy platform. The two alpaca control samples had 294 165 SNPs with positive signals in the bovine microarray. After eliminating the most common hamster SNPs and the most-likely false positives, 2924 SNPs remained. These were assigned to 33 linkage groups encompassing 216 SNPs. The estimated distance between these SNPs ranged from 65.3 and 671.9 cR, with a logarithm of the odds (LOD)  $\geq 6.0$ . Finally, a total of 31 SNPs was found in the reference genome (E-value < 0.05).

**Key words:** alpaca, genomic, SNPs, microarray

## INTRODUCCIÓN

La crianza de alpacas es una de las actividades de mayor importancia e impacto en el desarrollo socioeconómico de la población altoandina del Perú, por su capacidad de adaptación a condiciones ambientales extremas, por ser una fuente alimenticia proteica para la población humana, y por su producción de fibra.

Los polimorfismos de nucleótido simple (PNSs) son variaciones en la secuencia de ADN que afectan una base nitrogenada. Es-

tán siendo muy utilizados como marcadores moleculares con la finalidad de investigar procesos genéticos evolutivos, estudios de asociación a caracteres de interés productivo y para incrementar el conocimiento en especies no modelo donde no se conocen detalles sobre su genoma (Seeb *et al.*, 2011).

Existen estrategias para estimar la localización y ordenamiento de marcadores moleculares a lo largo del genoma de cualquier especie, siendo una de ellas el uso de paneles de células híbridas irradiadas, cuyo fundamento es la fusión de células de la especie de interés (donadora), conteniendo frag-

mentos de ADN causados por la irradiación, con células inmortales de otra especie (receptora) que retiene los fragmentos de ADN donados. El análisis de retención de marcadores moleculares en estas células híbridas irradiadas determina el ordenamiento de marcadores moleculares en grupos ligados para la especie de interés (Faraut *et al.*, 2009). El análisis de co-segregación de PNSs en las células híbridas irradiadas se facilita mediante el genotipaje de cada uno de los clones irradiados con una micromatriz de PNSs, que permite hacer genotipados de miles de PNSs en un solo análisis (Wiggans *et al.*, 2017).

Johnson y Perelman (2007) reportaron la creación de un panel celular de radiación híbrida (5000 Rad) de una alpaca macho llamada «Limerick», conformado por 92 clones. El panel está compuesto por clones con frecuencias de retención mayores al 40%, donde se llegaron a mapear 428 marcadores en base a la homología encontrada con ZooFISH existente entre los cromosomas de dromedario, bovino y humano (Balmus *et al.*, 2007). De la misma manera, Perelman *et al.* (2018) reportaron la creación de dos paneles híbridos irradiados (5000 y 15000 Rad) para el dromedario. El primero está constituido por 93 clones y su frecuencia de retención es de 47.7%, y el segundo compuesto por 90 clones y una frecuencia de retención de 39.9%. Estos paneles de radiación híbrida constituyen un recurso importante para la construcción de mapas genómicos de mediana y alta resolución en camélidos, respectivamente.

Se cuenta con un mapa preliminar citogenético integrado del genoma de alpaca, el cual está conformado por 230 marcadores moleculares ordenados a lo largo de los 37 pares de cromosomas de la alpaca (Ávila *et al.*, 2014). Asimismo, 86 genes de alpacas han sido mapeados en el dromedario, observándose localizaciones sinténicas entre ambos, confirmando que ambas especies son genómicamente muy similares. Mendoza *et al.* (2019) localizaron seis genes candidatos para el crecimiento de fibra (*COL1A1*,

*CTNNB1*, *DAB2IP*, *KRT15*, *KRTAP13-1*, y *TNFSF12*) y cinco genes para el color de fibra (*ALX3*, *NCOA6*, *SOX9*, *ZIC1*, y *ZIC5*). También se identificaron 48 PNSs ubicados en regiones intrónicas y exónicas de 22 genes de queratina (Fernández *et al.*, 2019).

No se han diseñado micromatrices para análisis genómico en alpacas, por lo que existe la posibilidad de utilizar micromatrices ya existentes en otras especies para identificar PNSs en alpacas. Estos PNSs serían conservados entre alpaca y bovino. Estudios en venados (Haynes y Latch, 2012) y bisonte (Pertoldi *et al.*, 2010) han logrado identificar PNSs con el uso de la micromatriz de bovino, evaluando más de 50 000 PNSs (BovineSNP50 BeadChip, Illumina). Sin embargo, cuando se analizan dos especies hay que considerar las distancias evolutivas entre ellas, dado que podría existir una relación inversa entre distancia evolutiva y el número de PNSs conservados. Por lo tanto, el uso de micromatrices de bovino para la identificación de PNSs comunes con la alpaca podría ser beneficioso para incrementar el número de PNSs de alpaca.

Por este motivo, el objetivo del estudio fue identificar PNSs conservados entre bovino y alpaca y desarrollar un mapa físico preliminar de estos PNSs usando un panel celular híbrido irradiado alpaca/hámster y una micromatriz de genotipado de alta densidad de bovino.

## MATERIALES Y MÉTODOS

Se utilizó como material experimental un panel de 92 clones celulares híbridos irradiados alpaca/hámster y cuatro muestras control adicionales al panel que contienen ADN genómico de alpaca Huacaya macho (Limmerick), alpaca huacaya hembra (Carlotta), la línea celular de hámster A23 y una mezcla de alpaca macho y hámster en una proporción de 1:10, respectivamente. Estas 96 muestras fueron genotipadas con la

micromatriz de alta densidad para bovinos (BovineHD BeadChip-Illumina) que contiene 777 962 PNSs.

Los resultados del genotipado de las cuatro muestras controles fueron analizados con los programas GenomeStudio, Excel y R. El parámetro utilizado para seleccionar PNSs fue la frecuencia de señal positiva (Call Frequency) equivalente a 1. Solo se aceptaron PNSs que fueron identificados en las dos muestras de alpaca con frecuencia 1, eliminando del análisis el resto de PNSs que no cumplieron con este criterio. Las dos muestras de ADN de alpaca, la muestra de hámster y la muestra mezcla de alpaca macho/hámster 1:10 se evaluaron de forma independiente y se generaron bases de datos en formato *.txt* y *.xlsx*. La información obtenida de estas muestras sirvió para, por comparación, retener PNSs con frecuencia 1 en las muestras de ADN de alpaca y hámster, para luego eliminar los conservados comunes entre ambas muestras y solo se tabularon los PNSs únicos en alpaca.

Los resultados del genotipado de los 92 clones celulares híbridos irradiados alpaca/hámster fueron evaluados en el GenomeStudio 2011.1 (Genotyping Module v. 1.9.4, Illumina, USA), considerando la frecuencia de señal positiva  $\geq 0.20$  y  $\leq 0.80$ . Luego los PNSs comunes con los que estuvieron presentes en el ADN de alpaca y ausentes en el ADN de hámster evaluados anteriormente, fueron tabulados con su genotipo respectivo. Finalmente, como el panel fue realizado con el ADN de una alpaca, el genotipado de cada PNS en los 92 clones celulares híbridos podría ser solo homocigoto (AA o BB), o heterocigoto (AB) o ausencia de señal (no genotipado, NC); donde ningún PNS debe haber sido homocigoto para ambos alelos o la combinación de homocigotos con heterocigotos. Por esta razón, solo se consideraron los PNSs cuyos genotipos fueron similares en todas las muestras evaluadas. De esta manera se eliminaron PNSs en los que se asume señal positiva errónea (falsos positivos).

Para calcular las distancias y el orden de los PNSs encontrados previamente se utilizó el programa Carthagene (Schiex *et al.*, 2009) para lo cual se tuvo que transformar la información obtenida al formato MapMaker, codificando con el valor 1 los genotipos AA, AB y BB, los no genotipados (NC) con el valor 0, en un archivo de texto delimitado por tabulaciones. Luego en base a la localización cromosómica de PNSs bovinos, los PNS positivos de alpaca se catalogaron en 30 grupos, facilitando de esta manera el análisis de ligamiento con el programa Carthagene. Una vez identificados los grupos de PNSs ligados se crearon tablas de datos donde se registraron con su secuencia bovina total de 121 nucleótidos, ubicándose el polimorfismo en el medio. Esta información fue extraída de la base de datos de PNSs del BovineHD BeadChip usando el programa GenomeStudio.

De la misma manera, se crearon diferentes tablas que incluían la secuencia de 20 y 25 nucleótidos anteriores al nucleótido polimórfico en dirección 5' - 3', ubicándose de esta manera el polimorfismo como el nucleótido final de la secuencia. El genoma de referencia *Vicugna pacos-2.0.2* del NCBI (National Center for Biotechnology Information) fue descargado en formato FASTA del GenBank (GCF\_000164845.1) y procesado con el software REPEAT-MASKER, con la finalidad de eliminar secuencias repetitivas del genoma de la alpaca. Este último file se subió a la plataforma Galaxy (Cock *et al.*, 2015) alojado en el Minnesota Supercomputing Institute (University of Minnesota) con la finalidad de localizar los PNSs bovinos que resultaron positivos en alpaca en los *scaffolds* correspondientes al genoma de alpaca, mediante el programa BLAST y los comandos SHORT BLAST a un valor esperado (E-value) de 0.05 (NCBI, 2016).

El software Carthagene utiliza un modelo probabilístico denominado Modelo Oculto de Markov y para el cálculo de la frecuencia de retención de los PNSs evaluados utiliza

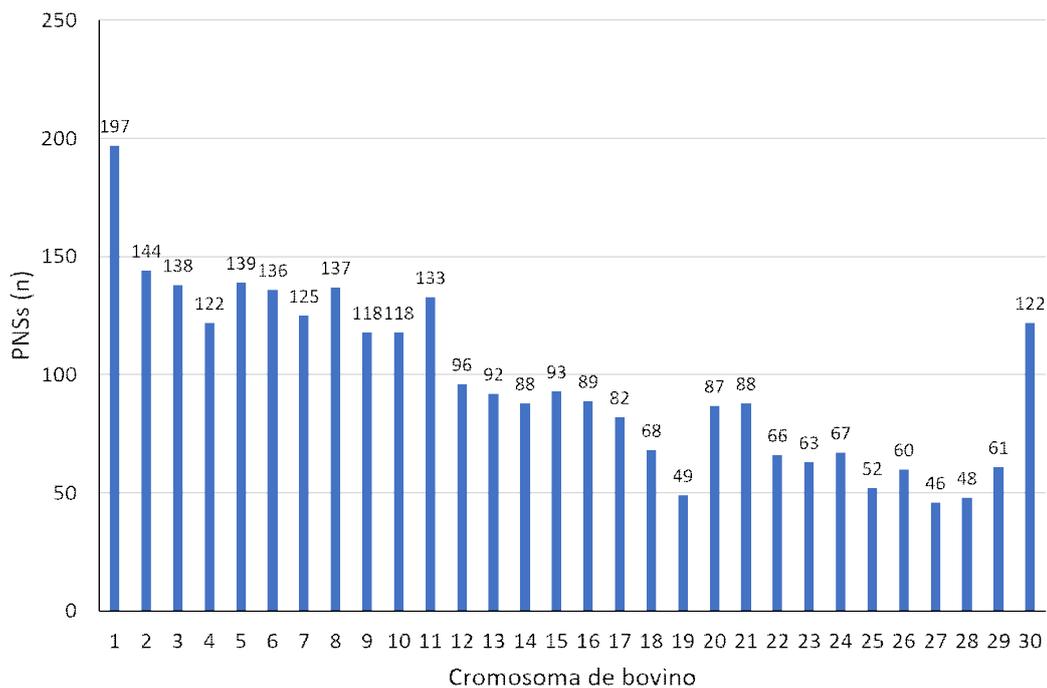


Figura 1. Distribución de polimorfismos de nucleótido simple (PNSs) conservados entre alpaca y bovino

un estimador de máxima verosimilitud cuyo algoritmo es el de esperanza-maximización (EM). Esta probabilidad calculada es transformada a un parámetro llamado LOD score (logaritmo de las probabilidades), el cual indica cuando dos loci o marcadores están ligados y se van a heredar juntos. Un LOD score de 3 indica que por cada 1000 análisis uno (1) sea diferente de lo observado al azar (De Givry *et al.*, 2004).

Para el caso de la localización de los PNSs en el genoma de la alpaca se calculó el E-value, el cual hace referencia al número de veces (probabilidad) que ocurran excelentes alineamientos comparados a que ocurran al azar en una base de datos. Las alineaciones biológicamente significativas, que indican que las secuencias de ADN son significativamente similares, tienen E-value < 1 (Pesvener, 2015).

## RESULTADOS

El esquema de análisis e identificación de PNSs que emitieron señal positiva y estuvieron presentes solo en el genoma de alpaca se realizó de la manera que fue descrito por Mamani *et al.* (2017). Brevemente, de los 777 962 PNSs analizados mediante la micromatriz de alta densidad, 294 165 PNSs fueron positivos en las dos muestras control de alpaca, lo que representa el 37.81% del total. En la muestra de hámster se identificaron 453 699 PNSs, y luego de realizar filtros y eliminar los PNSs comunes con alpaca, se encontró un total de 72 964 PNSs conservados entre la alpaca y el bovino. De la misma manera, la muestra que contenía ADN de alpaca y hámster en la proporción 1:10 se comparó con el resto de las muestras control para poder descartar PNSs falsos positivos y

Cuadro 1. Identificación de grupos de ligamiento en base a su orden, distancia, logaritmo de probabilidades y probable homología entre cromosomas de bovino y alpaca

Grupo de ligamiento	Cromosoma bovino	Probable cromosoma de alpaca	Total de PNSs	Distancia total (cR)	Promedio LOD score	Probabilidad de retención
1	1	1	18	671.9	7.03	0.47
2	2	5	4	91.6	9.07	0.49
3	2	13	7	158.41	9.55	0.52
4	3	9, 13, 21	7	285.6	6.08	0.47
5	3		4	115.4	7.73	0.45
6	4	7	6	203.5	6.88	0.49
7	5	12, 34	15	425.0	8.26	0.45
8	6	2	9	333.6	6.70	0.48
9	7	3, 22	6	202.2	6.80	0.48
10	8	4, 31	7	242.1	6.70	0.50
11	9	8	5	118.0	8.85	0.51
12	10	6, 27	8	271.1	7.10	0.48
13	11	4, 15, 28	10	355.0	7.39	0.45
14	12	14	4	108.8	7.00	0.50
15	13	19, 35	6	181.7	8.16	0.49
16	14	25, 29	6	176.1	7.30	0.50
17	15	10, 33	7	243.8	7.20	0.46
18	16	13, 23	6	156.8	8.04	0.48
19	17	2, 32	8	266.4	6.80	0.50
20	18	9	6	151.7	8.62	0.44
21	19	16	3	104.5	5.40	0.46
22	20	3	6	143.7	9.26	0.47
23	21	6, 27	7	163.1	9.48	0.50
24	22	17	5	133.2	7.95	0.44
25	23	20	7	180.3	8.35	0.52
26	24	24	4	96.2	8.83	0.44
27	24	30	4	85.9	9.37	0.49
28	25	18	6	190.0	6.90	0.47
29	26	11	6	174.7	7.60	0.50
30	27	26	8	271.7	8.11	0.45
31	28	11	3	65.3	8.15	0.48
32	29	10, 33	4	90.0	7.90	0.47
33	X	X	4	120.8	7.00	0.42
Total			216			

## Mapa de polimorfismos de nucleótido simple en alpaca

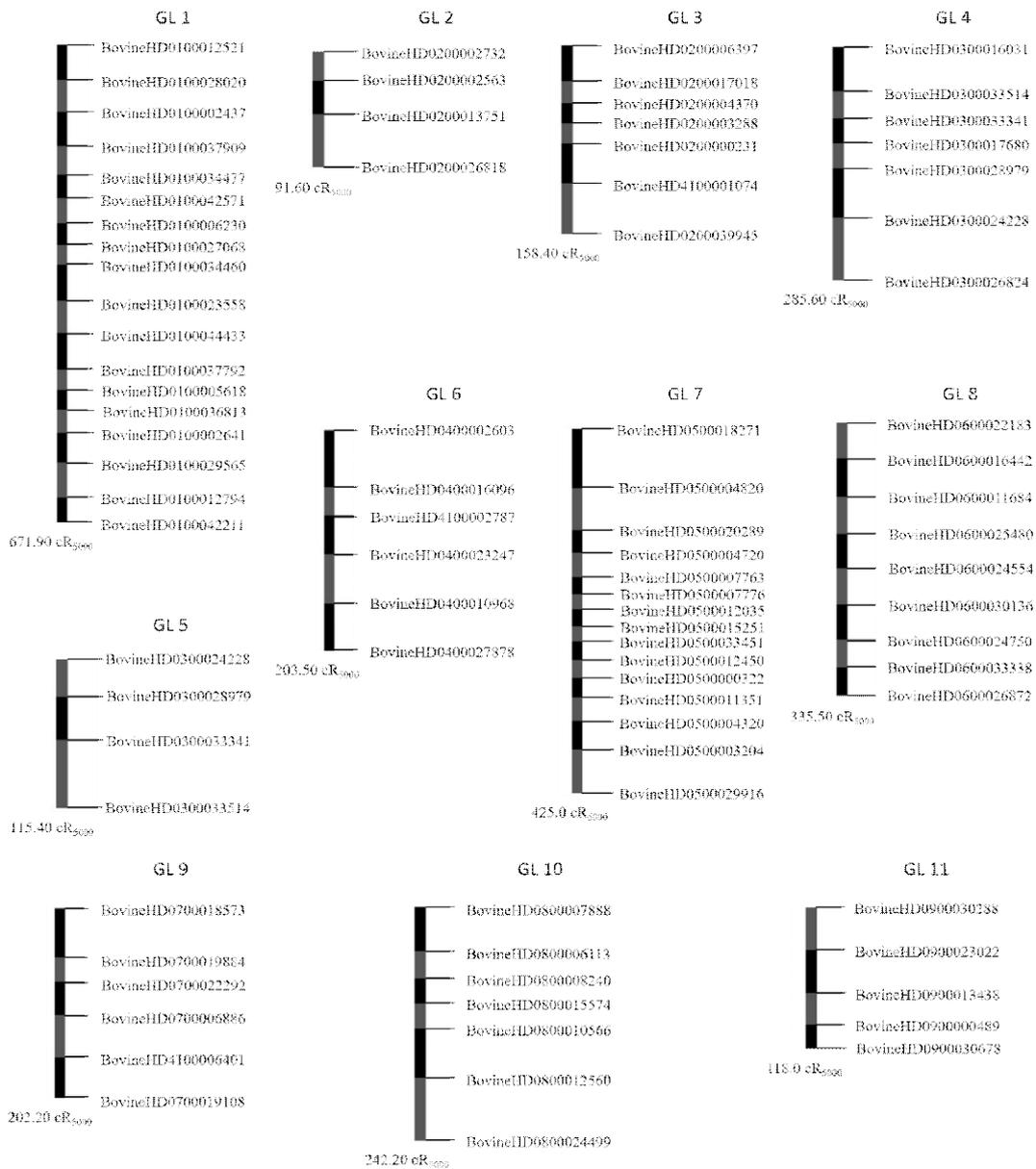


Figura 2. Mapas de referencias con los grupos de ligamiento (GL) del 1 al 11, con el orden y distancias relativas

de esta manera seleccionar con mayor asringencia PNSs específicos de alpaca. Esta última comparación identificó 50 686 PNSs, que representan el 6.5% del total.

De los 50 686 PNSs identificados, solo quedaron 2924 PNSs que cumplieron el requisito de ser AA y NC o BB y NC o AB y NC. La distribución de estos PNSs por

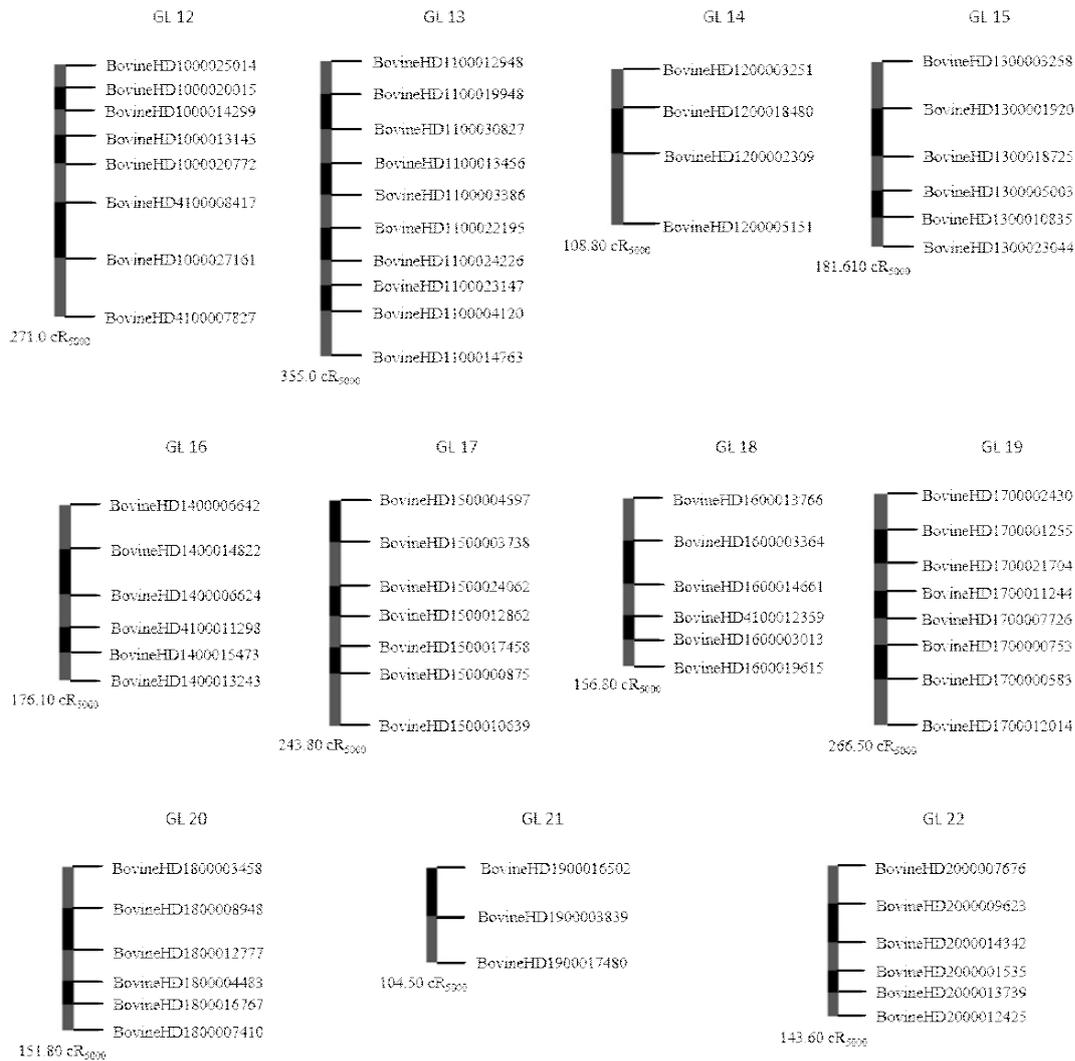


Figura 3. Mapas de referencias con los grupos de ligamiento (GL) del 12 al 22, con el orden y distancias relativas

cromosoma bovino se muestra en la Figura 1. Cada uno de estos grupos de PNSs fue evaluado con el programa Carthagene y los resultados se muestran en el Cuadro 1, donde se identifican 33 grupos ligados con un total de 216 PNSs presentes en el genoma de la alpaca. Los mapas de referencia (framework maps) generados con el orden, nombre y distancias entre PNSs se presentan en las Figuras 2, 3 y 4.

Los cuadros 2 y 3 muestran la ubicación de 23 y 8 PNS por comparación de las secuencias de 20 y 25 nucleótidos, respectivamente, en el *Vicugna pacos-2.0.2*, todos con E-value <0.05. Se consideró este valor porque es más restrictivo al momento de inferir si los alineamientos son producto del azar (Pevsener, 2015).

## Mapa de polimorfismos de nucleótido simple en alpaca

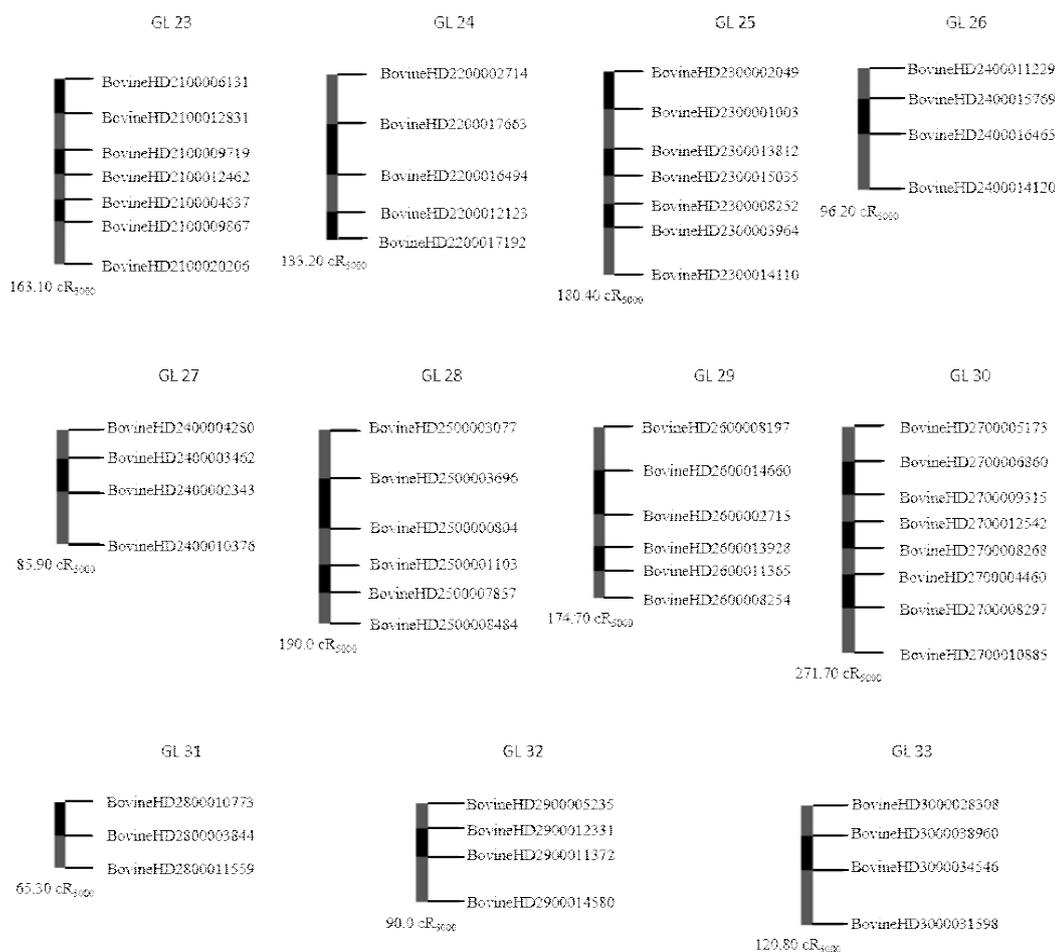


Figura 4. Mapas de referencias con los grupos de ligamiento (GL) del 23 al 33, con el orden y distancias relativas

## DISCUSIÓN

Con el fin de incrementar el número de marcadores moleculares identificados para especies con información genómica insuficiente, se han usado micromatrices de marcadores moleculares disponibles de otras especies que son evolutivamente cercanas. Por ejemplo, en venados usando la micromatriz de 50K disponible para bovinos se pudo obtener 38.7% de PNSs conservados entre estas especies (Haynes y Latch, 2012). Este resultado concuerda con el valor de 37.81% de PNSs que emitieron señal positiva, utilizando la

micromatriz de 777K, reportada en el presente estudio. Sin embargo, en otro estudio realizado en dos poblaciones de bisonte se reportó 96.7 y 98.7% de PNSs conservados usando la micromatriz bovina de 50K (Pertoldi *et al.*, 2010). Las distancias evolutivas entre el venado-bovino (Hassanin y Douzery, 2003), camélidos-bovinos (Wu *et al.*, 2014) y bisonte-bovino (MacEachern *et al.*, 2009) fueron estimadas en 25.1 a 30.1, 42.7 y 1.2 a 2.1 millones de años, respectivamente. Por lo tanto, al existir una mayor divergencia evolutiva entre camélidos y bovinos, se espera obtener una menor cantidad de PNSs identificados usando la micromatriz del bovino.

Cuadro 2. Secuencias de polimorfismos de nucleótido simple (PNSs) localizadas en el *Vicugna pacos-2.0.2* con un tamaño de 20 nucleótidos (E-value <0.05)

Grupo de ligamiento y nombre del PNS	% Identidad	E-value	Secuencia nucleotídica	Número del scaffold	Numero de accesión
3BovineHD0200017018	100 (19)	0.027	AAAGGGCAGAAAACCTCTT	49	NW_005882751.1
4BovineHD0300033341	100 (20)	0.007	CATTTGGTTTAAATTTTAT	88	NW_005882790.1
7BovineHD0500004820	100 (19)	0.027	TCACAAAGGAAGACAGTAA	235	NW_005882937.1
				255	NW_005883154.1
				452	NW_005882957.1
7BovineHD0500012035	100 (20)	0.007	AATCTGTCATCATACTAAGA	120	NW_005882822.1
8BovineHD0600024750	100 (20)	0.007	CATATGAAATCATTCTTTT	224	NW_005882926.1
9BovineHD0700019108	100 (19)	0.027	TGTTAATATCCTATAGAAA	33	NW_005882735.1
10BovineHD0800010566	100 (20)	0.007	AAAAAGCAATTGGTTATTAA	61	NW_005882763.1
10BovineHD0800012560	100 (20)	0.007	TTGAAGAATTTAGAAAGTA	8	NW_005882710.1
11BovineHD0900030678	100 (19)	0.027	CTTCCCCACCAAGGCAGA	279	NW_005882981.1
11BovineHD1000014299	100 (20)	0.007	GCTGTTGGCCAAGAGTCAA	114	NW_005882816.1
13BovineHD1100022195	100 (19)	0.027	TGATGATTTGAAAAGTTTT	29	NW_005882731.1
13BovineHD1100014763	100 (19)	0.027	TCAGTATGTTGAAGACTGA	104	NW_005882806.1
19BovineHD1700002430	100 (19)	0.027	TGTATTATGACTAAGAAAA	0	NW_005882702.1
19BovineHD1700012014	100 (20)	0.007	GCATTTAAGGAGAATTCTGT	348	NW_005883050.1
20BovineHD1800016767	100 (19)	0.027	ACAAATCAGGAGATGATTT	103	NW_005882805.1
22BovineHD2000007676	100 (19)	0.027	GATTGTTCCAGTTTTTCAT	15	NW_005882717.1
				30	NW_005882732.1
				119	NW_005882821.1
24BovineHD2200017192	100 (19)	0.027	TATCACCTCAAAAAAAAAA	5	NW_005882707.1
				60	NW_005882762.1
				66	NW_005882768.1
25BovineHD2300008252	100 (20)	0.007	TGTCACTGCTTGTGCCTCCA	18	NW_005882720.1
				77	NW_005882779.1
26BovineHD2400015769	100 (19)	0.027	TCTTTAGCAAAGATTTTCT	0	NW_005882702.1
28BovineHD2500003077	100 (19)	0.027	AAGGGGTGGGGAGGAAGAA	58	NW_005882760.1
29BovineHD2600014660	100 (20)	0.007	AAATATCAAAAATAGGCAGA	57	NW_005882759.1
31BovineHD2800011559	100 (19)	0.027	TGAATTCAGATGCTCCAGT	31	NW_005882733.1
32BovineHD2900005235	100 (20)	0.007	TGTCTGATACCAGAGCTCAT	57	NW_005882759.1

Bertolini *et al.* (2016) detectaron 10.5% PNSs en ocho alpacas usando la misma micromatriz de 777K empleada en la presente investigación, donde se detectó el 6.5% del total de PNSs analizados. Las razones para estas diferencias podrían ser porque en este estudio se utilizó el genoma de una alpaca en la construcción de los clones de células híbridas irradiadas y se excluyeron los comunes con hámster.

De los 50 686 PNSs detectados en el ADN de alpaca y ausentes en el hámster se utilizaron solo 2924 PNSs para el análisis de ligamiento. Esta disminución se debió al criterio que se impuso para incrementar la asringencia del análisis descartando los posibles falsos positivos. En el Cuadro 1 se presenta la relación de los 33 grupos de ligamiento que incluyen 216 PNSs de los 2924 resultantes del análisis con el programa

Cuadro 3. Secuencias de polimorfismos de nucleótido simple localizadas en el *Vicugna pacos-2.0.2* con un tamaño de 25 nucleótidos (E-value <0.05)

Grupo de ligamiento y nombre del PNS	% Identidad	E-value	Secuencia nucleotídica	Mismatch	N.º del scaffold	Numero de accesión
1BovineHD0100027068	95.8 (24)	0.014	GGGAACTCCA <b>A</b> G GTTAGGAGTCCA	A (Transición)	24	NW_005882726.1
1BovineHD0100023558	95.8 (24)	0.014	TTCAAAATTAGTG GGTGGTAATAT	G (Transversión)	138	NW_005882840.1
3BovineHD0200003288	96.0 (25)	0.003	ATGATGTCTCTCA ATGCCAAAGTTT	G (Transición)	293	NW_005882995.1
8BovineHD0600011684	96.0 (25)	0.003	CCTGGA <del>A</del> CTCCTA GACGGGAAACTT	C (Transversión)	80	NW_005882782.1
11BovineHD0900030288	95.8 (24)	0.014	CCAGACAGAAAA TCAATCAATAAA	G (Transición)	277	NW_005882979.1
13BovineHD1100019948	95.7 (23)	0.042	CTGCAGAAGTTC ACCAGCATGCA	G (Transición)	83	NW_005882784.1
18BovineHD1600003364	95.7 (23)	0.042	TCTCTGCTTTTA TGAGTTCGATT	A (Transición)	6	NW_005882708.1
22BovineHD2000012425	95.7 (23)	0.042	AGTGTGAGAAAA GCATGAAGCGA	G (Transversión)	23	NW_005882725.1

Letra en negrita representa al nucleótido que varió en referencia al *Vicugna pacos-2.0.2*

Carthagene. Esta disminución podría explicarse por el nivel de radiación (5000 rad) utilizado para generar el panel celular híbrido irradiado, porque a menor dosis de radiación los fragmentos de ADN que se generan son más grandes y por lo tanto el análisis de ligamiento de un número muy grande de marcadores PNSs tiene una menor resolución. La decisión de analizar grupos de PNSs por cromosoma de bovino también disminuyó el número total de PNSs que podrían ser parte de los grupos ligados.

Según los estudios realizados por Balmus *et al.* (2007), existen correspondencias cromosómicas entre el bovino y dromedario que pueden ayudar a ubicar cromosómicamente marcadores moleculares, tal como Ávila *et al.* (2014) que evidenciaron con sus estudios citogenéticos y mapeo por hibrida-

ción fluorescente *in situ* (FISH) que existe un escaso o ningún desarreglo cromosómico entre las especies de camélidos.

En el Cuadro 1 se muestra la probable homología entre cromosomas del bovino y la alpaca, donde se aprecia que el cromosoma 2 del bovino tiene sintenia compartida con los cromosomas 5 y 13 de la alpaca, el cromosoma 3 del bovino tiene sintenia compartida con los cromosomas 9, 13 y 21 de la alpaca y el cromosoma 24 del bovino tiene sintenia compartida con los cromosomas 24 y 30 de la alpaca. En tanto, estos cromosomas de alpaca han sido clasificados como subtelocéntricos o acrocéntricos (5, 9, 13), submetacéntricos (21, 24) y metacéntrico (30), por lo que explicaría la presencia de dos grupos de ligamiento en alpacas que podrían estar presentes en los brazos cromosómicos *p* y *q* respectivamente (Ramos, 2014).

Cabe resaltar que este trabajo sería el primero en identificar grupos de ligamiento usando la micromatriz de alta densidad del bovino para el genoma de alpaca. La metodología de mapeo por radiación híbrida ha permitido incrementar el número de marcadores moleculares en especies como el vacuno (Zimin *et al.*, 2009), cerdo (Servin *et al.*, 2012) y caprino (Du *et al.* 2014), debido a que en combinación con la información de secuenciamiento y otras metodologías de mapeo físico producen mayor precisión en el ensamblaje de genomas de referencia para lograr mapas de mayor resolución a nivel cromosómico. Por esta razón y considerando que en alpacas aún no existe una micromatriz de ADN para identificar PNSs, su utilidad será muy importante para complementar la información de su genoma.

Las distancias medidas en cR en los 33 grupos de ligamiento varían desde los 65.3 cR hasta los 671.9 cR. Para efectos del presente estudio son mediciones relativas entre PNSs en los grupos de ligamiento respectivos. Por lo tanto, la resolución expresada como kilobases (Kb) por cR no fue estimada en el presente trabajo. Sin embargo, en otras especies, como cerdos, se calculó una resolución de 8.6 Kb/cR y 5.3 Kb/cR usando paneles de 7,000 y 12,000 rads, respectivamente (Servin *et al.*, 2012). En el caso del caprino se estimó una resolución del mapa de 32.6 Kb/cR usando el panel de 5000 rad (Du *et al.*, 2014). Como se aprecia en los casos del porcino y caprino, a medida que se incrementa el nivel de radiación en la elaboración del panel, la resolución del mapa generado va a ser mayor, entonces la cantidad de marcadores moleculares que se podrían mapear con mejor precisión sería también mayor.

Al momento de analizar grupos de marcadores y verificar si están o no ligados se estima la estadística  $LOD \geq 3.0$  y los valores resultantes en los 33 grupos de ligamiento estuvieron por encima del  $LOD \geq 6$ , lo que indica que existen altos niveles de probabilidad de que el orden y las distancias sean correctos. En el presente estudio se decidió usar

el valor  $LOD=6$  para incrementar la astringencia de asociación entre los marcadores que componen el mapa, lo que a su vez reduce la posibilidad de incrementar el número de marcadores por grupo de ligamiento.

Los porcentajes de identidad mostrados en el Cuadro 2, producto de secuencias *query* de 20 nucleótidos, son calculados sumando las correspondencias idénticas entre secuencias alineadas y dividiendo entre el número total de bases alineadas. En el presente caso, todos los PNSs en los mapas desarrollados fueron del 100%; sin embargo, los emparejamientos idénticos (*matches*) en algunos casos fueron de 19 nucleótidos y otros de 20 (Klug *et al.*, 2016). Los valores del E-value en los 23 PNSs fueron menores a 0.05; por lo tanto, los alineamientos son significativos y tienen una mayor probabilidad que las secuencias ubicadas en el *Vicugna pacos-2.0.2* no sean productos del azar.

El Cuadro 3 muestra un total de ocho PNSs identificados en el *Vicugna pacos-2.0.2*. al analizar las secuencias *query* de 25 nucleótidos. A diferencia del Cuadro 2, los porcentajes de identidad son mayores al 95%, en donde aparecen alineamientos de 23 hasta 25 nucleótidos y la presencia de un *mismatch* (no emparejamiento) en todos los casos. En la columna de *mismatch* se considera el nucleótido que no alineó correctamente y se verifica si el cambio de base fue por transición o transversión. Los valores del E-value en los ocho PNSs fueron menores a 0.05, por lo tanto, los alineamientos son significativos.

Los resultados del presente análisis solo ubicaron 23 y 8 PNSs de los 216 PNSs en el *Vicugna pacos-2.0*, que se puede deber a una serie de factores como: la mayor divergencia evolutiva que existen entre el vacuno y la alpaca (Wu *et al.*, 2014) y que también explicaría la aparición de *mismatch* entre ambas especies, el valor más restringido del E-value que se consideró, la poca información que se tenía en cuanto al número exacto de nucleótidos que reacciona con la micro-

matriz y hace positiva su aparición en los resultados, y a la cobertura de 22X que no cubre el genoma completo de la alpaca.

## CONCLUSIONES

Se identificaron 50 686 polimorfismos de nucleótido simple utilizando la micromatriz de alta densidad del bovino y con un nivel más alto de astringencia, solo 2924 se estimaron como verdaderos PNSs de alpacas. Con este último dato se desarrolló un primer mapa referencial que resultó en 33 grupos de ligamiento que incluyen 216 PNSs, de los cuales fueron ubicados 31 PNSs en el genoma referencial evaluado.

## Agradecimientos

El presente trabajo fue financiado por el Proyecto N.º 125-2015-FONDECYT CONCYTEC. La Alpaca Research Foundation financió y brindó el panel celular híbrido irradiado alpaca/hámster usado en este estudio. PLM fue financiada por la Russian Science Foundation grant RSF 19-14-00034.

## LITERATURA CITADA

1. **Avila F, Baily MP, Perelman P, Das PJ, Pontius J, Chowdhary R, Owens E, et al. 2014.** A comprehensive whole-genome integrated cytogenetic map for the alpaca (*Lama pacos*). *Cytogenet Genome Res* 144: 193-204. doi: 10.1159/000370329.
2. **Balmus G, Trifonov VA, Biltueva LS, O'Brien PC, Alkalaeva ES, Fu B, Skidmore JA, Allen T, et al. 2007.** Cross-species chromosome painting among camel, cattle, pig and human: further insights into the putative Cetartiodactyla ancestral karyotype. *Chromosome Res* 15: 499-515. doi: 10.1007/s10577-007-1154-x
3. **Bertolini F, Elbeltagy A, Ponce de Leon FA, Gutierrez GA, Rothchild M. 2016.** P4013 Applicability of using bovine, ovine and caprine SNP chips for alpaca and dromedary genomic studies. *J Anim Sci* 94: 85. doi: 10.2527/jas2016.94supplement485x
4. **Cock PJ, Chilton JM, Gruning B, Johnson JE, Soranzo N. 2015.** NCBI Blast+ integrated into galaxy. *GigaScience* 4: 1-9. doi: 10.1186/s13742-015-0080-7
5. **De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T. 2005.** Carthagene: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* 21: 1703-1704. doi: 10.1093/bioinformatics/bti222
6. **Du X, Servin B, Womack JE, Cao J, Yu M, Dong Y, Wang W, Zhao S. 2014.** An update of the goat genome assembly using dense radiation hybrid maps allows detailed analysis of evolutionary rearrangements in Bovidae. *BMC Genomics* 15: 1-16. doi: 10.1186/1471-2164-15-625
7. **Faraut T, de Givry S, Hitte C, Lahbib-Mansais Y, Morisson M, Milan D, Schiex T, et al. 2009.** Contribution of radiation hybrids to genome mapping in domestic animals. *Cytogenet Genome Res* 126: 21-33. doi: 10.1159/000245904
8. **Fernandez AG, Gutierrez GA, Ponce de Leon FA. 2019.** Identificación bioinformática de polimorfismos de nucleótido simple (PNSs) en genes candidatos para las características de la fibra en alpacas (*Vicugna pacos*). *Rev Peru Biol* 26: 087-094. doi: 10.15381/rpb.v26i1.15911
9. **Hassanin A, Douzery EJ. 2003.** Molecular and morphological phylogenies of ruminantia and the alternative position of the moschidae. *Syst Biol* 52: 206-228. doi:10.1080/10635150390192726
10. **Haynes GD, Latch EK. 2012.** Identification of novel single nucleotide polymorphisms (SNPs) in deer (*Odocoileus* spp) using the BovineSNP50 BeadChip. *PLoS One* 7: e36536. doi: 10.1371/journal.pone.0036536

11. **Johnson W, Perelman P. 2007.** Development of a radiation-hybrid map. *Alpacas Magazine*. 234-239. [Internet]. Available in: [http://www.alpaca-research.org/resources/dyn/files/319762zc142794f\\_fn/the-alpaca-enters-the-genomic-era.pdf](http://www.alpaca-research.org/resources/dyn/files/319762zc142794f_fn/the-alpaca-enters-the-genomic-era.pdf)
12. **Klug W, Cummings M, Spencer C, Palladino M. 2016.** Concepts of genetics. 11<sup>th</sup> ed. UK: Pearson Education. 889 p.
13. **MacEachern S, McEwan J, Goddard M. 2009.** Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae). *BMC Genomics* 177 doi: 10.1186/1471-2164-10-177
14. **Mamani C, Gutierrez G, Ponce de Leon A. 2017.** Identificación de polimorfismos de nucleótido simple en alpaca (*Vicugna pacos*) usando un panel de células híbridas irradiadas alpaca/hámster. *RICBA* 1: 92-95. doi: 10.25127/ricba-20171.242.
15. **Mendoza MN, Raudsepp T, Alshanbari F, Gutierrez G, Ponce de Leon FA. 2019.** Chromosomal localization of candidate genes for fiber growth and color in alpaca (*Vicugna pacos*). *Front Genet* 10: 583. doi: 10.3389/fgene.2019.00583
16. **NCBI Resource Coordinators. 2016.** Database resources of the national center for biotechnology information. *Nucleic Acids Res* 44: D7-D19 doi: 10.1093/nar/gkv1290
17. **Perelman PL, Pichler R, Gaggi A, Larkin DM, Raudsepp T, Alshanbari F, Holl HM, et al. 2018.** Construction of two whole genome radiation hybrid panels for dromedary (*Camelus dromedarius*): 5000 RAD and 15000 RAD. *Sci Rep* 8: 1982. doi: 10.1038/s41598-018-20223-5
18. **Pertoldi C, Wókcik JM, Tokarska M, Kawalko A, Kristensen TN, Loeschke V, Gregersen VR, et al. 2010.** Phylogenetic relationships among the European and American bison and seven cattle breeds reconstructed using the BovineSNP50 Illumina Genotyping BeadChip. *Conserv Genet* 11: 627-634. doi: 10.4098/j.at.0001-7051.002.2010
19. **Pesvener J. 2015.** Bioinformatics and functional genomics. 3<sup>rd</sup> ed. USA: Wiley Blackwell. 1124 p.
20. **Ramos M. 2014.** Descripción del patrón de diferenciación longitudinal de los cromosomas de alpacas y llamas. Tesis de Maestría. Lima, Perú. Univ. Nacional Mayor de San Marcos. 110 p.
21. **Schiex T, De Givry S, Chabrier P, Bouchez M. 2009.** Carthagene-User documentation. INRA. 192 p. [Internet]. Available in: <https://www.animal-genome.org/bioinfo/resources/manuals/carthagene/>
22. **Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb W. 2011.** Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Mol Ecol Resour* 11(Suppl 1): 1-8. doi: 10.1111/j.1755-0998.2010.02979.x
23. **Servin B, Faraut T, Lannuccelli N, Zelenika D, Milan D. 2012.** High-resolution autosomal radiation hybrid maps of the pig genome and their contribution to the genome sequence assembly. *BMC Genomics* 13: 585. doi: 10.1186/1471-2164-13-585
24. **Wiggans G, Cole J, Hubbard S, Sontegard T. 2017.** Genomic selection in dairy cattle: The USDA experience. *Annu Rev Anim Biosci*. 5: 309-327. doi: 10.1146/annurev-animal-021815-111422
25. **Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, Zhang L, et al. 2014.** Camelid genomes reveal evolution and adaptation to desert environments. *Nat Commun* 5: 5188. doi: 10.1038/ncomms6188
26. **Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, Hanrahan F, et al. 2009.** A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol* 10: R42. doi: 10.1186/gb-2009-10-4-r42