

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

**ESCUELA DE POSGRADO
MAESTRÍA EN PRODUCCIÓN ANIMAL**



**“MAPA FÍSICO DE POLIMORFISMOS DE NUCLEÓTIDO
SIMPLE EN ALPACA (*Vicugna pacos*) USANDO UN PANEL DE
CÉLULAS HÍBRIDAS IRRADIADAS ALPACA/HÁMSTER”**

Presentada por:

CAMILO VICENTE MAMANI MONDRAGÓN

**TESIS PARA OPTAR EL GRADO DE MAESTRO
MAGISTER SCIENTIAE EN PRODUCCIÓN ANIMAL**

Lima – Perú

2018

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

**ESCUELA DE POSGRADO
MAESTRÍA EN PRODUCCIÓN ANIMAL**

**“MAPA FÍSICO DE POLIMORFISMOS DE NUCLEÓTIDO SIMPLE EN
ALPACA (*Vicugna pacos*) USANDO UN PANEL DE CÉLULAS
HÍBRIDAS IRRADIADAS ALPACA/HÁMSTER”**

**TESIS PARA OPTAR EL GRADO DE MAESTRO
MAGISTER SCIENTIAE**

Presentado por:

CAMILO VICENTE MAMANI MONDRAGÓN

Sustentada y aprobada ante el siguiente jurado:

Mg.Sc. María Elisa García Salas
PRESIDENTE

Ph.D. Gustavo Augusto Gutiérrez Reynoso
PATROCINADOR

Ph.D. Federico Abel Ponce de León Bravo
CO-PATROCINADOR

Ph.D. Juan Francisco Chávez Cossío
MIEMBRO

Mg.Sc. Enrique Alvarado Malca
MIEMBRO

DEDICATORIA

A Dios todopoderoso, por iluminarnos diariamente por el camino de la sabiduría y el éxito. Al Señor de los Milagros que en momentos difíciles supo darme la tranquilidad para seguir adelante.

A mis padres, Vicente Mamani y Blanca Mondragón, por su incondicional y abnegado esfuerzo para brindarme todos los medios necesarios para ser una persona responsable en la vida.

A toda la familia Mondragón y Mamani, con quienes comparto siempre momentos alegres en cada reunión familiar y permiten que me olvide por un instante de los estudios, que muchas veces era estresante.

No podía olvidarme de mis tres hermanos políticos, Luis Alonso, Luis David, Víctor. Gracias por estar presente cuando necesitaba hablar con alguien y que parte de la culminación de mi maestría es debido al apoyo brindado en su momento.

AGRADECIMIENTO

- Al Ph.D. Abel Ponce de León y al Ph.D. Gustavo Gutiérrez por la confianza brindada en la presente investigación y valoro mucho las constantes correcciones y consejos para crecer cada día como profesional e investigador.
- Al programa de becas FONDECYT, la cual subvencionó mis estudios de maestría en Producción Animal.
- Al proyecto de investigación “Establecimiento de un Mapa Físico Preliminar de Marcadores Moleculares de Polimorfismo de Nucleótido simple (SNP) en Alpaca (*Vicugna pacos*) en base a Información obtenida con un chip de Bovinos (Bovine HD Genotyping Beadchip)" - Convenio de Subvención N° 125 – 2015 - FONDECYT CONCYTEC, por brindar la mayor parte del financiamiento.
- Al Proyecto “Investigación en Sistemas Agrarios”, Subproyecto “Mejorando los sistemas de producción de alpacas en pastizales de la Sierra Central del Perú” del Convenio VLIR-UNALM.
- A University of Minnesota, por brindar acceso a la plataforma Galaxy, bajo la supervisión del Ph.D. Abel Ponce de León.
- Al Smithsonian Conservation Biology Institute, por la colaboración con la remisión de las muestras del panel celular híbrido.
- Al Programa de Ovinos y Camélidos Sudamericanos (POCA) y al Programa de Mejoramiento Animal (PMA) de la Facultad de Zootecnia por las facilidades brindadas en el uso de sus instalaciones.
- A la Universidad Nacional Agraria La Molina (UNALM), por darme la oportunidad de conocer a un gran grupo humano, involucrados en diversos temas de investigación.

RESUMEN

El objetivo del presente trabajo de investigación fue desarrollar un mapa físico 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. Los objetivos específicos fueron: a) identificar los PNSs de la micromatriz de alta densidad de bovinos detectados solo en ADN de alpaca y ausentes en el ADN de hámster; b) determinar las distancias entre grupos de PNSs coheredados en el panel de células híbridas irradiadas alpaca/hámster; y, c) determinar la localización cromosómica de los grupos ligados de PNSs. Se genotiparon las 96 muestras de ADN presentes en el panel (92 clones celulares híbridos alpaca/hámster, 2 muestras de alpaca, 1 muestra de hámster, 1:10 muestra alpaca/hámster) evaluadas con los programas GenomeStudio, R y Excel. Los PNSs identificados con la micromatriz en el ADN de alpaca se analizaron con el programa Carthagene para identificar los grupos de ligamiento. Los grupos ligados fueron ubicados en la base de datos *VicPac 2.0.2*, aplicando el programa Blast y el comando ShortBlast, en la plataforma Galaxy. Se identificó un 50,686 PNSs del total de polimorfismos de nucleótido simple, analizados en la micromatriz de bovino, presentes en el genoma de la alpaca y ausentes en el genoma del hámster; se determinaron 33 grupos de ligamiento, con un total de 216 PNSs, con distancias calculadas que oscilaron entre 65.3 y 671.9 cR, con un LOD>3.0; y, se halló la secuencia de 31 PNSs en el *VicPac 2.0.2*, y se infirió la localización de tres PNSs (1BovineHD0100027068, 8BovineHD0600011684 y 13BovineHD1100019948) en los cromosomas 1, 2 y 4 de alpaca.

Palabras clave: Alpaca, PNS, híbrido irradiado, micromatriz.

SUMMARY

The aim of this research study was to develop a physical map of single nucleotide polymorphisms (SNPs) in alpaca using the alpaca/hamster radiation hybrid panel and a Bovine high-density SNP genotyping microarray. The specific objectives were: a) to identify SNPs on the bovine microarray present only in alpaca DNA and absent in hamster DNA; b) to determine the distances between SNP groups co-inherited in the alpaca/hamster radiation hybrid panel and; c) to determine the chromosomal location of the SNP linked groups. The 96 DNA samples present in the panel were genotyped (92 hybrid cells clones, 2 alpaca samples, 1 hamster sample, 1:10 alpaca/hamster sample) and were evaluated with the GenomeStudio, R and Excel programs. The SNPs found in alpacas were analyzed with the Carthagene software to identify linkage groups. These linked groups were located in the *VicPac 2.0.2* database using the Blast software and the ShortBlast command with the help of the Galaxy platform. A total of 50,686 SNPs from the bovine microarray were detected in the alpaca genome and absent in the hamster genome. Thirty-three linkage groups that include 216 SNPs were identified. Their calculated distances that ranged between 65.3 and 671.9 cR, with a LOD > 3.0. The sequence of 31 of these SNPs was found in the *VicPac 2.0.2* and the location of 3 SNPs (1BovineHD0100027068, 8BovineHD0600011684 and 13BovineHD1100019948) was inferred to be on alpaca chromosomes 1, 2 and 4.

Key words: Alpaca, SNP, radiation hybrid, microarray.

ÍNDICE GENERAL

I.	INTRODUCCIÓN.....	1
II.	REVISIÓN DE LITERATURA	3
2.1.	CLASIFICACIÓN TAXONÓMICA DE LA ALPACA	3
2.2.	GENOMA DE LA ALPACA	4
2.3.	MAPAS GENÓMICOS	6
2.3.1.	Mapas genéticos	7
2.3.2.	Mapas físicos.....	7
–	Híbridos por radiación	8
2.4.	POLIMORFISMOS DE NUCLEÓTIDO SIMPLE.....	11
2.5.	MICROMATRICES DE ADN	13
III.	MATERIALES Y MÉTODOS.....	16
3.1.	LUGAR DE ESTUDIO	16
3.2.	MATERIAL EXPERIMENTAL	16
3.3.	METODOLOGÍA	17
3.3.1.	Genotipado de polimorfismos de nucleótido simple.....	17
3.3.2.	Identificación de polimorfismos de nucleótido simple.....	17
3.3.3.	Determinación de grupos ligados de polimorfismos de nucleótido simple .	18
3.3.4.	Localización de grupos ligados de polimorfismos de nucleótido simple.....	20
3.4.	ANÁLISIS ESTADÍSTICO	21
IV.	RESULTADOS Y DISCUSIÓN	23
4.1.	POLIMORFISMOS DE NUCLEÓTIDO SIMPLE.....	23
4.2.	DETERMINACIÓN DE GRUPOS LIGADOS	25
4.3.	LOCALIZACIÓN DE GRUPOS LIGADOS	30
V.	CONCLUSIONES.....	38
VI.	RECOMENDACIONES	39
VII.	REFERENCIAS BIBLIOGRÁFICAS	40
VIII.	ANEXOS.....	47

ÍNDICE DE TABLAS

Tabla 1: Paneles celulares híbridos irradiados en animales domésticos	10
Tabla 2: Relación entre el E-value y el p-value.....	21
Tabla 3: Total de grupos ligados	26
Tabla 4: Secuencias de polimorfismos de nucleótido simple localizadas en el <i>VicPac 2.0.2</i> con un tamaño de 20 nucleótidos	33
Tabla 5: Secuencias de polimorfismos de nucleótido simple localizadas en el <i>VicPac 2.0.2</i> con un tamaño de 25 nucleótidos	37

ÍNDICE DE FIGURAS

Figura 1: Relación filogenética entre bovinos y camélidos.....	4
Figura 2: Evento de recombinación entre dos alelos.....	7
Figura 3: Representación de la construcción del panel celular híbrido irradiado	9
Figura 4: Polimorfismo de nucleótido simple	12
Figura 5: Genotipado del genoma con una micromatriz	14
Figura 6: Distribución de las 96 muestras del panel.....	16
Figura 7: Identificación de polimorfismos de nucleótido simple en las cuatro muestras controles	23

ÍNDICE DE ANEXOS

Anexo 1. Comandos utilizados en R	47
Anexo 2. Formato Mapmaker para encontrar grupos de ligamiento	48
Anexo 3. Correspondencia cromosómica Bovino-Dromedario	49
Anexo 4. Relación del cromosoma con scaffolds del <i>VicPac 2.0.2</i>	51
Anexo 5. Gráficos de los grupos de ligamiento.....	54
Anexo 6. Glosario.....	57

I. INTRODUCCIÓN

La crianza de alpacas es una de las actividades de mayor importancia e impacto en el desarrollo socioeconómico de la población altoandina de nuestro país, no solo por su capacidad de adaptación a condiciones ambientales extremas y como fuente alimenticia proteica para la población humana, sino también porque es un recurso para la producción de fibra de buena calidad.

En el transcurso de los años se han venido implementando planes de mejoramiento genético de alpacas en distintos departamentos del Perú como son: Puno, Cusco, Arequipa, Huancavelica, etc., con la finalidad de mejorar la producción de fibra y poder aumentar los ingresos de los pequeños productores (Renieri *et al.* 2007). Sin embargo, se observan algunas limitaciones como, el manejo incorrecto de registros productivos y la baja capacidad organizacional de los productores. Esta situación y el hecho de que la mejora genética representa cambios acumulativos que toman años para observar su efecto en la economía del criador, hace que muchos de ellos no se interesen en aplicar las diferentes estrategias en la mejora genética de sus rebaños.

Para sostener y mejorar la producción de fibra de forma eficiente en la crianza de alpacas es importante identificar genes y marcadores moleculares asociados a caracteres de interés ganadero, con fines de tener mayor precisión en la selección de los mejores animales.

Los avances logrados en el mejoramiento genético en animales de granja con la aplicación de la genética molecular, específicamente su aporte al conocimiento de marcadores moleculares están generando mejoras en la aplicación de técnicas y estrategias para obtener un mejor conocimiento del genoma y el mejoramiento de la alpaca. Un marcador molecular como el polimorfismo de nucleótido simple (PNS), muy utilizado en el ganado bovino de leche, se viene aplicando para el conocimiento de la información genotípica de diversas especies.

Existen estrategias para estimar distancias entre marcadores moleculares y ordenarlos 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 fragmentos 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 y el desarrollo de mapas físicos de marcadores moleculares para la especie de interés. El ordenamiento se realiza con programas estadísticos que estiman las distancias físicas entre los marcadores analizados.

Otra estrategia empleada en animales de granja son las micromatrices de alta densidad para analizar PNSs. La ventaja de esta técnica molecular es la rapidez con la cual en un solo análisis se pueden determinar miles de PNSs. En bovinos existen varias micromatrices comerciales, con diferentes densidades, las cuales están siendo utilizadas para programas de selección genómica.

Hasta el momento la información genómica en los camélidos sudamericanos es limitada; sin embargo, poco a poco se está adaptando metodologías empleadas en otros animales de granja (bovino, ovino, equino) para encontrar polimorfismos, dentro de secuencias nucleotídicas, y realizar comparaciones con otras especies relativamente cercanas como son los camellos. Por lo tanto, el uso de paneles celulares híbridos irradiados empleados en el genoma humano y otros animales de granja, aportaría al conocimiento del genoma de la alpaca, incrementado su información genómica, que podría ser utilizada en estudios de asociación a caracteres productivos.

Por tal motivo, el objetivo general del presente estudio fue desarrollar un mapa físico de marcadores PNSs en alpaca usando un panel celular híbrido irradiado alpaca/hámster y micromatriz de alta densidad del bovino. Los objetivos específicos fueron: 1) identificar los PNSs de la micromatriz de alta densidad de bovinos presentes solo en ADN de alpaca y ausentes en ADN de hámster, 2) determinar las distancias entre grupos de PNSs coheredados en el panel de células híbridas irradiadas alpaca/hámster; y, 3) determinar la localización cromosómica de los grupos ligados de PNSs.

II. REVISIÓN DE LITERATURA

2.1. CLASIFICACIÓN TAXONÓMICA DE LA ALPACA

La alpaca es uno de los cuatro camélidos sudamericanos reconocidos actualmente y pertenece al grupo de animales domesticados junto a la llama, ambos llamados también camélidos del nuevo mundo. Existen 2 variedades fenotípicas bien definidas: Huacaya y Suri.

El origen de la alpaca hasta el siglo XX se consideraba que pertenecía al género *Lama*, debido a hallazgos encontrados en restos zooarqueológicos hace 6000 – 7000 años y a la alta tasa de hibridación existente entre la alpaca y llama; sin embargo, con el auge de la era genómica, estudios cromosómicos de bandeado G y análisis de ADN nuclear y mitocondrial han servido para sustentar que la alpaca es el resultado de la domesticación de la vicuña, debido a su cercanía filogenética encontrada en estas especies domésticas (Kadwell *et al.* 2001; Marín *et al.* 2007). Por este motivo, hasta el momento la alpaca es taxonómicamente ubicada en el Orden *Artiodactyla*, Suborden *Tylopoda*, Familia *Camelidae* y al Género y especie *Vicugna pacos* (Fowler, 2010).

Adicionalmente se sabe que las alpacas, camellos y dromedarios pertenecen a la misma familia, pero divergieron en la tribu Lamini (camélidos del nuevo mundo) y la tribu Camelini (camélidos del viejo mundo), respectivamente. La importancia de su cercanía filogenética radica en que se pueden realizar estudios evolutivos comparativos entre la familia *Camelidae* y evaluar genes, marcadores moleculares o regiones genómicas sinténicas (Ávila, 2014). Asimismo, también se puede resaltar la relación que existe entre los bovinos y camélidos, ambos pertenecientes al orden *Artiodactyla* (Figura 1) y se pueden realizar comparaciones entre sus genomas, encontrándose regiones conservadas que nos pueden servir para el mapeo de marcadores moleculares o genes de importancia económica.

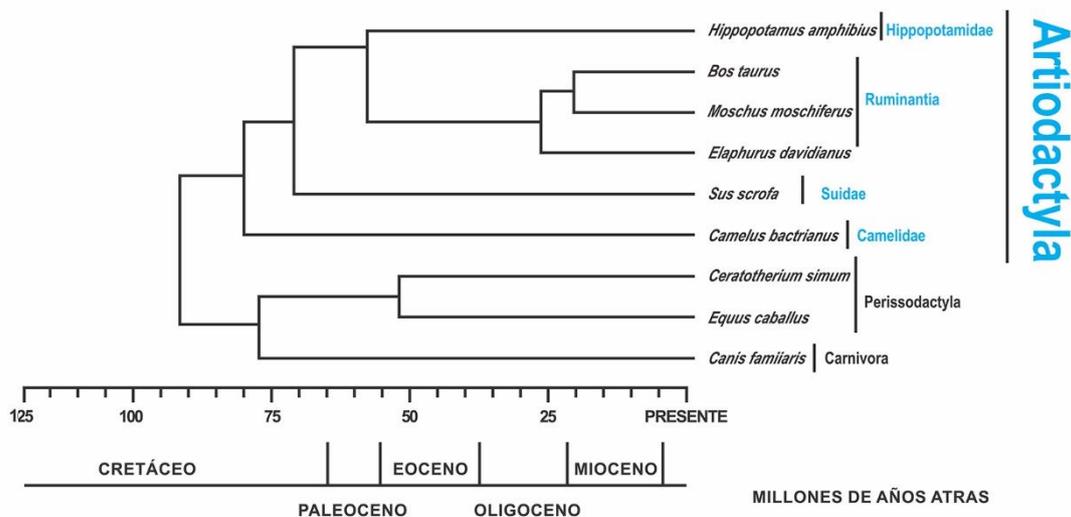


Figura 1: Relación filogenética entre bovinos y camélidos

FUENTE: Zhou *et al.* 2011

2.2. GENOMA DE LA ALPACA

El genoma se define como la totalidad de ADN (secuencia de nucleótidos) que está presente dentro de cada célula nucleada incluyendo el ADN mitocondrial de un organismo. Los fragmentos de ADN que contienen información genética para la formación de diferentes variedades de ARN, se denominan genes. En particular, el ARNm va ser el encargado de, por un proceso de traducción, la formación de un conjunto de polipéptidos para que finalmente conforme las proteínas (Griffiths *et al.* 2015).

El análisis del genoma (mapa físico y citogenética molecular) de la alpaca ha progresado lentamente en comparación con otras especies domésticas. Estudios citogenéticos llegaron a definir el cariotipo de la alpaca, el cual está conformado por $2n = 74$ cromosomas, siendo similar a otras especies de la familia *Camelidae* (Taylor *et al.* 1968). La aplicación de técnicas por bandeos cromosómicos G, C y NOR determinaron que los pares cromosómicos autosómicos del 1 al 20 son de tipo acrocéntrico-subterminal y del 21 al 36 son metacéntrico- submetacéntrico. Los cromosomas sexuales X e Y son de tipo metacéntrico y acrocéntrico, respectivamente (Bianchi *et al.* 1985). Sin embargo, otro estudio más reciente usando técnicas de bandeo cromosómico con bandas R, C y coloración con Giemsa se observaron 18 pares cromosómicos subtelo-céntricos, 10 pares submetacéntricos y 9 metacéntricos incluyendo el par sexual (Ramos, 2014).

El proyecto del genoma de la alpaca se inició en el año 2002, con el secuenciamiento y el ensamblado a una profundidad de 2X. Esta información fue utilizada principalmente para realizar genómica comparativa con el genoma humano. Adicionalmente, se tuvo en cuenta la familia *Camelidae* con la finalidad de incrementar el número de especies con un genoma disponible para estudios evolutivos y comparativos (Raudsepp, 2012); en el año 2008, estuvo disponible el ensamblado *VicPac 1.0*, secuenciado a 2.5X de una alpaca hembra llamada Carlotta con la tecnología de Sanger; el año 2011 se lanzó el ensamblado a 22X (NCBI, *VicPac 2.0.2*) realizado por la universidad de Washington mediante la combinación del secuenciamiento realizado en el 2008 y la adición de la tecnología de secuenciamiento de segunda generación (NGS). Este ensamblado cuenta con 276,726 scaffolds y un total de 2171.21 Mb; en el año 2014 se lanzó el ensamblado a 72.5X (NCBI, *Vi_Pacos 1.0*), elaborado por la Universidad de Mongolia (IMAU) y BGI (China) utilizando la plataforma HiSeq 2000 de Illumina; el ensamblado cuenta con 52,275 scaffolds y un total de 2013.70 Mb (Wu *et al.* 2014). Adicionalmente se cuenta con una biblioteca BAC CHORI-246, realizada con la alpaca Carlotta y está compuesta por un total de 202,752 clones (Ávila, 2014).

Para un mayor conocimiento del genoma se desarrollaron diferentes metodologías para encontrar genes asociados a caracteres productivos o enfermedades en diferentes especies, conocidos como marcadores moleculares. Estos hacen referencia a una secuencia de ADN utilizada para la ubicación de una región de interés dentro del genoma. Permite identificar diferencias (polimorfismos) a nivel de ADN entre miembros de la misma especie, y que pueden estar ubicados en regiones codificantes o no codificantes. Entre los marcadores moleculares que han sido utilizados son: Polimorfismos de longitud en fragmentos de restricción (RFLP), microsatélites y actualmente los polimorfismos de nucleótido simple (Al-Samarai y Al-Kazaz, 2015).

Los microsatélites son marcadores moleculares que nos permiten detectar y/o diferenciar genotipos. Un estudio efectuado en alpacas para encontrar marcadores asociados al fenotipo de ojo zarco observó que dos alelos relacionados al gen *KIT* (*hew1* y *hew2*) están asociados significativamente, en desequilibrio de ligamiento, con el fenotipo zarco (Jackling *et al.* 2012). También se buscó asociar los microsatélites con la característica diámetro de fibra en 140 alpacas (razas huacaya y suri). Se evaluaron 69 marcadores microsatélites y 599 alelos, y se identificaron cuatro microsatélites (LCA68, VOLP59, LCA90 y GLM6) y 11 alelos asociados significativamente con valores de cría para el

carácter diámetro de fibra (Paredes *et al.* 2014). Asimismo, los microsatélites se usaron en detección de tráfico ilegal de camélidos sudamericanos y diferenciarlos entre especies domésticas y silvestres (Di Rocco *et al.* 2011). Para la determinación de paternidad se han validado 15 marcadores microsatélites en alpacas de la región de Pasco, donde se obtuvo un 66.4 por ciento de 116 casos de verificación de paternidad en concordancia con los registros obtenidos en campo (Morón, 2016).

Además del diámetro de fibra, se vienen realizando investigaciones del color de fibra de alpacas, debido al aumento de la demanda por parte de la industria por fibra de colores naturales. La caracterización del gen *MC1R*, sus transcritos y el polimorfismo de nucleótido simple asociado con el color de fibra blanca, se encontró la relación del transcrito *MC1R-F1* con los colores blancos, negro y café; mientras el transcrito *MC1R-F2* fue específico para el color blanco. Además, al secuenciar el gen *MC1R* en alpacas con fibra blanca y de color se identificó un total de 12 polimorfismos de nucleótido simple, de los cuales nueve se encontraron en zonas codificantes y tres en zonas no codificantes (Chandramohan *et al.* 2015).

Los avances en las distintas técnicas moleculares, como el cultivo celular, permitió elaborar el primer mapa citogenético integrado del genoma completo de alpaca, mediante el uso de técnicas como la Hibridización fluorescente *in situ* (FISH) y Cromosoma artificial bacteriano (BAC). El mapa está conformado por 230 marcadores moleculares ordenados a lo largo de los 37 pares de cromosomas. Adicionalmente, 86 genes de alpacas fueron mapeados en el dromedario, observándose secuencias sinténicas entre ambos, afirmando que ambas especies son muy similares en su genoma (Ávila *et al.* 2015).

2.3. MAPAS GENÓMICOS

La importancia del mapeo del genoma es complementaria una vez que se tiene la secuencia completa del genoma de una especie, con la finalidad de ubicar puntos de referencia dentro de la secuencia del ADN y poder asignarlos a los cromosomas. Existen diversas técnicas para mapear el genoma de cualquier especie y se clasifican principalmente en dos tipos de mapas:

2.3.1. Mapas genéticos

El mapeo genético se basa en estimar la posición relativa de genes o marcadores moleculares a lo largo de los cromosomas. El fundamento de los mapas genéticos o ligamiento (Figura 2) está en la meiosis del ciclo celular donde los loci ubicados en diferentes cromosomas se segregan de manera independiente en los gametos, mientras los que se encuentran en un mismo cromosoma y están cercanos tienden a co-segregarse. La probabilidad que ocurra un evento de recombinación entre 2 genes depende de la distancia entre ellos y su unidad de medida es el centiMorgan (cM) que corresponde al uno por ciento de probabilidad de producir un gameto recombinante. La limitación que se presenta para desarrollar estos mapas de alta resolución en animales de granja es que se necesita una gran cantidad de cruza experimentales, los cuales son muy fáciles en microorganismos, y obtener una gran cantidad de progenie (Griffiths *et al.* 2015). Actualmente no se cuenta con un mapa genético en alpacas.

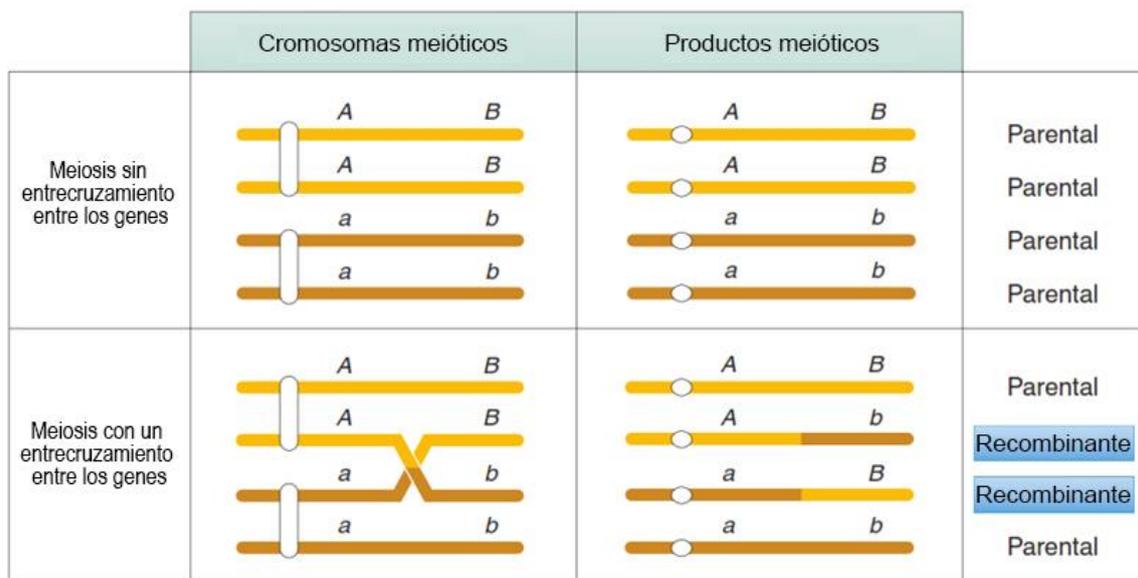


Figura 2: Evento de recombinación entre dos alelos

FUENTE: Griffiths *et al.* 2015

2.3.2. Mapas físicos

El mapeo físico consiste en el ordenamiento lineal de genes o marcadores a lo largo de los cromosomas utilizando técnicas en biología molecular como el FISH, clones, fragmentos de restricción de longitud polimórfica (RFLP) y sitio de etiquetado de secuencia (STS). La diferencia con el mapeo anterior se da al momento de calcular las

distancias físicas entre genes, y se miden en pares de bases (pb). La técnica STS consiste en una secuencia de ADN entre 100 y 500 pb, la cual es fácilmente reconocida y está presente solo una vez en el genoma en estudio. Para la obtención de un conjunto de secuencias para ser mapeadas se utilizan paneles de híbridos celulares por radiación (Brown, 2007).

– **Híbridos por radiación**

Esta tecnología fue desarrollada por Goss y Harris (1975), cuyas variaciones han derivado en la construcción de mapas físicos para genomas de varias especies. La ventaja de esta técnica es la posibilidad de mapear marcadores o genes no polimórficos (ya que se evalúa la presencia o ausencia del marcador) a una resolución mayor que la que se consigue con los mapas genéticos desarrollados con la técnica de análisis de ligamiento. También la unidad de medida de los mapas que se generan es el centiRay (cR). Esta metodología tiene la ventaja de complementar las técnicas de mapeos genéticos y físicos, porque tiene un enfoque de mapeo comparativo entre mamíferos y contribuye a mejorar el mapeo del genoma completo (Faraut *et al.* 2009).

El fundamento de la técnica (Figura 3) es la creación de clones celulares conformados por fragmentos de cromosomas de dos especies diferentes. Por un lado, están las células donantes (de la especie que se quiere mapear), que han sido irradiadas letalmente (mediante rayos X, o mediante exposición a algún elemento radioactivo) para causar la fragmentación del ADN; las células receptoras son líneas celulares deficientes para un marcador de selección, como puede ser la timidina quinasa (TK⁻) o la hipoxantina fosforibosil transferasa (HPRT⁻). Luego ambas células son fusionadas y se dejan crecer en medios selectivos como el HAT (hipoxantina-aminopterina-timidina) para que las células que contienen algún fragmento de las células donadoras logren sobrevivir. Muchos de los híbridos celulares obtenidos con esta metodología retienen cantidades del ADN donante entre el 20 a 30 por ciento del genoma. Por lo tanto, estas células contienen solo un pequeño fragmento del genoma donante (irradiado) incorporados en los cromosomas de la especie receptora (no irradiada) (Walter y Goodfellow, 1995) (Faraut *et al.* 2009).

El mapeo utilizando híbridos por radiación tiene un rol muy importante durante el proceso del secuenciamiento y ensamblado del genoma completo (Hitte *et al.* 2008). Proporciona una fuente de información independiente para validar ensamblados de genomas, porque

la comparación de mapas producidos por otras fuentes (genéticos, basados en secuencia) brinda evidencias sobre su precisión. Además, es una herramienta muy útil para la localización de genes de interés en animales. La integración de la genómica comparativa en el mapeo por híbridos por radiación permite mejorar el proceso de mapeo debido al aprovechamiento de colinearidad importante (sintenia conservada) entre genomas de especies relacionadas o cercanas (Servin *et al.* 2010).

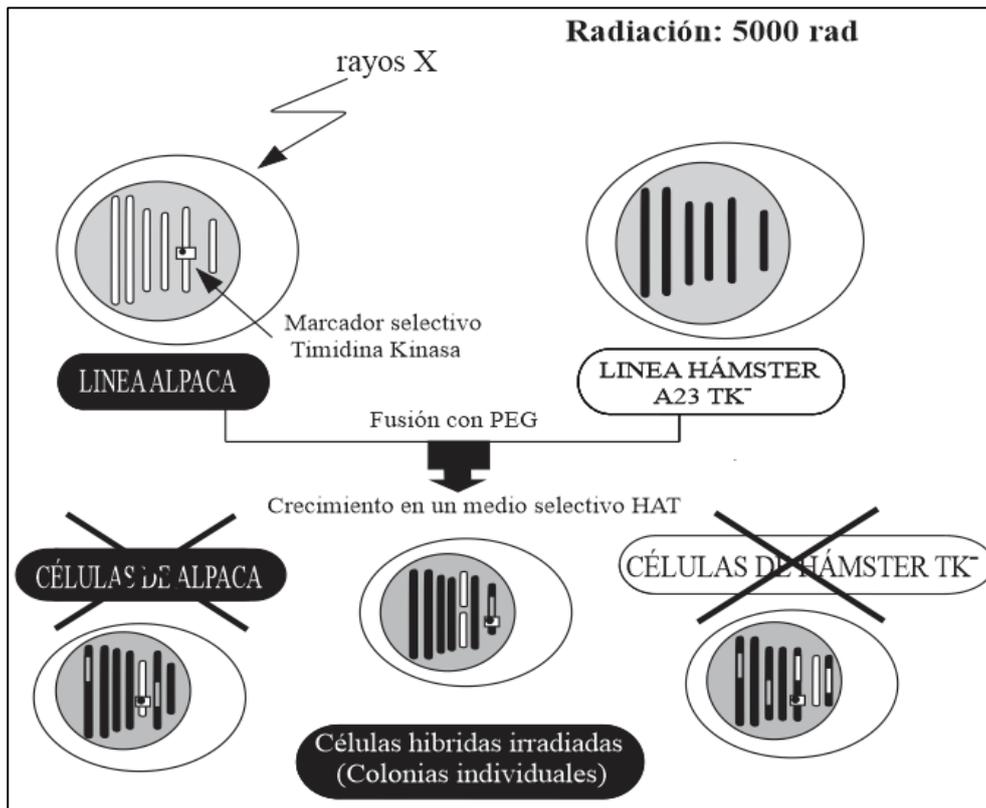


Figura 3: Representación de la construcción del panel celular híbrido irradiado

FUENTE: Elaboración propia

Existen una serie de paneles celulares híbridos irradiados en animales de granja, con lo cual se ha conseguido un mayor conocimiento de sus genomas, en donde se tienen mayores dosis de radiación para conseguir mapas de mayor resolución (Tabla 1). En el 2007 se realizó un panel celular de una alpaca macho llamado “Limerick” conformado por 92 clones híbridos con una radiación de 5000 Rad. El panel está compuesto por clones con frecuencias de retención mayores al 40 por ciento y se llegaron a mapear 428 marcadores en base a la homología existente entre los cromosomas dromedario-humano (Balmus *et al.* 2007; Jhonson y Perelman, 2007). Adicionalmente, se mapeo 50 y 60 marcadores dentro del cromosoma 16 y X, respectivamente. Perelman *et al.* (2018) construyeron dos paneles híbridos irradiados en el dromedario, de 5000 y 15000 Rad. El

primero está constituido por 93 clones y su frecuencia de retención es de 47.7 por ciento; el segundo está compuesto por 90 clones y una frecuencia de retención de 39.9 por ciento. Estos resultados obtenidos van a generar un recurso genético importante para la construcción de un mapa genómico de alta resolución, debido a que es un paso importante en la calidad del ensamblado a nivel de cromosomas.

La resolución de los mapas físicos depende del nivel de radiación aplicada a las células donantes y la frecuencia de retención de los fragmentos involucrados (Gyapay, 1996). Para poder realizar análisis sobre los datos existe programas como el RHMAP (Boehnke *et al.* 1991) y Carthagene (De Givry *et al.* 2005) para poder discriminar, en general, si dos marcadores se encuentran localizados en un mismo fragmento de ADN, o no. Esta tecnología, entonces, nos genera la posibilidad de realizar un mapeo físico global y poder determinar si dos marcadores se encuentran cercanos uno de otro y determinar las distancias físicas entre ellos.

Tabla 1: Paneles celulares híbridos irradiados en animales domésticos

Especie	Nombre del panel	Dosis de radiación (Rad)	Línea celular hámster	Total de híbridos	Frecuencia de retención
Vacuno	ResGenRH10	3,000	Hámster Wg3H	94	25
	BovR5	5,000	Hámster A23	90	28
	BovR12	12,000	Hámster A23	88	30.6
	SUNbRH	7,000	Hámster Wg3H	91	17.5
Porcino	SSRH	5,000	Ratón L-M (TK ⁻)	110	30.6
	IMpRH	6,000-7,000	Hámster Wg3H	118	27
	IMNpRH2	12,000	Hámster Wg3H	90	35

Continuación...

Especie	Nombre del panel	Dosis de radiación (Rad)	Línea celular hámster	Total de híbridos	Frecuencia de retención
Perro	RHDF5000	5,000	Hámster HTK3-1	126	21
	RHDF9000	9,000	Hámster HTK3-1	121	25.8
Ovino	USUoRH5000	5,000	Hámster A23	88	25
	INRA-SheepRH	12,000	Hámster Wg3H	90	31.8
Equino	TM99	3,000	Hámster A23	94	28
	EquineRH	5,000	Hámster A23	93	26

FUENTE: Faraut *et al.* 2009

2.4. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE

Los polimorfismos de nucleótido simple (PNSs) son variaciones en la secuencia de ADN que afectan una base nitrogenada (Figura 4) y se pueden ubicar en zonas codificantes o no codificantes. En términos poblacionales, se podría presentar en cualquiera de los cuatro nucleótidos, lo que generaría 4 alelos. Teóricamente sería posible, pero generalmente se consideran a los PNSs bialélicos porque en la población existiría la secuencia original que tendría una frecuencia observada más alta que la mutada. Esto se explica debido a la naturaleza del PNS, el cual ocurre debido a una mutación puntual en el genoma. Si esta mutación se da en los gametos, pasaría a la siguiente generación y se podría transmitir durante muchas generaciones para establecerse en la población. Para que se desarrolle un tercer alelo tendría que ocurrir una mutación en el mismo lugar donde paso la primera mutación y la descendencia debe reproducirse para incrementar la frecuencia del nuevo alelo. En consecuencia, este evento sería poco probable (Brown, 2007).

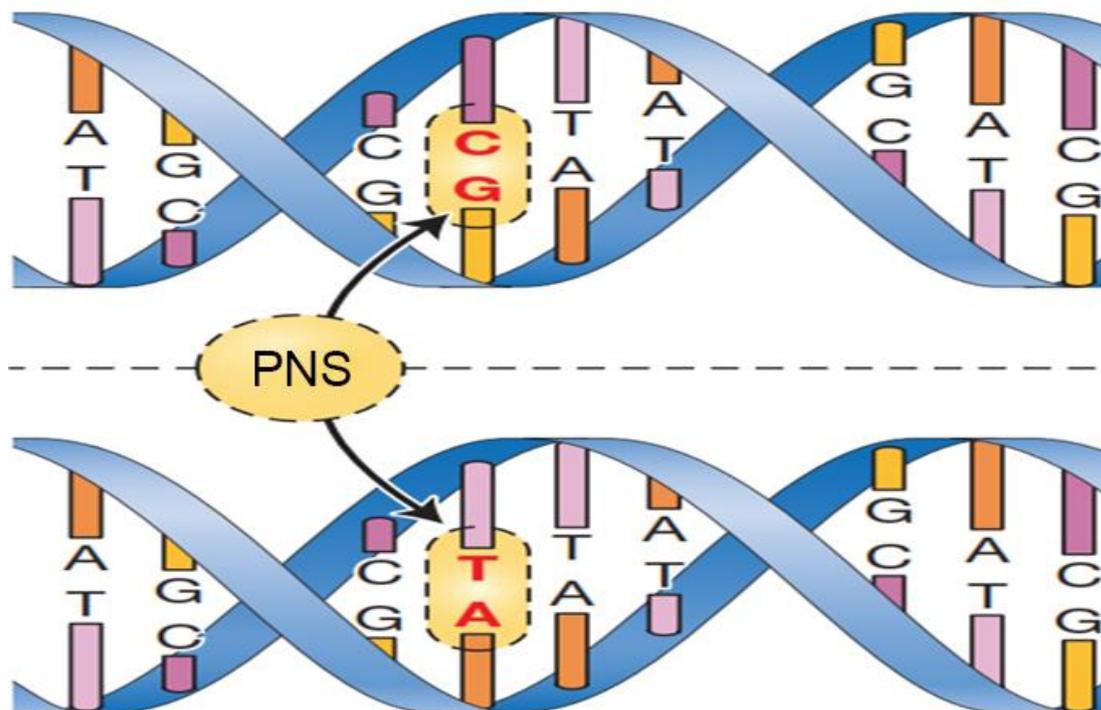


Figura 4: Polimorfismo de nucleótido simple

FUENTE: Griffiths *et al.* 2015

Los PNSs están adquiriendo mayor importancia como marcadores moleculares porque son fáciles de evaluar e interpretar ampliamente por su distribución a lo largo del genoma, ya que se calcula que en células eucariotas se presentan uno por cada kilobase (kb) de ADN. Además, los PNSs adquirieron relevancia porque se desarrollaron micromatrices con las cuales se pueden analizar simultáneamente miles de ellos en una sola reacción. Consecuentemente, se incrementó la eficiencia y se redujo el costo de genotipificación por marcador y por animal consiguiéndose una cobertura mayor del genoma de cada animal. Además, ello facilita el análisis una gran cantidad de haplotipos (combinación de PNSs en un determinado cromosoma estadísticamente asociados) y poder evaluar el grado de diferenciación entre poblaciones (Vignal *et al.* 2002).

Los PNSs han sido ampliamente usados en la identificación de animales y determinación de parentescos en diferentes especies de animales, como *Bos taurus* (Heaton *et al.* 2002), y en asociación a genes de interés en *Ovis aries* (Kijas *et al.* 2009) y *Sus scrofa* (Hidalgo *et al.* 2013). Adicionalmente, se han reportado en ganado Simmental la asociación de PNSs a caracteres relacionados al carácter carcasa y calidad de carne (Wu *et al.* 2014); mapeo de genes relacionados al color de fibra en equinos (McCue *et al.* 2012); y, la

generación de borradores de genomas en especies relacionadas a las ovejas (Miller *et al.* 2015). También contribuyen con el mejor entendimiento de la evolución de las especies, formación y domesticación de las razas, elaboración de nuevas teorías en genética de poblaciones y mejorando los métodos de selección genómica para un mejor desarrollo del mejoramiento genético (Fan *et al.* 2010).

Existen escasos estudios de PNSs en alpacas, entre los cuales se pueden destacar el realizado por Paucar, (2011) donde bajo un estudio de simulación pudo determinar que un panel de 100 PNSs con MAF de 0.35 proporciona una asignación de paternidad casi perfecta para los sistemas de apareamiento tradicionales (controlado, alternado y múltiple) y que se requieren 40 PNSs aproximadamente para dar un poder de exclusión comparable a un panel de 10 microsatélites. Adicionalmente, debido a la importancia de la fibra en alpacas y el interés en la identificación de marcadores moleculares y/o genes relacionados a sus características que nos permitan implementar programas de mejoramiento genético, se reportó un estudio sobre dos PNSs (A57G y A458G), ubicados en el extremo 3' terminal (3'UTR), que fueron altamente informativos ($PIC=0.374$, $H_0>0.351$, $H_e=0.5$) y que se encuentran en desequilibrio de ligamiento con el gen de la trichialina en alpacas (Delgado, 2014). Pero no se encontró una asociación entre estos marcadores con los fenotipos Suri y Huacaya. Otro estudio en 96 alpacas de las regiones de Puno, Cusco, Huancavelica, Apurímac, Arequipa y Junín identificó 27 marcadores PNSs en cinco genes de la familia de proteínas asociadas a la queratina (KRTAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1, KRTAP13-1), y al realizar un estudio de asociación entre el gen KRTAP11-1 con el diámetro de fibra no evidenció una asociación significativa en muestras de 152 alpacas huacaya blancas de la región Puno (Foppiano, 2016).

2.5. MICROMATRICES DE ADN

Las micromatrices o microarreglos de ADN son pequeños soportes sólidos de vidrio o cuarzo pulido, de más o menos dos cm^2 de área, en los que se anclan oligonucleótidos conocidos como sondas, siguiendo un patrón de ordenamiento conocido. Estas sondas son de cadena sencilla y pueden ser sintetizadas *in vitro*, clonadas o amplificadas por PCR. El fundamento de la técnica es la hibridación del ADN de la muestra a analizar con las sondas ancladas que contiene la micromatriz (Figura 5). Cada sonda anclada representa un PNS y contiene por lo menos 50 nucleótidos que terminan en el nucleótido

polimórfico del PNS en la posición 3'. El ADN genómico que requiere ser analizado primero será cortado en pequeños fragmentos y las regiones de interés serán amplificadas con PCR. Estas secuencias amplificadas serán marcadas con un colorante fluorescente para que a continuación sean hibridizadas con la micromatriz. La emisión de fluorescencia será de mayor a menor intensidad conforme la secuencia analizada corresponda con las sondas de la micromatriz. Muchos PNSs pueden ser analizados en una simple reacción por animal (Brown, 2007).

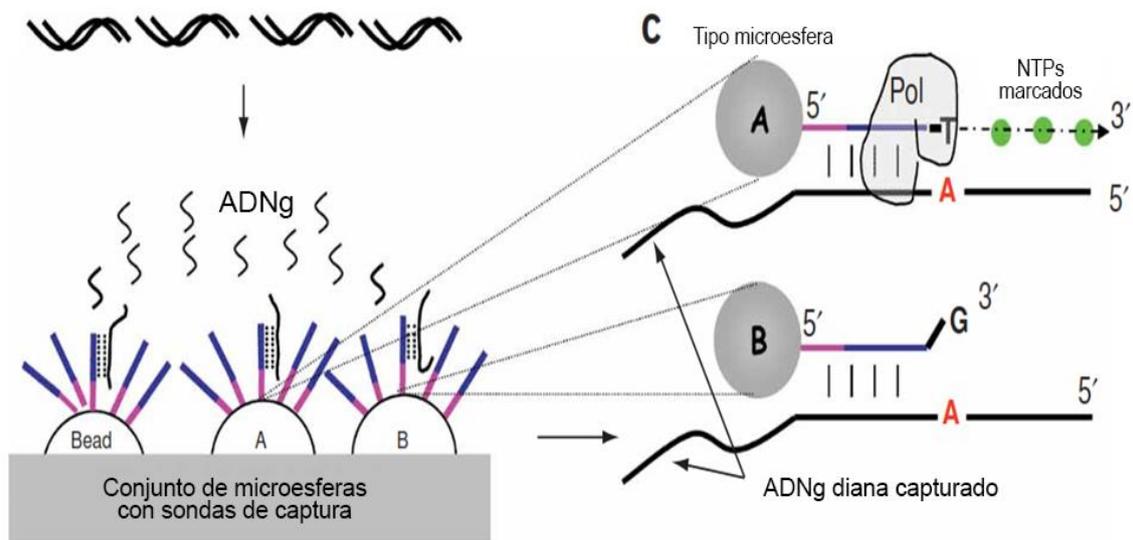


Figura 5: Genotipado del genoma con una micromatriz

FUENTE: Gunderson *et al.* 2005

Con el avance de las técnicas moleculares y el deseo de predecir las características fenotípicas de un animal, se elaboraron las llamadas micromatrices de alta densidad en el ganado bovino, llegando a analizar en un solo experimento 777,962 PNS que son usados de manera rutinaria en la selección genómica del ganado lechero (Su *et al.* 2012). Estos marcadores moleculares PNSs están en desequilibrio de ligamiento con genes que afectan los caracteres de interés y nos permiten realizar un cribado de los animales y predecir su performance fenotípica en base a la información genómica obtenida.

El análisis de expresión génica por la técnica de micromatriz se ha establecido como uno de los mejores métodos de cribado en numerosas especies. Numerosos estudios de este tipo han sido la base de identificación de genes con funciones conocidas e implicados en el control del ciclo celular. El concepto micromatriz se ha extendido a otros campos y no sólo a la expresión génica, y ya es habitual encontrar numerosas referencias de estudios

de metilación, PNSs o microARNs utilizando soportes de micromatriz. Dependiendo del tipo de micromatriz que se utilice, de expresión, de metilación o microARNs la técnica puede variar ampliamente, siendo constante la necesidad de un soporte donde se producirán los miles de micro-ensayos, y donde cada micro-ensayo, corresponde a una secuencia, normalmente perteneciente a un gen (Martínez, 2012).

La similitud filogenética que existe entre los rumiantes y camélidos ha sido motivo de diversas investigaciones en el área de la genética molecular, siendo uno de esos estudios el genotipado que se realizó en alpaca y dromedario utilizando 3 micromatrices de alta densidad comerciales de PNSs de bovino (777K BeadChip), ovino (600K BeadChip) y caprino (53K BeadChip) con el fin de identificar polimorfismos de nucleótido simples que podrían estar presente en el genoma de la alpaca. Se analizaron muestras de ADN de 20 alpacas y 25 dromedarios con las micromatrices mencionadas y los resultados demostraron que la micromatriz de bovino pudo genotipar con éxito un 10.5 por ciento y 20.7 por ciento de los PNSs en alpacas y dromedario, respectivamente, en todos los animales evaluados (Bertolini *et al.* 2016).

III. MATERIALES Y MÉTODOS

3.1. LUGAR DE ESTUDIO

El presente estudio se desarrolló en dos fases: la primera fue mediante el genotipado del panel celular alpaca/hámster en los laboratorios de Neogen-Geneseek (USA) y la segunda consistió en el análisis bioinformático realizado en la Universidad Nacional Agraria La Molina.

3.2. MATERIAL EXPERIMENTAL

El análisis se realizó sobre un panel de 96 muestras de ADN de alpaca y hámster (Figura 6), generados por el Smithsonian Conservation Biology Institute (USA) (Johnson y Perelman, 2007). Está formado por 92 clones celulares híbridos irradiados alpaca/hámster y cuatro muestras de ADN genómico control de: alpaca huacaya macho (Limmerick), alpaca huacaya hembra (Carlotta), línea celular de hámster A23 y una mezcla de ADN de alpaca macho y hámster en una proporción 1:10, respectivamente.

Los 92 clones celulares fueron generados con líneas celulares donadoras de fibroblastos de alpaca (Limmerick) y la línea celular receptora de hámster A23, utilizando una intensidad de radiación de 5000 rad. Cada clon celular presenta una frecuencia de retención mayor al 40 por ciento del genoma de alpaca.

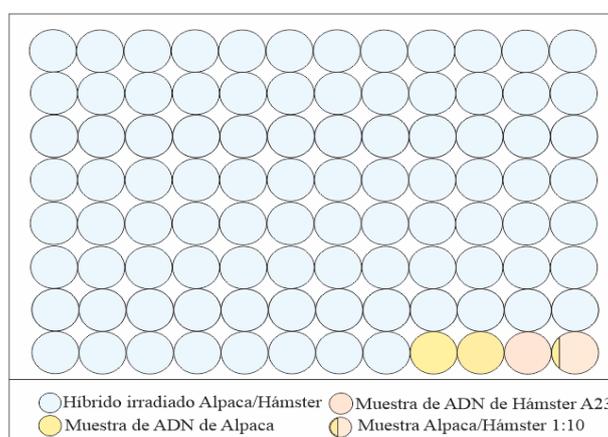


Figura 6: Distribución de las 96 muestras del panel

FUENTE: Elaboración propia

3.3. METODOLOGÍA

3.3.1. Genotipado de polimorfismos de nucleótido simple

El panel fue genotipado mediante el uso de una micromatriz de alta densidad para bovinos (BovineHD BeadChip – Illumina), la cual contiene 777,962 PNSs. El análisis fue realizado por el laboratorio Neogen-Geneseek, bajo los lineamientos de la compañía Illumina (Illumina, 2015).

3.3.2. Identificación de polimorfismos de nucleótido simple en alpacas

La información obtenida del genotipado del panel fue almacenada en el programa GenomeStudio (Illumina, 2016), donde se tabuló la información genotípica de las 96 muestras analizadas con la micromatriz. Se identificaron señales positivas cuando hubo una hibridación específica entre las sondas de la micromatriz y la muestra evaluada.

Las cuatro muestras no híbridadas de ADN genómico fueron analizadas en el GenomeStudio para determinar el número total de polimorfismos de nucleótido simple que emitieron señal positiva en alpacas macho y hembra, así como también en hámster. La muestra 1:10 alpaca: hámster nos permitió identificar, por comparación, señales positivas débiles en clones de células híbridadas en que pudiera existir un porcentaje bajo de ADN de alpaca con respecto al hámster. Para lograrlo se utilizó el parámetro Call Frequency con frecuencia igual a uno, el cual nos indica que la señal positiva estuvo presente simultáneamente en todas las muestras evaluadas.

Para la identificación de PNSs presentes en el genoma de la alpaca, no comunes con las de hámster, se generaron tres bases de datos en formato “.xlsx” con la información de Call Frequency igual a uno, las cuales fueron: las dos muestras de alpaca, la muestra de hámster y la muestra de alpaca/hámster 1:10. Cada base de datos consta de tres columnas que son el número, nombre y el número de cromosoma donde se ubica el PNS en el bovino. Luego se verificaron los PNSs positivos comunes usando el programa R (v. 3.3.2) mediante los comandos “merge” y “cbind” (Anexo 1), que permitieron realizar tres comparaciones:

- La base de datos de las dos alpacas con la de hámster, para obtener los PNSs comunes entre ambos. Estos fueron retirados y se generó una base de datos con los PNSs que quedaron de las dos alpacas.

- La base de datos de hámster y alpaca/hámster 1:10, para obtener los PNSs comunes entre ambos. Estos fueron retirados y se generó una base de datos con los PNSs que quedaron de la alpaca/hámster 1:10.
- Las dos bases de datos generadas en los apartados previos, consideramos los PNSs comunes, con lo cual se obtuvo una base de datos final de PNSs que fue evaluada posteriormente.

Estas comparaciones nos aseguran que los polimorfismos de nucleótido simple evaluados identificados sean solo aquellos detectados en el genoma de la alpaca.

3.3.3. Determinación de grupos ligados de polimorfismos de nucleótido simple

Se realizó un filtrado final en el GenomeStudio considerando un Call Frequency \geq a 0.20 y \leq 0.80, luego se identificaron los genotipos obtenidos de los 92 clones celulares híbridos irradiados alpaca/hámster. Esta base de datos se comparó con la generada en el apartado 3.3.2 de PNSs presentes en las dos alpacas, para identificar los PNSs comunes mediante el software R. Finalmente, como el panel fue realizado con el ADN de una alpaca, el resultado de cada PNS en los 92 clones celulares híbridos podría ser solo homocigoto (AA, BB), o heterocigoto (AB) o no genotipado (NC); donde ningún PNSs puede haber sido como 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 para encontrar los grupos de ligamiento.

Para calcular las distancias y orden de los PNSs encontrados previamente se utilizó el programa Carthagene (Schiex *et al.* 2009). Sin embargo, antes de introducir los datos al programa para su análisis se realizaron dos pasos previos:

- Se identificó la información de los genotipos de los 92 clones celulares híbridos irradiados en los PNSs filtrados previamente. Los genotipos AA, AB y BB fueron codificados con uno y los no genotipado (NC) con cero en un archivo de extensión “.xlsx”. Luego este archivo fue dividido en 30 grupos de PNSs en referencia al número de cromosomas que presenta el genoma del bovino, debido a que nos facilita el análisis de ligamiento en el programa Carthagene.
- El programa Carthagene se ejecutó creando un formato Mapmaker, el cual consiste en convertir los datos a un bloc de notas con la siguiente información: la 1era línea indica el tipo de dato que se va analizar, en nuestro caso fue “data type

radiated hybrid diploid”; la 2da línea indica el número de clones del panel celular híbrido (92), la cantidad de PNSs a analizar y la nomenclatura de presencia o ausencia del marcador en cada clon celular (1=H; 0=A); y la 3ra línea empieza a indicar el nombre del PNS con los resultados del genotipado de los 92 clones celulares definidos con el número uno para el resultado positivo y el número cero para resultado negativo. De la 4ta línea para adelante se continúan indicando los PNSs y sus resultados uno por línea. Según la especificación del formato Mapmaker, el número del PNS debe ser antecedido por “*S” y el número 0 hasta alcanzar ocho caracteres en total (Anexo 2).

A continuación, se detallan los comandos ejecutados en el programa Carthagene para la obtención de los mapas de ligamiento entre los PNSs evaluados.

- Comando “load”, para subir el archivo creado en formato Mapmaker;
- Comando “group” 0.5 y 7.0, que permitió seleccionar los marcadores que se encuentren a una distancia de 50 centiRay (cR) y con un valor de comparación entre la probabilidad de que dos loci estén ligados, a la probabilidad de que no lo estén (LOD) igual o mayor de 7;
- Comando “groupget”, para separar los posibles marcadores que cumplieron con ese requerimiento del resto de marcadores que no lo hicieron;
- Comando “mrkselset [groupget]”, para trabajar con los grupos de marcadores seleccionados;
- Comando “mrklod2p”, para calcular la matriz de Lod score para los PNSs analizados;
- Comandos “sem”, “nicemapl” y “nicemapd”, para estimar el orden y distancia de los PNSs analizados;
- Comando “heaprintd”, para ver los resultados de los comandos ejecutados previamente;
- Comando “buildFW”, para estimar el mejor mapa con los grupos de marcadores con altas probabilidades de estar ligados. En base al orden y las distancias se crearon los mapas de ligamiento.

3.3.4. Localización de grupos ligados de polimorfismos de nucleótido simple

Una vez identificados los grupos de PNSs ligados se crearon tablas de datos donde se registraron cada PNS con su secuencia 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. Sin embargo, se formaron diferentes bases de datos de grupos ligados, cada una con el PNSs y un fragmento de 15, 20 y 25 nucleótidos (de derecha a izquierda), ubicándose en la posición final el polimorfismo. Este procedimiento se realizó debido a que la cantidad de 121 nucleótidos es para analizar bovinos y no alpacas, y que la polimerasa es muy específica en los fragmentos cercanos al PNS.

La información resultante fue guardada en formato FASTA, que consiste en un bloc de notas donde por cada línea se incluye el símbolo “>” continuado con el número de ligamiento y el nombre del PNSs, seguido por la secuencia de nucleótidos del PNS; este procedimiento se repite con todos los PNSs para cada grupo evaluado.

El genoma de referencia *VicPac 2.0.2* del NCBI (National Center for Biotechnology Information) fue descargada en formato FASTA del GenBank (GCF_000164845.1) para luego analizarlo con el software REPEATMASKER, con la finalidad de eliminar las secuencias repetitivas dentro del genoma de la alpaca. Esta nueva base de datos contiene en su mayor parte secuencias genómicas únicas debido a que el programa REPEATMASKER reemplaza las secuencias repetidas con la letra “N” la que es ignorada por otros programas de análisis cuando se hacen comparaciones entre secuencias genómicas y fragmentos de secuencias incógnitas.

Las dos bases de datos creadas previamente se subieron a la plataforma Galaxy (Cock *et al.* 2015); con la finalidad de localizar los PNSs en los scaffolds correspondientes, mediante el programa BLAST y los comandos SHORT BLAST (Basic Local Alignment Search Tool) y un E-value de 0.01 (NCBI, 2016).

Finalmente, se realizó una estimación de las posibles correspondencias de los PNSs identificados en los grupos de ligamiento hacia cromosomas de alpacas. Este procedimiento se llevó a cabo en referencia a literatura científica, donde se definió una correspondencia macroscópica entre cromosomas de bovino – dromedario (Anexo 3) (Balmus *et al.* 2007) y por la relación filogenética perteneciente a la misma familia se

asume que es similar en casos de alpacas. Asimismo, Ávila *et al.* (2015) utilizando el *VicPac 2.0.2* lograron asignar determinados scaffolds a cromosomas en alpacas, lo cual contribuyó en el presente estudio para afirmar que determinados PNSs estén localizados en los cromosomas de alpacas (Anexo 4).

3.4. ANÁLISIS ESTADÍSTICO

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

Para el caso de la localización de los PNSs en el genoma de la alpaca se calculó el valor esperado (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 inferiores a uno. En la tabla 2 se muestran valores equivalentes con el p-value, usado en muchas pruebas estadísticas.

Tabla 2: Relación entre el E-value y el p-value

E-value	p-value
10	0.9999
5	0.9933
2	0.8647
1	0.6321
0.1	0.0952

Continuación...

E-value	p-value
0.05	0.0488
0.001	0.0009
0.0001	0.0001

FUENTE: Pesvener, 2015

IV. RESULTADOS Y DISCUSIÓN

4.1. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE

El esquema de análisis e identificación de PNSs que emitieron señal positiva y estuvieron presentes solo en el genoma de la alpaca se muestra en la Figura 7. Como se aprecia de los 777,962 PNSs analizados mediante la micromatriz de alta densidad, se pudo seleccionar 294,165 PNSs en las 2 muestras de alpaca, siendo un 37.81 por ciento del total. En la muestra de hámster se identificaron 453,699 PNSs, y luego de realizar filtrados y eliminar los PNSs comunes, se encontró un total de 72,964 PNSs en ambas alpacas. La muestra que contenía ADN de alpaca y hámster en la proporción 1:10 se comparó con el resto de muestras controles para encontrar la cantidad de PNSs final presentes en alpaca y ausentes en el hámster, resultando en 50,686 PNSs, siendo un 6.5 por ciento de los PNSs analizados.

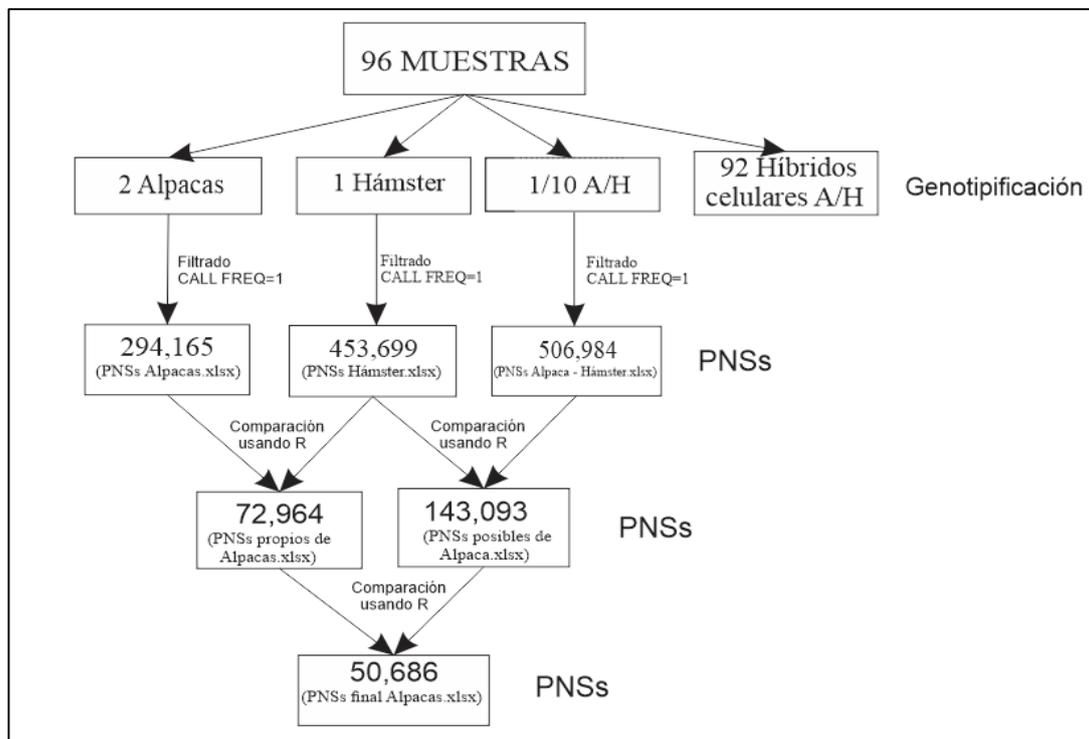


Figura 7: Identificación de polimorfismos de nucleótido simple en las cuatro muestras de control

FUENTE: Elaboración propia

Estos resultados demuestran la posibilidad de usar la micromatriz de bovino para la identificación de PNSs en el genoma de la alpaca. En el presente estudio solo se obtuvo un 37.81 por ciento de PNSs positivos en alpaca del total de PNS analizados, lo cual concuerda con los resultados del trabajo realizado en venados usando la micromatriz de 50K disponibles para bovinos, en donde se pudo obtener un 38.7 por ciento de PNSs positivos (Haynes y Latch, 2012), pero difiere totalmente con una similar metodología aplicada en el bisón, donde se obtuvo un 96.7-98.7 por ciento de PNSs positivos con la misma micromatriz (Pertoldi *et al.* 2010). La explicación para esta diferencia está en la mayor divergencia que existen entre familias filogenéticas, donde la distancia evolutiva entre bovinos y los venados es de 25.1-30.1 millones de años (Hassanin y Douzery, 2003), en comparación con los 1.2-2.1 millones de años entre el bovino y el bisón (MacEachern *et al.* 2009). En el caso de la divergencia estimada entre los camélidos y vacunos es de 42.7 millones de años (Wu *et al.* 2014), por lo tanto, se espera y se observa una ligera menor cantidad de PNSs identificados en la alpaca con la micromatriz de bovino.

Por otra parte, el genotipado del panel celular irradiado más las muestras controles nos permitió identificar aquellos PNSs que son detectados solo en el genoma de la alpaca. En tanto, esta cantidad se redujo a 6.5 por ciento del total de PNSs analizados. Bertolini (2016) logró genotipar un 10.5 por ciento de PNSs en 8 alpacas, lo cual comparado con el presente estudio difiere porque, en nuestro caso, se utilizó solo una alpaca y además se tuvo que excluir PNSs comunes con hámster. Sin embargo, esto no quiere decir que los PNSs excluidos no sean datos importantes que podrían generar mayor información del genoma, sino que para fines del presente estudio y evitar posibles errores al momento de realizar los análisis de ligamiento, solo nos interesó analizar los PNSs ubicados en ambas muestras de las alpacas genotipadas y observados también en los 92 clones celulares híbridos.

Los polimorfismos de nucleótido simple están siendo muy utilizados como marcadores moleculares para investigar procesos genéticos evolutivos, estudios de asociación a caracteres de interés productivos y para incrementar el conocimiento en especies no modelo donde no se conocen detalles sobre su genoma (Seeb *et al.* 2011). Para los fines de la presente investigación el tener una mayor cantidad de PNSs referenciados a lo largo

del genoma va permitir mejorar el ensamblado actual del genoma de la alpaca y empezar a asignar los scaffolds a cromosomas.

Los métodos utilizados para la genotipificación de PNSs en genomas diploides de gran tamaño dependen de la amplificación por PCR de las regiones genómicas que abarcan los polimorfismos de nucleótido simple antes de la reacción de genotipificación real. La PCR proporciona la sensibilidad y especificidad requeridas para distinguir entre genotipos heterocigotos y homocigotos en genomas grandes y complejos. La dificultad de diseñar y llevar a cabo reacciones de PCR múltiple es un factor importante que limita el rendimiento de los ensayos de genotipificación (Steemers y Gunderson, 2007).

4.2. DETERMINACIÓN DE GRUPOS LIGADOS

Una vez identificados los 50,686 PNSs distribuidos en los 30 cromosomas del bovino, el resultado del filtrado de un Call Frequency ≥ 0.2 y ≤ 0.8 presentes en los PNSs identificados en el genoma de la alpaca y que cumplen el requisito de ser AA y NC o BB y NC o AB y NC fueron un total de 2924 PNSs. Como se puede observar y deducir que el panel celular híbrido irradiado para un determinado PNS solo podría tener un genotipo, porque su preparación estuvo basada en el genoma de un solo animal. Por lo tanto, el PNS que presenta ambos genotipos dominantes o la combinación de dominante y heterocigoto a lo largo de los 92 clones indica que la genotipificación de esos PNSs es inconsistente para el análisis. Las razones por las cuales hubo una disminución considerable de los polimorfismos de nucleótido simple a ser analizados por el Carthagene fueron: el nivel de radiación (5000 rad) utilizada en el presente estudio, porque a menor dosis de radiación los fragmentos de ADN son más grandes y menor la resolución del análisis; y la decisión de analizar grupos de PNSs por cromosoma de bovino, ya que podrían existir asociaciones si analizáramos PNSs de varios cromosomas de bovinos, lo que podría aumentar el número de PNSs asociados.

Al realizar los análisis en el Carthagene se determinó un total de 33 grupos de ligamiento con un total de 216 PNSs (Tabla 3) presentes en el genoma de la alpaca. Los mapas de referencia (framework) generados con el orden, nombre y distancias de los polimorfismos de nucleótido simple se observan en el Anexo 5.

Tabla 3: Total de grupos ligados

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		13	7	158.41	9.55	0.52
4	3	9, 13, 21	7	285.6	6.08	0.47
5			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

Continuación...

Grupo de ligamiento	Cromosoma bovino	Probable cromosoma de alpaca	Total de PNSs	Distancia total (cR)	Promedio LOD score	Probabilidad de retención
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		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			

FUENTE: Elaboración propia

Como se puede observar en la Tabla 3, de los 30 grupos analizados inicialmente basados en el número de cromosomas del bovino, se estimaron 33 grupos de ligamiento con los PNSs que se identificaron en el genoma de la alpaca. La razón por la cual existen un mayor número de grupos ligados es porque podrían estar en ambos brazos de los cromosomas identificados o también en el mismo brazo del cromosoma pero suficientemente distante entre grupos ligados. Según los estudios realizados por Balmus *et al.* (2007) y la tabla mostrada en el Anexo 3 existen correspondencias cromosómicas entre el bovino y dromedario que nos pueden ayudar a ubicar marcadores moleculares, y tal como evidenció Ávila *et al.* (2015) en base a estudios citogenéticos y mapeo por FISH donde indicaron que existe un escaso o ningún rearrreglo cromosómico entre las especies de camélidos.

En la Tabla 3 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 2 grupos de ligamiento que podrían estar presentes a ambos lados de los brazos del cromosoma (Ramos, 2014). Los demás cromosomas solo presentaron un grupo de ligamiento y existe una mayor probabilidad que los PNSs identificados se encuentren en el brazo q. Adicionalmente como se puede ver en los gráficos del Anexo 5, los polimorfismos de nucleótido simple BovineHD0300033514, BovineHD0300033341, BovineHD0300028979 y BovineHD0300024228 están presentes en los grupos de ligamiento 4 y 5, siendo este último conformado solo por estos cuatro PNSs. La explicación podría ser que las mismas secuencias estén duplicadas en regiones distintas del genoma de la alpaca o que el ensamblaje del genoma es aún incipiente.

Cabe resaltar que este trabajo sería el primero en identificar estos 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 los ensamblados de referencia para lograr mapas de mayor resolución a nivel de cromosomas. 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 cada PNS en los grupos de ligamiento respectivos, debido a que por el momento no se tiene información completa del ensamblado del genoma de la alpaca, así como tampoco asignación específica de “scaffolds” a cromosomas. 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 rad, 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 5,000 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 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 $LOD = 6$, lo que nos indica que existen altos niveles de probabilidad de que el orden y las distancias en los gráficos del Anexo 5 sean correctos. El motivo por el cual se decidió usar un valor alto de LOD es porque para el genotipado del panel se usó la micromatriz de bovino y se necesitaba ser más restrictivo al momento de evaluar la exclusión de PNSs.

Una importante característica observada en la Tabla 3 es la probabilidad de retención calculada en el Carthagene para los 33 grupos de ligamiento fue de 47 por ciento, la cual es muy similar a la obtenida en el panel celular híbrido irradiado del dromedario, con un 47.7 por ciento con una dosis de radiación de 5000 rad. Esto resulta muy diferente a las

frecuencias de retención obtenidas en la Tabla 1, donde se observan valores muchos menores a los reportados en el presente estudio. Las razones para la explicación de este resultado aún son desconocidos. Sin embargo, una posible explicación podría ser que las células de los camélidos presentan una mayor eficiencia de fusión (número de colonias híbridas por millón de células donadoras fusionadas irradiadas) con las células de hámster, como se puede apreciar en el panel elaborado para el dromedario donde la eficiencia de fusión fue de 588.9×10^{-6} y que en el caso de alpacas no se ha reportado (Perelman, datos no publicados), comparado con las especies domésticas. Adicionalmente, se puede mencionar que debido a la alta cantidad de bloques de heterocromatina que posee el genoma de la alpaca (Ávila *et al.* 2015; Raudsepp, 2014) y al momento de ocurrir la radiación en el panel, los fragmentos de cromosomas fueron más afines con las de hámster y se consiguió mayor frecuencia de retención (Perelman *et al.* 2018).

4.3. LOCALIZACIÓN DE GRUPOS LIGADOS

La Tabla 4 muestra un total de 23 PNSs ubicados en el *VicPac 2.0.2* al analizar las secuencias de 20 nucleótidos, todos con E-value menor a 0.05. Se consideró este valor porque es más restrictivo al momento de inferir si el alineamiento es producto del azar. Hay que tener en cuenta que el valor del E-value va a depender del tamaño de la secuencia que se va a analizar (query) y de la base de datos del genoma. Para nuestro caso las secuencias analizadas fueron de 15, 20 y 25 nucleótidos, siendo el nucleótido final el polimorfismo de nucleótido simple; y la base de datos del *VicPac 2.0.2* está conformada por 2.1 Gbp pares de bases. Por lo tanto, a medida que la secuencia de búsqueda (query) disminuye en el número de nucleótidos, va aumentar el E-value y habrá mayores probabilidades de que la identificación de secuencias sean productos del azar. Esta sería la razón porque al analizar con 15 nucleótidos el E-value fue mayor a 1 (Pesvener, 2015).

El porcentaje de identidad mostrado en la Tabla 4 es determinado sumando las correspondencias idénticas entre secuencias alineadas y dividiendo entre el número total de bases alineadas, y fueron todos del 100 por ciento; sin embargo, los emparejamientos idénticos (matches) en algunos casos fueron 19 y otros de 20 (representados entre paréntesis en la columna de % identidad) (Klug *et al.* 2013). 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 esas secuencias ubicadas en el *VicPac 2.0.2* no sean productos de azar.

Cada PNSs fue ubicado en los scaffolds correspondientes producto del BLAST. Estos scaffolds hasta el momento no han sido asignados a cromosomas de la alpaca, por lo que su asignación solo se podría hacer en nuestro estudio por inferencias realizadas en trabajos previos donde se realizó una correspondencia macroscópica entre el dromedario y vacuno (Balmus *et al.* 2007) y un mapa citogenético integrado utilizando la técnica FISH en alpaca (Ávila *et al.* 2015). La ventaja de ambos trabajos realizados es que determinados scaffolds fueron asignados a cromosomas de alpaca. Para fines del presente estudio solo se puede afirmar que las secuencias encontradas tienen altas probabilidades de encontrarse en los scaffolds respectivos. Cabe resaltar que los scaffolds son secuencias de nucleótidos por encima de los 10 Mbp y que es posible que en un scaffold correspondiente se ubiquen secuencias de dos cromosomas distintos, debido a ensamblaje imperfecto, como se puede apreciar en el Anexo 4.

La Tabla 5 nos muestra un total de 8 PNSs identificados en el *VicPac 2.0.2*. al analizar las secuencias query de 25 nucleótidos. A diferencia de la tabla anterior los porcentajes de identidad son mayores al 95 por ciento, en donde aparecen emparejamientos 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 alineo correctamente y se verifica si el cambio de base fue por transición o transversión. Los valores del E-value en los 8 PNSs fueron menores a 0.05, por lo tanto, los alineamientos son biológicamente significativos.

Los motivos por el cual los resultados del presente análisis solo ubicaron 31 PNS (14.35 por ciento) de los 216 PNSs en el *VicPac 2.0.2* se pueden deber a una serie de factores como son: 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 y a la poca información que se tenía en cuanto al número exacto de nucleótidos que reacciona con la micromatriz y hace positiva su aparición en los resultados.

Considerando los Anexos 3 y 4 se pueden realizar las siguientes aseveraciones: el 1BovineHD0100027068, 8BovineHD0600011684 y 13BovineHD1100019948 podrían estar asignados a los cromosomas de alpaca 1, 2 y 4, respectivamente, debido al estudio realizado por Ávila *et al.* (2015), en donde asignó a los scaffolds 24, 80 y 83 a los cromosomas mencionados. Por lo tanto, de los 31 PNSs ubicados en la base de datos *VicPac 2.0.2*, solo se pudo inferir 3 PNSs mencionados anteriormente a cromosomas de alpaca. La razón por la cual no se pudo inferir el resto de PNSs es porque existen miles de scaffolds no asignados a cromosomas en alpaca. Esta aseveración debería ser corroborada, en un futuro, con un mapeo físico mediante la técnica FISH.

La formación de grupos de ligamiento fue realizada con la intención de buscar un conjunto de PNSs que puedan ser ubicados en un cromosoma y generar información genómica para poder aumentar el conocimiento de marcadores moleculares dispersos a lo largo de todo el genoma para posteriores trabajos de investigación en cuanto a estudios de asociación a caracteres productivos. Sin embargo, debido al ensamblado aun inicial, a nivel de scaffold, y la poca asignación de estos a cromosomas hace difícil complementar esta información por medio solo de la bioinformática. Esto se debería complementar con mapeos genéticos y físicos para que puedan servir de guía para mejorar el ensamblado que se tiene actualmente (Ávila, 2014).

Tabla 4: Secuencias de polimorfismos de nucleótido simple localizadas en el *VicPac 2.0.2* con un tamaño de 20 nucleótidos

Grupo de ligamiento y Nombre del PNS	% Identidad	E- value	Secuencia nucleotídica	Número del scaffold	Numero de Accesión
3BovineHD0200017018	100 (19)	0.027	AAAGGGCAGAAACTCTTT	49	NW_005882751.1
4BovineHD0300033341	100 (20)	0.007	CATTTGGTTTTAATTTTTAT	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	CATATGAAATCATTCCTTTT	224	NW_005882926.1
9BovineHD0700019108	100 (19)	0.027	TGTTAATATCCTATAGAAA	33	NW_005882735.1
10BovineHD0800010566	100 (20)	0.007	AAAAAGCAATTGGTTATTAA	61	NW_005882763.1

Continuación...

Grupo de ligamiento y Nombre del PNS	% Identidad	E- value	Secuencia nucleotídica	Número del scaffold	Numero de Accesión
10BovineHD0800012560	100 (20)	0.007	TTGAAGAATTTTAGAAAGTA	8	NW_005882710.1
11BovineHD0900030678	100 (19)	0.027	CTTTCCCCACCAAGGCAGA	279	NW_005882981.1
11BovineHD1000014299	100 (20)	0.007	GCTGTTGGCCAAGAGTTCAA	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

Continuación...

Grupo de ligamiento y Nombre del PNS	% Identidad	E- value	Secuencia nucleotídica	Número del scaffold	Numero de Accesión
22BovineHD2000007676	100 (19)	0.027	GATTGTTTCCAGTTTTTCAT	15 30 119	NW_005882717.1 NW_005882732.1 NW_005882821.1
24BovineHD2200017192	100 (19)	0.027	TATCACCTCAAAAAAAAAAA	5 60 66	NW_005882707.1 NW_005882762.1 NW_005882768.1
25BovineHD2300008252	100 (20)	0.007	TGTCACTGCTTGTGCCTCCA	18 77	NW_005882720.1 NW_005882779.1
26BovineHD2400015769	100 (19)	0.027	TCTTTAGCAAAGATTTTCT	0	NW_005882702.1
28BovineHD2500003077	100 (19)	0.027	AAGGGGTGGGGAGGAAGAA	58	NW_005882760.1

Continuación...

Grupo de ligamiento y Nombre del PNS	% Identidad	E- value	Secuencia nucleotídica	Número del scaffold	Numero de Accesión
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

FUENTE: Elaboración propia

Tabla 5: Secuencias de polimorfismos de nucleótido simple localizadas en el *VicPac 2.0.2* con un tamaño de 25 nucleótidos

Grupo de ligamiento y Nombre del PNS	% Identidad	E-value	Secuencia nucleotídica	Mismatch	Número del scaffold	Numero de Acceso
1BovineHD0100027068	95.8 (24)	0.014	GGGAACTCCAA G GTTAGGAGTCCA	A (Transición)	24	NW_005882726.1
1BovineHD0100023558	95.8 (24)	0.014	TTCAAAA T TAGTGGGTGGTAATAT	G (Transversión)	138	NW_005882840.1
3BovineHD0200003288	96.0 (25)	0.003	ATGATGTCTCTCAATGCCA A AGTTT	G (Transición)	293	NW_005882995.1
8BovineHD0600011684	96.0 (25)	0.003	CCTGGAAC T CCTAGAC G GGAAACTT	C (Transversión)	80	NW_005882782.1
11BovineHD0900030288	95.8 (24)	0.014	CCAGACAGAAAATCAATCA A TAAA	G (Transición)	277	NW_005882979.1
13BovineHD1100019948	95.7 (23)	0.042	CTGCAGAAG T TCA A CCAGCATGCA	G (Transición)	83	NW_005882784.1
18BovineHD1600003364	95.7 (23)	0.042	TCTCTGCTTTTATGA G TTTCGATT	A (Transición)	6	NW_005882708.1
22BovineHD2000012425	95.7 (23)	0.042	AGT G T G AGAAAAGCATGAAGCGA	G (Transversión)	23	NW_005882725.1

FUENTE: Elaboración propia

V. CONCLUSIONES

- Se identificaron 50,686 polimorfismos de nucleótido simple utilizando la micromatriz de alta densidad del bovino BovineHD BeadChip.
- Se desarrolló un primer mapa referencial con 33 grupos de ligamiento que incluyen 216 polimorfismos de nucleótido simple.
- Se halló la secuencia de 31 polimorfismos de nucleótido simple en el genoma de referencia *VicPac 2.0.2*, y la localización de tres PNSs (1BovineHD0100027068, 8BovineHD0600011684 y 13BovineHD1100019948) fueron inferidos a los cromosomas 1, 2 y 4 de alpaca.

VI. RECOMENDACIONES

- Realizar el secuenciamiento completo de “Limmerick” para relacionarlo con el panel celular híbrido irradiado e incrementar el número de marcadores moleculares.
- Elaborar una micromatriz de polimorfismos de nucleótido simple específica para alpacas y aplicar la metodología de híbridos celulares irradiados alpaca/hámster.
- Mapeo por Hibridación fluorescente *in situ* de las secuencias de polimorfismos de nucleótido simple que fueron identificadas en el genoma de la alpaca.

VII. REFERENCIAS BIBLIOGRÁFICAS

Al-Samarai, FR, Al-Kazaz, AA. 2015. Applications of Molecular Markers in Animal Breeding: A review. American Journal of Applied Scientific Research 1(1): 1-5.

Ávila, F. 2014. Comparative mapping of the alpaca genome. PhD Dissertation in Biomedical Sciences. Texas A&M University, EE.UU. 179 p.

Ávila, F; Baily, MP; Perelman, P; Das, PJ; Pontius, J; Chowdhary, R; Owens, E; Johnson, WE; Merriwether, DE; Raudsepp, T. 2015. A comprehensive whole-genome integrated cytogenetic map for the alpaca (*Lama pacos*). Cytogenetic Genome Research 144: 193-204.

Balmus, G; Trifonov, VA; Biltueva, LS; O'brien, P; Alkalaeva, ES; Fu, B; Skidmore, JA; Allen, T; Graphodatsky, AS; Yang, F; Ferguson-Smith, MA. 2007. Cross-species chromosome painting among camel, cattle, pig and human: further insights into the putative Cetartiodactyla ancestral karyotype. Chromosome Research 15: 499-514.

Bertolini, F; Elbeltagy, A; Ponce de Leon, FA; Gutierrez, GA; Rothschild, M. 2016. Applicability of using bovine, ovine and caprine SNP chips for alpaca and dromedary genomic studies. 35th International Society for Animal Genetics Conference. Utah. p. 85.

Bianchi, NO; Larramendy ML; Bianchi MS; Cortes, L. 1986. Karyological conservatism in South American camelids. Experientia 42: 622-624.

Boehnke, M; Lange, K; Cox, DR. 1991. Statistical methods for multipoint radiation hybrid mapping. Am. J. Hum. Genet. 49: 1174-1188.

Brown, TA. 2007. Genomes 3. Third ed. New York. Wiley. 69-70 p.

Chandramohan, B; Renieri C; La Manna, V; La Terza, A. 2015. The Alpaca Melanocortin 1 Receptor: Gene Mutations, Transcripts, and Relative Levels of Expression in Ventral Skin Biopsies. Hindawi Publishing Corporation. Scientific World Journal. p. 1-9.

- Cock, PJ; Chilton, JM, Gruning, B; Johnson, JE; Soranzo, N. 2015. NCBI Blast + integrated into galaxy. *GigaScience* 4(39): 1-9
- De Givry, S; Bouchez, M; Chabrier, P; Milan, D; Schiex, T. 2005. Carthagene: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* 21(8): 1703-1704.
- Delgado de la Flor, I. 2014. Microsatélites e identificación de SNP en el gen de Trichohialina en alpacas (*Vicugna pacos*). Tesis Magíster en Bioquímica y Biología Molecular. Lima, Perú, UPCH. 78 p.
- Di Rocco, F; Posik, D; Ripoli, M; Díaz, S; Maté, M; Giovambattista, G; Vidal-Rioja, L. 2011. South American camelid illegal traffic detection by means of molecular markers. *Legal Medicine* 13(6): 289- 292.
- 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(625): 1-16.
- Fan, B; Du, ZQ; Gorbach, DM; Rothschild, MF. 2010. Development and application of high-density SNP array in genomic studies of domestic animals. *Asian-Ast J Anim Sci* 23 (7): 833-847.
- Faraut, T; De Givry, S; Hitte, C; Lahbib-Mansais, Y; Morisson, M; Milan, D; Schiex, T; Servin, B; Vignal, A; Galibert, F; Yerle, M. 2009. Contribution of radiation hybrids to genome mapping in domestic animals. *Cytogenetic Genome Research* 126: 21-33.
- Foppiano, F. 2016. Caracterización de marcadores genéticos en genes que codifican a proteínas asociadas a queratina y evaluación de la asociación del gen KRTAP11-1 al diámetro de fibra en alpaca (*Vicugna pacos*) siguiendo una aproximación de gen candidato. Tesis de Maestría. Universidad Peruana Cayetano Heredia, Perú.
- Fowler, ME. 2010. *Medicine and Surgery Camelids*. Third ed. Iowa. Blackwell Publishing. p. 3-4.
- Goss, SJ; Harris, H. 1975. New method for mapping genes in human chromosomes. *Nature* 255: 680-684.

Griffiths, A; Wessler, S; Carroll, S; Doebley, J. 2015. Introduction to Genetic Analysis. Eleventh ed. New York, USA, W. H. Freeman and Company. 868 p.

Gunderson, KL; Steemers, FJ; Lee, G; Mendoza, LG; Chee, MS, 2005. A genome-wide scalable SNP genotyping assay using microarray technology. *Nature Genetics* 37(5): 549-554.

Gyapay, G; Schmitt, K; Fizames, C; Jones, H; Vega-Czarny, N; Spillett, D; Muselet, D; Prud'homme, JF; Dib, C; Auffray, C; Morissette, J; Weissenbach, J; Goodfellow, PN. 1996. A radiation hybrid map of the human genome. *Human Molecular Genetics* 5(3): 339-346.

Hassanin, A; Douzery, EJP. 2003. Molecular and morphological phylogenies of Ruminantia and the alternative position of the Moschidae. *Systematic Biology* 52: 206-228.

Haynes, GD; Latch, EK. 2012. Identification of novel single nucleotide polymorphisms (SNPs) in Deer (*Odocoileus* spp.) using the BovineSNP50 BeadChip. *PLoS ONE* 7 (5): 1-11.

Heaton, MP; Harhay, GP; Bennett, GL; Stone, RT; Grosse, WM; Casas, E; Keele, JW.; Smith, TP.; Chitko-Mckown, CG; Laegreid, WW. 2002. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. *Mammalian Genome* 13: 272-281.

Hidalgo, AM; Lopes, PS; Paixão, DM. 2013. Fine mapping and single nucleotide polymorphism effects estimation on pig chromosomes 1, 4, 7, 8, 17 and X. *Genetics and Molecular Biology* 36(4): 511-519.

Hitte, C; Kirkness, EF; Ostrander, EA; Galibert, F. 2008 Facilitating genome navigation: survey sequencing and dense radiation-hybrid gene mapping. *Nat Rev Genet* 6: 643–648.

Illumina. 2015. BovineHD Genotyping BeadChip. Data Sheet: Agrogenomics. 4 p.

Illumina. 2016. GenomeStudio® 2.0 Data analysis software. Designs Markers. Process BeadChip. Analyze Data. 4 p.

Jackling, FC; Johnson, WE; Appleton BR. 2012. The genetic inheritance of the blue-eyed white phenotype in alpacas (*Vicugna pacos*). J Heridity American Genetic Association. 11 p.

Johnson, W; Perelman, P. 2007. Development of a radiation-hybrid map. Alpacas Magazine. 234-239.

Kadwell, M; Fernandez, M; Stanley, H; Baldi, R; Wheeler, J; Rosadio, R; Bruford, M. 2001. Genetic analysis reveals the wild ancestors of the llama and the alpaca. Proceedings of the Royal Society 268: 2575-2584.

Kijas, JW; Townley, D; Dalrymple, BP; Heaton, MP. 2009. A Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds. PLoS ONE 4(3): 1-13.

Klug, W; Cummings, M; Spencer, C; Palladino, M. 2016. Concepts of Genetics. Eleventh edition. Editorial Pearson Education. Inglaterra. 889 p.

MacEachern, S; McEwan, J; Goddard, M. 2009. Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (*Bovidae, Bovinae*). BMC Genomics 10: 177.

Marín, JC; Zapata, B; Gonzalez, BA; Bonacic, C; Wheeler, JC; Casey, C; Bruford, MW; Palma, RE; Poulin, E; Alliende, MA; Spotorno, AE. 2007. Sistemática, taxonomía y domesticación de alpacas y llamas: nueva evidencia cromosómica y molecular. Revista Chilena Historia Natural 80:121-40.

Martínez, J; Ayala, RM. 2012. Divergencias genéticas y proteómicas en la fisiopatogenia de las neoplasias mieloproliferativas (NMP). Tesis Doctoral. Madrid. 110 p.

McCue, ME; Bannasch, DL; Petersen, JL. 2012. A High Density SNP Array for the Domestic Horse and Extant Perissodactyla: Utility for Association Mapping, Genetic Diversity, and Phylogeny Studies. PLoS Genetics 8(1): 1-14.

Miller, JM; Moore, SM; Stothard, P; Liao, X; Coltman, DW. 2015. Harnessing cross-species alignment to discover SNPs and generate a draft genome sequence of a bighorn sheep (*Ovis canadensis*). BMC Genomics 16(1): 397.

Morón, JA. 2016. Verificación de la genealogía en alpacas (*Vicugna pacos*) usando un panel de microsatélites. Tesis Mag. Sc. Producción Animal. Lima, Perú, UNALM. 79 p.

NCBI Resource Coordinators. 2016. Database resources of the national center for biotechnology information. *Nucleic Acids Research* 45: 12-17.

Paredes, MM; Membrillo, A; Gutiérrez, JP; Cervantes, I; Azor, PJ; Morante, R; Alonso, A; Molina, A; Muñoz-Serrano, A. 2014 Association of microsatellite markers with fiber diameter trait in Peruvian alpacas (*Vicugna pacos*). *Livestock Science* 161: 6-16.

Paucar, R. 2011. Utilidad de marcadores SNP en la mejora genética de poblaciones altoandinas de alpacas. Tesis Máster Agrobiotecnología. Universidad Pública de Navarra. 40 p.

Perelman, P. 2011. Second generation of alpaca radiation hybrid map: use of high-throughput genotyping. Conference on Camelid Genetics & Reproduction, September 15 – 18, Houston, Texas, USA.

Perelman, PL; Pichler, R; Gaggi, A; Larkin, DM., Raudsepp, T; Alshanbari, F; Holl, HM; Brooks, SA; Burger, PA; Periasamy, K. 2018. Construction of two whole genome radiation hybrid panels for dromedary (*Camelus dromedarius*): 5000 RAD and 15000 RAD. *Scientific Reports* 8: 1982.

Pertoldi, C; Wókcik, JM; Tokarska, M; Kawalko, A; Kristensen, TN; Loeschke, V; Gregersen, VR; Coltman, D; Wilson, GA; Randi, E; Henryon, M; Bendixen, C. 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.

Pesvener, J. 2015. Bioinformatics and Functional Genomics. Third Edition. Editorial Wiley Blackwell. USA. 1124 p.

Ramos, M. 2014. Descripción del patrón de diferenciación longitudinal de los cromosomas de alpacas y llamas. Tesis Magíster en Genética. Lima, Perú. UNMSM. 110 p.

Raudsepp, T. 2012. Camelid genomics and its applications in clinical cases. Proceedings from the International Camelid Health Conference for Veterinarians. Marzo 21-25, Columbus, Ohio, USA.

Raudsepp, T. 2014. Cytogenetics and Infertility. In *Llama and Alpaca Care: Medicine, Surgery, Reproduction, Nutrition, and Herd Health*. First ed. St Louis, Canada, Elsevier. p. 243-249.

Renieri, C; Pacheco, C; Valbonesi, A; Frank, E; Antonini, M. 2007. Programa de mejoramiento genético en camélidos domésticos. *Arch. Latinoam. Prod Anim* 15 (1): 205-210.

Schiex, T; De Givry, S; Chabrier, P; Bouchez, M. 2009. Carthagene-User documentation. Univ Kansas INRA. 192 p.

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. *Molecular Ecology Resources* 11(1): 1-8.

Servin, B; De Givry, S; Faraut, T. 2010. Confidence measures for genome maps: application to the validation of genome assemblies. *Oxford University Press* 26(24): 3035-3042.

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): 1-12.

Stemers, FJ; Gunderson, KL. 2007. Whole genome genotyping technologies on the BeadArray platform. *Biotechnol* 2: 41-49.

Su, G; Brondum, RF; Ma, P; Guldbrandtsen, B; Aamand, GP; Lund, MS. Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *Journal Dairy Science* 95: 4657-4665.

Taylor, KM; Hungerford, DA; Snyder, RC; Ulmer, FA. 1968. Uniformity of karyotypes in the Camelidae. *Cytogenetics* 7: 8-15.

Vignal, A; Milan, D; Sancristobal, M; Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol.* 34: 275-305.

Wu, H; Guang, X; Al-Fageeh, MB; Cao, J; Pan, S; Zhou, H; Zhang, L; Zhang, S; Wang, J; Yin, Y; Yang, H; Al-Swailem AM; Wang, J. 2014. Camelid genomes reveal evolution and adaptation to desert environments. *Nature Communications* 5: 5188.

Wu, Y; Fan, H; Wang, Y; Zhang, L; Gao, X; Chen, Y; Li, J. 2014. Genome-Wide Association Studies Using Haplotypes and Individual SNPs in Simmental Cattle. *PLoS ONE* 9(10): 1-11

Zhou, X; Xu, S; Yang, Y; Zhou, K; Yang, G. 2011. Phylogenomic analyses and improved resolution of cetartiodactyla. *Mol Phyl Evol* 61: 255-264.

Zimin, AV; Delcher, AL; Florea, L; Kelley, DR; Schatz, MC; Puiu, D; Hanrahan, F; Pertea, G; Van Tassell, CP; Sonstegard, TS; Marçais, G; Roberts, M; Subramanian, P; Yorke, JA; Salzberg, SL. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biology* 10(4): 1-10.

VIII. ANEXOS

Anexo 1. Comandos utilizados en R

```
macho=read.table("macho.txt",header=FALSE)
```

```
hembra=read.table("hembra.txt",header=FALSE)
```

```
merge<-merge(hembra,macho,all=FALSE)
```

```
write.csv(merge,"PNSs.csv")
```

```
alpacas=read.table("alpacas.txt",header=FALSE)
```

```
hamster=read.table("hamster.txt",header=FALSE)
```

```
merge<-merge(alpacas,hamster,all=FALSE)
```

```
write.csv(merge,"PNScomunes.csv")
```

```
hamster=read.table("hamster.txt",header=FALSE)
```

```
alpacahamster=read.table("alpacahamster.txt",header=FALSE)
```

```
merge<-merge(alpacas,hamster,all=FALSE)
```

```
write.csv(merge,"possiblealpaca.csv")
```

```
alpacas=read.table("alpacas.txt",header=FALSE)
```

```
merge=read.table("PNScomunes.txt",header=FALSE)
```

```
cbind(merge,1)
```

```
merge1=cbind(merge,1)
```

```
merge1
```

```
merge2=merge(merge1,alpacas,all=TRUE)
```

```
write.csv(merge2,"PNSspropios.csv")
```

Anexo 2. Formato Mapmaker para encontrar grupos de ligamiento

data type radiated hybrid diploid

92 197 0 symbols 1=H 0=A

*S025038 0 1 1 1 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 0 0 1 0 0 0 0 1 0 1 0 1 0 1 1 1 1 0 1 1 1 1
0 1 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 0 0 0 0
0 0 1 1

*S025127 0 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1
1 1 0 0 0 0 0 1 0 1 1 0
0 1 1 1

*S025329 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 0 1 1 0 0 0 0 0 0
1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 1 0 1 1
0 0 0 0

*S025493 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 0 0 1 0 0 0 0 1 0 0 1 1 0 1 1 1 1 1 1
1 1 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 0 0 0 1
0 1 1 1

*S025612 0 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 1 1 1 1 0 1 0 1 0 1 0 1 1 1 1 1 0 1 1 1 1
1 1 0 0 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 0 0 0 0
0 0 1 1

*S025688 0 1 1 0 0 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 0 0 1 0 0 1 0 1 1 1 1 1 0 0 1 1 1 0 1 1 1 1
1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 0 0 0 0
0 0 0 0

*S025700 1 0 1 0 0 1 1 0 1 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 1 1 0 0 1 0 1 1 0 0 0 0 0 0
1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 1 1 0 0 0 0 0 1 1 0 1 0 1 1 0 0 0 1 0 0 0 0
0 0 0 0

*S025862 0 1 1 1 1 0 1 1 1 1 1 1 1 1 0 1 1 1 0 1 0 0 1 1 0 1 0 1 0 1 1 1 0 1 1 0 1 0 1 1 1 1
1 1 0 1 1 0 0 0 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0 0 0
0 1 0 1

*S025864 1 0 1 0 1 1 1 0 0 1 1 1 1 1 1 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 1
1 1 0 0 0 1 0 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0
0 0 1 0

*S025924 1 1 1 0 1 1 1 1 0 1 1 1 1 1 0 0 1 1 1 1 0 1 1 0 0 1 0 1 0 1 1 0 0 1 1 1 1 1 1 1 1 1
1 1 0 1 0 1 0 0 1 1 0 1 0 1 0 0 0 1
1 1 1 1

*S025944 1 0 0 0 0 1 1 0 0 1 1 1 1 1 1 0 0 1 1 0 1 0 0 1 0 0 1 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0
1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 0 1 1 0 0 0 0 0 0 0 0 1 1 1 0 1 1 0 1 1 0 0 1
1 0 0 0

.
. .
.

Anexo 3. Correspondencia cromosómica Bovino-Dromedario

Bos Taurus	Camelus dromedarius
1	1
2	5, 13
3	9, 13, 21
4	7
5	12, 34
6	2
7	3, 22
8	4, 31
9	8
10	6, 27
11	4, 15, 28
12	14
13	19, 35
14	25, 29
15	10, 33

Continuación...

Bos Taurus	Camelus dromedarius
16	13, 23
17	2, 32
18	9
19	16
20	3
21	6, 27
22	17
23	20
24	24, 30
25	18
26	11
27	26
28	11
29	10, 33
X	X

FUENTE: Balmus *et al.* 2007

Anexo 4. Relación del cromosoma con scaffolds del *VicPac 2.0.2*

N° cromosoma de alpaca	N° scaffold
1	24, 25, 50, 221, 234 419
2	9, 80, 102, 168
3	5, 16, 44, 73, 143, 202, 217
4	42, 52, 83, 127, 343, 354
5	0, 67, 76, 92, 104
6	17, 90, 98, 183, 285
7	11, 106, 112, 258, 276
8	6, 19, 21, 53, 199, 237
9	4, 8, 178, 195, 203
10	4, 10, 133, 271
11	20, 61, 85, 291, 379, 420
12	7, 55, 370
13	2, 22, 51, 157, 275, 523
14	89, 106, 110, 4376
15	3, 410, 662

Continuación...

N° cromosoma de alpaca	N° scaffold
16	15, 145, 171, 181, 188, 251, 266, 370, 439
17	1, 23
18	27, 39, 43
19	34, 186
20	31, 210
21	123, 140, 197, 220
22	14, 60
23	13, 39, 213
24	48, 176
25	78, 99, 124, 129, 133, 148
26	8, 26
27	11, 54, 86, 131, 179
28	40, 382
29	33, 45
30	54, 314, 39552

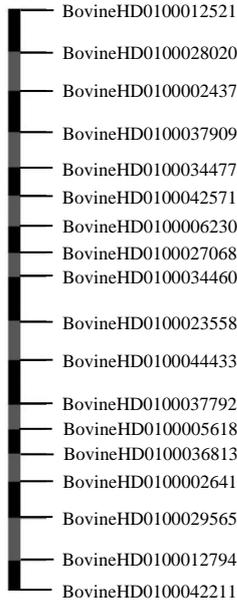
Continuación...

N° cromosoma de alpaca	N° scaffold
31	2, 37
32	97, 411
33	74, 101, 151
34	18, 358
35	101, 137, 230, 231, 406
36	-
X	2, 86, 141, 367, 452, 573, 652

FUENTE: Ávila *et al.* 2015

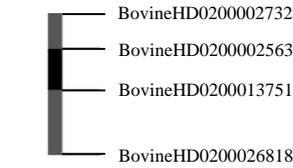
Anexo 5. Gráficos de los grupos de ligamiento

Grupo de ligamiento 1



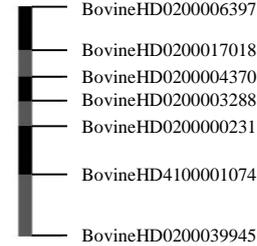
671.90 cR₅₀₀₀

Grupo de ligamiento 2



91.60 cR₅₀₀₀

Grupo de ligamiento 3



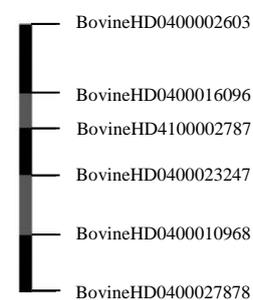
158.40 cR₅₀₀₀

Grupo de ligamiento 5



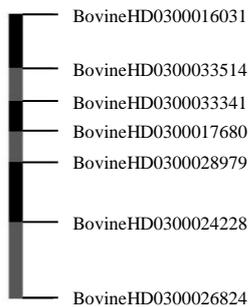
115.40 cR₅₀₀₀

Grupo de ligamiento 6



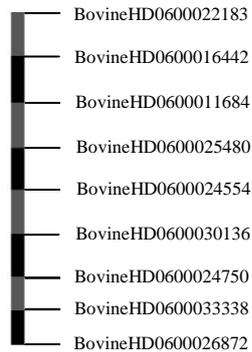
203.50 cR₅₀₀₀

Grupo de ligamiento 4



285.60 cR₅₀₀₀

Grupo de ligamiento 8



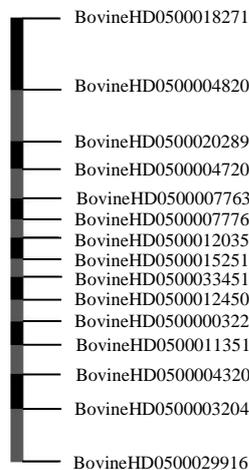
335.50 cR₅₀₀₀

Grupo de ligamiento 9



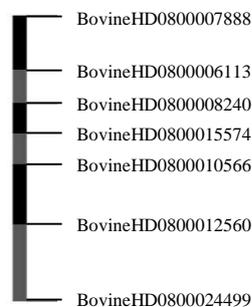
202.20 cR₅₀₀₀

Grupo de ligamiento 7



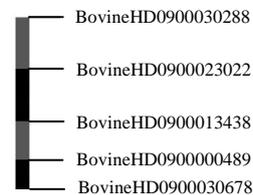
425.0 cR₅₀₀₀

Grupo de ligamiento 10



242.20 cR₅₀₀₀

Grupo de ligamiento 11

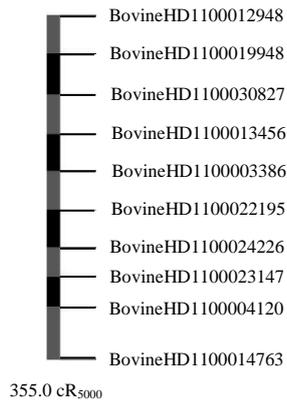


118.0 cR₅₀₀₀

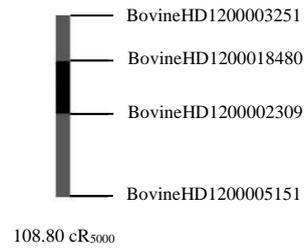
Grupo de ligamiento 12



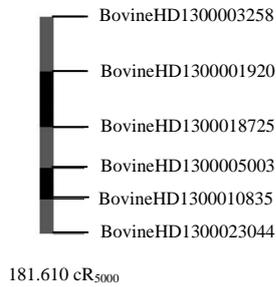
Grupo de ligamiento 13



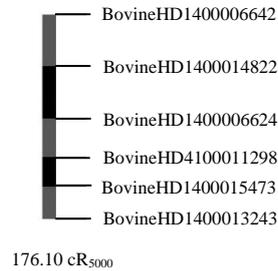
Grupo de ligamiento 14



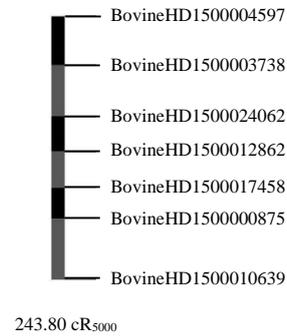
Grupo de ligamiento 15



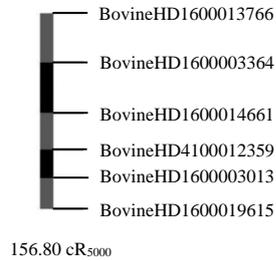
Grupo de ligamiento 16



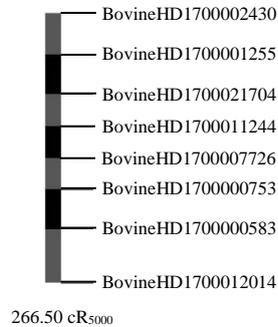
Grupo de ligamiento 17



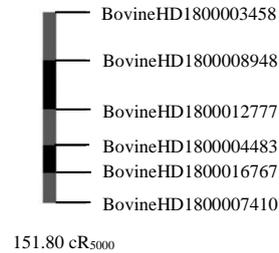
Grupo de ligamiento 18



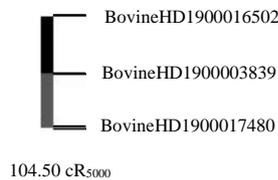
Grupo de ligamiento 19



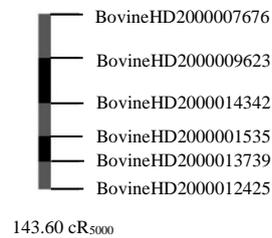
Grupo de ligamiento 20



Grupo de ligamiento 21



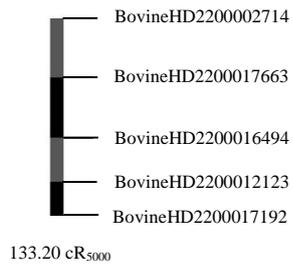
Grupo de ligamiento 22



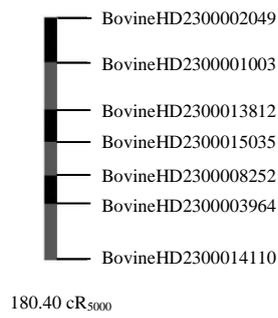
Grupo de ligamiento 23



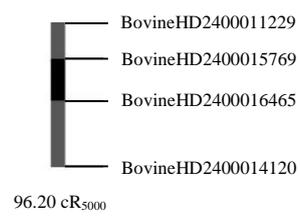
Grupo de ligamiento 24



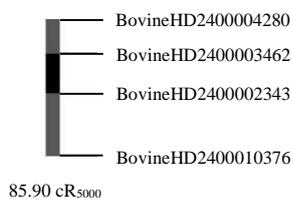
Grupo de ligamiento 25



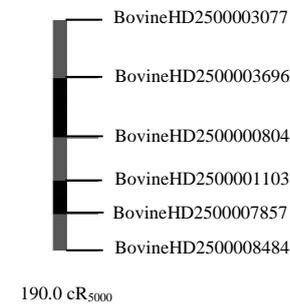
Grupo de ligamiento 26



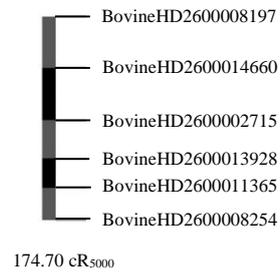
Grupo de ligamiento 27



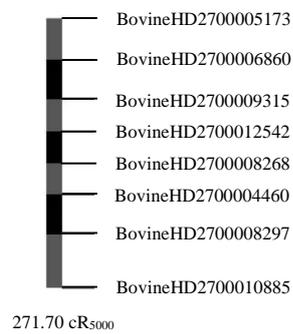
Grupo de ligamiento 28



Grupo de ligamiento 29



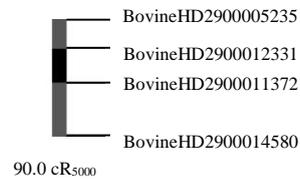
Grupo de ligamiento 30



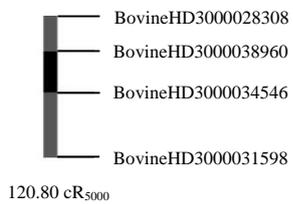
Grupo de ligamiento 31



Grupo de ligamiento 32



Grupo de ligamiento 33



Anexo 6. Glosario

Alineamiento: proceso bioinformático de alinear dos o más secuencias para alcanzar niveles máximos de identidad.

Bandeo cromosómico: Técnica para teñir diferencialmente los cromosomas mitóticos o meióticos, que da lugar a un patrón de bandas característico de ciertas regiones cromosómicas, como centrómeros, regiones del organizador nucleolar y regiones ricas en GC o AT.

Biblioteca BAC: conjunto de clones bacterianos celulares que contienen fragmentos de ADN de cualquier especie asociado a un vector como el factor F de bacterias *Escherichia coli*.

Cobertura o profundidad: el número de nucleótidos secuenciados en un experimento de secuenciamiento dividido por el total de pares de bases que contiene el genoma de la especie que está siendo secuenciada.

Contig: secuencia de ADN continua reconstruida a partir de secuencias de ADN solapadas obtenidas por clonación o análisis de secuencias.

Cribado: separación (selección) rigurosa de animales en grupos basados en sus características genotípicas de asociación de fragmentos genómicos a caracteres fenotípicos.

Cromosoma: estructura lineal alargada compuesta de ADN, ARN y proteínas (histonas) que contiene información genética y son visibles durante la mitosis o meiosis.

Desequilibrio de ligamiento: asociación no aleatoria de alelos en 2 o más loci en un cromosoma, es decir, poseen una frecuencia de recombinación menor al 50 por ciento; por lo tanto, están siendo segregados juntos en la siguiente generación.

Ensamblado: proceso mediante el cual las lecturas de secuencia de ADN solapantes de un experimento de secuenciación se ordenan correctamente a lo largo del genoma.

Hibridación: proceso por el cual se combinan dos cadenas de ácidos nucleicos antiparalelas y complementarias en una única de doble cadena.

Hibridación fluorescente *in situ* (FISH): técnica de localización citológica de secuencias de ADN complementarias utilizando sondas marcadas con señal fluorescente, observadas con luz ultravioleta.

Identidad: el grado en que secuencias comparadas tienen bases o residuos idénticos en posiciones equivalentes, generalmente expresada en porcentaje.

Metilación: transferencia enzimática de grupos metilos (-CH₃) desde la S-adenosilmetionina a moléculas biológicas, incluyendo fosfolípidos, proteínas, ARN y ADN. Está asociada a la regulación de la expresión génica.

Polimorfismo: hacen referencia a distintas formas (alelos) que se presentan en la secuencia del ADN.

Región conservada: una secuencia de nucleótidos (ADN) o residuos de aminoácidos (proteínas) que permanecen invariables a través de la evolución en diferentes especies.

Regiones o secuencias sinténicas: una secuencia de nucleótidos (ADN) donde aparecen genes o loci en el mismo cromosoma, independientemente si están o no ligados.

Scaffold: conjunto ordenado de contigs que están en el orden correcto, pero que no necesariamente forman un tramo contiguo de secuencia.

Secuenciamiento: proceso mediante el cual por diferentes métodos o técnicas se determinan las secuencias de ácidos nucleicos o proteínas.

Selección genómica: metodología utilizada para la predicción del mérito genético de un animal en base a la información de su ADN. En otras palabras, es poder predecir el fenotipo de un animal en base al genotipo.

Sonda: secuencia de ADN o ARN marcada, que puede detectarse mediante microscopio de fluorescencia, u otras formas de detección, con la finalidad de identificar genes o productos génicos.