



Facultad de Veterinaria
Universidad Zaragoza



Trabajo Fin de Grado en Veterinaria

SEÑALES DE SELECCIÓN EN LOS CROMOSOMAS SEXUALES EN
VACUNO DE CARNE AUTÓCTONO

SIGNATURES OF SELECTION IN SEXUAL CHROMOSOMES IN
AUTOCHTHONOUS BEEF CATTLE

Autor

Adrián López Fuertes

Director

Luis Varona Aguado

Facultad de Veterinaria

Curso 2017-2018



Índice

1. Resumen	3
1.1. Abstract	4
2. Introducción	5
2.1. El sector de vacuno de carne en España	5
2.2. Evolución del vacuno de carne	6
2.3. Métodos de detección de huella de selección	8
2.3.1. Diferencias entre poblaciones	9
2.3.2. Análisis de reducción de la variabilidad	9
2.3.3. Modificaciones del espectro de frecuencias	9
2.3.4. Extensión del desequilibrio de ligamiento	10
2.4. El cromosoma X bovino	10
3. Justificación y Objetivos	11
4. Metodología	12
4.1. Material	12
4.2. Métodos	13
4.2.1. Diferenciación entre poblaciones	13
4.2.2. Análisis de reducción de la variación local	13
4.2.3. Extensión del desequilibrio de ligamiento	14
4.2.4. Identificación de genes candidatos	15
5. Resultados y Discusión	16
5.1. Diferencias entre Poblaciones	16
5.2. Análisis de Reducción de Variabilidad	19
5.3. Extensión del desequilibrio de ligamiento	21
5.4. Concordancia con QTL en vacuno	23
6. Conclusiones	24
6.1. Conclusions	25
7. Valoración personal	26
8. Bibliografía	27



1. Resumen

Las poblaciones autóctonas de vacuno de carne español presentan una amplia distribución geográfica a lo largo de la Península Ibérica. Pese a las aparentes diferencias morfológicas, la notable variabilidad en caracteres de crecimiento y desarrollo muscular o de calidad de canal y de la carne, las razas autóctonas españolas se encuentran agrupadas genéticamente. Por lo tanto, es plausible que procedan de un tronco genético común que se ha diferenciado mediante procesos de adaptación y selección, acontecido recientemente en términos evolutivos. Estos procesos afectan de manera desigual al genoma bovino. Como consecuencia, es esperable que las regiones genómicas que contienen genes afectados por estos procesos presenten, en mayor medida, un patrón diferente que aquellas que contienen genes neutrales. En este trabajo se han identificado regiones genómicas asociadas a los procesos de selección y diferenciación en el cromosoma X mediante tres procedimientos distintos: *SELESTIM*, *ROH* y *nSL*, basados en la diferenciación entre poblaciones, el análisis de reducción de la variación local y la extensión del desequilibrio de ligamiento, respectivamente.

El material utilizado consistió en el genotipado, con el Illumina BovineHD BeadChip, de 171 tríos (individuo-padre-madre) procedentes de 7 poblaciones autóctonas de vacuno de carne (Asturiana de los Valles, Avileña Negra-Ibérica, Bruna dels Pirineus, Morucha, Pirenaica, Retinta y Rubia Gallega). Los resultados del estudio han detectado varias regiones con una señal asociada a estos procesos de selección. En concreto, el procedimiento de diferenciación entre poblaciones identificó 4 regiones localizadas en torno a las posiciones 7, 35, 59 y 67 Mb, mientras que el método basado en la reducción de la variación local sugirió 3 regiones localizadas en las posiciones 52, 74 y 86 Mb. Por el contrario, el método basado en la extensión del desequilibrio de ligamiento no fue capaz de detectar ninguna región.



1.1. Abstract

The autochthonous populations of Spanish beef cattle have a wide geographical distribution throughout the Iberian Peninsula. Despite the apparent morphological differences, the remarkable variability in growth and muscle development characters, quality of carcass and meat, the Spanish autochthonous breeds are genetically grouped. Therefore, it is plausible that they come from a common genetic trunk that has differentiated through processes of adaptation and selection, recently happened in evolutionary terms. These processes affect the bovine genome unevenly. Therefore, it is expected that genomic regions containing genes affected by these processes have, largely, a different pattern than those that contain neutral genes. In this work, genomic regions associated with the processes of selection and differentiation in the X chromosome have been identified through three different procedures: *SELESTIM*, *ROH* y *nSL*, based on the differentiation between populations, the analysis of reduction of local variation and the extension of linkage disequilibrium respectively.

The material used consisted of genotyping, with the Illumina BovineHD BeadChip, of 171 trios (individual-father-mother) from 7 native populations of beef cattle (Asturiana de los Valles, Avileña Negra-Ibérica, Bruna dels Pirineus, Morucha, Pirenaica, Retinta and Rubia Gallega). The results of the study have detected several regions with a signal associated with these selection processes. In particular, the procedure of differentiation between populations identified 4 regions located around positions 7, 35, 59 and 67 Mb, while the method based on the reduction of local variation suggested 3 regions located at positions 52, 74 and 86 Mb. To the contrary, the method based in the extension of linkage disequilibrium was not able to detect any region.

2. Introducción

2.1. El sector de vacuno de carne en España

El sector de producción de carne de vacuno en España se sitúa en sexta posición en cuanto a relevancia respecto a los diferentes estados miembros de la Unión Europea (Subdirección General de Productos Ganaderos, 2017a). En nuestro país, la importancia económica de este sistema productivo supone la cuarta posición en la producción final ganadera, que representa en torno al 17,1% de la misma, justo por detrás del sector porcino, la producción láctea y producción avícola (incluyéndose en esta la producción final ganadera de huevos y carne de ave). En términos de producción final agraria, la producción de carne de vacuno representa alrededor del 6,0% (Subdirección General de Productos Ganaderos, 2016).

La evolución de la población total de vacuno en España presenta una ligera tendencia ascendente, pese a un fuerte receso del censo de vacuno lechero, lo que indica un aumento considerable de la cabaña de vacas nodrizas. De este modo nos encontramos una población de 6.609.991 animales, de las cuales 2.151.996 son nodrizas (Subdirección General de Productos Ganaderos, 2017b), que se distribuyen en parte entre las diferentes razas reconocidas oficialmente por el Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente. Esta institución las clasifica en un total de 8 *razas autóctonas de fomento*, 32 *razas autóctonas en peligro de extinción* y 6 *razas integradas en España*. Las poblaciones representadas por los animales incluidos en las *razas autóctonas de fomento* cuentan con un censo total de 510.887 ejemplares. En este grupo se encuentran 6 de las 7 poblaciones motivo de este estudio (Bruna dels Pirineus está catalogada como *raza autóctona en peligro de extinción*), todas ellas de aptitud cárnica, excluyéndose del mismo a la raza de lidia puesto que su finalidad de explotación es totalmente diferente. Las diferentes razas cárnicas incluidas en esta clasificación son, por orden de censo: Asturiana de los Valles (103.020 individuos), Avileña-Negra Ibérica (50.550), Pirenaica (39.372), Parda de la Montaña (37.504), Rubia Gallega (36.256), Retinta (26.111), Morucha (16.879) y Bruna dels Pirineus (12.368).

Por lo tanto, el ganado vacuno autóctono español se dedica casi exclusivamente al aprovechamiento cárnico. Sin embargo, la diversidad de ecosistemas presentes en la península genera una necesaria versatilidad en los sistemas de explotación y aprovechamiento de los recursos naturales, y da lugar a la amplia variedad de poblaciones. Las características principales de las ganaderías de vacuno autóctono español son su régimen extensivo y semiextensivo, la baja densidad ganadera, la dependencia de la irregular producción de pasto y su buena adaptación y rusticidad. Tradicionalmente, se han diferenciado tres grandes grupos de sistemas

productivos (Acero, 2009): *sistemas de la dehesa, sistemas de montaña y sistemas no montañosos de la España húmeda.*

El *sistema de la dehesa* (Aveleña-Negra Ibérica, Morucha y Retinta) está asociado a grandes terrenos, con una baja carga ganadera y una producción vegetal elevada en primavera con un pequeño repunte tras las lluvias del otoño, pero con un pastizal agostado en los meses del verano. Debido a la amplia distribución geográfica nos encontramos con dos variantes: dehesa fría asociada a Meseta Central y Serranía de la Ibérica; y una dehesa cálida en la zona de Extremadura y Andalucía.

Los *sistemas montañosos* (Bruna del Pirineus y Pirenaica) se asientan principalmente en los Pirineos, Picos de Europa o Sistema Central. Con la llegada de la primavera las vacas nodrizas se alimentan de los pastos de media montaña, ascendiendo en altitud hasta llegar a puerto en los meses de verano. En otoño vuelven a descender a la media montaña para estabularse durante el invierno o realizar trashumancia.

Finalmente, los sistemas asociados a la *España húmeda no montañosa* (Asturiana de los Valles y Rubia Gallega) se sitúan en torno a la cornisa cantábrica. Las condiciones climatológicas proveen una abundancia de pasto, con una producción muy baja en invierno lo que obliga a suplementar con silo o estabular. El aprovechamiento vegetal para alimentación se puede dar en prados o en praderas sembradas con especies como ray-grass y tréboles para mejorar la producción.

2.2. Evolución del vacuno de carne

El origen de la domesticación del ganado vacuno se remonta a dos procesos independientes entre sí, acontecidos hace unos 10.000 años (Götherström *et al.*, 2005) a partir de un ancestro salvaje extinto, *Bos primigenius*, que se distribuía por Europa, Asia y el Norte de África (Van Vuure, 2001). De esta forma en el Creciente Fértil se originó el *Bos Taurus*, que es el origen del actual vacuno europeo y africano (Helmer *et al.*, 2005). Sin embargo, los bovinos asiáticos, o *Bos indicus*, proceden de una domesticación más reciente en el valle del Indo (Ajmone-Marsan, 2010). Pese a que existe controversia en cuanto a un tercer evento de domesticación en África, investigaciones recientes de análisis del genoma mitocondrial revelan un origen común con el bovino europeo (Edwards *et al.*, 2010; Decker *et al.*, 2014).

Posteriormente a la domesticación, las poblaciones bovinas europeas han sufrido procesos de migración junto con los ganaderos, potenciales reintrogresiones de genes del Uro (*Bos*

promigenius) que sobrevivió hasta la Edad Media (Upadhyay *et al.*, 2017), aislamiento reproductivo y procesos de adaptación y de selección al medio geográfico y a los sistemas productivos. Todo este conjunto de fenómenos ha generado la variabilidad entre las poblaciones actuales de vacuno, y ha permitido su estandarización mediante el concepto de raza, definida como *grupo subespecífico de animales domésticos con características externas definidas e identificables que le permiten, por apreciación visual, ser diferentes de otros grupos definidos de la misma especie* (FAO, 1999).

En cuanto a las poblaciones españolas, tiene un origen europeo, aunque la presencia de introgresión genética por parte de ejemplares africanos es notable (Decker *et al.*, 2014). La historia evolutiva de las razas actuales ha pasado por diferentes etapas. Inicialmente, todas ellas fueron destinadas a una triple aptitud (trabajo, carne y leche), y, más recientemente se han especializado en la producción de carne. Pese a ello, el grado de diferenciación genética entre las razas autóctonas españolas tiene un rango de pequeño a moderado, relacionado con una limitada presión de selección y un moderado flujo genético entre las diferentes poblaciones (Cañas-Álvarez *et al.*, 2015). Sin embargo, parece estar claro que las principales razas autóctonas españolas presentan tres agrupaciones distinguidas como se muestra en la Figura 1.

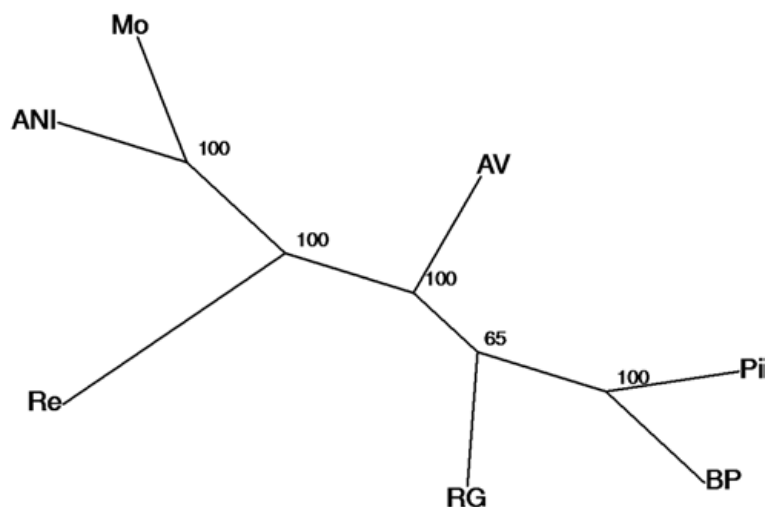


Figura 1. Distribución evolutiva de las razas de vacuno de carne autóctono español. Tomado de Cañas-Álvarez *et al.*, 2015.

Un primer tronco sería el Ibérico, en el que se encuentran agrupadas las razas características de la dehesa como son la Morucha, la Avileña Negra Ibérica y la Retinta. El segundo grupo estaría constituido por el grupo Turdetano, albergando a las razas Pirenaicas, Bruna dels Pirineus y Rubia Gallega. En una posición intermedia se sitúa la Asturiana de los Valles representando el tronco Cantábrico (Jordana *et al.*, 1991; Martín-Burriel *et al.*, 2011; Cañas-Álvarez *et al.*, 2015).

2.3. Métodos de detección de huella de selección

La investigación genómica ofrece la posibilidad de indagar sobre la localización de genes en los cromosomas, identificando polimorfismos genéticos asociados a características productivas, reproductivas y de salud (Womak, 2005). Entre la amplia posibilidad de marcadores disponibles, en la actualidad, los más comunes son los marcadores SNP (*Single Nucleotide Polymorphism*), que son polimorfismos producidos por un cambio nucleotídico que aparecen con una frecuencia elevada a lo largo del genoma. Otra de las posibles aplicaciones de esta información molecular es la detección de “señales” o “huellas” de la selección, tanto natural como artificial. Así, se define “huella de selección” como aquellas regiones del genoma que contiene variantes de secuencia funcionalmente importantes donde la selección ha dejado patrones especiales en la estructura del ADN (Qanbari y Simianer, 2014; Utsonomiya *et al.*, 2015).

La detección de las regiones asociadas a huellas de selección permite la reconstrucción evolutiva de las poblaciones y la identificación de genes con importancia productiva, económica o involucrados en procesos de adaptación ambiental. Los procedimientos de detección de estas señales de selección están basados en la Teoría de la Evolución Neutral (Kimura, 1983) que postula:

- Muchas de las mutaciones que surgen en una población son deletéreas y, en consecuencia, son rápidamente eliminadas.
- Una ínfima proporción de las mutaciones que surgen son ventajosas y rápidamente fijadas en la población
- La amplia mayoría de las mutaciones que se observan son selectivamente neutrales y, en consecuencia, son éstas las que contribuyen al polimorfismo y divergencia.

Esta teoría nos permite generar una hipótesis nula, en el sentido estadístico, donde las regiones del genoma sometidas a selección natural o artificial por presentar genes ventajosos mostrarán una desviación respecto a la estructura del genoma patrón, definido previamente. Por tanto, la hipótesis nula puede ser rechazada cuando la información molecular observada no cumpla con los patrones esperados por la hipótesis nula.

En lo referente a procedimientos enfocados a estudiar procesos evolutivos recientes Oleksyk *et al.* (2010) los clasifica en cuatro categorías: *Diferencias entre poblaciones*, *Análisis de la reducción de la variación local*, *Modificaciones del espectro de frecuencias* y *Extensión del desequilibrio de ligamiento*.

2.3.1. Diferenciación entre poblaciones

El fundamento de estos métodos consiste en que las poblaciones se diferencian entre sí por los procesos de deriva genética. Por lo tanto, las frecuencias alélicas para los SNP difieren entre las poblaciones en mayor medida cuanto más alejadas genéticamente se encuentren. Además, si una determinada región genómica se ha visto sometida a selección, mostrará una mayor diversidad entre poblaciones. Los métodos más representativos de este tipo son el *Estadístico Fst* (Wright en 1943), que calcula la variabilidad genética entre subpoblaciones con respecto a la variabilidad genética total, asumiendo que todas subpoblaciones provienen de una misma población ancestral; y *SELESTIM* (Vitalis *et al.*, 2014) que utiliza una aproximación bayesiana en un modelo jerárquico.

2.3.2. Análisis de reducción de la variabilidad

Este grupo de procedimientos se fundamenta en la hipótesis de que los procesos de selección positiva albergarían regiones concretas del genoma con una marcada reducción de la variabilidad. Estas señales de selección se diferenciarían de otros procesos, como puedan ser cuellos de botella o un número reducido de individuos fundadores, porque estos últimos afectan a todo el genoma y no a zonas específicas. Algunos de los métodos más importantes son la *Medida de Heterocigosidad* (Oleksyk *et al.*, 2008) o la identificación de *Segmentos de Homocigosidad -ROH-* (McQuillan *et al.*, 2008). Los *ROH* son segmentos del genoma donde los individuos son homocigotos para todos sus marcadores SNP. Inicialmente fueron descritos para el estudio de la consanguinidad, pero el recuento del número de individuos que, en una población, tiene un determinado SNP incluido en un *ROH* nos permite identificar regiones con una disminución de la variabilidad genética y, por lo tanto, susceptibles de ser identificadas como una señal de selección.

2.3.3. Modificaciones del espectro de frecuencias

Estos métodos utilizan la teoría de evolución neutral para definir la hipótesis nula a partir de la distribución esperada de las frecuencias alélicas de una determinada región. Por tanto, pese a la baja ocurrencia de mutación, se verá un incremento de la frecuencia para aquellas zonas con ventaja selectiva o la afectada por el efecto de arrastre de otra mutación ventajosa. Entre ellos, vale la pena mencionar el *Estadístico D de Tajima* (Tajima, 1989), el *Estadístico D de Fu y Li* (Fu y Li, 1993) y el *Estadístico H de Fay y Wu* (Fay y Wu, 2000).

Sin embargo, estos métodos están descritos para el análisis de datos de secuenciación y no ofrecen resultados precisos con genotipado menos denso, como el que proporcionan los

marcadores de este estudio. Además, un estudio reciente (González-Rodríguez *et al.*, 2016) sugirió que sus resultados son similares a los obtenidos por los procedimientos de análisis de reducción de la variabilidad.

2.3.4. Extensión del desequilibrio de ligamiento

Para finalizar, existe un conjunto de métodos que estudian patrones en el genoma conocidos como barrido selectivo. Para ello se fundamentan en el desequilibrio de ligamiento asociado a locus detectándose un aumento de frecuencia de haplotipos de mayor longitud. Algunos de los test más importantes son la *Extensión de la Homocigosidad Haplotípica -EHH-* (Sabeti *et al.*, 2002), *Integral Bajo la Curva de Homocigosidad -iHS-* (Voight *et al.*, 2006), *Test de Homocigosidad Haplotípica* (Hanchard *et al.*, 2006) y *nSL* (Ferrer-Admedlla *et al.*, 2014). La principal ventaja de este último método es que no necesita de la fracción de recombinación entre marcadores y utiliza exclusivamente la distancia de mapa. Además, el trabajo de González-Rodríguez *et al.* (2016) mostró que obtenía resultados equivalentes a otros procedimientos de esta categoría como iHS.

2.4. El cromosoma X bovino

El cromosoma X bovino tiene una longitud de 148,823,899 pares de bases (pb) (<https://www.ensembl.org>) y se han localizado 833 genes codificantes, 221 genes no codificantes y 74 pseudogenes. El cromosoma X recombina exclusivamente en la formación de los gametos femeninos, excepto una pequeña región denominada como “pseudoautosómica” que hibrida y recombina con el cromosoma Y en la formación de los gametos masculinos. Esta región se localiza en la zona terminal del cromosoma (Das *et al.*, 2009), a partir de la localización del gen GPR143 (*G protein-coupled receptor 143*), situado entre las pares de bases 143,861,798 y 143,891,357 (<https://www.ensembl.org>).



3. Justificación y Objetivos

A lo largo del proyecto de investigación AGL2010-15903 se identificaron las regiones genómicas asociadas a los procesos de selección y adaptación entre las poblaciones de vacuno de carne autóctono español en los cromosomas autosómicos. Sin embargo, el material experimental procedente de este proyecto también incluía el genotipado para los cromosomas sexuales. Por lo tanto, el objetivo principal del Trabajo de fin de Grado en Veterinaria es la detección de las regiones genómicas localizadas en el cromosoma X, que puedan ser asociadas a procesos de selección y/o a la adaptación a los diferentes sistemas de producción y la identificación de los genes localizados en ellas, junto con una descripción de su implicación metabólica y biológica.

4. Metodología.

4.1. Material

Como material experimental utilizado en este Trabajo fin de Grado en Veterinaria, se han utilizado los genotipados de 171 tríos (individuo-padre-madre) de siete poblaciones de vacuno autóctono español, que cubren prácticamente toda la península ibérica (Figura 2). En concreto se utilizaron individuos de las siguientes poblaciones: Asturiana de los Valles -AV-, n=25; Avileña-Negra Ibérica -ANI, n=24; Bruna dels Pirineus -BP-, n=25; Morucha -Mo-, n=25; Pirenaica -Pi-, n=24; Retinta -Re-, n=24; y Rubia Gallega -RG-, n=24, que fueron genotipados con el Illumina BovineHD BeadChip (777.962 marcadores SNP). Este material procede del desarrollo experimental del proyecto AGL2010-15903.

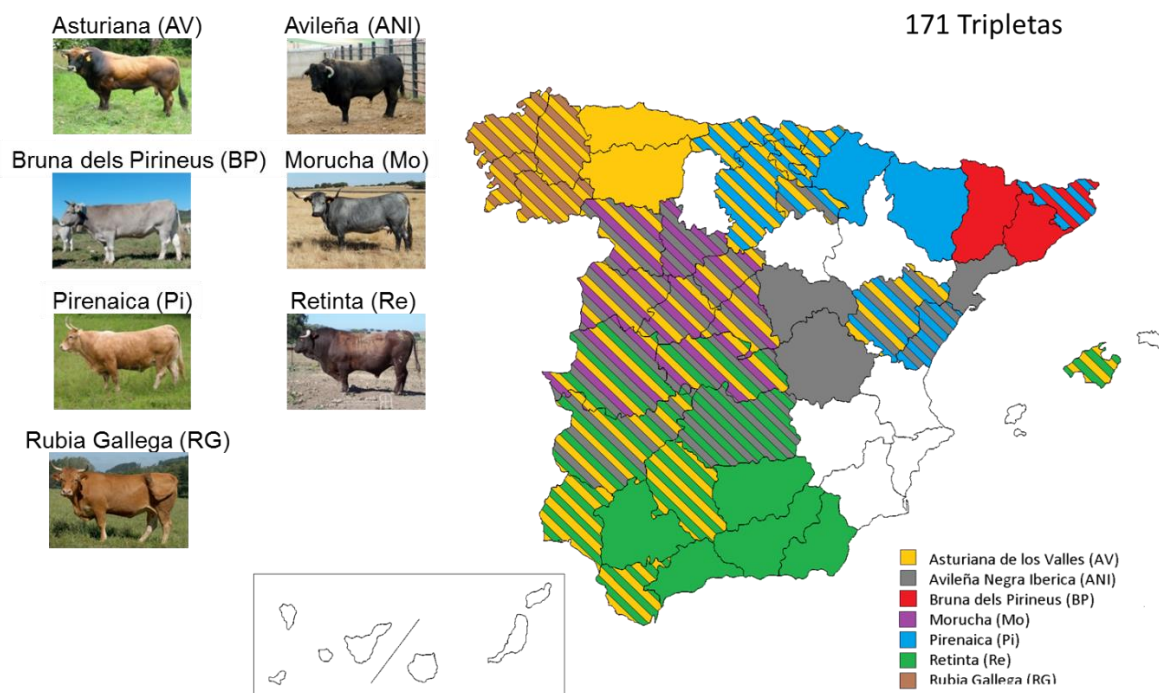


Figura 2. Distribución geográfica de las diferentes razas de carne autóctonas en España. Adaptado de González-Rodríguez, 2015.

De los 777,962 marcadores SNP, 39,367 correspondían al cromosoma X. Posteriormente, se realizó un filtrado standard mediante el software *PLINK* (Purcell *et al.*, 2007) con los siguientes criterios:

- *Call Rate* superior al 95%.
- Frecuencia Alélica menor del 0.01

Después del filtrado se utilizaron 36,734 marcadores SNP. A continuación, se llevó a cabo una imputación de los genotipos faltantes y una reconstrucción de las fases haplotípicas parentales mediante el software *BEAGLE*, con la opción *TRIO* (Browning and Browning, 2009). El proceso de reconstrucción de fases haplotípicas se ilustra con la Figura 3, que muestra como la información familiar permite obtener las fases de los individuos parentales.



Figura 3. Modelo de faseado haplotípico (<http://www.chromosomechronicles.com/>)

4.2. Métodos

4.2.1. Diferenciación entre poblaciones.

El primer procedimiento que se ha utilizado está basado en el estudio de las diferencias entre las frecuencias alélicas de las poblaciones. De entre los métodos disponibles para este enfoque, se ha seleccionado el análisis bayesiano basado en un modelo jerárquico que proporciona el programa *SELESTIM* (Vitalis *et al.*, 2014). Para la implementación de este programa se han calculado las frecuencias alélicas para los 36,734 marcadores SNP utilizando los dos alelos de las madres y uno del macho. El genotipado de los hijos no se ha utilizado ya que se trata de información redundante con la de sus padres.

4.2.2. Análisis de reducción de la variación local.

El estudio de reducción de la variación local se ha realizado mediante el análisis de *Segmentos de Homocigosis* o *ROH* (*Runs of Homocigosity*). Los *ROH* (McQuillan *et al.*, 2008) son segmentos del genoma donde un individuo es homocigoto para todos sus marcadores y pueden ser

definidos en cuanto al número de marcadores y en cuanto a la distancia de mapa entre los marcadores (Figura 4).

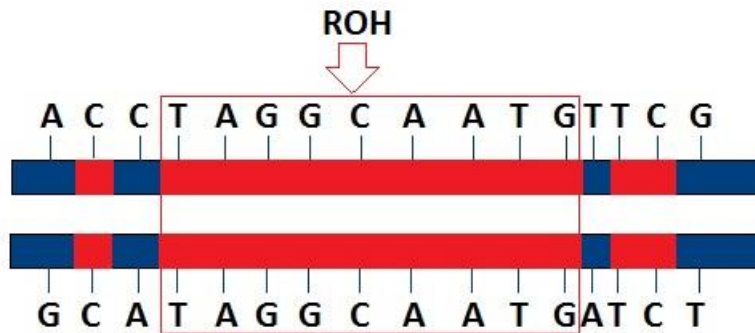


Figura 4. Ilustración de un *ROH* a lo largo de un cromosoma.

Para el cálculo de los *ROH* se ha utilizado el programa *detectRUNS* (Biscarini *et al.*, 2018), desarrollado en entorno R. En concreto, se ha usado la función *consecutiveRUNS* para el cálculo de *ROH* de al menos 50 SNP y, posteriormente, se ha usado la función *SNPinRUNS* para cuantificar el porcentaje de individuos que tiene cada uno de los SNP incluidos dentro de *ROH*. La hipótesis de trabajo es que las zonas del genoma con una reducción de la variación genética local presentarán un mayor porcentaje de individuos homocigotos. Los *ROH* se tienen que calcular en individuos diploides, por lo tanto, el análisis se ha restringido a las hembras fundadoras de cada población.

4.2.3. Extensión del desequilibrio de ligamiento

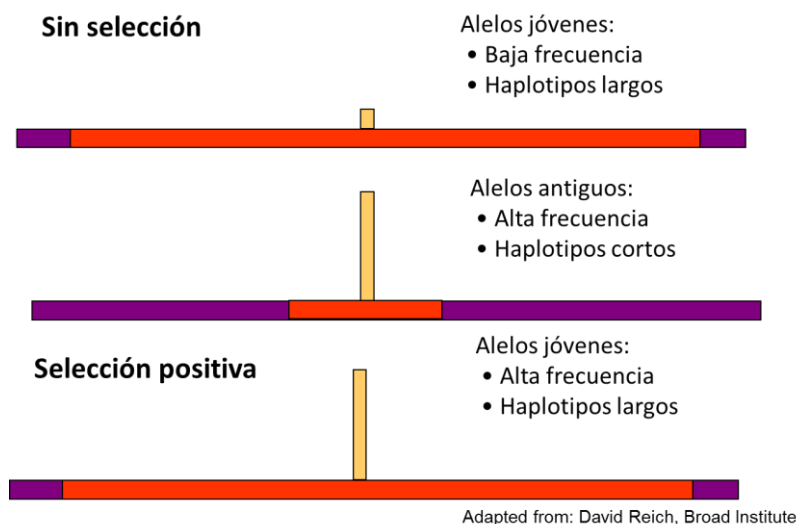


Figura 5. Diferencias entre segmento de cromosoma sometido a selección positiva y sin selección. Adaptado de la presentación de David Reich (University of Harvard) - <http://slideplayer.com/slide/4639927/>.

El tercer procedimiento que se ha utilizado es *nSL* (Ferrer-Admetlla et al, 2014). Este método compara la extensión de los haplotipos asociados a cada uno de los dos alelos de cada SNP y los pone en relación con su frecuencia alélica. Si identifica un alelo con alta frecuencia y asociado a un haplotipo de gran longitud le asocia una señal de selección (Figura 5). El procedimiento se implementó con el programa *Selescan* (Szpiech y Hernández, 2014).

4.2.4. Identificación de genes candidatos.

A partir de los resultados de los tres procedimientos se han identificado los genes que contienen las regiones genómicas definidas por la posición con señal máxima y una Megabase (Mb) por encima y por debajo en el caso de análisis de reducción de la variabilidad local y media Mb desde la señal en el caso de diferenciación entre poblaciones. Para ello, se ha utilizado la aplicación *Biomart* del navegador genómico *Ensembl* (www.ensembl.org). Además, se ha realizado una búsqueda de QTL (*Quantitative Trait Loci*) en estas regiones mediante el mapa genético disponible en la base de datos *cattleQTLdb* (www.animalgenome.org) como se muestra en la Figura 6 con el fin de detectar secuencias que pudieran estar asociadas a caracteres continuos productivos.

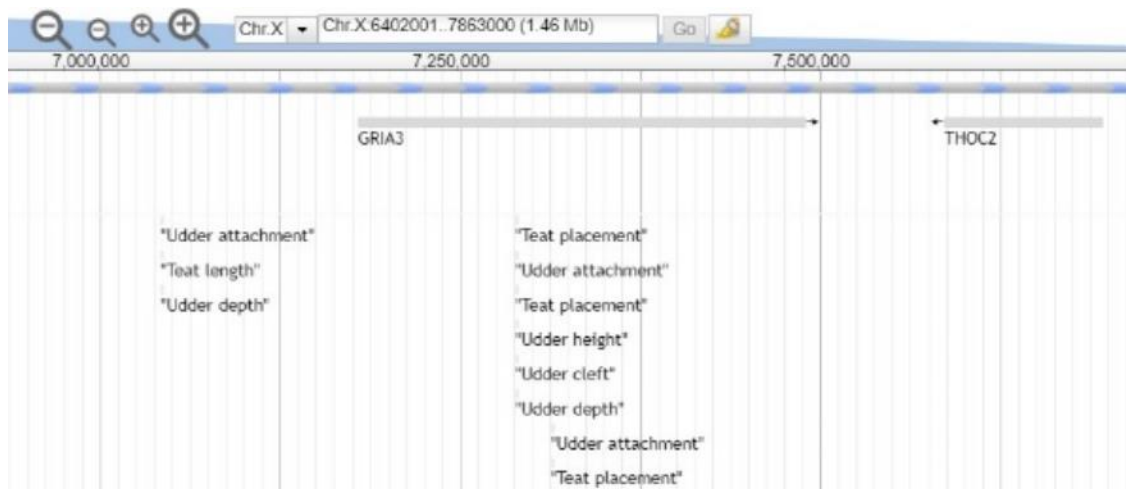


Figura 6. Mapa de identificación de QTL en bovino (www.animalgenome.org).

5. Resultados Y Discusión

5.1. Diferenciación entre poblaciones.

El programa informático *SELESTIM* proporciona un parámetro (KLD) asociado a cada marcador SNP, que identifica la intensidad de selección a la que se ha visto sometida. El valor medio de este parámetro fue 0.174 y su desviación típica 0.277. Sin embargo, el resultado mostró diferencias en la región pseudoautosómica. En esta región la media del parámetro es de 0.433 y la desviación típica 0.445. Este resultado es esperable ya que la oportunidad de recombinación en la región pseudoautosómica tiene lugar en machos y hembras, mientras que en el resto del cromosoma X solamente ocurre en las segundas.

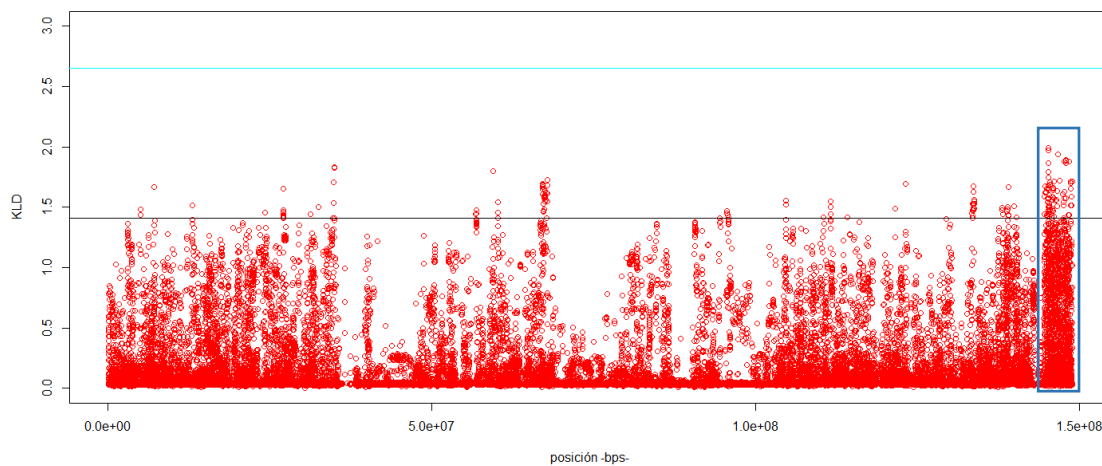


Figura 7. Resultados del barrido genómico con el método *SELESTIM*. Línea horizontal superior umbral de la región pseudoautosómica (azul cian) y línea horizontal inferior umbral de la región no pseudoautosómica (gris).

En la Figura 7 se presenta un barrido genómico de este parámetro a lo largo del mapa genómico del cromosoma X. Se han definido umbrales de relevancia a partir de la media de los estimadores dentro de cada segmento (pseudoautosómico o no) y 5 desviaciones típicas. El umbral de relevancia de la región pseudoautosómica (2.60) es comparable con el obtenido por González-Rodríguez (2015) en los cromosomas autosómicos, sin embargo, en la zona no pseudoautosómica el umbral de relevancia es considerablemente inferior (1.46). En este sentido, en la región pseudoautosómica no hay ningún SNP asociado a una señal de selección relevante, mientras que en la región no pseudoautosómica aparecen varias regiones que merecen ser destacadas.

En concreto los marcadores seleccionados se muestran en la Tabla 1.

Tabla 1. Marcadores con una señal de selección con el procedimiento SELESTIM.

Marcador	Posición
BovineHD3000045971	7148233
BovineHD3000011436	34831053
BovineHD3000011442	34851400
BovineHD3000011449	34873203
BovineHD3000011456	34894695
BTA-103030-no-rs	59388233
BovineHD3000046704	67026447
BovineHD3000019564	67068760
BovineHD3000019566	67076569
BovineHD3000019570	67088500

Estos marcadores con un intervalo de media Mb por encima y por debajo delimitan cuatro regiones genómicas que se presentan en la Tabla 2 y el Anexo 1.

Tabla 2. Regiones genómicas con una señal de selección mediante el procedimiento SELESTIM.

Start	End	GENES
6648233	7648233	GRIA3 y THOC2
34331053	35394695	PRRG3, FATE1, CNGA2, GABRE, GABRA3, GABRQ, CETN2 y NSDHL
58888233	59888233	TSC22D3, PRPS1 y FRMPD3
66526447	67588500	-

En la primera de las regiones genómicas (6,648,233-7,648,233 bps) se localizan dos genes con diferentes funciones biológicas (GRIA3 y THOC2). Entre ellos, es especialmente interesante el gen GRIA3, que es un receptor del glutamato y actúa como neurotransmisor. Su activación estimula la secreción de LHRH/LH y afecta a la edad de pubertad en primates (Terasawa y Fernández, 2001) y ratones (Zhu *et al.*, 2008). También ha sido identificado como un gen candidato que explica las diferencias en el comportamiento agresivo de ratones (Brodkin *et al.*, 2002) y la expresión génica relacionada con el estrés en gallinas (Fallahsharoudi *et al.*, 2017). El segundo gen es el THOC2 que codifica una proteína que forma parte del complejo THO (Jimeno *et al.*, 2002). Este conjunto molecular desempeña un papel importante en la transcripción del mRNA actuando en el transporte de este hacia el citoplasma. Pese a todo, no se ha encontrado ninguna asociación con caracteres de interés en la bibliografía.

En la segunda de las regiones genómicas (34,331,053-35,394,695 bps) se localizan 8 genes (PRRG3, FATE1, CNGA2, GABRE, GABRA3, GABRQ, CETN2 y NSDHL). Entre ellos, merecen ser destacados los genes GABRE, GABRA3 y GABRQ que codifican distintas subunidades del receptor GABA tipo A. El ácido γ -aminobutírico (GABA) es el mensajero químico de tipo inhibitor más abundante en el sistema nervioso central (Cortes-Romero *et al.*, 2011) y el gen GABRA3 localizado en esta región ha sido identificado como gen candidato para la regulación del comportamiento en ratones (Miller *et al.*, 2010). Asimismo, el gen PRRG3 codifica un dominio carboxilado/gamma-carboxiglutámico dependiente de vitamina K que forma parte de una proteína transmembrana receptora del glutamato (Kulman *et al.*, 2001) y, por lo tanto, puede formar parte de la misma ruta metabólica que el GRIA3, mencionado anteriormente.

Por otra parte, el gen FATE1 parece desempeñar un papel importante en el control del desarrollo testicular temprano y la proliferación celular. Igualmente, el gen CNGA2 codifica una de las subunidades que forman parte de canales controlados por nucleótidos cíclicos. Su función se limita a los fotorreceptores retinianos y a neuronas sensoriales olfatorias (Kaupp *et al.* 2002). Finalmente, el gen CETN2 codifica una proteína de los centriolos involucrada en su división y requerida para la formación del huso acromático (Salisbury *et al.* 2002) y el NSDHL codifica un enzima localizado en el retículo endoplásmico e involucrada en la biosíntesis del colesterol (Dara, 2017).

En la tercera región (58,888,233-59,888,233 bps) se localizan 3 genes (TSC22D3, PRPS1 y FRMPD3). El gen TSC22D3 codifica una proteína con función antiinflamatoria e inmunosupresora para que glucocorticoides e interleucina 10 desempeñen sus funciones en los macrófagos (Asselin-Labat *et al.*, 2004). También, el gen PRPS1 codifica un enzima que se utiliza para la síntesis de *de novo* de nucleótidos purínicos y pirimidínicos (Becker *et al.*, 1990) y el gen FRMPD3 da lugar dos proteínas (PDZ y FERM) de unión de la célula. La primera está involucrada en la sinapsis neuronal y la segunda se relaciona con el ensamblaje entre la membrana plasmática y el citoesqueleto (Nagase *et al.*, 2001).

Por último, la cuarta de las regiones genómica (66,526,447-67,588,500 bps) no contenía ningún gen en el estado actual del mapa genético bovino. Pese a todo, las causas de la diferenciación entre poblaciones asociadas a esta región puede que tengan como origen secuencias de ADN no codificantes, que pueden actuar como reguladoras de la expresión génica.

5.2. Análisis de reducción de la variabilidad.

Los segmentos de homocigosidad o *ROH* (*Runs of Homocigosity*) se definieron inicialmente con el objetivo de estudiar la consanguinidad individual (McQuillan *et al.*, 2008). Si se estudia la variabilidad regional del porcentaje de individuos que presentan homocigosis para cada SNP, es plausible identificar regiones con una reducción significativa de la variabilidad genética. Los resultados de este análisis para las siete poblaciones analizadas a lo largo del cromosoma X se muestran en la Figura 8.

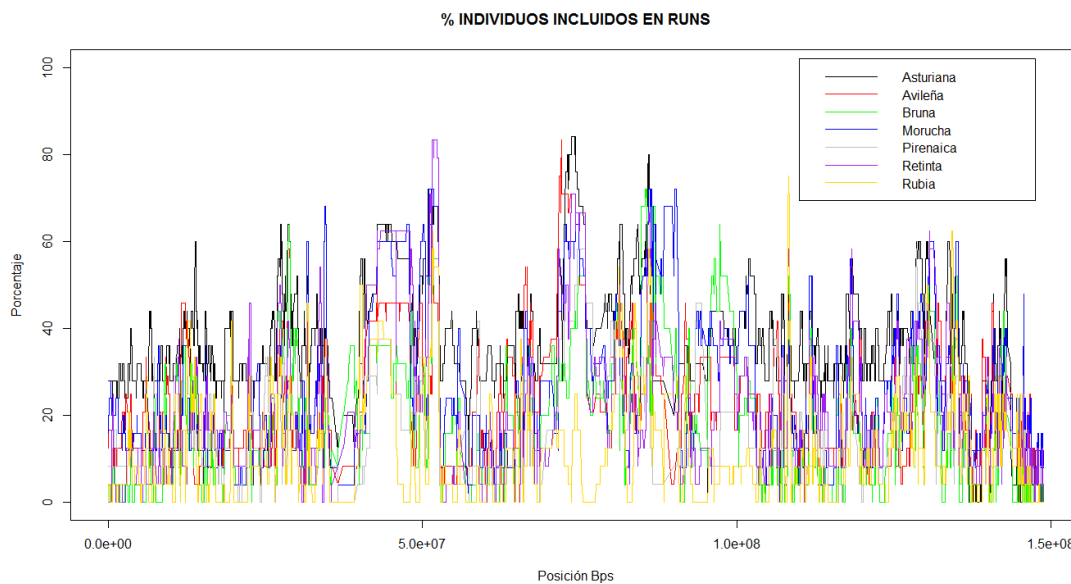


Figura 8. Resultados del porcentaje de individuos incluidos en el *ROH* para las 7 poblaciones analizadas.

En ella se puede observar como la que la mayor parte del cromosoma X presenta porcentajes de inclusión en los segmentos de homocigosis inferiores al 40%, con un valor medio del 32.3 % y una desviación típica del 9.1 %. Sin embargo, se pueden identificar 3 regiones genómicas dentro del cromosoma X en las que aparecen un mayor porcentaje de individuos con homocigosis, que se localizan en torno a las Mb 52, 74 y 86. Es destacable que la presencia de un mayor porcentaje de individuos con homocigosis aparece en la mayoría de las poblaciones analizadas. Por lo tanto, este hecho nos hace pensar que se trata de una causa estructural del genoma, que por algún motivo es menos variable en estas regiones genómicas en todas las poblaciones analizadas. Los genes presentes en cada una de estas regiones se presentan en la Tabla 3 y en el Anexo 2.

En la primera de las regiones (52 Mb) se localiza el gen *PCDH19*. Las protocaderinas juegan un papel importante en las conexiones del sistema nervioso (Sano *et al.*, 1993) y se trata de proteínas con alto grado de conservación evolutiva. De hecho, Mouresan *et al.* (2017)

identificaron una región en el cromosoma 7 altamente conservada que contenía los genes codificantes de otras proteínas de esta familia.

Tabla 3. Regiones genómicas con una señal de selección mediante el procedimiento ROH.

Start	End	GENES
51000000	53000000	PCDH19
73000000	75000000	POU3F4, CHM, POF1B y ZNF711
85000000	87000000	SLC7A3, TEX11, DLG3, GDPD2, KIF4A, PDZD11, ARR3, P2RY4, AWAT1, DGAT2L6, IGBP1, OTUD6A, AWAT2, EDA, RPL23A, FAM155B y EFNB1

Entre los genes localizados en la segunda de las regiones (74 Mb) se encuentran el gen POU3F4, perteneciente a una familia muy conservada entre homólogos de humana y ratón (Douville *et al.*, 1994), lo que evidencia una gran importancia funcional. También se ha asociado al desarrollo del oído interno, así como en el correcto desarrollo de la corteza auditiva derecha (Cai *et al.*, 2014) y que en humana su defecto implica sordera mixta no sindrómica (Bitner-Glindzicz *et al.*, 1995). Asimismo, también se localizan en esta región el gen CHM que está implicado en el tráfico intracelular (Preising y Ayuso, 2004), el gen POF1B del que se ha descrito su función en el desarrollo ovárico (Bione *et al.*, 2004) y el gen ZNF711 pese a pertenecer a un grupo muy conservado se asocia en humana con una discapacidad cognitiva (Tarpey *et al.*, 2009). Este último, corresponde a la misma familia que el gen ZNF185 (*zinc finger protein 185 with LIM domain*) situado entre las bases 35452632 y 35491147, no descrito en el apartado anterior pero que se sitúa justo al límite con la segunda región identificada con el procedimiento de diferenciación entre poblaciones.

Por último, en la última región se identifican hasta 17 genes. Entre ellos merecen ser destacados los genes que presentan alguna similitud funcional o pertenecen al mismo grupo familiar que los obtenidos en los resultados en el análisis de diferenciación entre poblaciones. El primero con estas características es el TEX11, que corresponde con un gen de la misma familia que el gen FATE1. El TEX11 está altamente conservado en mamíferos y su defecto se ha asociado a infertilidad, puesto que desempeña un papel importante en el desarrollo germinal de diferentes etapas espermatogónicas y es fundamental para la meiosis (Tang *et al.*, 2011).

El segundo de los genes emparentados con los anteriormente descritos es el PDZD11, que se relaciona con el FRMPD3. El PDZD11 es un dominio de interacción con otras proteínas, y se ha demostrado su implicación en la absorción intestinal de vitaminas (Nabokina *et al.*, 2011) y su

papel como eficaz estabilizador de las nectinas de células epiteliales (Guerrera *et al.*, 2016). Del mismo modo, en esta misma región se encuentra el gen *DLG3* estrechamente ligado a el *GRIA3* por su papel en la sinapsis excitatoria, concretamente en los receptores NMDA del glutamato y que en humana la alteración de este curso con retraso mental no sindrómico (Tarpey *et al.*, 2004).

Seguidamente, el gen *ARR3* se puede ligar biológicamente, pese a no estar integrado en el mismo grupo, con el gen *CNGA2*. Este gen juega un papel en la adaptación de los conos frente a la luz y la oscuridad (Sterne-Marr, *et al.* 1993) y se ha asociado, en humana, a una mutación que cursa con miopía alta de comienzo temprano en mujeres (Xiao *et al.*, 2016). Si nos referimos a enfermedades oculares, se puede establecer un nexo con el gen *CHM* citado con anterioridad donde su alteración causa una enfermedad de degeneración retiniana llamada coroideremia (Preising y Ayuso, 2004).

Para terminar, otro gen que merece la pena destacar es el gen *KIF4A* que actúa en la regulación de la condensación cromosómica y en la segregación durante la división por mitosis (Hou *et al.*, 2018) y que por lo tanto interviene en funciones biológicas relacionadas con el gen *CETN2* comentado con anterioridad. También con implicación nerviosa aparecen *IGBP1* que cuyo defecto causa agenesia del cuerpo calloso con retraso mental, coloboma ocular y micrognatia (Graham *et al.*, 2003) y *EFNB1* que desarrolla un papel en la adhesión celular y desarrollo o mantenimiento del sistema nervioso (Twigg *et al.*, 2004). Finalmente, un gen que puede tener una relación estrecha con los QTL de estas regiones es el gen *RPL23A*. Se puede relacionar con la conformación de la ubre puesto que forma parte de la subunidad mayor ribosomal (60S) (Pool *et al.*, 2002) y se le ha establecido como candidato a afectar al porcentaje de proteína y grasa en la leche (Cui *et al.*, 2014).

5.3. Extensión del desequilibrio de ligamiento

El valor medio del estadístico *nSL* fue de -0.338 y su desviación típica de 0.618. Si seguimos la misma estrategia que en los apartados anteriores, los umbrales de relevancia serían 2.75 y -3.42. Como se observa en la Figura 9, no existe ninguna región genómica que supere estos criterios de relevancia. Este resultado es coherente con el obtenido por González-Rodríguez *et al.* (2016), que apenas detecta dos regiones asociadas a señales de selección con este procedimiento en los cromosomas autosómicos. Según González-Rodríguez *et al.* (2016), los métodos basados en la extensión del desequilibrio de ligamiento detectan señales de selección recientes.

Por lo tanto, o bien, la selección reciente no ha afectado ninguna región genómica del cromosoma X, o los genes implicados en los caracteres sometidos a selección tienen un efecto pequeño, debido a que se encuentran integrados en procesos de interacción epistática entre genes que impiden la detección de estas señales.

Los resultados del análisis de la longitud de los haplotipos o de extensión del desequilibrio de ligamiento se presentan en la Figura 9.

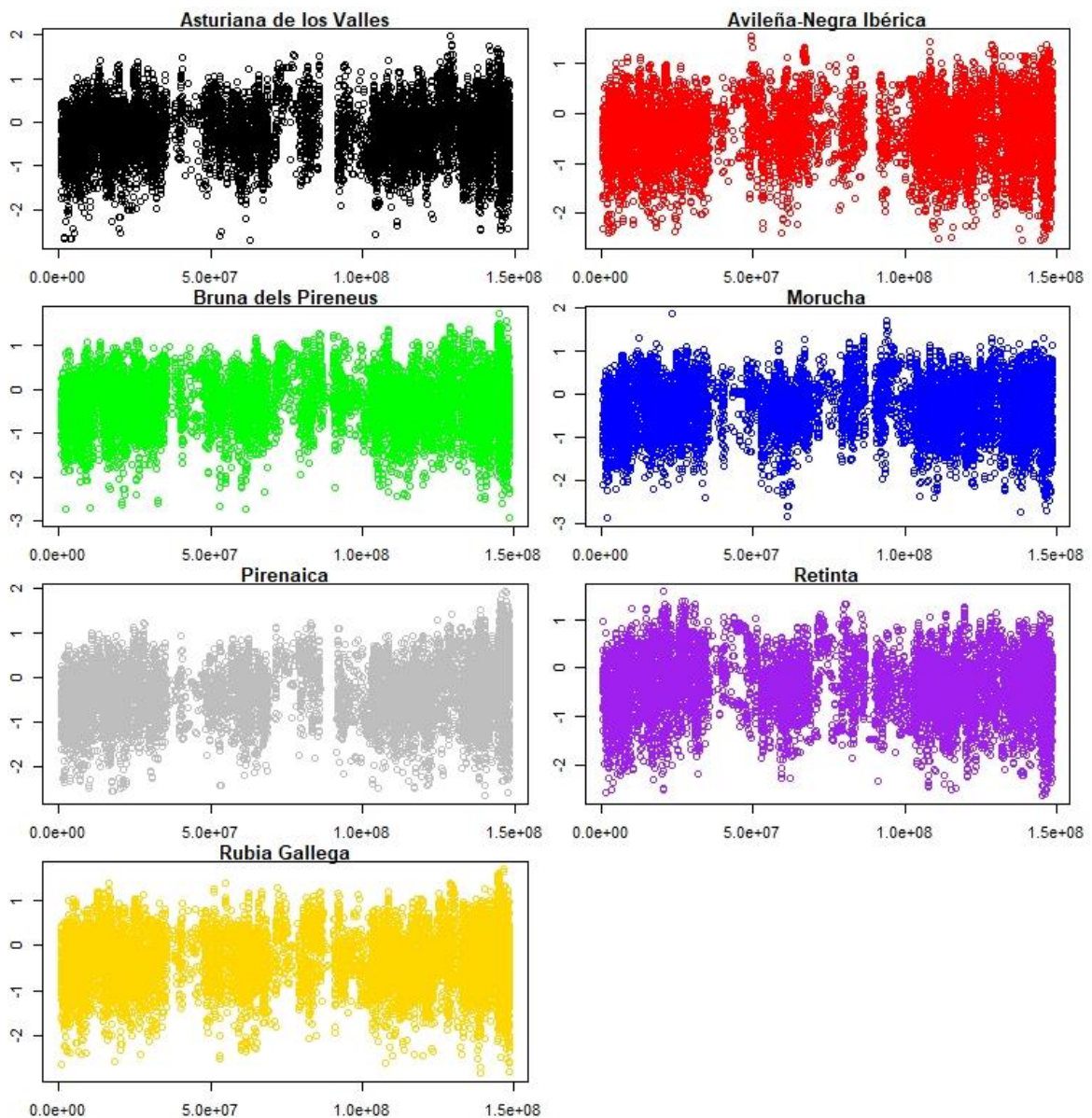


Figura 9. Resultados obtenidos para las 7 poblaciones con *nSL*.



5.4. Concordancia con QTL en vacuno.

Finalmente, se ha procedido a comparar los resultados de las regiones anteriormente mencionadas con la base de datos de *cattleQTLdb* depositada en la página web (<https://www.animalgenome.org/QTLdb/>). En concreto, se han encontrado varios QTL en las regiones asociadas a ambos métodos que se resumen en el Anexo 3.

Los caracteres afectados por estos QTL son la morfología de la ubre (Cole *et al.*, 2011), el recuento de células somáticas (Cole *et al.*, 2011; Wang *et al.*, 2015), la profundidad del cuerpo y anchura de la grupa (Cole *et al.*, 2011), la precocidad (Crispim *et al.*, 2015), la duración de la vida productiva y el mérito total (Cole *et al.*, 2011) y la calidad seminal, diámetro escrotal y edad a la pubertad (Fortes *et al.*, 2013).

Estos resultados sugieren que las regiones en las que se han detectado señales de selección contienen genes que afectan a caracteres de interés en la producción de vacuno de carne, sobre todo en los aspectos morfológicos o de crecimiento y en los reproductivos. Estos resultados son coherentes con los genes descritos en las secciones anteriores y con los resultados de González-Rodríguez *et al.* (2016) para los cromosomas autosómicos en estas mismas poblaciones.



6. Conclusiones

Tras la finalización de este Trabajo Fin de Grado en Veterinaria pueden ser destacadas una serie de conclusiones:

- Las regiones y señales detectadas por los tres métodos utilizados son diferentes entre sí, posiblemente debido a que identifican procesos de selección en diferente escala evolutiva.
- Se encuentran diferencias en la magnitud de las señales entre la región pseudoautosómica y no pseudoautosómica, puesto que la primera actúa como un cromosoma autosómico.
- La gran cantidad de genes identificados implicados en el desarrollo del sistema nervioso hace pensar en su posible relación de estas regiones con el temperamento del animal.
- Se han detectado genes implicados en el correcto desarrollo genital, tanto masculino como femenino.
- Los datos analizados de las siete poblaciones muestran unos valores similares debido a su proximidad evolutiva.



Conclusions

After the completion of this End of Degree Project in Veterinary can be highlighted a series of conclusions:

- The regions and signals detected by the three methods used are different from each other, possibly because they identify selection processes in different evolutionary scale.
- Differences are found in the magnitude of the signals between the pseudoautosomal and non-pseudoautosomal regions, since the former acts as an autosomal chromosome.
- The large number of identified genes involved in nervous system development suggests the possible relationship of these regions with the animal's temperament.
- Genes involved in the correct genital development, both male and female, have been detected.
- The analyzed data of the seven populations show similar values due to their evolutionary proximity.



7. Valoración personal

La valoración global tras la realización de este Trabajo Fin de Grado en Veterinaria integrado en la dotación de una Beca de Colaboración la puedo considerar muy gratificante. El esfuerzo realizado para llevarlo a cabo ha supuesto un reto personal del que se puede destacar la gran cantidad de conocimientos y oportunidad de análisis de datos a los que he tenido alcance. En primer lugar, valoro especialmente el afianzar y reforzar las bases en genética, con especial hincapié en aspectos relacionados con la consanguinidad, así como también, el retomar el uso de la estadística. Integrando a estas dos se basa la elaboración de este trabajo, por ello, el haber tenido la oportunidad de tener acceso a manuales de programas con gran capacidad de desarrollo en el análisis de datos como R y haber aprendido las nociones básicas, me ha abierto la posibilidad del uso de estos en el futuro. Posiblemente, el aspecto de mayor dificultad haya sido cuando en la elaboración de las gráficas que representen los resultados obtenidos, invertías muchas horas y la utilización inadecuada de los comandos te bloqueaba en un paso, pero por el contrario, una vez que se conseguía finalizar con éxito suponía una gran satisfacción personal y la mejor forma de poder aplicar los conocimientos adquiridos.

Para terminar, agradecer la gran disponibilidad y colaboración de mi director Luis Varona Aguado ante cualquier problema surgido en el desarrollo de este trabajo y que sin él no hubiera sido posible. De igual modo, reconocer en esta primera toma de contacto con el análisis genómico a todo el departamento de Genética Cuantitativa y Mejora Genética, por permitirme participar en este proyecto en un ámbito de la genética que desconocía totalmente y que no me ha dejado indiferente.

8. Bibliografía

Acero, P. (2009). Planificación y manejo de la explotación de vacuno de carne (Tomo II). *Prácticas agropecuarias*, 39-48.

Ajmone-Marsan, P. (2010). On the origin of cattle: how aurochs became cattle and colonized the world. *Evolutionary Anthropology: Issues, News, and Reviews*, 19(4), 148-157.

Asselin-Labat, M.L., David, M., Biola-Vidamment, A., Lecoeuche, D., Zennaro, M.C., Bertoglio, J., y Pallardy, M. (2004). GILZ, a new target for the transcription factor FoxO3, protects T lymphocytes from interleukin-2 withdrawal-induced apoptosis. *Blood*, 104, 215-223.

Becker, M. A., Heidler, S. A., Bell, G. I., Seino, S., Le Beau, M. M., Westbrook, C. A.,...Palella, T. D. Cloning of cDNAs for human phosphoribosylpyrophosphate synthetases 1 and 2 and X chromosome localization of PRPS1 and PRPS2 genes. *Genomics*, 8, 555-561.

Bione, S., Rizzolio, F., Sala, C., Ricotti, R., Goegan, M., Manzini, M. C.,...Toniolo, D. (2004). Mutation analysis of two candidate genes for premature ovarian failure, DACH2 and POF1B. *Hum. Reprod.*, 19, 2759-2766.

Biscarini, F., Cozzi, P., Gaspa, G., y Marras, G. (2018). detectRUNS: Detect Runs of Homozygosity and Runs of Heterozygosity in Diploid Genomes. <https://rdr.io/cran/detectRUNS/>

Bitner-Glindzicz, M., Turnpenny, P., Höglund, P., Kääriäinen, H., Sankila, E.M., van der Maarel, S.M.,...Pembrey, M. (1995). Further mutations in Brain 4 (POU3F4) clarify the phenotype in the X-linked deafness, DFN3. *Human Molecular Genetics*, 4(8),1467-1469.

Brodkin, E.S., Goforth, S.A., Keene, A.H., Fossella, J.A., y Silver, L.M. (2002). Identification of quantitative trait Loci that affect aggressive behavior in mice. *The Journal of Neuroscience*, 22(3), 1165–1170.

Browning, B.L., y Browning, S.R. (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *American Journal of Human Genetics*, 84(2), 210-223.

Cai, D.C., Fonteijn, H., Guadalupe, T., Zwiers, M., Wittfeld, K., Teumer, A.,...Hagoort, P. (2014). A genome-wide search for quantitative trait loci affecting the cortical surface area and thickness of Heschl's gyrus. *Genes, Brain and Behavior*, 13, 675-685.

Cañas-Álvarez, J.J., González-Rodríguez, A., Munilla, S., Varona, L., Díaz, C., Baro, J.A.,... Piedrafita, J.(2015). Genetic diversity and divergence among Spanish beef cattle breeds assessed by a bovine high-density SNP chip. *American Society of Animal Science*, 93, 5164–5174.

Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., Lawlor, T.J., Crooker, B.A.,...Da, Y. (2011). Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics*, 12, 408.

Cortes-Romero, C., Galindo, F., Galicia-Isasmendi, S., y Flores, A. (2011). GABA: ¿dualidad funcional? Transición durante el neurodesarrollo. *Rev. Neurol*, 52 (11), 665-675.

Crispim, A.C., Kelly, M.J., Guimaraes, S.E.F., E Silva, F.F., Fortes, M.R.S., Wenceslau, R.R., y Moore, S. (2015). Multi-Trait GWAS and New Candidate Genes Annotation for Growth Curve Parameters in Brahman Cattle. *PLoS one*, 10(10): e0139906.

Cui, X., Hou, Y., Yang, S., Xie, Y., Zhang, S., Zhang, Y.,...Sun, D. (2014). Transcriptional profiling of mammary gland in Holstein cows with extremely different milk protein and fat percentage using RNA sequencing. *BMC Genomics*, 15, 226.

Dara, S. (2017). Child syndrome: a rare genetic disorder. *American Journal of Advances in Nursing Research*, 4(2), 80-83.

Das, P.J., Chowdhary, B.P., y Raudsepp, T. (2009). Characterization of the Bovine Pseudoautosomal Region and Comparison with Sheep, Goat, and Other Mammalian Pseudoautosomal Regions. *Cytogenet Genome Res.*,126(1-2), 139-147.

Decker, J.E., McKay, S.D., Rolf, M.M., Kim, J.W.W., Molina-Alcalá, A., Sonstegard, T.S.,...Taylor J.F. (2014). Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. *PLoS Genetics*, 10(3): e1004254.

Douville, P.J., Atanasoski, S., Tobler, A., Fontana, A., y Schwab, M.E. (1994). The brain-specific POU-box gene *Brn4* is a sex-linked transcription factor located on the human and mouse X chromosomes. *Mammalian Genome*, 5(3), 180-182.

Edwards, C.J., Magee, D.A., Park, S.D.E., McGettigan, P.A., Lohan, A.J., y Murphy, A. (2010). A Complete Mitochondrial Genome Sequence from a Mesolithic Wild Aurochs (*Bos primigenius*). *PLoS ONE*, 5(2): e9255.

Fallahsharoudi, A., de Kock, N., Johnsson, M., Bektic, L., Ubhayasekera, S.J., Bergquist, J., Wright, D., y Jensen, P. (2017). QTL mapping of stress related gene expression in a cross between domesticated chickens and ancestral red junglefowl. *Molecular and Cellular Endocrinology*, 446, 52-58.

FAO. (1999). The global strategy for the management of farm animal genetic resources. *Informe ejecutivo*.

Fay, J.C., y Wu, C.I. (2000). Hitchhiking under positive Darwinian selection. *Genetics*, 155,1405-1413.

Ferrer-Admetlla, A., Liang, M., Korneliussen, T.,y Nielsen, R. (2014). On Detecting Incomplete Soft or Hard Selective Sweeps Using Haplotype Structure. *Molecular Biology and Evolution*, 31(5), 1275–1291.

Fortes, M.R.S., Reverter, A., Kelly, M., McCulloch, R., y Lehnert, S.A. (2013). Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species. *Andrology*, 1(4), 644-650.

Fu, Y.X. y Li, W.H., (1993). Statistical tests of neutrality of mutations. *Genetics*, 133(3), 693–709.

González-Rodríguez, A. (2015). *Detección de regiones del genoma involucradas en los procesos de selección y diferenciación en las poblaciones autóctonas de vacuno de carne*. Universidad de Zaragoza, España.

González-Rodríguez, A., Munilla, S., Mouresan, E.F., Cañas-Alvarez, J.J., Díaz, C., Piedrafita, J.,... Varona, L. (2016). On the performance of tests for detection of signatures of selection: a case study with the Spanish autochthonous beef cattle populations. *Genetics, Selection, Evolution*, 48, 81.

Götherström, A., Anderung, C., Hellborg, L., Elburg, R., Smith, C., Bradley, D.G., y Ellegren, H. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc Biol Sci*, 272, 2345–2350.

Guerrera, D., Shah, J., Vasileva, E., Sluysmans, S., Méan, I., Jond, L.,...Citi, S. (2016). PLEKHA7 Recruits PDZD11 to Adherens Junctions to Stabilize Nectins. *The journal of biological chemistry*, 291(21), 11016–11029.

Graham Jr., J. M., Wheeler, P., Tackels-Horne, D., Lin, A. E., Hall, B. D., May, M.,...Cox, T. C. (2003). A new X-linked syndrome with agenesis of the corpus callosum, mental retardation, coloboma, micrognathia, and a mutation in the alpha 4 gene at Xq13. *Am. J. Med. Genet.*, 123a, 37-44.

Hanchard, N.A., Rockett, K.A., Spencer, C., Coop, G., Pinder, M., Jallow, M.,...Kwiatkowski, D.P. (2006). Screening for Recently Selected Alleles by Analysis of Human Haplotype Similarity. *American Journal of Human Genetics*, 78(1), 153–159.

Helmer, D., Gourichon, L., Monchot, H., Peters, J., y Saña-Segui, M. (2005). *Identifying early domestic cattle from Pre-Pottery Neolithic sites on the middle Euphrates using sexual dimorphism*. En: *The First Steps of Animal Domestication*. Oxbow Books: Oxford, Reino Unido. Eds: Vigne, J.D., Peters, J., y Helmer, D.

Hou, P., Jiang, T., Chen, F., Shi, P., Li, H., Bai, J., y Song, J. (2018). KIF4A facilitates cell proliferation via induction of p21-mediated cell cycle progression and promotes metastasis in colorectal cancer. *Cell Death and Disease*, 9, 477.

Jimeno, S., Rondón, A.G., Luna, R., y Aguilera A. (2002). The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO journal*, 21(13), 3526–3535.

Jordana, J., Pelegrín, M., y Piedrafita, J. (1991). Relaciones genéticas en bovinos españoles obtenidas a partir del estudio de caracteres morfológicos. *ITEA*, 87A, 50–64.

Kaupp, U.B., y Seifert, R. (2002). Cyclic nucleotide-gated ion channels. *Physiological rev.*, 82(3), 769-824.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*, Cambridge University Press, Reino Unido.

Kulman, J.D., Harris, J.E., Xie, L., y Davie, E. W. (2001). Identification of two novel transmembrane gamma-carboxyglutamic acid proteins expressed broadly in fetal and adult tissues. *Proc. Nat. Acad. Sci.*, 98(4), 1370-1375.

Martín-Burriel, I., Rodellar, C., Cañón, J., Cortés, O., Dunner S., Landi, V.,...Delgado, J.V. (2011). Genetic diversity, structure, and breed relationships in Iberian cattle. *J. Anim. Sci*, 89, 893–906.

- Maxfield, K.E., Taus, P.J., Corcoran, K., Wooten, J., Macion, J., Zhou, Y.,...Whitehurst, A.W. (2015). Comprehensive functional characterization of cancer–testis antigens defines obligate participation in multiple hallmarks of cancer. *Nature Communications*, 6, 8840.
- McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L.,... Wilson, J.F. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics* 83(3), 359–372.
- Miller, B.H., Schultz, L.E., Long, B.C., y Pletche, M.T. (2010). Quantitative trait locus analysis identifies *Gabra3* as a regulator of behavioral despair in mice. *Mamm Genome*, 21, 247–257.
- Mouresan, E.F., González-Rodríguez, A., Munilla, S., Moreno, C., Altarriba, J., Díaz, C.,...Varona, L. (2017). Detección de regiones genómicas con elevado desequilibrio de ligamiento en poblaciones de vacuno de carne españolas con análisis de BovineHD BeadChip. *Archivos de Zootecnia*, 66(253), 59-65.
- Nabokina, S., Subramanian, V., y Said, H. (2011). Association of PDZ-containing protein PDZD11 with the human sodium-dependent multivitamin transporter. *Am J Physiol Gastrointest Liver Physiol*, 300, G561–G567.
- Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R., y Ohara, O. (2001). Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res*, 8, 85-95.
- Oleksyk, T.K., Zhao, K., De La Vega, F.M., Gilbert, D.A., O'Brien, S.J., y Smith, M.W. (2008) Identifying Selected Regions from Heterozygosity and Divergence Using a Light-Coverage Genomic Dataset from Two Human Populations. *PLoS ONE*, 3(3): e1712.
- Oleksyk, T.K, Smith, M.W., y O'Brien, S.J. (2010). Genome-wide scans for footprints of natural selection. *Philosophical transaction of the Royal Society of London. Series B, Biological sciences*, 365, 185-205.
- Peña, A. (2010). *Caracterización estructural y funcional del complejo THO de Saccharomyces Cerevisiae*. Universidad Autónoma de Madrid, España.
- Pool, M. R., Stumm, J., Fulga, T. A., Sinning, I., y Dobberstein, B. (2002) Distinct modes of signal recognition particle interaction with the ribosome. *Science*, 297, 1345-1348.
- Preisig, M., y Ayuso, C. (2004). Rab escort protein 1 (REP1) in intracellular traffic: a functional and pathophysiological overview. *Ophthalmic Genetics*, 25(2), 101–110.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D.,...Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Human Genetics*, 81, 559-575.
- Qanbari, S., y Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livestock Science*, 116, 133-143.
- Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F.,...Lander, E.S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419(6909), 832–837.

Salisbury, L., Suino, K., Busby, R., y Springett, M. (2002). Centrin-2 Is Required for Centriole Duplication in Mammalian Cells. *Current Biology*, 12, 1287–1292.

Sano, K., Tanihara, H., Heimark, R.L., Obata, S., Davidson, M., John, T., Taketani, S., y Suzuki, S. (1993). Protocadherins: a large family of cadherin-related molecules in central nervous system. *The EMBO journal*, 12(6):2249-56.

Sterne-Marr, R., Gurevich, V.V., Goldsmith, P., Bodine, R.C., Sanders, C., Donoso, L.A., y Benovic, J.L. (1993). Polypeptide variants of beta-arrestin and arrestin3. *J Biol Chem.*, 68(21), 15640-15648.

Subdirección General de Productos Ganaderos. (2016). *Caracterización del sector vacuno de carne en España*. Ed: Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente.

Subdirección general de productos ganaderos. (2017a). *El sector de la carne de vacuno en cifras*. Ed: Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente.

Subdirección General de Productos Ganaderos. (2017b). *Situación de mercado sector vacuno de carne*. Ed: Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585–595.

Tang, L., Zeng, W., Clark, R.K., y Dobrinski, I. (2011). Characterization of the porcine testis-expressed gene 11 (Tex11). *Spermatogenesis* 1(2), 147-151.

Tarpey, P. S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy,...Stratton, M. (2009). A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nature Genet.*, 41, 535-543.

Tarpey, P., Parnau, J., Blow, M., Woffendin, H., Bignell, G., Cox, C.,...Raymond, F.L. (2004). Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation. *Am J Hum Genet.*, 75(2), 318-24.

Terasawa, E., y Fernandez, D.L. (2001). Neurobiological mechanisms of the onset of puberty in primates. *Endocrine Reviews*, 22(1), 111-151.

Twigg, S.R.F., Kan, R., Babbs, C., Bochukova, E.G., Robertson, S.P., Wall, S. A., Morriss-Kay, G.M., y Wilkie, A.O.M. (2004). Mutations of ephrin-B1 (EFNB1), a marker of tissue boundary formation, cause craniofrontonasal syndrome. *Proc. Nat. Acad. Sci.*, 101, 8652-8657.

Upadhyay, M.R., Chen, W., Lenstra, J.A., Goderic, C.R.J., MacHugh, D.E., Park, S.D.E.,... Crooijmans, R.P.M.A. (2017). Genetic origin, admixture and population history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity*, 118, 169–176.

Utsonomiya, Y.T., Pérez-O’Brien, A.M., Sonstegard, T.S., Solkner, J., y García, J.F. (2015). Genomic data as the “hitchhiker’s guide” to cattle adaptation: tracking the milestones of past selection in the bovine genome. *Frontiers in Genetics*, 6, 36.

Van Vuure, C. (2001). *Retracting the Auroch*. Universitat Paderborn: Hoxter, Alemania.



Vitalis, R., Gautier, M., Dawson, K.J., y Beaumont, M.A. (2014). Detecting and Measuring Selection from Gene Frequency Data. *Genetics*, 196, 799–817.

Voight, B.F., Kudravalli, S., Wen, X., y Pritchard, J.K. (2006). A map of recent positive selection in the human genome. *PLOS Biology*, 4(3), 0446–0458.

Wang, X., Ma, P., Liu, J., Zhang, Q., Zhang, Y., Ding, X,...Yu, Y. (2015). Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC genetics*, 16, 111.

Womack, J.E. (2005). Advances in livestock genomics: opening the barn door. *Genome Res*, 15, 1699-1705.

Wright, S. (1943). Isolation by distance. *Genetics*, 28(2), 114–138.

Xiao, X., Li, S., Jia, S., Guo, X., y Zhang, Q. (2016). X-linked heterozygous mutations in ARR3 cause female-limited early onset high myopia. *Molecular Vision*, 22, 1257-1266.

Zhu, W., Fan, Z., Zhang, C., Guo, Z., Zhao, Y., Zhou, Y.,...Xiao, J. (2008). A Dominant X-Linked QTL Regulating Pubertal Timing in Mice Found by Whole Genome Scanning and Modified Interval-Specific Congenic Strain Analysis. *PLOS ONE*, 3(8): e3021.

ANEXO 1. Genes detectados mediante el procedimiento *SELESTIM*.

SIMBOLO	NOMBRE	START	END
GRIA3	glutamate ionotropic receptor AMPA type subunit 3	7180735	7482667
THOC2	THO complex subunit 2	7580441	7696156
PRRG3	proline rich and Gla domain 3	34402712	34407507
FATE1	fetal and adult testis expressed 1	34423522	34429968
CNGA2	cyclic nucleotide gated channel alpha 2	34433641	34450338
GABRE	gamma-aminobutyric acid type A receptor epsilon subunit	34660685	34674488
GABRA3	gamma-aminobutyric acid type A receptor alpha3 subunit	34852875	35094442
GABRQ	gamma-aminobutyric acid type A receptor theta subunit	35299889	35314450
CETN2	centrin 2	35382935	35386969
NSDHL	sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating	35387262	35413263
TSC22D3	TSC22 domain family protein 3	59707292	59772737
PRPS1	ribose-phosphate pyrophosphokinase 1	59804943	59827385
FRMPD3	FERM and PDZ domain containing 3	59847736	59915953



ANEXO 2. Genes detectados mediante el procedimiento ROH.

SIMBOLO	NOMBRE	START	END
PCDH19	protocadherin 19	51118208	51246606
POU3F4	POU class 3 homeobox 4	73224661	73225752
CHM	rab escort protein 1	74152508	74398550
POF1B	actin binding protein	74869019	74967178
ZNF711	zinc finger protein 711	74973931	74999240
SLC7A3	solute carrier family 7 member 3	85010185	85015644
TEX11	testis expressed 11	85042821	85229518
DLG3	discs large MAGUK scaffold protein 3	85245062	85298606
GDPD2	glycerophosphodiester phosphodiesterase domain containing 2	85309754	85319162
KIF4A	chromosome-associated kinesin KIF4A	85321685	85449499
PDZD11	PDZ domain containing 11	85450103	85453009
ARR3	arrestin 3	85458940	85470775
P2RY4	pyrimidinergic receptor P2Y4	85481339	85482436
AWAT1	acyl-CoA wax alcohol acyltransferase 1	85499886	85508546
DGAT2L6	diacylglycerol O-acyltransferase 2 like 6	85556424	85577736
IGBP1	immunoglobulin binding protein	185588917	85635721
OTUD6A	OTU deubiquitinase 6A	85677239	85678114
AWAT2	acyl-CoA wax alcohol acyltransferase 2	85700206	85702893
EDA	ectodysplasin A	85708003	86099973
RPL23A	60S ribosomal protein L23a	85898898	85899424
FAM155B	family with sequence similarity 155 member B	86163690	86189860
EFNB1	ephrin B1	86944028	86956186



ANEXO 3. QTL asociados a las regiones que muestran huella de selección.

REGIÓN	TIPO DE QTL	RELACIÓN
6648233 - 7648233	UBRE	Fijación de la ubre, Altura de la Ubre, Hendidura de la Ubre, Profundidad de la Ubre, Fijación de la Ubre, Longitud de pezones y Localización de pezones
	MAMITIS	Recuento de Células Somáticas
	CRECIMIENTO	Profundidad del Cuerpo y Anchura de la Grupa
	HISTORIAL DE VIDA	Capacidad de Transmisión Predicha y Merito Total
	VIDA PRODUCTIVA	Duración de la Vida Productiva
34331053 - 35394695	REPRODUCTIVO	Porcentaje Normal de Espermatozoides, Diámetro Escrotal, Edad de pubertad
51000000 - 53000000	UBRE	Profundidad de la Ubre y Longitud de pezones
	MAMITIS	Recuento de Células Somáticas
	HISTORIAL DE VIDA	Merito Total
	VIDA PRODUCTIVA	Duración de la Vida Productiva
	REPRODUCTIVO	Porcentaje Normal de Espermatozoides, Diámetro Escrotal, Edad de pubertad
58888233 - 59888233	UBRE	Fijación de la ubre, Altura de la Ubre, Hendidura de la Ubre y Profundidad de la Ubre
	MAMITIS	Recuento de Células Somáticas
	HISTORIAL DE VIDA	Merito Total
	VIDA PRODUCTIVA	Duración de la Vida Productiva
	REPRODUCTIVO	Porcentaje Normal de Espermatozoides, Diámetro Escrotal, Edad de pubertad
66526447 - 67588500	UBRE	Fijación de la ubre, Altura de la Ubre, Hendidura de la Ubre y Profundidad de la Ubre
	CRECIMIENTO	Profundidad del Cuerpo, Anchura de la Grupa y Precocidad
	HISTORIAL DE VIDA	Capacidad de Transmisión Predicha
	REPRODUCTIVO	Porcentaje Normal de Espermatozoides, Diámetro Escrotal, Edad de pubertad
73000000 - 75000000	-	
85000000 - 87000000	UBRE	Profundidad de la Ubre, Longitud de pezones y Localización de pezones
	MAMITIS	Recuento de Células Somáticas
	CRECIMIENTO	Profundidad del Cuerpo
	VIDA PRODUCTIVA	Duración de la Vida Productiva
	HISTORIAL DE VIDA	Merito Total
	REPRODUCTIVO	Porcentaje Normal de Espermatozoides, Diámetro Escrotal, Edad de pubertad