



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

UAB
Universitat Autònoma
de Barcelona



**MÁSTER EN MEJORA GENÉTICA ANIMAL Y BIOTECNOLOGÍA
DE LA REPRODUCCIÓN**

Análisis genómico de la depresión endogámica para caracteres reproductivos en dos variedades de cerdo ibérico

Trabajo final de máster

Carlos Hervás Rivero

Universidad Politécnica de Valencia,
Julio 2020

Director: Luis Varona Aguado

Análisis genómico de la depresión endogámica para caracteres reproductivos en dos variedades de cerdo ibérico

Resumen

La depresión endogámica afecta en mayor medida a los caracteres reproductivos, como el tamaño de camada en porcino. Recientemente, se ha demostrado que el determinismo de la depresión endogámica es heterogéneo a lo largo del genoma en varios caracteres de interés. Por este motivo, el objetivo de este trabajo es realizar un análisis genómico a lo largo del genoma autosómico porcino para detectar las regiones genómicas que controlan la depresión endogámica para tamaño de camada en dos variedades de cerdo ibérico (Entrepelado y Retinto). La base de datos ha consistido en 2.167 (367 hembras) y 2.101 (345 hembras) registros de tamaños de camada (número total de nacidos y número de nacidos vivos) para las variedades Entrepelado y Retinto respectivamente. Todas las hembras fueron genotipadas con el chip GENESEEK GGP Porcine HD 70K (*Illumina*), en el que se ha efectuado un filtrado standard que eliminó a los marcadores SNP con una frecuencia alélica menor a 0.05. El estudio genómico se realizó utilizando dos procedimientos: 1) la resolución de un modelo mixto single-step con el programa BLUPF90 para cada región genómica, que incluye dos covariadas, la primera con la homocigosidad de la región genómica considerada y la segunda con la homocigosidad del resto del genoma y 2) el programa HAPLOFINDER que identifica regiones de homocigosidad (ROH) asociadas a la reducción del rendimiento fenotípico mediante la implementación de un modelo mixto para cada ROH detectada en la población. Los resultados del estudio han permitido detectar varias regiones con mayor influencia en la depresión endogámica en el tamaño de camada del porcino, aunque no se han encontrado coincidencias relevantes ni entre las poblaciones ni entre los procedimientos.

Palabras clave: cerdo ibérico, ROH, tamaño de camada, depresión endogámica

Genomic analysis of the inbreeding depression for reproductive traits in two varieties of iberian pig

Abstract

Inbreeding depression is expected to be more noticeable in fitness related traits, such as pig litter size. Recent studies have suggested that the genetic determinism of inbreeding depression may be heterogeneous along the genome. Therefore, the objective of this study is to perform a genomic scan along the pig autosomal genome to detect the genomic regions that controls the inbreeding depression for litter size in two varieties of Iberian pig (Entrepelado and Retinto). The datasets consist of 2,167 (367 sows) and 2,101 (345 sows) litter size (Total Number Born and Number Born Alive) records for the Entrepelado and Retinto varieties, respectively. All sows were genotyped with the GENESEEK GGP Porcine HD 70K (Illumina), where a standard filtering was practiced retaining only SNP markers with minor allele frequency over 0.05. The genomic scan was developed with two alternative strategies: 1) a procedure that involves solving a sequence of mixed model equations that include covariates for the homozygosity at each specific genomic region and at the rest of the genome, and 2) with the HAPLOFINDER software that identifies Runs of Homozygosity (ROH) associated with a reduction of the phenotypic performance by solving a mixed model for each detected ROH. The results of the study have been able to identify several genomic regions associated with inbreeding depression of litter size in the Iberian pigs, although very low coincidence between populations or methods was found.

Key words: iberian pig, ROH, litter size, inbreeding depression

Anàlisi genòmic de la depressió endogàmica per caràcters reproductius en dues varietats de porc ibèric

Resum

La depressió endogàmica afecta en major mesura als caràcters reproductius, com la mida de ventrada en porcí. Recentment, s'ha demostrat que el determinisme de la depressió endogàmica és heterogeni al llarg de l'genoma en diversos caràcters d'interès. Per aquest motiu, l'objectiu d'aquest treball és realitzar un anàlisi genòmic al llarg de l'genoma autosòmic porcí per detectar les regions genòmiques que controlen la depressió endogàmica per mida de ventrada en dues varietats de porc ibèric (entrepelat i Retinto). La base de dades ha consistit en 2.167 (367 femelles) i 2.101 (345 femelles) registres de mides de ventrada (nombre total de nascuts i nombre de nascuts vius) per a les varietats entrepelat i Retinto respectivament. Totes les femelles van ser genotipades amb el xip GENESEEK GGP Porcine HD 70k (Illumina), en el qual s'ha efectuat un filtrat estàndard que elimina als marcadors SNP amb una freqüència al·lèlica menor a 0.05. L'estudi genòmic es va realitzar utilitzant dos procediments: 1) la resolució d'un model mixt single-step amb el programa BLUPF90 per a cada regió genòmica, que inclou dues covariades, la primera amb la homocigosidad de la regió genòmica considerada i la segona amb la homocigosidad de la resta de l'genoma i 2) el programa HAPLOFINDER que identifica regions de homocigosidad (ROH) associades a la reducció de l'rendiment fenotípic mitjançant la implementació d'un model mixt per a cada ROH detectada en la població. Els resultats de l'estudi han permès detectar diverses regions amb una forta influència en la depressió endogàmica en la mida de ventrada de l'porcí, encara que no s'han trobat coincidències rellevants ni entre les poblacions ni entre els procediments.

Paraules clau: porc ibèric, ROH, mida de ventrada, depressió endogàmica

ÍNDICE

1. INTRODUCCIÓN	6
1.1. El cerdo Ibérico	6
1.1.1. <i>Diversidad Genética</i>	7
1.1.2. <i>Características Reproductivas del Cerdo Ibérico</i>	8
1.2. Consanguinidad y Depresión Consanguínea.	9
1.2.1. <i>Consanguinidad genealógica</i>	10
1.2.2. <i>Coefficientes de consanguinidad y de parentesco moleculares</i>	11
1.2.3. <i>Depresión consanguínea o endogámica.</i>	13
2. OBJETIVOS	15
3. MATERIAL Y MÉTODOS	16
3.1. Material Biológico y Datos Fenotípicos.	16
3.2. Datos genotípicos.	16
3.4. Segmentos de Homocigosidad y Depresión Endogámica.	20
3.5. Búsqueda de genes candidatos	21
4. RESULTADOS Y DISCUSIÓN	22
4.1. Heterocigosidad y Depresión Endogámica	22
4.2. Segmentos de Homocigosidad y Depresión endogámica	37
4.3. Consideraciones Finales	46
5. CONCLUSIÓN	48
6.- BIBLIOGRAFÍA	49

1. INTRODUCCIÓN

1.1. El cerdo Ibérico

El cerdo ibérico es el principal representante del cerdo negro mediterráneo (*Sus scrofa meridionalis*), tanto en cuanto a su censo como a su impacto cultural y económico. Su hábitat natural se localiza en el sureste de la península ibérica en las zonas de bosque mediterráneo o Dehesa, y se caracteriza por ser una de las poblaciones de mayor calidad de la carne a nivel mundial (Serra et al., 1998; Gandemer, 2009). Su importancia económica radica en esta calidad de carne y en su elevada proporción de grasa intramuscular con un alto contenido en ácidos grasos insaturados (López-Bote, 1998). Además, el cerdo ibérico tiene una gran adaptación a su entorno productivo y a posibles cambios del clima y de la alimentación gracias a esta alta acumulación de grasa (Silio et al., 2001). Sin embargo, presenta un menor crecimiento, eficiencia alimentaria y prolificidad que las poblaciones comerciales de cerdo blanco (Silio et al., 2001).

A partir de los 60, el censo del cerdo ibérico se redujo de manera considerable, ya que se produjo la introducción de razas extranjeras con alto nivel de mejora genética. Este fenómeno, junto con la progresiva destrucción de su entorno pusieron en peligro la existencia de esta raza ibérico (López-Bote, 1998). Sin embargo, la alta calidad de la carne y la apertura al mercado internacional conseguido revalorizar los productos procedentes del cerdo ibérico (López-Bote, 1998), y permitió el incremento del censo a partir de los años 80.

En la actualidad, según la encuesta ganadera porcina realizada por MAPA, existen un total de 3,375,281 cerdos de raza ibérica, mientras que el número de cerdos de raza no ibérica es de 27,870,760. La comunidad autónoma con mayor producción de cerdo ibérico es Extremadura, aunque también existe un porcentaje importante de la producción en Castilla- León y Andalucía.

El sistema de producción tradicional del cerdo ibérico se base en la cría de individuos en pureza de forma extensiva o semi-extensiva, con pasto de bellotas hasta los 155-165 kg y unos 18 meses. Sin embargo, los datos del Ministerio de Agricultura, Pesca y Alimentación (MAPA) para el año 2017 muestran que únicamente el 10% de los cerdos sacrificados fueron de raza pura ibérica (100 % Ibéricos) y que tan solo un 20% de los animales han sido criados del modo tradicional (Ministerio de Agricultura, Pesca y Alimentación, 2017). De hecho, la mayor parte de la producción de cerdo ibérico se realiza mediante la producción intensiva de engorde de cruces de cerdo ibérico con poblaciones de origen Duroc, para ser sacrificados en torno a los 10 meses de edad (Nieto et al., 2019). El cruzamiento con Duroc incrementa la precocidad, prolificidad y rendimiento reproductivo con respecto a la población ibérica pura (Miao et al., 2004).

La normativa aprobada en el Real Decreto 4/2014, de 10 de enero, por el que se aprueba la norma de calidad para la carne, el jamón, la paleta y la caña de lomo ibérico (2014), indica que para que un cerdo pueda ser considerado 100% ibérico se requiere que tanto el cómo sus progenitores tengan un 100% de pureza racial y estén inscritos en el correspondiente libro genealógico. Por otra parte, se considera Ibérico cuando se trata de productos procedentes de animales con al menos el 50% de su porcentaje genético correspondiente a la raza porcina ibérica cuyos progenitores deben poseer las siguientes características:

-Para obtener animales del 75% ibérico se emplearán hembras de raza 100% ibérica inscritas en libro genealógico y machos procedentes del cruce de madre de raza 100% ibérica y padre de raza 100% Duroc, ambos inscritos en el correspondiente libro genealógico de la raza.

-Para obtener animales del 50% ibérico se emplearán hembras de raza 100% ibérica y machos de raza 100% Duroc, ambos inscritos en el correspondiente libro genealógico de la raza.

1.1.1. Diversidad Genética

El Ministerio de Agricultura, Pesca y Alimentación (MAPA) otorga al cerdo ibérico la denominación de raza autóctona de fomento, aunque según la Asociación Española de Criadores de Cerdo Ibérico (AECERIBER) se distinguen hasta cinco estirpes diferenciadas denominadas Retinto, Entrepelado, Lampiño, Torbiscal y Manchado de Jabugo. Las variedades Entrepelado y Retinto tienen un mayor censo y se consideran como poblaciones de fomento por parte del Ministerio de Agricultura, Pesca y Alimentación (MAPA), mientras que el resto son clasificadas como razas en peligro de extinción (Ministerio de Agricultura, Pesca y Alimentación, 2019). Estas variedades se diferencian en función de su pelaje, aunque también se aprecian diferencias entre ellas en cuanto a sus rendimientos productivos y reproductivos (Silio et al., 2001).

El análisis de la variabilidad genética entre poblaciones de cerdo ibérico se ha estudiado tanto con microsatélites (Martínez et al., 2000; Alves et al., 2006), como a partir de información proporcionada por chips de genotipado (Alonso et al., 2020). Algunos de estos estudios concluyen que la variabilidad genética entre las variedades del cerdo ibérico es tan grande como entre las poblaciones de cerdo blanco (Laval *et al.*, 2000; Martínez *et al.*, 2000; Fabuel *et al.*, 2004).

1.1.2. Características Reproductivas del Cerdo Ibérico

De modo general, el primer parto de las cerdas ibéricas se produce entre los 10 y los 16.5 meses (Rodríguez et al., 1994), la media partos por año es de aproximadamente 2.2, con una prolificidad media de unas 7,5 crías por parto (Benito et al., 2006). Estos índices reproductivos son sustancialmente inferiores a otras poblaciones porcinas, que obtienen generalmente prolificidades por encima de las 10 crías (Andersson et al, 2015) y llegan en ocasiones a medias superiores a los 15, como en las poblaciones de origen asiático como la Meishan (Jiang et al., 2019).

Esta limitación ha intentado solventar por parte de la empresa Inga Food S. A. mediante la generación de una cerda híbrida denominada “Castua” (Inga Food, 2016) con el objetivo de obtener una mayor eficiencia productiva y

homogeneidad tanto a nivel morfológico como de calidad de producto. El desarrollo de esta cerda híbrida implicó la implementación de un cruce dialélico entre 3 estirpes de cerdo ibérico (Retinta, Entrepelado y Torbiscal) y concluyó que el cruce que proporcionaba la mayor eficiencia productiva era el Retinto x Entrepelado, que genera las cerdas híbridas a partir de machos de la población Retinto y hembras entrepeladas (Noguera et al., 2019). Finalmente, las cerdas “Castua” son cruzadas con verracos de origen “Duroc” para dar lugar a un producto final 50% ibérico.

1.2. Consanguinidad y Depresión Consanguínea.

El manejo genético de poblaciones de tamaño reducido implica la aparición de consanguinidad, que se ocasiona como consecuencia del emparejamiento entre individuos emparentados, aunque también puede ocurrir en poblaciones grandes, cuando se aparean parientes entre sí. La consanguinidad se mide mediante el coeficiente de consanguinidad (F) de Wright (1922,1969), que se define como la probabilidad de que los dos alelos de un determinado *locus* en un individuo sean idénticos por descendencia (IBD), es decir, que procedan de un mismo ancestro (Falconer y Mckay, 1996). En este sentido, es importante distinguir entre los conceptos de identidad por estado (IBS) e identidad por descendencia (IBD). Un locus es idéntico por estado (IBS) cuando los dos alelos presentes en un individuo son iguales, aunque procedan de dos orígenes genéticos distintos, mientras que un loci es idéntico por descendencia (IBD) cuando los dos alelos son iguales porque proceden de un mismo alelo en un antecesor común. La diferencia entre estos dos conceptos se ilustra en la Figura 1.

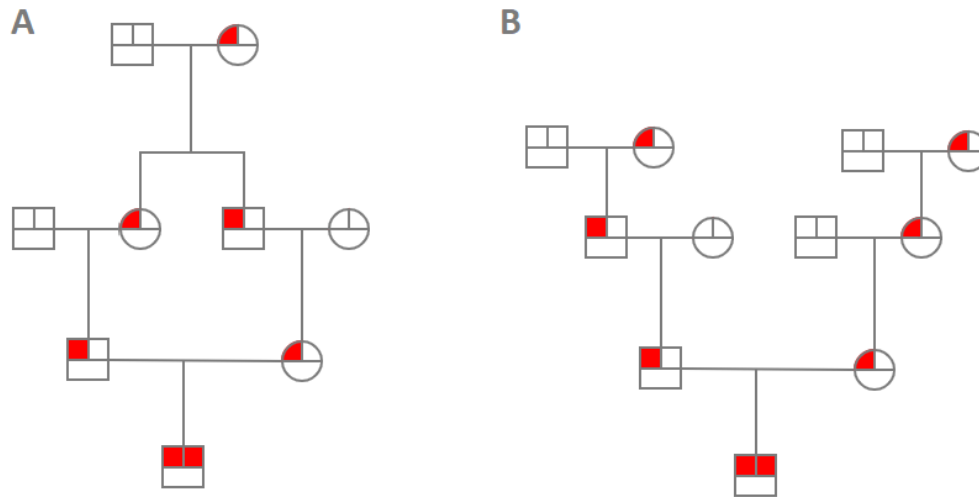


Figura 1. Diferencia entre homocigotos idénticos por descendencia (IBD) , donde ambos alelos proceden de un mismo ancestro (A), y homocigotos idénticos por estado (IBS), en los que proceden de diferentes ancestros (B).

Pese a todo, la distinción entre IBD e IBS es sutil, ya que es posible plantear que todos los alelos actuales de un mismo tipo se sean copias de un mismo alelo ancestral, más o menos alejado en el tiempo. Sin embargo, el cálculo de consanguinidad en la práctica está referido a una población base inicial, y si en ella existen varias copias de un determinado alelo, se asume que no proceden de un mismo origen genético, aunque es plausible que sí que procedan de un ancestro desconocido.

El coeficiente de consanguinidad se define como la probabilidad de identidad por descendencia de un alelo y, por lo tanto, varía entre 0 y 1, y puede generalizarse a todos los loci autosómicos del individuo como la proporción de los mismo que sean homocigotos por descendencia. La definición de consanguinidad como una probabilidad está estrechamente ligada al coeficiente de parentesco o de coascendencia (f) propuesto por Malecot (1948), que se define como la probabilidad de que, tomando un alelo al azar en el mismo locus en dos individuos, estos sean idénticos por descendencia.

1.2.1. Consanguinidad genealógica

Tradicionalmente, los coeficientes de consanguinidad y de parentesco pueden calcularse a partir de la información genealógica mediante dos procedimientos alternativos. Por un lado, el método de los senderos de Wright (1922) que exige identificar las posibles rutas de los alelos que generan IBD a partir de individuos presentes simultáneamente entre los ancestros por vía paterna y materna. Por otro, es posible utilizar el método tabular, descrito por Malecot (1948), que se basa en el cálculo de coeficiente de coascendencia entre los progenitores. Mas recientemente, se han descrito procedimientos alternativos (Meuwissen y Luo, 1992) que permiten el cálculo de la consanguinidad de todos los individuos de la población de una manera mucho más eficiente.

1.2.2. Coeficientes de consanguinidad y de parentesco moleculares

La disponibilidad de polimorfismos genéticos localizados en el genoma autosómico permite el cálculo de los coeficientes de consanguinidad y parentesco moleculares, que son de mayor utilidad cuando la información genealógica es escasa o nula. La definición más sencilla del coeficiente de consanguinidad molecular para un locus dado es 0 si los dos alelos son diferentes y 1 si ambos son idénticos. En este sentido, el coeficiente de parentesco molecular entre dos individuos se calcularía comparando dos a dos los alelos de cada individuo. Si se dispone de más de un locus en el cálculo, los valores de consanguinidad y parentesco moleculares se obtienen promediando sobre los loci.

El mayor inconveniente que encontramos con el uso de los marcadores moleculares es la imposibilidad de hacer una distinción entre IBD e IBS, por lo que tanto coeficientes moleculares de consanguinidad como de parentesco son generalmente más elevados que los obtenidos a partir de genealogías, ya que incluyen ambos tipos de identidad. Esta disparidad entre IBD e IBS se ha intentado corregir mediante varios procedimientos, entre ellos la medida (F_{snp}) (Keller et al. 2011) se basa en la siguiente expresión:

$$F_{snp(i)} = \frac{O(H_i) - E(H)}{s - E(H)}$$

donde $O(H_i)$ es la suma de la homocigosidad observada a lo largo de todos los SNPs de un individuo i , $E(H)$ es la homocigosidad esperada para todos los individuos de la muestra y s el número de SNPs estudiados (Keller et al. 2011). La homocigosidad esperada se calcula mediante:

$$E(H) = \sum_i 1 - 2p_i(1 - p_i)$$

Donde p_i es la frecuencia del alelo menor del SNP i .

La hipótesis subyacente de este método se basa en que para un determinado locus se espera un porcentaje de individuos homocigotos en función de las frecuencias alélicas, y que la presencia de un mayor número de homocigotos es consecuencia de la consanguinidad.

Por otra parte, otro grupo de métodos que tratan de discriminar entre IBS e IBD son los basados en regiones de homocigosidad (runs of homozygosity - ROH). Los ROH son regiones del genoma de un individuo en las que solo se encuentran SNP homocigotos (McQuillan et al. 2008). El punto de partida de estos métodos es que un determinado marcador puede ser homocigoto por IBS, pero si encontramos un grupo grande de SNP homocigotos es más plausible que su origen sea IBD. Además, la longitud de la región de homocigosidad ofrece una indicación de la antigüedad de la consanguinidad. Los segmentos de mayor tamaño corresponden a consanguinidad reciente, mientras que los segmentos de tamaño menor se atribuyen a consanguinidad antigua, como consecuencia de la recombinación. Así, una vez decidido el tamaño mínimo del ROH, el coeficiente de consanguinidad F_{ROH} mide como:

$$F_{ROH_j} = \frac{\sum_k length(ROH_{jk})}{L}$$

donde $length(ROH_{jk})$ es la longitud de la k-esima región ROH del individuo j y L es la longitud del genoma de la especie (Keller et al 2011).

1.2.3. Depresión consanguínea o endogámica.

La principal consecuencia de la consanguinidad a nivel de la media de los caracteres fenotípicos es la aparición de la depresión consanguínea (Falconer and Mckay, 1996), que se define como la pérdida de adaptación (vigor, viabilidad, fecundidad...) producida por la pérdida de variabilidad genética como consecuencia del incremento de la homocigosidad. En concreto, la magnitud de la depresión consanguínea para un carácter regulado por n loci bi-alélicos y en ausencia de epistasis se describe como:

$$-2F \sum_{i=1}^n p_i q_i d_i,$$

donde F es el coeficiente de consanguinidad, p_i y q_i son las frecuencias alélicas y d_i es la desviación del fenotipo de los heterocigotos sobre el promedio de los dos tipos de homocigotos para un locus bialélico i (Crow y Kimura, 1970). Tradicionalmente, se han definido dos hipótesis principales para explicar el fenómeno de la depresión endogámica. Por una parte, la hipótesis de dominancia parcial y, por otra, la hipótesis de la sobredominancia (Charlesworth y Willis, 2009). La primera de las hipótesis asume que la depresión endogámica está provocada por la expresión de los alelos deletéreos recesivos en individuos homocigotos. Al aumentar la consanguinidad, se genera una mayor proporción de alelos recesivos en homocigosis que ocasiona la depresión consanguínea (Kristensen & Sorensen 2004). En la segunda hipótesis, se asume que los heterocigotos tienen un rendimiento superior a ambos homocigotos o sobredominancia, y con la aparición de la consanguinidad, la reducción del porcentaje de heterocigotos impide la expresión de esta sobredominancia (Kristensen & Sorensen 2004).

En porcino, la presencia de depresión endogámica asociada con la consanguinidad genealógica en los caracteres reproductivos ha sido descrita tanto en cerdo blanco (Farkas et al., 2007), como en cerdo Ibérico (Silio et al., 2015). Estos estudios utilizan los coeficientes de consanguinidad calculados a partir de la información genealógica. Sin embargo, la disponibilidad de dispositivos de genotipado de marcadores SNP permite estudiar la distribución de la depresión endogámica a lo largo del genoma (Howard et al., 2017) bajo la hipótesis de que las regiones del genoma contribuyen de manera heterogénea a la depresión endogámica. En concreto, Saura et al. (2015) analizaron el efecto de las regiones del genoma en la depresión endogámica en una población de la estirpe Guadyerbas concluyendo que la depresión endogámica en los caracteres reproductivos de esta estirpe del cerdo ibérico se distribuía de manera heterogénea a lo largo del genoma.

2. OBJETIVOS

El objetivo principal de este trabajo es estudiar la arquitectura genética de la depresión endogámica en los caracteres reproductivos en dos estirpes de cerdo ibérico: Entrepelado y Retinto. De este objetivo principal se desprenden los siguientes objetivos secundarios.

- El estudio de la distribución de las regiones de homocigosidad a largo del genoma en Entrepelado y Retinto.
- La identificación de genes candidatos localizadas en las regiones del genoma asociadas con una reducción o incremento productivo en función del grado de homocigosidad.

3. MATERIAL Y MÉTODOS

3.1. Material Biológico y Datos Fenotípicos.

El desarrollo de este trabajo ha utilizado la información procedente de los proyectos de investigación IDI-20170304 y CGL-2016-80155, desarrollados en cooperación por la empresa INGA FOOD, S. A., el Institut de Recerca i Tecnologia Agroalimentaria (IRTA), la Universitat Politècnica de Valencia (UPV), la Universitat Autònoma de Barcelona (UAB) y la Universidad de Zaragoza (UNIZAR).

El fichero de datos contenía información de Nacidos Totales y Nacidos Vivos de 14,991 partos procedentes de 436 cerdas (7,308 partos) de la variedad Entrepelado y de 397 cerdas (7,638 partos) de la variedad Retinto. Adicionalmente, el fichero de datos incluye el número de ciclo, la granja y la fecha de parto. La descripción fenotípica de los datos se presenta en la Tabla 1.

Tabla 1. Número de datos (N), media, mediana y desviación típica (DT) de Nacidos Totales (NT) y Nacidos Vivos (NV) en las dos poblaciones estudiadas (Entrepelado y Retinto).

Población	Carácter	N	Media	Mediana	DT
Entrepelado	NT	7308	7,87	8	2,02
	NV		7,59	8	2,02
Retinto	NT	7683	8,24	8	2,16
	NV		7,89	8	2,14

Por otra parte, se extrajo de la información genealógica con relación genética con las cerdas genotipadas y fenotipadas que consistió en 710 entradas individuo – padre – madre en la población Entrepelado y en 654 entradas en la población Retinto.

3.2. Datos genotípicos.

Adicionalmente, se dispuso del genotipado de las 833 cerdas con el chip GENESEEK GGP Porcine HD 70K (Illumina), que contiene 63,072 polimorfismos de nucleótido simple, o SNPs (Single Nucleotide Polymorphisms). Los SNP autosómicos se filtraron utilizando el software PLINK v1.9. (Chang et al., 2015). Los criterios de filtrado fueron:

- Los individuos deben tener al menos un 90% de los marcadores genotipados.
- Los SNP deben estar genotipados en al menos un 95% de los individuos.
- La frecuencia alélica menor (MAF) debe ser superior al 5%.

El resultado del proceso de filtrado seleccionó 36,449 SNP de 427 animales de la población entreladada y de 392 de la población retinta. De los 36.440 SNP, un total de 34,815 se localizaron en los cromosomas autosómicos distribuyéndose del siguiente modo:

Tabla 2. Cromosoma (Crom), número de marcadores del cromosoma (Nº mar) y longitud en pares de bases del cromosoma (Pb).

Crom	Nº mar	Pb
1	3166	274315671
2	2043	151610480
3	1951	132657669
4	2164	130773976
5	1696	104477606
6	2429	170802600
7	2315	121758423
8	2234	138930735
9	2263	139386589
10	1581	69319537
11	1416	79072521
12	1214	60834034
13	2640	208240759
14	1971	141719266
15	2105	140404164
16	1356	79282526
17	1167	63391207
18	1104	55752892

Una vez seleccionados los SNP, se detectó la presencia de 65,843 genotipados faltantes (79.04 en promedio por cerda genotipada – 0.23 %-), que fueron imputados mediante el software *FindHap* (VanRaden and Sun, 2014) en función de la información disponible a nivel poblacional y en los individuos emparentados.

Una vez imputados los marcadores SNP faltantes, se utilizó el software *ADMIXTURE* (Alexander et al., 2011) para determinar la existencia de posibles “outliers”, es decir, individuos con un valor distinto al esperado para sus características, y posteriormente se realizó un análisis de componentes principales. Para ello se utilizó el paquete *ggfortify* del software R (R Core Team, 2016). Se generó un gráfico con los dos primeros componentes principales (PC1 y PC2) y se determinaron los “clusters” correspondientes a las 2 líneas estudiadas (EE y RR), que quedaron representados en la figura 2. Los individuos de una variedad se localizarán dentro del “cluster” asignado a la población deberían haber sido excluidas del estudio. Por lo tanto, no se eliminó ningún individuo.

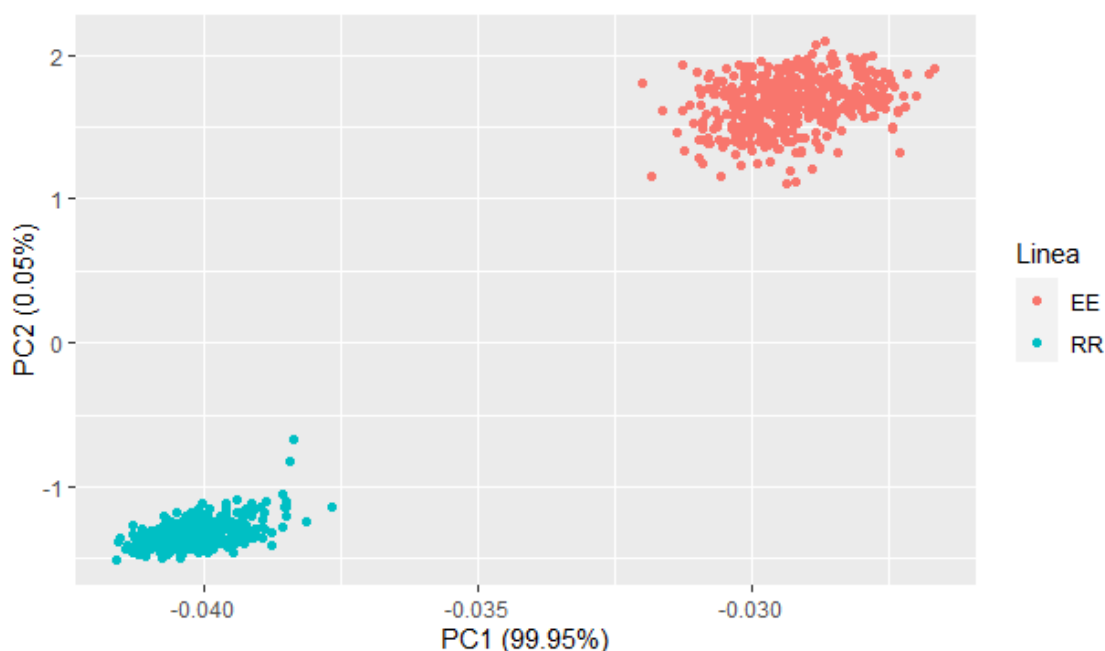


Figura 2. Estructura de la población identificada por análisis de componentes principales (PCA). Los dos primeros componentes principales PC1 y PC2 explican el 99,5 y 0,5%, respectivamente. El PC1 corresponde al eje x PC2 al eje y.

3.3 Homocigosidad y Depresión Endogámica

En primer lugar, se calculó la homocigosidad de cada uno de los individuos de la población, dividiendo el número de marcadores SNP heterocigotos con respecto al total de marcadores, para realizar el estudio de la relación entre el rendimiento fenotípico y la homocigosidad individual. Para ello se utilizó el paquete de programas BLUPf90 (Misztal et al., 2002) y se asumió el siguiente modelo estadístico:

$$y = hd + Xb + Zu + Wp + e$$

Donde d es la covariada asociada a la homocigosidad individual, b es el vector de efectos sistemáticos (orden de parto – 1, 2, 3, 4 y más de 4-), línea genética del verraco – Duroc, Entrepelando o Retinto -, Granja-Año-Estación –(96 niveles en Entrepelado y 113 niveles en Retinto), u es el vector de efectos genéticos aditivos, p es el efecto permanente asociado a la cerda y e es el vector de residuos. X , Z y W son las matrices de incidencia y h es el vector de heterocigosidades individuales. Los componentes de varianza aditiva, ambiental permanente y residual se estimaron mediante máxima verosimilitud restringida (REML) (Patterson y Thomson, 1971) utilizando el paquete airemlf90 (Misztal et al., 2002).

A continuación, se calculó la homocigosidad individual de cada uno de los individuos para los marcadores localizados en regiones específicas del genoma. En concreto, estas regiones se definieron mediante dos procedimientos:

1. La región comprendida dentro de 1 Mb centrada en cada uno de los marcadores SNP.
2. La región comprendida por los 15 marcadores SNP anteriores y posteriores al cada marcador SNP.

Estos resultados de homocigosidad a nivel regional se utilizaron para analizar la arquitectura de la depresión endogámica mediante la aplicación de un modelo mixto para cada segmento del genoma centrado en cada uno de los SNP. El modelo de análisis fue:

$$y = h_1d_1 + h_2d_2 + Xb + Zu + Wp + e$$

Donde h_1 es el vector de la homocigosidad en la región del genoma analizada y h_2 es el vector de la homocigosidad en el resto del genoma. Además, d_1 y d_2 son los estimadores de la depresión endogámica asociada a la región del genoma analizada y a la asociada al resto del genoma, respectivamente. Para el análisis se utilizó el software blupf90 (Miszta et al., 2002), con los componentes de la varianza obtenidos en el estudio inicial.

3.4. Segmentos de Homocigosidad y Depresión Endogámica.

El estudio de la arquitectura genética de la depresión endogámica se llevó a cabo mediante un segundo procedimiento, que se basó en el análisis de los segmentos de homocigosidad o ROH (*run of homocigosity*). Una primera aproximación trató de asociar el porcentaje del genoma autosómico incluido en segmentos de homocigosidad de longitud variable con el rendimiento fenotípico de las cerdas para nacidos vivos y nacidos totales. Para ello, se utilizó el software detectRUNS (Biscarini et al., 2018) mediante la opción *consecutiveRuns()* con el objetivo de detectar las regiones de homocigosidad presentes en cada cerda. Posteriormente, se utilizó la opción *FrohinbreedingClass()* para calcular el porcentaje del genoma incluido en segmentos ROH de tamaños crecientes (>0 Mb, >0.5 Mb, > 1 Mb, >2 Mb, >3Mb, >4 Mb, >6 Mb, >8 Mb, >12 Mb y >16 Mb) para cada uno de las cerdas genotipadas. A continuación, estos porcentajes se asociaron con el rendimiento fenotípico para Nacidos Totales y Nacidos Vivos, mediante la aplicación del siguiente modelo mixto:

$$y = hd + Xb + Zu + Wp + e$$

Donde, en esta ocasión, \mathbf{h} es el vector de los porcentajes del genoma incluidos en segmentos ROH de tamaños

Sin embargo, es esperable que el efecto asociado a los segmentos ROH no sea el mismo a lo largo del genoma, sino que es plausible que exista variabilidad en cuanto a la asociación con el rendimiento fenotípico. Por este motivo, se utilizó el programa *Haplofinder* (Baxer et al., 2008), que realiza el siguiente proceso. En primer lugar, calcula los segmentos de homocigosidad presentes en cada individuo de la población y, posteriormente, para cada segmento de homocigosidad, utiliza un modelo mixto donde asigna un nuevo efecto sistemático con 2 categorías (ROH y No ROH), mediante la clasificación de los individuos entre aquellos que son homocigotos para el segmento y los que no lo son. El modelo mixto utilizado fue equivalente al descrito en el apartado anterior, además de la inclusión de este nuevo efecto sistemático adicional. La aplicación del programa se restringió a segmentos de homocigosidad (ROH) de más de 20 marcadores y con una frecuencia superior al 5% en la población. Los componentes de varianza aditiva, genética y residual que se utilizaron fueron los mismos que se estimaron en el apartado anterior.

3.5. Búsqueda de genes candidatos

Finalmente, se procedió a la localización de posibles genes candidatos en las regiones genómicas asociadas con una reducción o incremento del rendimiento fenotípico tanto por el grado de homocigosidad regional o por la presencia de un segmento de homocigosidad. Para ello, se utilizó aplicación BIOMART dentro de la base de datos Ensembl (Hunt et al., 2018). En concreto, se usó la versión Sscrofa 11.1 del genoma porcino.

4. RESULTADOS Y DISCUSIÓN

4.1. Heterocigosidad y Depresión Endogámica

En primer lugar, se analizó la homocigosidad de los 34,815 SNP autosómicos en las poblaciones Entrepelado y Retinto. Los resultados medios de la homocigosidad poblaciones se presentan en la tabla 3, así como en la figura 3.

Tabla 3. Media, mediana, desviación típica, máximo y mínimo de la homocigosidad observada en las poblaciones Entrepelado y Retinto.

	Media	Mediana	Desv. típica	Máximo	Mínimo
Entrepelado	0.642	0.642	0.024	0.732	0.591
Retinto	0.722	0.720	0.017	0.795	0.645

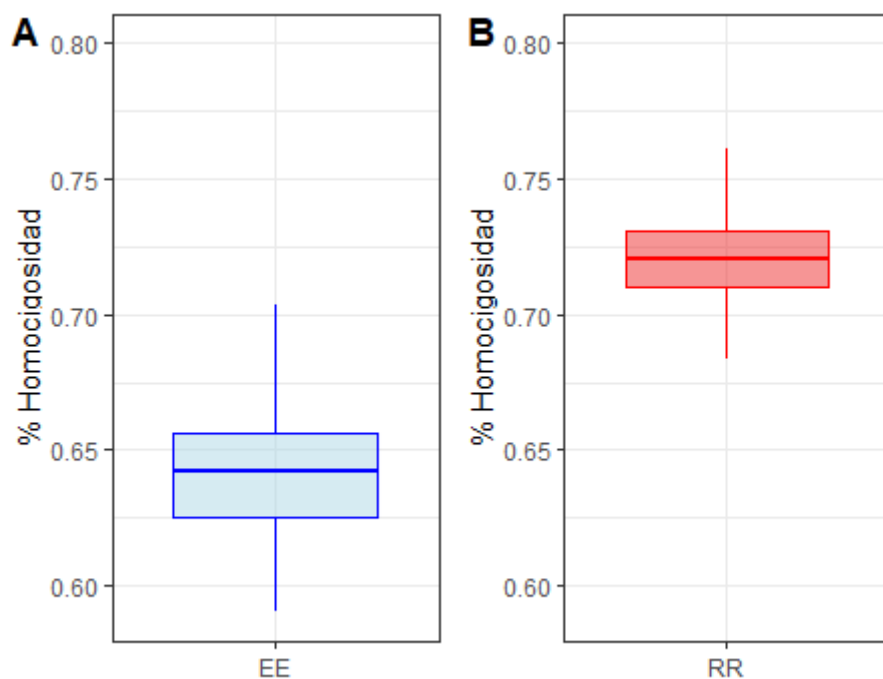


Figura 3. Distribución de la homocigosidad observada de cada individuo para las líneas EE (entrepelado) y RR (retinto).

Los individuos de la población Retinto mostraron un grado de homocigosidad muy superior a la población Entrepelado (0.722 vs. 0.642). Este resultado es lógico, ya que la variedad Entrepelado tiene una fundación más reciente, ya

que se trata de un híbrido fijado entre las variedades Lampiño y Retinto (<https://feagas.com/razas/porcino/iberico-entrepelado/>). Sin embargo, el grado de homocigosidad observado en las dos poblaciones fue considerablemente inferior al que reportó Saura et al. (2015) en el análisis de la variedad Guadyerbas (0.868). Este resultado es esperable ya que los individuos Guadyerbas analizados en este estudio procedía de la población experimental de *El Dehesón del Encinar (Oropesa, Toledo)*, que se ha mantenido aislada desde 1944. En esta población se han descrito niveles de consanguinidad elevados en varias publicaciones (Alves et al., 2008; Saura et al., 2013).

Las homocigosidades esperadas a partir de las frecuencias alélicas estimadas fueron 0.644 en la población Entrepelado y 0.719 para la población Retinta. Por lo tanto, si se utilizan estos valores para estimar la consanguinidad según el procedimiento descrito por Keller et al. (2011), los valores medios fueron de -0.006 en Entrepelado y de 0.007 en Retinto, y oscilaron entre 0.246 y -0.151 en Entrepelado y entre -0.266 y 0.270 en Retinto. Pese a todo, debe tenerse en cuenta que el cálculo de la consanguinidad a partir de las homocigosidades observadas y esperadas exige el conocimiento de las frecuencias alélicas en la generación base (Toro et al., 2002), que no se encuentran disponibles. Por lo tanto, los estimadores obtenidos deben tomarse con mucha cautela.

A partir del cálculo de la homocigosidad observada para cada uno de los individuos, se estudió el efecto de la homocigosidad en el rendimiento fenotípico para Nacidos Vivos y Nacidos Totales en las dos poblaciones. Para ello, se utilizó un modelo mixto que incluía la homocigosidad individual como covariada. Los resultados de las estimaciones de máxima verosimilitud restringida (REML) y de la solución para la pendiente de la covariada con el porcentaje de homocigosidad individual se presentan en la Tabla 4 para la población Entrepelada y en la Tabla 5 para la población Retinta.

Tabla 4. Estimación de componentes de la varianza y pendiente con la covariada con la heterocigosidad individual en la población Entrepelado.

	NT	NV
σ_a^2	0.075 (0.097)	0.056 (0.070)
σ_p^2	0.364 (0.077)	0.327 (0.090)
σ_e^2	2.932 (0.104)	2.834 (0.103)
b	-4.615 (2.500)	-4.201 (2.396)
Valor p	0.032	0.040

Tabla 5. Estimación de componentes de la varianza y pendiente con la covariada con la heterocigosidad individual en la población Retinto.

	NT	NV
σ_a^2	0.196 (0.104)	0.281 (0.126)
σ_p^2	0.257 (0.110)	0.214 (0.115)
σ_e^2	3.822 (0.134)	3.733 (0.131)
b	-4.389 (3.521)	-5.664 (3.541)
Valor p	0.106	0.054

En estas tablas se puede observar como en la población Entrepelado la varianza aditiva es de muy pequeña magnitud, dando lugar a heredabilidades de 0.022 y 0.017 para Nacidos Totales y Nacidos Vivos, respetivamente. Por el contrario, en la población Retinto, la heredabilidad estimada para nacidos totales fue de 0.045 y 0.066 para nacidos totales y nacidos vivos, respectivamente. Estas estimas de heredabilidad fueron inferiores a las obtenidas tanto en poblaciones de cerdo blanco (Chen et al., 2003; Putz et al., 2015; López y Seo ,2019) como ibérico (Fernández et al., 2008; Noguera et al., 2019). Los resultados contrastan especialmente con los obtenidos por Noguera et al. (2019), que encontró heredabilidades superiores a 0.10 en la población Entrepelado y de entre 0.08 y 0.09 en la población Retinto. La causa de esta

posible divergencia puede ser que el análisis efectuado se ha restringido a los animales con el genotipado conocido, mientras que Noguera et al. (2019) se utilizó toda la información disponible en la base de datos de la población y, que, además se tuvo en cuenta la información fenotípica de los animales cruzados.

Las pendientes de las covariadas con la homocigosidad fueron siempre negativas y oscilaron entre -4.201 lechones (Nacidos Vivos en Entrepelado) y -5.664 (Nacidos Totales en Entrepelado). Los valores p de un test en una cola oscilaron entre 0.032 (Nacidos Totales en Entrepelado) y 0.106 (Nacidos Vivos en Retinto). La reducción del número de lechones nacidos vivos y totales a medida que se incrementa la homocigosidad es un resultado esperable, que concuerda con estudios previos en otras poblaciones de origen ibérico (Toro et al., 1988; Saura et al., 2015; Casellas et al., 2019). Sin embargo, no se encuentra una diferencia relevante entre las poblaciones, en contraste con los resultados que obtuvo Casellas et al. (2019), donde se encontró una mayor carga de depresión endogámica en la población Retinto.

En un segundo paso, se calculó la homocigosidad observada en regiones específicas del genoma, que se definieron centradas en cada SNP y con un tamaño definido por 1 Megabase o por 15 marcadores SNP en ambas direcciones del genoma. En concreto, se realizaron 34,815 resoluciones de las ecuaciones de modelo mixto BLUP para cada carácter (Nacidos Totales y Nacidos Vivos), procedimiento de definición de las regiones genómicas (1 Mb o 15 SNP) y población (Entrepelado y Retinto). Para ello se utilizaron los componentes de varianza estimados en el análisis anterior y se obtuvo como resultado la estimación de la pendiente asociada a una covariada con la homocigosidad observada en cada una de las regiones específicas del genoma. Los resultados para ambas estrategias (1 Mb y 15 SNP) y caracteres (Nacidos Totales y Nacidos Vivos) en la población Entrepelado se presentan en la Figura 4.

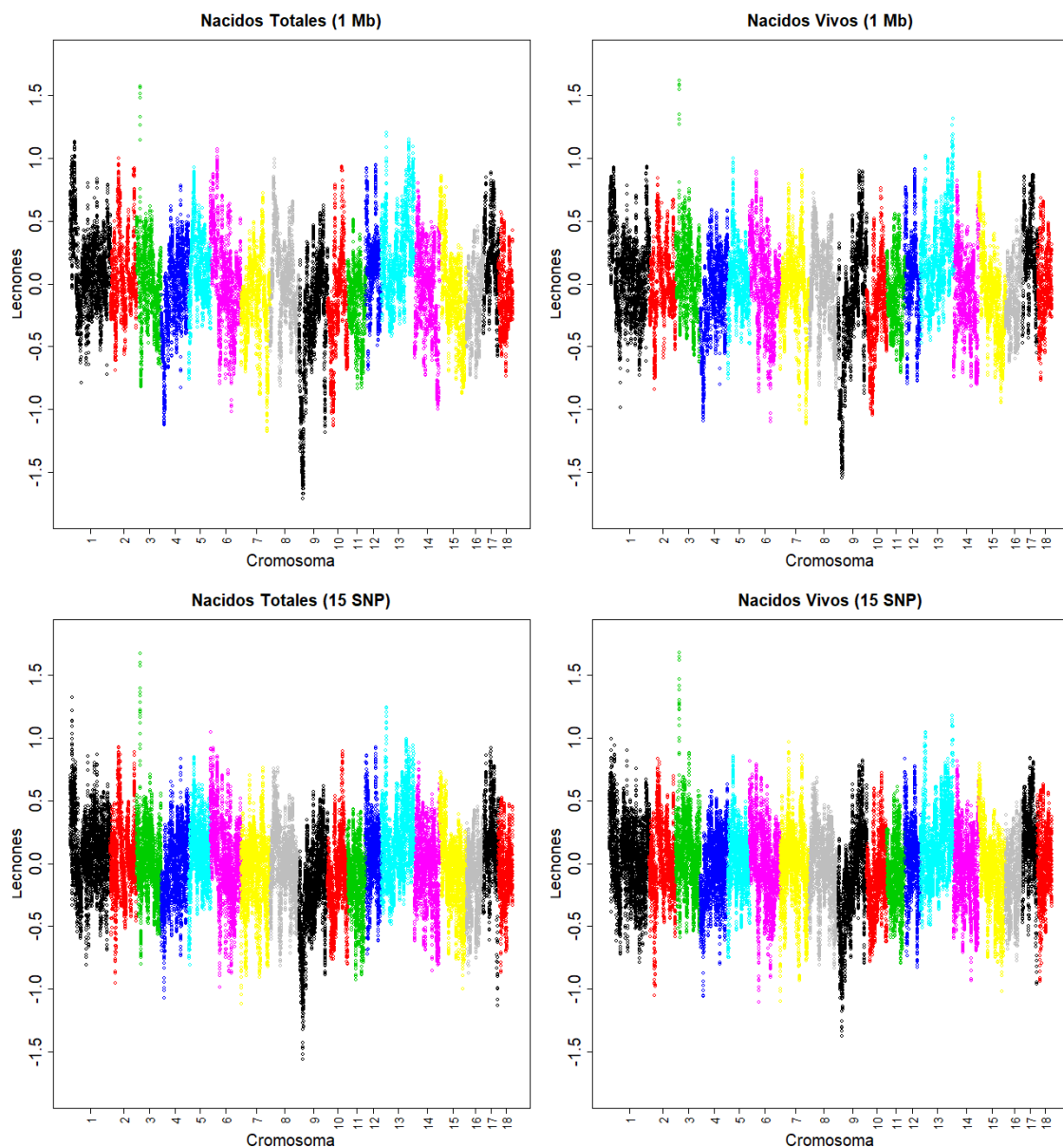


Figura 4. Representación gráfica de la estimación de la covariada con el grado de homocigosidad observada regional para los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Entrepelado.

En la figura se puede obtener como los estimadores de la pendiente asociada con la homocigosidad observada regional obtenidos para Nacidos Totales y Nacidos Vivos con las dos estrategias son muy similares. De hecho, las correlaciones entre estos estimadores oscilaron entre 0.79 (Nacidos Totales con 1 Mb y Nacidos Vivos con 15 SNP) y 0.92 (Nacidos Totales con 1 Mb y Nacidos Totales con 15 SNP). En cada una de las posiciones se calculó el valor p asociado mediante un test de t de una sola cola, tanto para la reducción del

rendimiento fenotípico (o depresión endogámica) o para su incremento. El porcentaje de regiones del genoma que mostraron un valor p nominal oscilaron entre el 5.08% y el 5.32% para el incremento del rendimiento fenotípico y entre el 5.23% y el 5.56% para su reducción. Estos valores fueron ligeramente superiores a los esperados por azar. La distribución a lo largo del genoma de los valores p asociados al incremento y a la reducción del rendimiento fenotípico se presentan en las figuras 5 y 6.

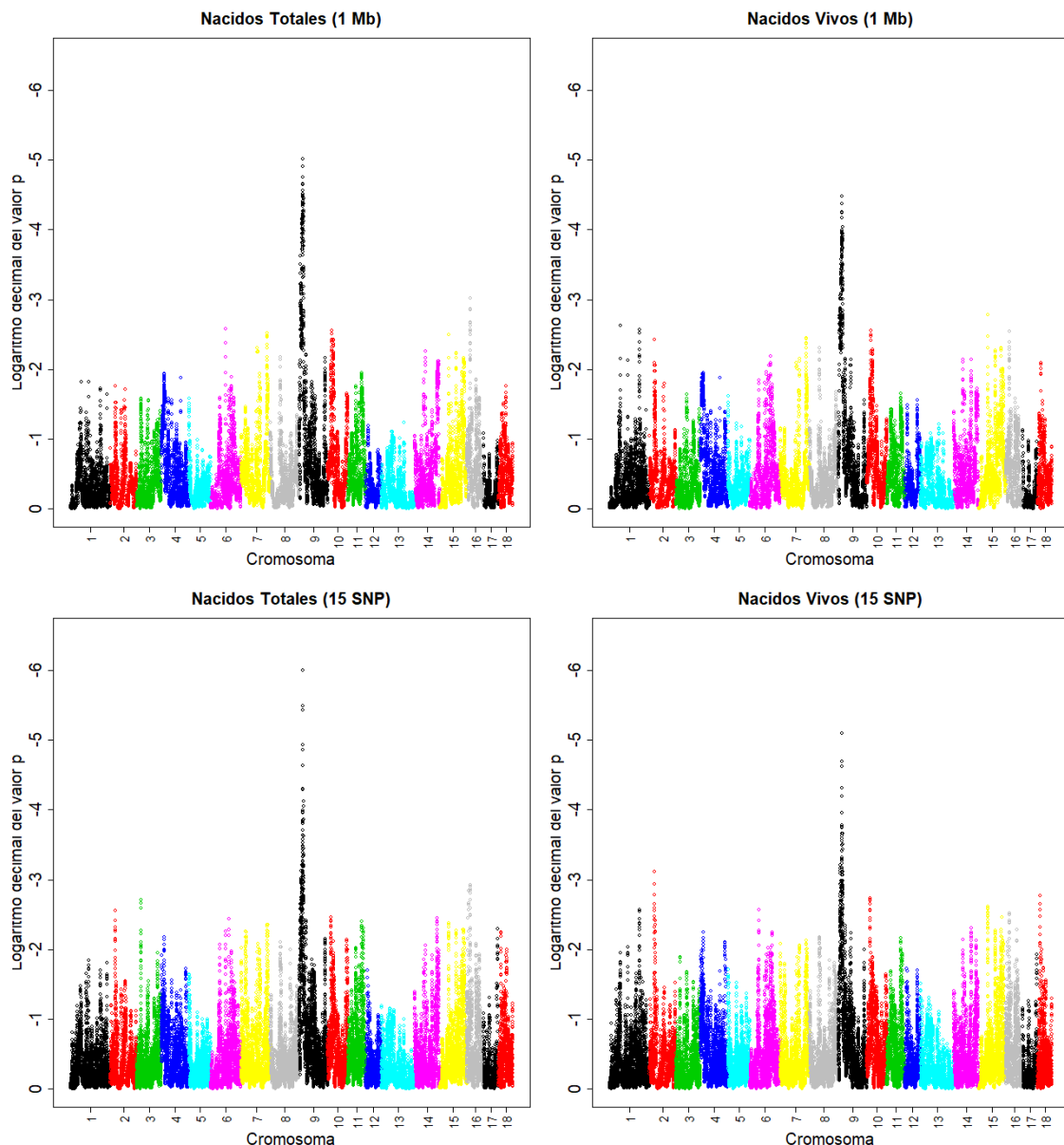


Figura 5. Representación gráfica del $-\log_{10}$ del p-valor asociado a la covariada con la heterocigosidad regional para el incremento de rendimiento fenotipo en los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Entrepelado.

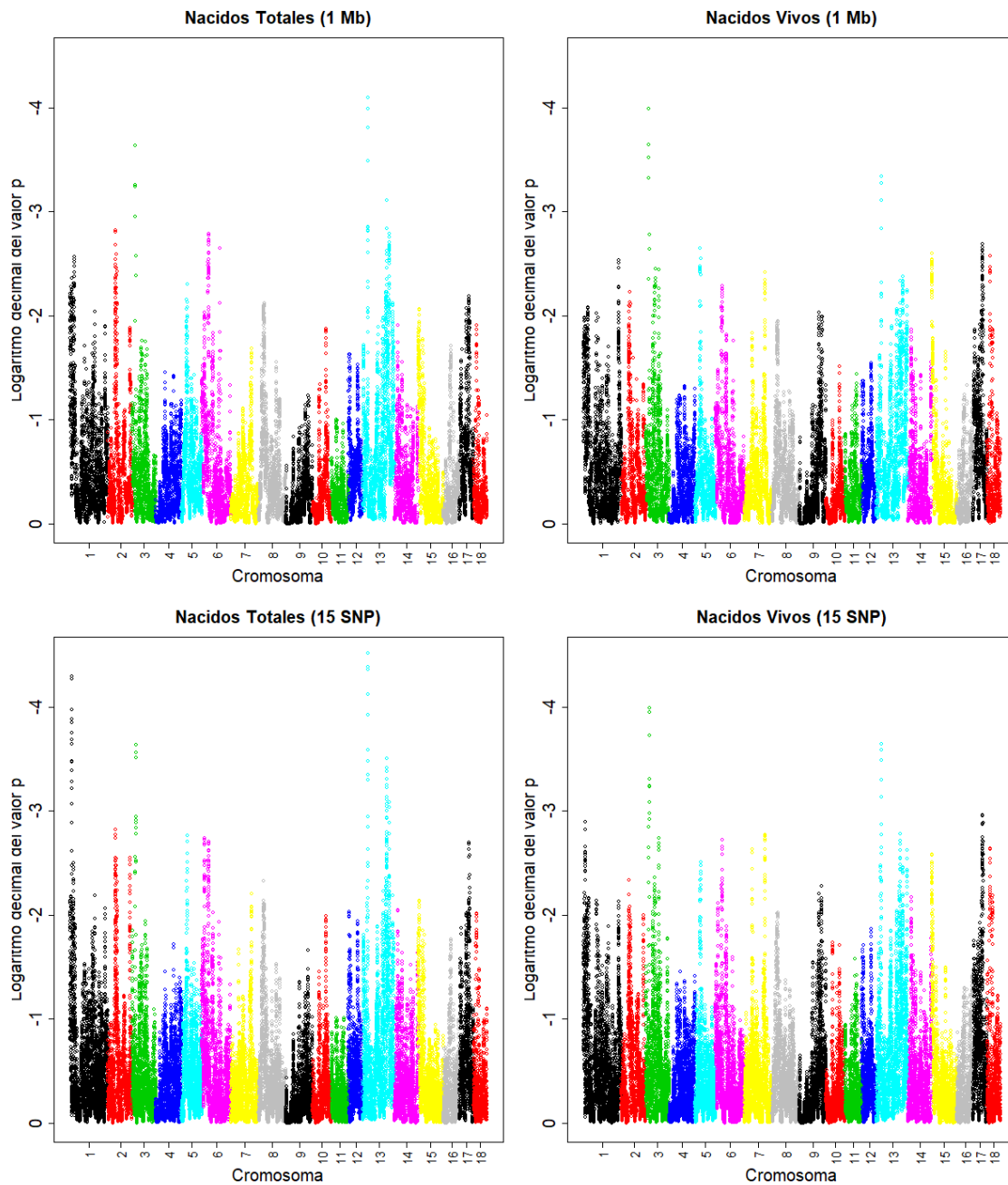


Figura 6. Representación gráfica del $-\log_{10}$ del p-valor asociado a la covariada con la heterocigosidad regional para la reducción de rendimiento fenotipo en los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Entrepelado.

Pese a que los valores p estimados lleguen a ser inferiores a 10^{-6} , si se aplican correcciones de comparaciones múltiples como Bonferroni (Bonferroni, 1936),

o la tasa de falso descubrimiento o *False Discovery Rate* (Benjamini y Hochberg, 1995), ninguna región puede definirse como significativa. Sin embargo, debe tenerse en cuenta que estos dos procedimientos asumen que las comparaciones son independientes y, en este caso, las comparaciones de regiones del genoma centradas en marcadores consecutivos están claramente relacionadas. De hecho, la correlación entre los estimadores centrados en marcadores adyacentes fue superior a 0.99. Por este motivo, se ha estimado el tamaño efectivo de comparaciones a partir de la siguiente expresión:

$$N_{eff} = \frac{N}{1 + 2 \sum_{i=1}^{\infty} \rho^i}$$

Donde ρ es la correlación entre muestras sucesivas. En este caso el resultado fue que el tamaño efectivo de comparaciones fue de 170. Por lo tanto, se aplicó la corrección de Bonferroni como si se hubiesen realizado 170 comparaciones. Por este motivo, se consideró como relevantes a aquellas regiones con un valor p inferior a 0.0003.

En el caso de la población Entrepelado y con asociación positiva con la homocigosidad regional, la única zona del genoma con un valor p inferior a 0.0003 se sitúa en el cromosoma 9 y con un nivel máximo de significación en el marcador rs328022440 que localiza en la posición 14090682. Los genes más cercanos a este marcador son GAB2 (*GRB2 Associated Binding Protein 2*) y NARS2 (*Asparaginyl-tRNA synthetase 2, mitochondrial*). El gen GAB2 juega un papel determinante en la ruta de señalización de la hormona FSH en células de granulosa ováricas (Hunzicker-Dunn et al., 2012) y el NARS2 actúa en la biosíntesis de proteínas y se ha encontrado que se encuentra diferencialmente expresado en vacas gestantes y no gestantes (Mayer et al., 2013).

Las regiones del genoma en las que se identificó un valor p inferior a 0.0003 para la reducción del rendimiento fenotípico y los genes presentes en un

intervalo de 1 Mb con el marcador con mayor valor p se presentan en la Tabla 6.

Tabla 6. Cromosoma (SSC), marcador (SNP), posición en pares de bases (Bp) y genes localizados en un intervalo de 1 Mb de las regiones del genoma con p-valor<0.0003 para la reducción del rendimiento fenotípico en la población Entrepelado.

SSC	SNP	Bp	Genes
1	rs80987387	6633776	QKI, PACRG, PRKN, SNORA70, AGPAT4, MAP3K4, PLG, SLC22A3, SLC22A2, SLC22A1, IGF2R, MAS1, PNLDC1, MRPL18, TCP1, SNORA29, ACAT2, SNORA20, WTAP, SOD2, FNDC1, TAGAP, RSPH3, EZR, SYTL3, DYNLT1, TMEM181
3	rs81371462	10855031	DTX2, ZP3, SSC4D, HSPB1, SRRM3, MDH2, STYXL1, TMEM120A, POR, RHBDD2, CCL26, HIP1, NSUN5, TRIM50, FKBP6, FZD9, BAZ1B, BCL7B, TBL2, MLXIPL, VPS37D, BUD23, STX1A, ABHD11, CLDN3, CLDN4, METTL27, TMEM270, ELN, LIMK1, EIF4H, LAT2, RFC2, GTF2IRD1, GTF2I, NCF1
13	rs81298310	17540185	TGFBR2, GADL1, STT3B, OSBPL10, GPD1L, CMTM8, CMTM7, DYNC1LI1

Entre los genes presentes en la región del cromosoma 1, merecen ser destacados por su potencial relación con la prolificidad porcina los siguientes genes: SOD2 (*Superoxide dismutase 2*), EZR (*Ezrin*), IGF2R (*insulin like growth factor 2 receptor*), MAP3K4 (*Mitogen-Activated Protein Kinase Kinase Kinase 4*), PLG (*Plasminogen*) y AGPAT4 (*1-Acylglycerol-3-Phosphate O-Acyltransferase 4*). El SOD2 codifica una enzima antioxidativa que se ha demostrado importante para la persistencia del cuerpo lúteo en linceos (Braun et al., 2020), y que ha sido relacionada con el tamaño de camada en porcino por Rothschild et al. (1996), ya que está implicada en la regulación del receptor de estrógeno (ESR). El gen EZR se expresa en endodermo y el trofoectodermo, y es necesario para la implantación de embriones en ratones (Matsumoto et al., 2004). El gene IGF2R codifica un receptor que regula la concentración de IGF2 (Spitschak y Hoeflich, 2018), con funciones en foliculogénesis ovárica y

esteroidogénesis. Además, el gen *MAP3K4* está implicado en la regulación positiva de secreción de la prostaglandina y en la determinación sexual (Mendonça et al., 2015), mientras que el gen *PLG* codifica el plasminógeno que contribuye en la actividad proteolítica que se requiere para la ruptura de la pared folicular en el momento de la ovulación (Peng et al., 1993). Finalmente, la función de la enzima codificada por el gen *AGPAT4* es la síntesis lipídica, principalmente en células de granulosa (Pangas et al., 2007).

En la región del cromosoma 3 se localizan varios posibles genes candidatos. Entre ellos, el gen *FZD9* (*Frizzled-9*) pertenece a la familia de genes *FZD*, que codifican proteínas receptoras de las vías de señalización *WNT* que regula el entorno uterino previo a la implantación en diferentes especies como bovino, ovejas o humanos (Neupane et al., 2017). El gen *ZP3* (*Zona Pellucida Glycoprotein 3*) codifica una glicoproteína involucrada con el crecimiento de los oocitos, así como en la foliculogénesis (El-Mestrah et al., 2002). El gen *HSPB1* (*heat shock protein family B (small) member 1*) ha sido ligado con el desarrollo folicular (Hao et al., 2019), ya que la expresión de las proteínas de la familia *HSP* son fundamentales en el sistema reproductivo de mamíferos (Velazquez et al., 2013). El gen *FKBP6* (*FK506 binding protein 6*) posee un rol de importancia en el emparejamiento de cromosomas homólogos en la meiosis (Crackower et al., 2003), además de estar implicado en la fertilidad masculina (Crackower et al., 2003; Schrimpf et al., 2015). Por último, el gen *BCL7B* (*BAF chromatin remodeling complex subunit BCL7B*) ha sido asociado en el estudio de asociación genómica realizado por Onteru et al. (2011) con el número de nacidos vivos y totales en cerdos y *CLDN4* (*Claudin 4*) y *CLDN3* (*Claudin 3*) juegan un papel importante en la implantación del embrión y en la regulación del exceso de testosterona en el ovario (Serafini et al., 2009).

Finalmente, en el cromosoma 13 se localizan tres potenciales genes de interés: *TGFBR2* (*transforming growth factor beta receptor 2*), *STT3B* (*Dolichyl-diphosphooligosaccharide—protein glycosyltransferase subunit STT3B*) y *OSBPL10* (*Oxysterol Binding Protein Like 10*). El gen *TGFBR2* ha sido

identificado como potencial candidato para caracteres reproductivos en un estudio desarrollando por Li et al. (2020) en poblaciones de cerdo Yorkshire, donde se relaciona este gen con el desarrollo del folículo primordial. Además, se ha asociado recientemente con el tamaño de camada en ovejas (Wang et al., 2019). El gen STT3B está involucrado en la estimulación de la hormona FSH (Bousfield y Harvey, 2019) y el gen OSBPL10 ha sido asociado con el peso al nacimiento y se encuentra implicado en transporte de lípidos materno fetales (Wang et al., 2016).

En la población Retinto, los resultados de la pendiente asociada a la homocigosidad observada regional para Nacidos Totales y Nacidos Vivos y con las dos estrategias (15 SNP y 1 Mb) para la definición de las regiones genómicas se presentan en la Figura 7.

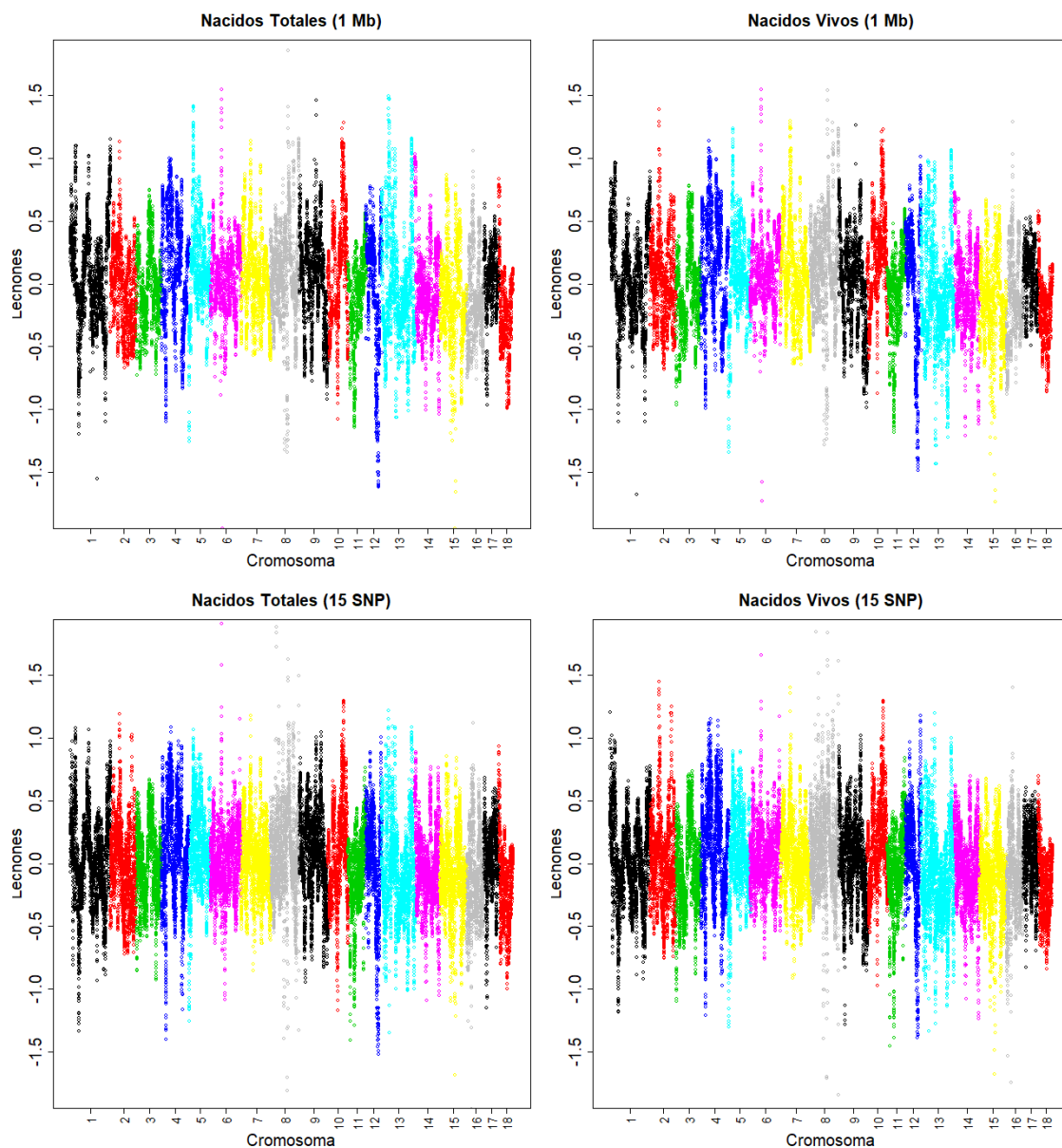


Figura 7. Representación gráfica de la estimación de la covariada con el grado de homocigosidad observada regional para los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Retinto.

Como en la población Entrepelado, los resultados son muy similares y las correlaciones entre los estimadores con la pendiente asociada con la homocigosidad regional oscilaron entre 0.80 (Nacidos Totales -1 Mb- y Nacidos Vivos -15 SNP-) y 0.95 (Nacidos Vivos -1 Mb- y Nacidos Vivos -15 SNP-). El porcentaje de marcadores SNP que mostraron un valor nominal inferior a 0.05 osciló entre el 4.99% y el 5.56% para la reducción del rendimiento fenotípico y entre el 3.34% y el 3.71% para su incremento. La presencia de un porcentaje

mayor de valores p nominales inferiores a 0.05 para la reducción del rendimiento fenotípico es coherente con la presencia de una pendiente negativa asociada a la homocigosidad global.

La distribución de los valores p asociados con un test de una cola para el incremento y la reducción del rendimiento fenotípico asociados a la homocigosidad observada regional se presentan en las Figura 8 y 9.

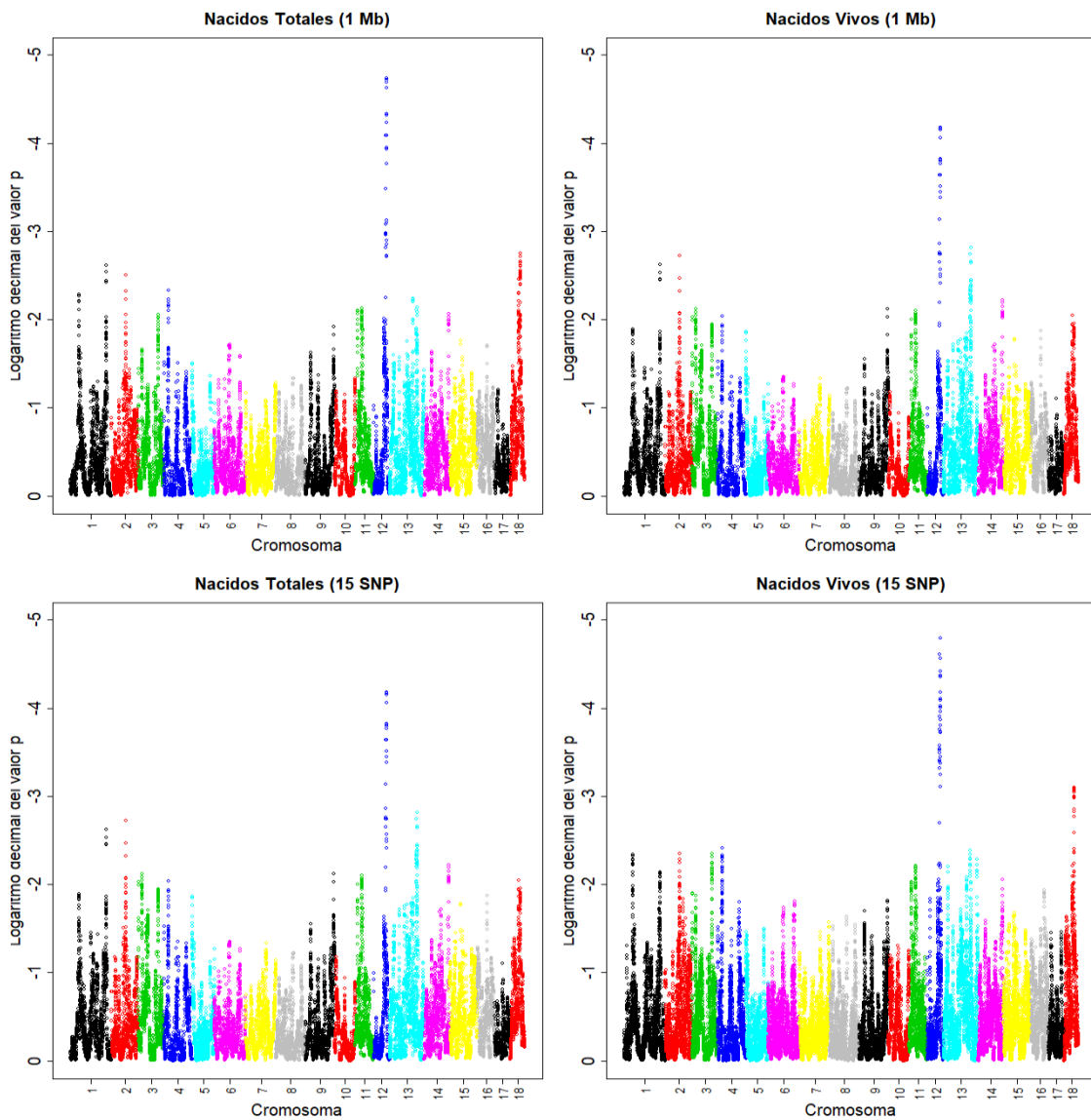


Figura 8. Representación gráfica del $-\log_{10}$ del p-valor asociado a la covariada con la heterocigosidad regional para el incremento de rendimiento fenotipo en los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Retinto.

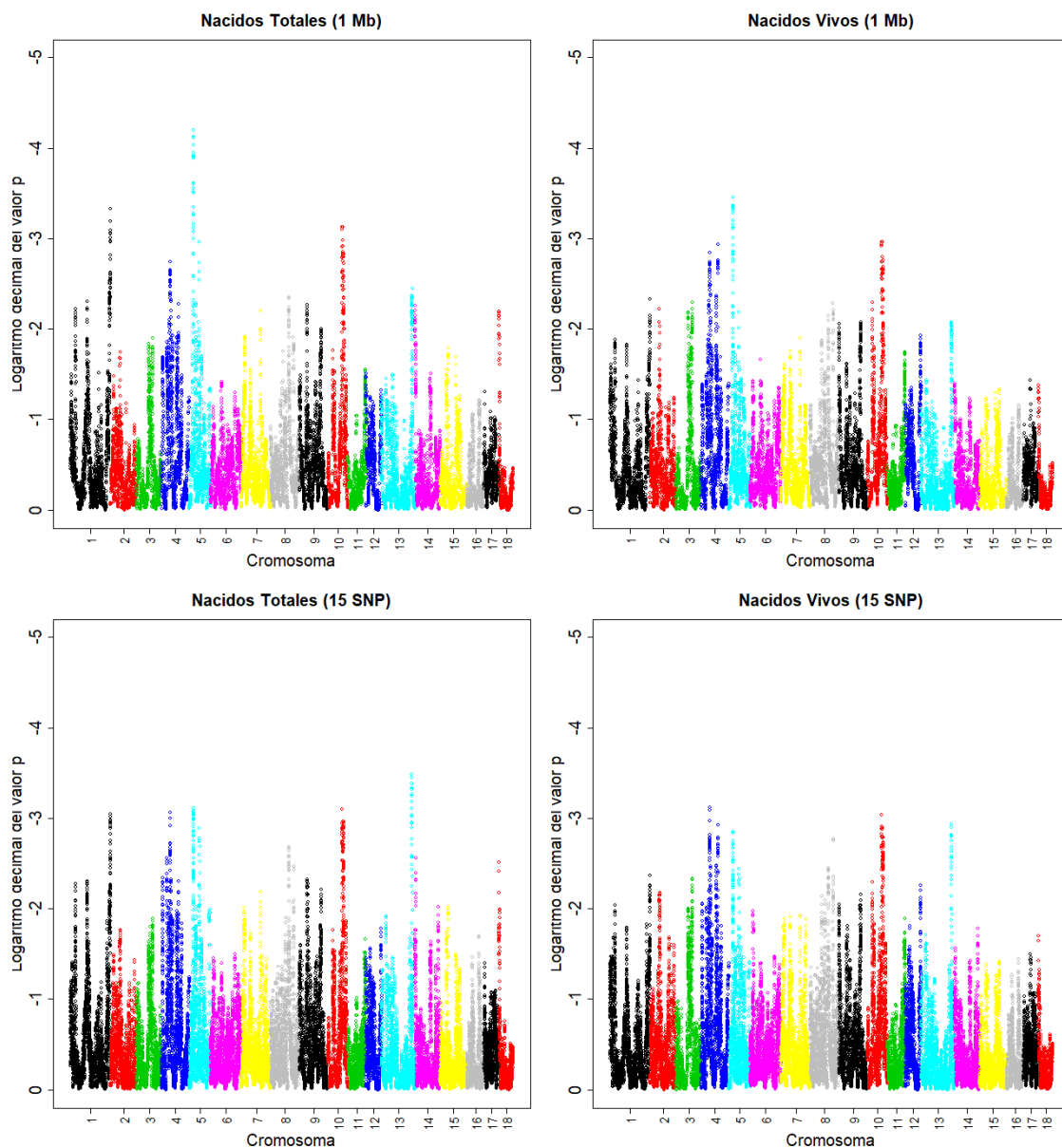


Figura 9. Representación gráfica del $-\log_{10}$ del p-valor asociado a la covariada con la heterocigosidad regional para la reducción de rendimiento fenotipo en los caracteres Nacidos Totales y Nacidos vivos con dos métodos de definición de regiones del genoma (1 Mb y 15 SNPs) en la población Retinto.

La autocorrelación entre los estimadores fue similar a la obtenida en la población Entrepelada (0.99), por lo tanto, se consideraron relevantes aquellas regiones del genoma asociadas a valores p inferiores a 0.0003, después de una transformación de Bonferroni corregida. En la población Retinto y en el caso de la asociación positiva con el grado de homocigosidad observada regional, se detectó una única región localizada en torno al marcador SNP rs330562274, localizado en el cromosoma 12 en la posición 48019618. Los genes localizados

en la cercanía de este marcador son NXN, MRM3, GEMIN4, TLCD3A, VPS53, RFLNB, RPH3AL, DOC2B, CRK, MYO1C, INPP5K, PITPNA, SLC43A2, SCARF1, RILP, PRPF8, TLCD2, WDR81, SERPINF2, SERPINF1, SMYD4, RPA1, RTN4RL1, DPH1, HIC1, SMG6, SGSM2, TSR1, MNT, METTL16, PAFAH1B1, CLUH y RAP1GAP2. Entre ellos, algunos merecen ser destacados como posibles genes candidatos relacionados con la prolificidad en porcino. El gen *SLC43A2* (*Solute Carrier Family 43 Member 2*) está asociado con el desarrollo uterino en rumiantes, con cambios según el día de ciclo estral, así como con en el embarazo temprano (Bazer et al., 2012). Los genes RPH3AL (*Rabphilin 3A Like (Without C2 Domains)*) y GEMIN4 (*Gem Nuclear Organelle Associated Protein 4*) han sido asociados a señales recientes de selección en poblaciones porcinas (Zhong et al., 2020) y ovinas (Dolebo et al., 2019). Además, la proteína codificada por el gen HIC1 (*Hypermethylated in cancer 1 protein*) interacciona con al complejo ARID1A (Van Rechem et al., 2009) que es esencial para la función del endometrio en el inicio de la gestación (Kim et al., 2015). Finalmente, se ha descrito el papel de la proteína codificada por el gen PAFAH1B1 (*Platelet Activating Factor Acetylhydrolase 1b Regulatory Subunit 1*) en el desarrollo embrionario (Yao et al., 2015).

En el caso de la reducción del rendimiento fenotípico, el umbral de significación definido por la corrección de Bonferroni solo fue superado por una región del cromosoma 5 en el caso de Nacidos Totales con regiones de 1 Mb. Esta región del genoma situada en torno al marcador rs335297714 que se localiza en la posición 12265777. Los genes localizados en un intervalo de 1 Mb alrededor de este marcador fueron los siguientes: TXN2, RBFOX2, SYN3, TIMP3, FBXO7, BPIFC, RTCB, ASCL4, PRDM4, PWP1 y BTBD11. Entre ellos, son destacables los genes *TXN2* (Thioredoxin 2) y *TIMP3* (Metalloproteínase inhibidor 3). El gen TXN2 ha sido propuesto como gen candidato para caracteres reproductivos en poblaciones Duroc (Zhang et al., 2019). Por otra parte, se ha demostrado que las proteínas asociadas al gen *TIMP3* median en la regulación de cambios y remodelación de la matriz extracelular folicular a medida que avanza el proceso de ovulación en humanos (Rosewell et al., 2013).

Si se comparan las estimaciones de la pendiente con la homocigosidad regional en las dos poblaciones (Retinto y Entrepelado), se observa que las zonas del genoma asociadas tanto al incremento como a la reducción del tamaño de camada son distintas. Además, cuando se calculó la correlación entre los estimadores de las pendientes asociadas con la homocigosidad regional entre las dos poblaciones oscilaron entre 0.071 (Nacidos Totales y regiones de 15 SNP) y 0.097 (Nacidos Totales y regiones de 1 Mb). Este resultado indica que la arquitectura genética de los efectos de la consanguinidad en el rendimiento fenotípico es específica de cada población, aunque se encuentre muy cercanas evolutivamente (Alonso et al., 2020). De hecho, en ninguno de las 2 poblaciones se encontró un efecto relevante en el cromosoma 13, en el que Saura et al., (2015) localizaron un mayor efecto de la depresión consanguínea en la variedad *Guadyervas*.

4.2. Segmentos de Homocigosidad y Depresión endogámica

Un segundo enfoque para el estudio de la arquitectura genética de la depresión endogámica se ha desarrollado a partir del estudio de los segmentos de homocigosidad o ROH (*Runs of Homozigosity*). El paquete R 'detectRUNS' (Biscarini et al. 2018) detectó 63,841 segmentos de homocigosidad entre los 427 individuos de la población Entrepelado y 82,102 segmentos de homocigosidad entre los 391 individuos de la población Retinta. El porcentaje del genoma cubierto por los segmentos de homocigosidad tuvo un valor medio del 24,2 % en la población Entrepelado y un 39,8% en la variedad Retinto, confirmando la presencia de un mayor grado de homocigosidad en la variedad Retinto.

Los ROH permiten distinguir entre consanguinidad reciente que se asocia a ROH de tamaño grande y la consanguinidad más antigua que está representada por ROH más pequeños. El fenómeno genético de la purga (Crnokrak y Barret, 2002) puede hacer que los efectos de la consanguinidad antigua sean menos nocivos en términos fenotípicos, como se ha comprobado en el tamaño de camada de ratones (Hinrichs et al., 2007). En este sentido, se ha implementado el mismo modelo mixto descrito en el apartado anterior incluyendo del porcentaje del

genoma cubierto por todos los ROH y por ROH de tamaño superior a 0.5, 1, 2, 3, 4, 6, 8, 12 y 16 Mb. Los resultados para las variedades Entrepelado y Retinto se presentan en las tablas 7 y 8, respectivamente. En ellas se presenta el porcentaje medio del genoma incluido en los ROH o F_{roh} (y su desviación típica entre paréntesis) y la pendiente de la covariada (y su error estándar entre paréntesis) de este porcentaje con los nacidos vivos y nacidos, así como su valor p de significación.

Tabla 7. Porcentaje medio del genoma incluido en los ROH (F_{roh} , con desviación típica entre paréntesis) y la pendiente de la covariada (d, con error estándar entre paréntesis) de este porcentaje con los Nacidos Vivos y Nacidos Totales, y valor p de significación para los diferentes tamaños de ROH (0.5, 1, 2, 3, 4, 6, 8, 12 y 16 Mb) en la población Entrepelado.

	F_{roh}	d (Nvivos)	Valor p	d (Ntotales)	Valor p
Todos	0.242 (0.052)	-1.911 (1.093)	0.040	-1.995 (1.128)	0.039
>0.5 Mb	0.238 (0.052)	-1.871 (1.086)	0.042	-1.957 (1.121)	0.041
>1 Mb	0.216 (0.053)	-1.829 (1.062)	0.043	-1.941 (1.096)	0.038
>2 Mb	0.179 (0.056)	-1.826 (1.012)	0.036	-1.885 (1.045)	0.036
>3 Mb	0.154 (0.058)	-1.773 (0.982)	0.036	-1.891 (1.014)	0.031
>4 Mb	0.136 (0.059)	-1.680 (0.963)	0.040	-1.767 (0.992)	0.037
>6 Mb	0.114 (0.059)	-1.689 (0.963)	0.040	-1.832 (0.994)	0.033
>8	0.098 (0.056)	-1.671 (1.013)	0.050	-1.831 (1.046)	0.040
>12	0.079 (0.052)	-1.841 (1.084)	0.044	-1.905 (1.119)	0.044
>16	0.063 (0.048)	-2.118 (1.139)	0.031	-2.37 (1.176)	0.022

Tabla 8. Porcentaje medio del genoma incluido en los ROH (F_{roh} , con desviación típica entre paréntesis) y la pendiente de la covariada (d , con error estándar entre paréntesis) de este porcentaje con los Nacidos Vivos y Nacidos Totales, y valor p de significación para los diferentes tamaños de ROH (0.5, 1, 2, 3, 4, 6, 8, 12 y 16 Mb) en la población Retinto.

	F_{roh}	d (Nvivos)	Valor p	d (Ntotales)	Valor p
Todos	0.396 (0.038)	-2.233 (1.596)	0.081	-1.811 (1.590)	0.127
>0.5 Mb	0.391 (0.038)	-2.282 (1.589)	0.075	-1.859 (1.583)	0.120
>1 Mb	0.371 (0.039)	-2.214 (1.556)	0.077	-1.821 (1.549)	0.120
>2 Mb	0.328 (0.042)	-2.180 (1.445)	0.066	-1.784 (1.438)	0.107
>3 Mb	0.295 (0.044)	-2.087 (1.378)	0.065	-1.745 (1.371)	0.102
>4 Mb	0.266 (0.045)	-1.940 (1.332)	0.072	-1.659 (1.326)	0.105
>6 Mb	0.227 (0.046)	-2.256 (1.285)	0.040	-1.919 (1.280)	0.067
>8 Mb	0.200 (0.046)	-2.516 (1.292)	0.026	-2.115 (1.287)	0.050
>12 Mb	0.154 (0.044)	-2.295 (1.323)	0.041	-2.164 (1.318)	0.050
>16 Mb	0.121 (0.041)	-1.933 (1.430)	0.088	-1.877 (1.426)	0.094

En las tablas se observa como todos los estimadores de la pendiente asociada al porcentaje del genoma incluido en ROH fueron negativos, y los valores p oscilaron en 0.022 (Nacidos Totales y regiones incluidas en ROH mayores de 16 Mb en Entrepelado) y 0.127 (Nacidos Totales y todas las regiones incluidas en ROH). Sin embargo, no se observa una tendencia clara que pueda sugerir que la consanguinidad más reciente tiene un mayor efecto en la depresión endogámica. Este resultado concuerda con el obtenido por Köck et al. (2009) en cerdos Landrace o Large White que observaron que tanto la consanguinidad reciente como la antigua contribuían a la depresión consanguínea en el tamaño de camada.

A continuación, se procedió al estudio del efecto de ROH específicos con el objetivo de abordar el estudio de la arquitectura genética de la depresión endogámica. Para ello, se utilizó el software *Haplofinder* (Howard et al., 2017), que fue capaz de localizar 6,172 y 6,572 ROH compartidos por al menos el 5% de los individuos y con más de 20 marcadores SNP en las variedades Entrepelado y Retinto, respectivamente. Para cada uno de estos ROH se realizó

un análisis de modelo mixto con los componentes de varianza obtenidos en el apartado anterior y se obtuvo un estimador asociado a la presencia o ausencia de cada ROH en los individuos. Los resultados de estos estimadores mostraron una gran dispersión como observa en la Figura 10.

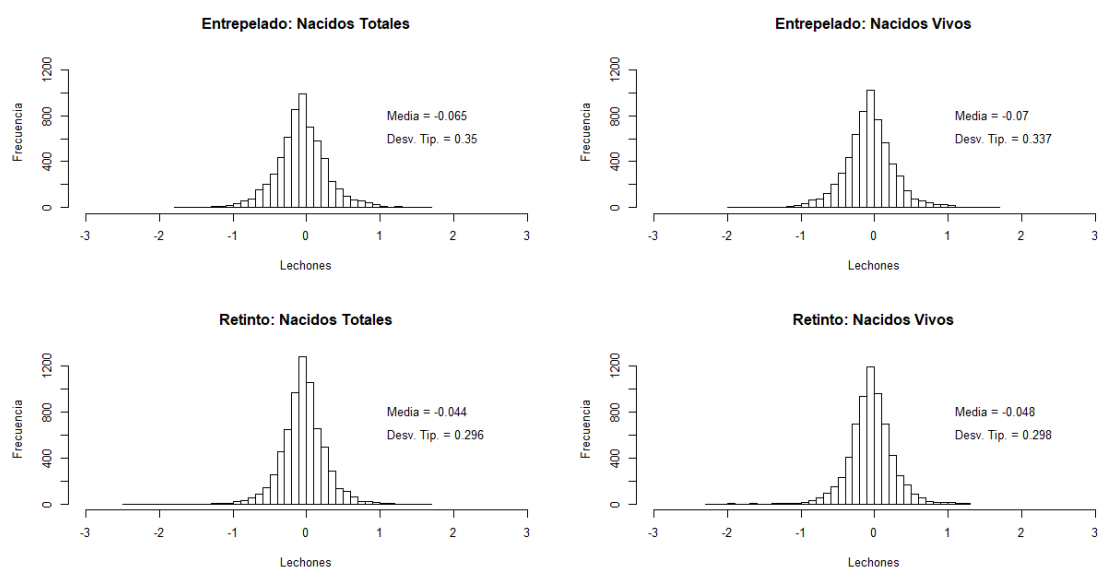


Figura 10. Distribución de los efectos asociados a los ROH para Nacidos Vivos y Nacidos Totales en las poblaciones Entrepelado y Retinto.

La media de los efectos asociados a los ROH fue ligeramente negativa, ya que osciló entre -0.048 lechones para Nacidos Vivos en Retinto hasta -0.070 lechones para Nacidos Vivos en Entrepelado. Sin embargo, se encuentra un reducido número de ROH con efectos mucho más relevantes. Si se toma el nivel de significación nominal ($p < 0.05$), el número de regiones con efecto significativa es muy elevada (Tabla 9).

Tabla 9. Porcentaje de segmentos de homocigosidad asociados significativamente para la reducción o incremento del fenotipo en los caracteres Nacidos Totales (NT) y Nacidos Vivos (NV) en las poblaciones Entrepelado y Retinto.

Población	Carácter	Reducción	Incremento
Entrepelado	NT	8.57 %	4.77 %
Entrepelado	NV	9.07 %	4.50 %
Retinto	NT	7.16 %	4.44 %
Retinto	NV	7.89 %	4.38 %

El porcentaje de ROH asociados con un valor $p < 0.05$ para la reducción del rendimiento fenotípico osciló entre el 7.16 y 9.07 % para Nacidos Totales y Retinto y Nacidos Vivos y Entrepelado, respectivamente. A su vez, el porcentaje de ROH con una valor $p < 0.05$ para el incremento de los carácter varió entre el 4.38 % para Nacidos Vivos en Retinto y el 4.77 % para Nacidos Totales en Entrepelado. Este resultado concuerda con los estimadores negativas de la covariadas con el porcentaje de homocigosidad y con el porcentaje del genoma incluido en ROH que han presentado en los apartados anteriores. Además, las ROH significativas se repartieron de manera prácticamente homogénea a lo largo del genoma, como se muestra en la figura 11.

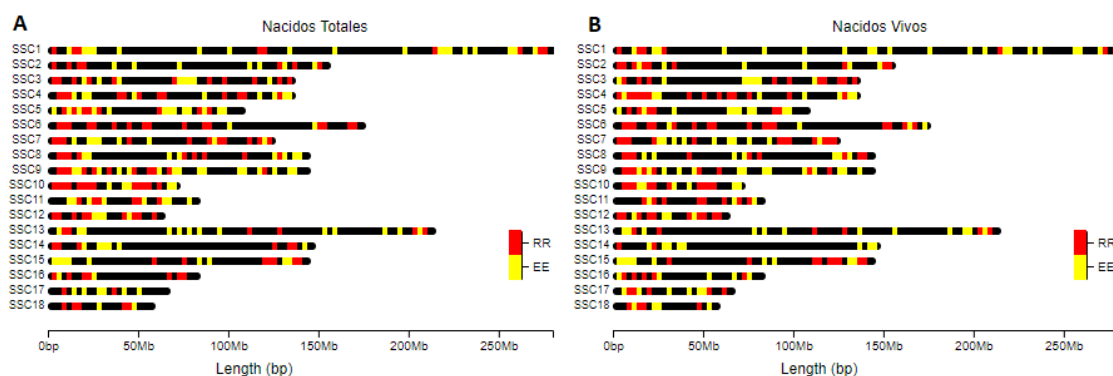


Figura 11. Distribución de los segmentos de homocigosidad asociados significativamente para los caracteres Nacidos Totales (A) y Nacidos Vivos (B) en las poblaciones Retinto (RR) y Entrepelado (EE).

Sin embargo, si se realiza una corrección por comparaciones múltiples mediante la corrección de Bonferroni (1936) ninguno de los ROH puede ser identificado como significativo. Pese a todo, al realizar una corrección para una tasa de falso descubrimiento o *False Discovery Rate -FDR-* (Benjamini y Hochberg, 1995) al 50% sí que fue posible identificar algunos ROH susceptibles de albergar genes relacionados con la depresión endogámica o con el incremento del rendimiento en homocigosidad. Los ROH asociados con la reducción de Nacidos Vivos o Nacidos Totales y con un FDR inferior a 0.50 en la población Entrepelado se presentan en la Tabla 10.

Tabla 10. Cromosoma (SSC), posición en pares de bases (Bp) de los segmentos de homocigosidad (ROH) asociadas con la reducción de Nacidos

Vivos (NV) y Nacidos Totales (NT) con una tasa de falso descubrimiento (FDR) inferior al 0.50 en la población Entrepelado y genes presentes en ellos.

SSC	Bp	Carácter	FDR	Genes
4	126099518- 128333002	NV	0.184	ZNF326, LRRC8D, LRRC8C, LRRC8B, GBP2, KYAT3, GTF2B, PKN2
4	40408391- 42374474	NV	0.475	MTERF3, GDF6, PLEKHF2, NDUFAF6, TP53INP1, CCNE2, INTS8, DPY19L4, ESRP1, VIRMA, GEM, RAD54B, FSBP
5	15997122- 17020150	NT	0.456	SMARCD1, GPD1, CERS5, LIMA1, LARP4, DIP2B, ATF1, TMPRSS12, SLC4A8, SCN8A
8	130854564- 131436362	NV	0.035	ABCG2, PKD2, SPP1, MEPE, IBSP, DMP1, DSPP, SPARCL1
13	191695303- 192188992	NT	0.443	

La región con un menor FDR (0.035) se localizó en el cromosoma 8 y contiene al gen SSP1 (*Secreted Phosphoprotein 1*), que tiene un rol importante en la gestación temprana en diversas especies (Hernández et al., 2013) y ha sido asociada con el tamaño de camada en cerdos (King et al., 2003). Además, esta misma región contiene al gen SPARCL1 (*Secreted Protein Acidic and Rich in Cysteine Like 1*), que está relacionado con la regulación de la progesterona en la implantación del embrión (Dassen et al, 2007). La segunda región con un FDR inferior (0.184) se situó en el cromosoma 4, y entre los genes presentes en su interior, se localiza el gen PKN2 (*Protein Kinase N2*), que, en ratones, se ha probado que es determinante para el desarrollo embrionario (Danno et al., 2017).

Adicionalmente, hay 3 regiones con FDR cerca de 0.50. En el cromosoma 4 se localiza y entre las pares de bases 40408391 y 42374474 se localiza una de ellas, que contiene al gen MTERF4 (*Mitochondrial Transcription Termination Factor 4*), que ha sido asociado con la mortalidad embrionaria en ratones (Camara et al., 2011). Por otra parte, el ROH localizado en el cromosoma 13 no contiene ningún gen en su interior, pero se encuentra cerca de un QTL para

tamaño de camada que fue propuesto por He et al. (2017) y que contiene al gen USP16 que regula la expresión de las células madre embrionarias (Yang et al., 2014). Finalmente, la última de las regiones se localiza en el cromosoma 5, y entre los genes presentes en su interior merece la pena destacar al gen ATF1 (*Activating Transcription Factor 1*), implicado en la ruta de señalización del estrógeno (Liu et al., 2019)

En el caso de los ROH asociados con el incremento de los fenotipos, las regiones con un FDR inferior a 0.50 están recogidas en la Tabla 11.

Tabla 11. Cromosoma (SSC), posición en pares de bases (Bp) de los segmentos de homocigosidad (ROH) asociadas con el incremento de Nacidos Vivos (NV) y Nacidos Totales (NT) con una tasa de falso descubrimiento (FDR) inferior al 0.50 en la población Entrepelado y genes presentes en ellos.

SSC	Bp	Carácter	FDR	Genes
5	74360252- 75644214	NV	0.273	ADAMTS20, PUS7L, IRAK4, TWF1, TMEM117, NELL2
9	42349931- 43868205	NV	0.404	CADM1
9	44558610- 45448558	NV-NT	0.114	SIDT2, TAGLN, PCSK7, RNF214, BACE1, CEP164, DSCAML1, TMPRSS13, IL10RA, SMIM35, TMPRSS4

La región con menor FDR (0.114) la encontramos en el cromosoma 9 y está asociada tanto a Nacidos Vivos como a Nacidos Totales. En esta región se encuentran los genes CEP164 (*Centrosomal Protein 164*), que se considera causante de la infertilidad (Siller et al., 2017), DSCAML1 (*DS Cell Adhesion Molecule Like 1*) altamente asociado con el tamaño de camada en cabra (Wang et al., 2020) y IL10RA (*Interleukin 10 Receptor Subunit Alpha*), que desempeña un papel importante en la remodelación vascular en el embarazo (Zeng et al., 2019). Las otras regiones con FDR inferior a 0.5 se sitúan en el cromosoma 5 (FDR = 0.273) y en el cromosoma 9 (FDR = 0.404). En la primera, se localizan el gen ADAMTS20 (*ADAM Metallopeptidase With Thrombospondin Type 1 Motif 20*), cuya expresión en el endometrio uterino porcino es significativamente mayor

el día 12 de embarazo que el día 12 del ciclo estral (Kim et al., 2012), mientras que en la segunda se localiza únicamente el gen *CADM1* (*Cell Adhesion Molecule 1*) que ha sido asociado con la espermatogénesis en ratones (Van der Weyden et al., 2006)

Las ROH asociado a una tasa de falso descubrimiento o FDR inferior a 0.5 para la reducción de Nacidos Vivos y Nacidos Totales en la población Entrepelado se presentan en la tabla 12.

Tabla 12. Cromosoma (SSC), posición en pares de bases (Bp) de los segmentos de homocigosidad (ROH) asociadas con la reducción de Nacidos Vivos (NV) y Nacidos Totales (NT) con una tasa de falso descubrimiento (FDR) inferior al 0.50 en la población Retinto y genes presentes en ellos.

SSC	Bp	Caracter	FDR	Genes
6	10415017- 10977337	NV-NT	0.310	ADAMTS18, MON1B
6	54054875- 55453781	NV	0.480	MAMSTR, IZUMO1, RASIP1, FUT1, FGF21, BCAT2, PLEKHA4, PPP1R15A, TULP2, NUCB1, DHDH, FTL, GYS1, RUVBL2, NTF4, KCNA7, SNRNP70, LIN7B, PPFIA3, HRC, TRPM4, CD37, TEAD2, DKKL1, CCDC155, PTH2, GFY, SLC17A7, PIH1D1, ALDH16A1, RPS11, FCGRT, RCN3, PRRG2, NOSIP, PRR12, RRAS, SCAF1, IRF3, BCL2L12, PRMT1, CPT1C, TSKS, AP2A1, MED25, FUZ, PTOV1, PNKP, AKT1S1, TBC1D17, IL4I1, NUP62, ATF5, VRK3, ZNF473, IZUMO2, MYH14, KCNC3, NAPSA, NR1H2, POLD1, MYBPC2, FAM71E1, EMC10, ASPDH, LRRC4B, SYT3
15	122076948- 122863462	NV	0.083	

De los 3 ROH identificados, el de menor FDR (0.083) se localizó en el cromosoma 15, en cuya cercanía se pueden encontrar genes como el *INHA*

(*Inhibin Subunit Alpha*) que se ha asociado con un el tamaño de la camada en cabras (Wu et al., 2009) o el EPAA4 (*EPH Receptor A4*), que ha sido propuesto por Fu et al. (2012) como gen candidato para la prolificidad porcina.

Las otras dos regiones se encuentran localizadas en el cromosoma 6 y, en ambas ocasiones, con un FDR próximo a 0.5. En la primera de ellas, situada en torno a la décima Megabase del cromosoma, merecen ser destacados el gen MON1B (*MON1 Homolog B, Secretory Trafficking Associated*), que ha sido asociado con el desarrollo de blastocistos en el ganado bovino (Cochran et al., 2013) y gen ADAMTS18 (*ADAM Metallopeptidase With Thrombospondin Type 1 Motif 18*) que juega un papel en el desarrollo del aparato reproductor femenino (Ataca et al., 2016). En la segunda, localizada entre los pares de bases 54054875 y 55453781, son destacables los genes PNKP (*Polynucleotide Kinase 3'-Phosphatase*), asociado con la provocación de abortos (Derks et al., 2019), el MYH14 (*Myosin Heavy Chain 14*) que produce una de las proteínas más abundantes en las secreciones oviductales (Saint-Dizier et al., 2019), el RCN3 (*Reticulocalbin 3*), que regula la formación del folículo primordial (Xu et al., 2017); y el PTOV (*PTOV1 Extended AT-Hook Containing Adaptor Protein*) que se ha asociado con la respuesta a la estimulación de estrógeno durante el desarrollo del ovocito en ratones (Yao et al., 2011).

Por último, para el incremento de la Nacidos Vivos, tan solo se obtuvo un único segmento con una tasa de falso descubrimiento inferior a 0.50 (FDR = 0.321) y se localiza en el cromosoma 4 entre los pares de bases 62640640 y 63563088. Este segmento contiene los genes SBSPON, TERF1, KCNB2, TRPA1, y entre ellos, merece la pena destacar al gen TRPA1 (*Transient Receptor Potential Cation Channel Subfamily A Member 1*), que interactúa con el estrógeno y está presente en el endometrio en ratas (Pohóczy et al., 2011)

Como en el análisis anterior, y tal y como se puede observar en la figura 11, ninguno de los ROH significativos es compartido entre ambas variedades.

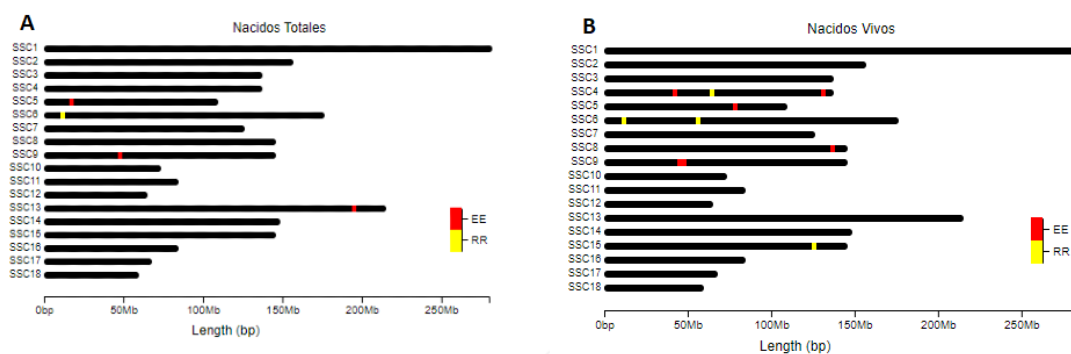


Figura 11. Distribución de los segmentos de homocigosidad asociados significativamente para los caracteres Nacidos Totales (A) y Nacidos Vivos (B) con FDR < 0.5 en las poblaciones Retinto (RR) y Entrepelado (EE).

4.3. Consideraciones Finales

Los resultados de las regiones del genoma asociadas con la homocigosidad regional o con los segmentos de homocigosidad son diferentes entre las variedades Entrepelado y Retino. Las causas de estas diferencias pueden deberse a la divergencia evolutiva entre las poblaciones. Sin embargo, al tratarse de población con un grado de diferenciación escaso (Alonso et al., 2020), es plausible que, aunque los efectos de la endogamia en el rendimiento reproductivo sean heterogéneos a lo largo del genoma, los efectos asociados a cada gen o región del genoma sean en sí mismos de pequeño efecto. Por lo tanto, estos efectos pueden ser indetectables con poblaciones de tamaño reducido, como la utilizadas en este estudio. De hecho, los resultados obtenidos con los dos procedimientos propuestos en este trabajo también proporcionan resultados no coincidentes incluso en la misma población (Retinto y Entrepelado) y carácter (Nacidos Vivos y Totales). Pese a todo, sí que se ha podido observar que las consecuencias de la endogamia en el rendimiento en prolificidad sí que son heterogéneas a lo largo del genoma y es destacable que, aunque se detectan en mayor medida efectos negativos, ha sido posible detectar regiones o segmentos con una asociación positiva con Nacidos Vivos o Totales. Este fenómeno es explicable con la presencia de genes cuyos efectos serán principalmente aditivos y estén presente con frecuencias alélicas superiores a 0.5 para el alelo favorable.

Finalmente, es necesario mencionar que las regiones del genoma o segmentos de homocigosidad detectados pueden considerarse exclusivamente como

sugestivos, aunque la búsqueda de los genes localizados en su interior ha permitido identificar numerosos genes implicados en los procesos biológicos asociados con la prolificidad, que deben ser confirmados en futuros estudios con una mayor cantidad de información fenotípica y genómica.

5. CONCLUSIONES

Los resultados de este trabajo han permitido extraer las siguientes conclusiones:

1. La variedad Retinto presenta un grado de homocigosidad más elevado que la variedad Entrepelado.
2. Los resultados del análisis con el grado de homocigosidad individual y el porcentaje del genoma comprendido en segmentos de homocigosidad sugieren la presencia de depresión endogámica
3. Los efectos de la endogamia en el rendimiento fenotípico se distribuyen de manera heterogénea a lo largo del genoma.
4. Las regiones del genoma asociadas con el incremento o reducción del tamaño de camada difieren entre variedades y métodos de detección.
5. Se han detectado numerosos genes potencialmente implicados en los efectos de la endogamia en el rendimiento reproductivos en cerdo ibérico que deben ser confirmados en futuros estudios.

6.- BIBLIOGRAFÍA

- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC bioinformatics*, 12, 246. <https://doi.org/10.1186/1471-2105-12-246>
- Alonso, I., Ibanez, N., Noguera, J.L., Casellas, J., Martín de Hijas-Villalba, M., Santanta, M. & Varona, L. (2020). Genomic differentiation among varieties of Iberian pig. *Spanish Journal Of Agricultural Research*. 18. <https://doi.org/10.5424/sjar/2020181-15411>.
- Alves, E., Barragán, C. & Toro, M. A. 2008. Short communication. Inbreeding and homozygosity in Iberian pigs. *Spanish Journal of Agriculture Research* 6: 248-251. <https://doi.org/10.5424/sjar/2008062-323>.
- Alves, E.C., Fernández, A.I., Barragán, C., Óvilo, C., Rodríguez, C., & Silio, L.F. (2006). Inference of hidden population substructure of the Iberian pig breed using multilocus microsatellite data. *Spanish Journal of Agricultural Research*, 4, 37-46. <https://doi.org/10.5424/sjar/2006041-176>
- Andersson, E., Frössling, J., Engblom, L., Algers, B., & Gunnarsson, S. (2016). Impact of litter size on sow stayability in Swedish commercial piglet producing herds. *Acta veterinaria Scandinavica*, 58(1), 31. <https://doi.org/10.1186/s13028-016-0213-8>
- Ataca, D., Caikovski, M., Piersigilli, A., Moulin, A., Benarafa, C., Earp, S. E., Guri, Y., Kostic, C., Arsenijevic, Y., Soininen, R., Apte, S. S., & Brisken, C. (2016). Adams18 deletion results in distinct developmental defects and provides a model for congenital disorders of lens, lung, and female reproductive tract development. *Biology open*, 5(11), 1585–1594. <https://doi.org/10.1242/bio.019711>
- Baxter, R., Hastings, N., Law, A., & Glass, E. J. (2008). A rapid and robust sequence-based genotyping method for BoLA-DRB3 alleles in large numbers of heterozygous cattle. *Animal genetics*, 39(5), 561–563. <https://doi.org/10.1111/j.1365-2052.2008.01757.x>
- Bazer, F., Kim, J., Song, G., Satterfield, M., Johnson, G., Burgardt, R. & Wu, Gem. (2012). Uterine environment and conceptus development in ruminants. *Animal reproduction*. 9. 297-304. Recuperado de <https://www.animal-reproduction.org/article/5b5a6058f7783717068b46e9/pdf/animreprod-9-3-297.pdf>
- Benito, J., Albarrán, A., & García Casco, J. M. (2006). Extensive Iberian pig production grazing systems. In *Proceedings of the 21st General Meeting of the European Grassland Federation, Badajoz, Spain* (pp. 3-6). Recuperado de <https://www.cabdirect.org/cabdirect/abstract/20063101368>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal*

Statistical Society. Series B (Methodological), 57(1), 289-300. Recuperado de www.jstor.org/stable/2346101.

- Biscarini, F., Cozzi, P., Gaspa, G., & Marras, G. (2018). detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes. Recuperado de <http://orca.cf.ac.uk/id/eprint/108906>
- Bonferroni, C.E. Teoria statistica delle classi e calcolo delle probabilita. (1936) (No. 8). Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze. <https://dx.doi.org/10.4135/9781412961288.n455>
- Braun, B. C. , Halaski, N. , Painer, J., Krause, E., Jewgenow, K. The antioxidative enzyme SOD2 is important for physiological persistence of corpora lutea in lynxes. *Scientific Reports*, 2020; 10 (1) <https://doi.org/10.1038/s41598-020-60634-x>
- Cámara, Y., Asin-Cayuela, J., Park, C. B., Metodiev, M. D., Shi, Y., Ruzzenente, B., Kukat, C., Habermann, B., Wibom, R., Hultenby, K., Franz, T., Erdjument-Bromage, H., Tempst, P., Hallberg, B. M., Gustafsson, C. M., & Larsson, N. G. (2011). MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. *Cell metabolism*, 13(5), 527–539. <https://doi.org/10.1016/j.cmet.2011.04.002>.
- Casellas, J., Ibáñez-Escriche, N., Varona, L., Rosas, J. P., & Noguera, J. L. (2019). Inbreeding depression load for litter size in Entrepelado and Retinto Iberian pig varieties1. *Journal of animal science*, 97(5), 1979–1986. <https://doi.org/10.1093/jas/skz084>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11), 783–796. <https://doi.org/10.1038/nrg2664>
- Chen, P., Baas, T. J., Mabry, J. W., Koehler, K. J., & Dekkers, J. C. (2003). Genetic parameters and trends for litter traits in U.S. Yorkshire, Duroc, Hampshire, and Landrace pigs. *Journal of animal science*, 81(1), 46–53. <https://doi.org/10.2527/2003.81146x>
- Cochran, S. D., Cole, J. B., Null, D. J., & Hansen, P. J. (2013). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC genetics*, 14, 49. <https://doi.org/10.1186/1471-2156-14-49>
- Crackower, M. A. (2003). Essential Role of Fkbp6 in Male Fertility and Homologous Chromosome Pairing in Meiosis. *Science*, 300(5623), 1291-1295. <https://doi.org/10.1126/science.1083022>

- Crow, J. F., & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. Blackburn Press, Caldwell, NJ.
<https://doi.org/10.1002/tera.1420050318>
- Danno, S., Kubouchi, K., Mehruba, M., Abe, M., Natsume, R., Sakimura, K., Eguchi, S., Oka, M., Hirashima, M., Yasuda, H., & Mukai, H. (2017). PKN2 is essential for mouse embryonic development and proliferation of mouse fibroblasts. *Genes to cells : devoted to molecular & cellular mechanisms*, 22(2), 220–236. <https://doi.org/10.1111/gtc.12470>
- Dassen, H., Punyadeera, C., Kamps, R., Klomp, J., Dunselman, G., Dijcks, F., de Goeij, A., Ederveen, A., & Groothuis, P. (2007). Progesterone regulation of implantation-related genes: new insights into the role of oestrogen. *Cellular and molecular life sciences : CMLS*, 64(7-8), 1009–1032.
<https://doi.org/10.1007/s00018-007-6553-9>
- Derks, M., Gjuvslund, A. B., Bosse, M., Lopes, M. S., van Son, M., Harlizius, B., Tan, B. F., Hamland, H., Grindflek, E., Groenen, M., & Megens, H. J. (2019). Loss of function mutations in essential genes cause embryonic lethality in pigs. *PLoS genetics*, 15(3), e1008055.
<https://doi.org/10.1371/journal.pgen.1008055>
- Dolebo, A. T. (2019, 22 noviembre). Genome-wide scans identify known and novel regions associated with prolificacy and reproduction traits in a sub-Saharan African indigenous sheep (*Ovis aries*). *Mammalian Genome*.
<https://doi.org/10.1007/s00335-019-09820-5>
- El-Mestrah, M., Castle, P. E., Borossa, G., & Kan, F. W. (2002). Subcellular distribution of ZP1, ZP2, and ZP3 glycoproteins during folliculogenesis and demonstration of their topographical disposition within the zona matrix of mouse ovarian oocytes. *Biology of reproduction*, 66(4), 866–876.
<https://doi.org/10.1095/biolreprod66.4.866>
- Fabuel, E., Barragán, C., Silió, L., Rodríguez, M. C., & Toro, M. A. (2004). Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity*, 93(1), 104–113.
<https://doi.org/10.1038/sj.hdy.6800488>
- Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics* (4th Edition) . Pearson..
- Farkas, J., Curik, I., Csató, L., Csörnyei, Z., Baumung, R., István, N. (2007). Bayesian inference of inbreeding effects on litter size and gestation length in Hungarian Landrace and Hungarian Large White pigs. *Livestock Science*. 112. 109-114. <https://doi.org/10.1016/j.livsci.2007.01.160>.
- Fernández, A., Rodrigáñez, J., Zuzúarregui, J., Rodríguez, M. C., & Silió, L. 2008. Genetic parameters for litter size and weight at different parities in Iberian pigs. *Spanish Journal of Agricultural Research* 6: 98-106.
<https://doi.org/10.5424/sjar/200806S1-378>

- Fu, Y., Wang, G., Fu, J., & Wang, A. (2012). Association of EphrinB2 Gene Polymorphism with Litter Size in Pigs. *Advances in Intelligent and Soft Computing*, 811-818. https://doi.org/10.1007/978-3-642-27537-1_95
- Gandemer, G. (2009). Dry cured ham quality as related to lipid quality of raw material and lipid changes during processing: A review. *Grasas y Aceites*, 60. <https://doi.org/10.3989/gya.130908>.
- Hao, Q., Zhu, Z., Xu, D., Liu, W., Lyu, L., & Li, P. (2019). Proteomic characterization of bovine granulosa cells in dominant and subordinate follicles. *Hereditas*, 156, 21. <https://doi.org/10.1186/s41065-019-0097-5>
- Horikoshi, M., & Tang, Y. (2018). ggfortify: Data Visualization Tools for Statistical Analysis Results. CRAN - Package ggfortify. <https://cran.r-project.org/web/packages/ggfortify/index.html>.
- Hou, J., An, X., Li, G., Wang, Y., Song, Y., & Cao, B. (2012). Exploring polymorphisms and their effects on reproductive traits of the INHA and INH β A genes in three goat breeds. *Animal science journal = Nihon chikusan Gakkaiho*, 83(4), 273–278. <https://doi.org/10.1111/j.1740-0929.2011.00968.x>
- Hunt, S. E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., Parton, A., Armean, I. M., Trevanion, S. J., Flicek, P., & Cunningham, F. (2018). Ensembl variation resources. *Database : the journal of biological databases and curation*, 2018, bay119. <https://doi.org/10.1093/database/bay119>
- Hunzicker-Dunn, M. E., Lopez-Biladeau, B., Law, N. C., Fiedler, S. E., Carr, D. W., & Maizels, E. T. (2012). PKA and GAB2 play central roles in the FSH signaling pathway to PI3K and AKT in ovarian granulosa cells. *Proceedings of the National Academy of Sciences of the United States of America*, 109(44), E2979–E2988. <https://doi.org/10.1073/pnas.1205661109>
- Inga Food (2016). La Cerda Castúa, un proyecto genético en cerdo ibérico. Recuperado de: <https://www.ingafood.es/Settings/News/divulgacion/cerda-castua-iberico-genetica/>
- Jiang, X., Lu, N., Xue, Y., Liu, S., Lei, H., Tu, W., Lu, Y., & Xia, D. (2019). Crude fiber modulates the fecal microbiome and steroid hormones in pregnant Meishan sows. *General and comparative endocrinology*, 277, 141–147. <https://doi.org/10.1016/j.ygcen.2019.04.006>
- Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, 189(1), 237–249. <https://doi.org/10.1534/genetics.111.130922>
- Kim, M., Seo, H., Choi, Y., Shim, J., Kim, H., Lee, C. K., & Ka, H. (2012).

Microarray Analysis of Gene Expression in the Uterine Endometrium during the Implantation Period in Pigs. *Asian-Australasian journal of animal sciences*, 25(8), 1102–1116. <https://doi.org/10.5713/ajas.2012.12076>

Kim, T. H., Yoo, J. Y., Wang, Z., Lydon, J. P., Khatri, S., Hawkins, S. M., Leach, R. E., Fazleabas, A. T., Young, S. L., Lessey, B. A., Ku, B. J., & Jeong, J. W. (2015). ARID1A Is Essential for Endometrial Function during Early Pregnancy. *PLoS genetics*, 11(9), e1005537. <https://doi.org/10.1371/journal.pgen.1005537>

King, A. H., Jiang, Z., Gibson, J. P., Haley, C. S., & Archibald, A. L. (2003). Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8. *Biology of reproduction*, 68(6), 2172–2179. <https://doi.org/10.1095/biolreprod.102.012955>

Kristensen, T. & Sørensen, C. (2005). Inbreeding - Lessons from animal breeding, evolutionary biology and conservation genetics. *Animal Science*. 80. 121 - 133. <https://doi.org/10.1079/ASC41960121>.

Laval G., Iannucelli N., Legault C., Milan D., Groenen MAM, Giuffra E., Andersson L., Nissen P H, Jorgensen CB., Beeckmann P., Geldermann H., Foulley JL., Chevalet C., & Ollivier L. (2000) Genetic diversity of eleven European pig breeds. *Genet Sel Evol* 32:187-203. <https://doi.org/10.1186/1297-9686-32-2-187>

Li, X., Ye, J., Han, X., Qiao, R., Li, X., Lv, G., & Wang, K. (2020). Whole-genome sequencing identifies potential candidate genes for reproductive traits in pigs. *Genomics*, 112(1), 199–206. <https://doi.org/10.1016/j.ygeno.2019.01.014>

Liu, C., Ran, X., Yu, C., Xu, Q., Niu, X., Zhao, P., & Wang, J. (2019). Whole-genome analysis of structural variations between Xiang pigs with larger litter sizes and those with smaller litter sizes. *Genomics*, 111(3), 310–319. <https://doi.org/10.1016/j.ygeno.2018.02.005>

Lopez, B., & Seo, K. (2019). Genetic parameters for litter traits at different parities in purebred Landrace and Yorkshire pigs. *Animal science journal*, 90(12), 1497–1502. <https://doi.org/10.1111/asj.13298>

Lopez-Bote, C.J. (1998). Sustained Utilization of the Iberian Pig Breed. *Meat science*. 49S1. S17-27. [https://doi.org/10.1016/S0309-1740\(98\)90036-5](https://doi.org/10.1016/S0309-1740(98)90036-5).

Malécot, G. (1948). *Les Mathématiques de L'hérédité*. Masson, Paris, France.

Martinez, A.M., Delgado, J.V., Rodero, A. & Vega-Pla, J.L. (2000). Genetic structure of the Iberian pig breed using microsatellites. *Animal Genetics*, 31: 295-301. <https://doi.org/10.1046/j.1365-2052.2000.00645.x>

Matsumoto, H., Daikoku, T., Wang, H., Sato, E., & Dey, S. K. (2004). Differential expression of ezrin/radixin/moesin (ERM) and ERM-associated adhesion

molecules in the blastocyst and uterus suggests their functions during implantation. *Biology of reproduction*, 70(3), 729–736.
<https://doi.org/10.1095/biolreprod.103.022764>

Mayer, J., Beck, J., Soller, J. T., Wemheuer, W., Schütz, E., & Brenig, B. (2013). Analysis of circulating DNA distribution in pregnant and nonpregnant dairy cows. *Biology of reproduction*, 88(2), 29.
<https://doi.org/10.1095/biolreprod.112.103168>

McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., Campbell, H., ... Wilson, J. F. (2008). Runs of Homozygosity in European Populations. *American Journal of Human Genetics*, 83(5), 658.
<https://doi.org/10.1016/j.ajhg.2008.10.009>

Mendonça, A., Guimarães, A. L., da Silva, N. M., Caetano, A. R., Dode, M. A., & Franco, M. M. (2015). Characterization of the IGF2 Imprinted Gene Methylation Status in Bovine Oocytes during Folliculogenesis. *PloS one*, 10(10), e0142072. <https://doi.org/10.1371/journal.pone.0142072>

Meuwissen, T., & Luo, Z. (1992). Computing inbreeding coefficients in large populations. *Genetics Selection Evolution*, 24(4), 305.
<https://doi.org/10.1186/1297-9686-24-4-305>

Miao, Z. H., Glatz, P. C., & Ru, Y. J. (2004). Review of Production, Husbandry and Sustainability of Free-range Pig Production Systems. *Asian-Australasian Journal of Animal Sciences*, 17(11), 1615-1634.
<https://doi.org/10.5713/ajas.2004.1615>

Ministerio de Agricultura, Pesca y Alimentación (MAPA)(2016). Catalogo oficial de razas. Recuperado de:
<https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/razas-ganaderas/razas/catalogo/default.aspx>

Ministerio de Agricultura, Pesca y Alimentación (MAPA)(2016). Indicadores económicos del sector porcino año 2017. Recuperado de :
https://www.mapa.gob.es/es/ganaderia/estadisticas/indicadoreseconomico ssectorporcinoano2017_tcm30-379728.pdf

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). BLUPF90 and related programs (BGF90). In *Proceedings of the 7th world congress on genetics applied to livestock production* (Vol. 33, pp. 743-744).

Neupane, M., Geary, T. W., Kiser, J. N., Burns, G. W., Hansen, P. J., Spencer, T. E., & Neibergs, H. L. (2017). Loci and pathways associated with uterine capacity for pregnancy and fertility in beef cattle. *PloS one*, 12(12), e0188997. <https://doi.org/10.1371/journal.pone.0188997>

- Noguera, J. L., Ibáñez-Escriche, N., Casellas, J., Rosas, J. P., & Varona, L. (2019). Genetic parameters and direct, maternal and heterosis effects on litter size in a diallel cross among three commercial varieties of Iberian pig. *Animal : an international journal of animal bioscience*, 13(12), 2765–2772. <https://doi.org/10.1017/S1751731119001125>
- Onteru, S. K., Fan, B., Du, Z. Q., Garrick, D. J., Stalder, K. J., & Rothschild, M. F. (2012). A whole-genome association study for pig reproductive traits. *Animal genetics*, 43(1), 18–26. <https://doi.org/10.1111/j.1365-2052.2011.02213.x>
- Pangas, S. A., Jorgez, C. J., Tran, M., Agno, J., Li, X., Brown, C. W., Kumar, T. R., & Matzuk, M. M. (2007). Intraovarian activins are required for female fertility. *Molecular endocrinology (Baltimore, Md.)*, 21(10), 2458–2471. <https://doi.org/10.1210/me.2007-0146>
- Patterson, H., & Thompson, R. (1971). Recovery of Inter-Block Information when Block Sizes are Unequal. *Biometrika*, 58(3), 545-554. <https://doi.org/10.2307/2334389>
- Peng, X. R., Hsueh, A. J., & Ny, T. (1993). Transient and cell-specific expression of tissue-type plasminogen activator and plasminogen-activator-inhibitor type 1 results in controlled and directed proteolysis during gonadotropin-induced ovulation. *European journal of biochemistry*, 214(1), 147–156. <https://doi.org/10.1111/j.1432-1033.1993.tb17907.x>
- Pohóczky, K., Kun, J., Szalontai, B., Szőke, É., Sággy, É., Payrits, M., Kajtár, B., Kovács, K., Környei, J., Garai, J., Garami, A., Perkecz, A., Czeglédi, L., & Helyes, Z. (2016). Estrogen-dependent up-regulation of TRPA1 and TRPV1 receptor proteins in the rat endometrium, *Journal of Molecular Endocrinology*, 56(2), 135-149. <https://doi.org/10.1530/JME-15-0184>
- Putz, A. M., Tiezzi, F., Maltecca, C., Gray, K. A., & Knauer, M. T. (2015). Variance component estimates for alternative litter size traits in swine. *Journal of animal science*, 93(11), 5153–5163. <https://doi.org/10.2527/jas.2015-9416>
- Real Decreto 4/2014, de 10 de enero, por el que se aprueba la norma de calidad para la carne, el jamón, la paleta y la caña de lomo ibérico. Boletín Oficial del Estado, Núm. 10, del Sábado 11 de enero de 2014 Sec. I. Pág. 1569 Recuperado de https://www.mapa.gob.es/es/ganaderia/temas/produccion-y-mercados-ganaderos/Real%20Decreto%204_2014%20%20Norma%20de%20calidad%20para%20la%20carne%20el%20jam%C3%B3n%20la%20paleta%20y%20la%20ca%C3%B1a%20de%20lomo%20ib%C3%A9rico_tcm30-58901.pdf
- Rodriguez, C., Rodrigañez, J., & Silio, L. (1994). Genetic analysis of maternal ability in Iberian pigs. *Journal of animal breeding and genetics*, 111(1-6), 220–227. <https://doi.org/10.1111/j.1439-0388.1994.tb00461.x>

- Rosewell, K. L., Li, F., Puttabyatappa, M., Akin, J. W., Brännström, M., & Curry, T. E., Jr (2013). Ovarian expression, localization, and function of tissue inhibitor of metalloproteinase 3 (TIMP3) during the periovulatory period of the human menstrual cycle. *Biology of reproduction*, 89(5), 121. <https://doi.org/10.1095/biolreprod.112.106989>
- Rothschild, M., Jacobson, C., Vaske, D., Tuggle, C., Wang, L., Short, T., Eckardt, G., Sasaki, S., Vincent, A., McLaren, D., Southwood, O., van der Steen, H., Mileham, A., & Plastow, G. (1996). The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1), 201–205. <https://doi.org/10.1073/pnas.93.1.201>
- Saint-Dizier, M., Schoen, J., Chen, S., Banliat, C., & Mermillod, P. (2019). Composing the Early Embryonic Microenvironment: Physiology and Regulation of Oviductal Secretions. *International journal of molecular sciences*, 21(1), 223. <https://doi.org/10.3390/ijms21010223>
- Saura, M., Fernández, A., Varona, L., Fernández, A. I., de Cara, M. Á., Barragán, C., & Villanueva, B. (2015). Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. *Genetics, selection, evolution : GSE*, 47(1), 1. <https://doi.org/10.1186/s12711-014-0081-5>
- Saura, M., Fernández, A., Rodríguez, M. C., Toro, M. A., Barragán, C., Fernández, A. I., & Villanueva, B. (2013). Genome-wide estimates of coancestry and inbreeding in a closed herd of ancient Iberian pigs. *PloS one*, 8(10), e78314. <https://doi.org/10.1371/journal.pone.0078314>
- Schrimpf, R., Metzger, J., Martinsson, G., Sieme, H., & Distl, O. (2015). Implication of FKBP6 for male fertility in horses. *Reproduction in domestic animals*, 50(2), 195–199. <https://doi.org/10.1111/rda.12467>
- Serafini, P. C., Silva, I. D., Smith, G. D., Motta, E. L., Rocha, A. M., & Baracat, E. C. (2009). Endometrial claudin-4 and leukemia inhibitory factor are associated with assisted reproduction outcome. *Reproductive biology and endocrinology : RB&E*, 7, 30. <https://doi.org/10.1186/1477-7827-7-30>
- Serra, X., Gil, F., Pérez-Enciso, M., Oliver, M.A., Vazquez, J.M., Gispert, M., Díaz, I.G., Moreno, F., Latorre, R., & Noguera, J.L. (1998). A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbás line) and Landrace pigs. [https://doi.org/10.1016/S0301-6226\(98\)00151-1](https://doi.org/10.1016/S0301-6226(98)00151-1)
- Silió L, Barragán C, Fernández AI, García-Casco J, Rodríguez MC. Assessing effective population size, coancestry and inbreeding effects on litter size using the pedigree and SNP data in closed lines of the Iberian pig breed. *J Anim Breed Genet*. 2016;133(2):145-154. doi:10.1111/jbg.12168.
- Siller, S. S., Sharma, H., Li, S., Yang, J., Zhang, Y., Holtzman, M. J.,

- Winuthayanon, W., Colognato, H., Holdener, B. C., Li, F. Q., & Takemaru, K. I. (2017). Conditional knockout mice for the distal appendage protein CEP164 reveal its essential roles in airway multiciliated cell differentiation. *PLoS genetics*, 13(12), e1007128. <https://doi.org/10.1371/journal.pgen.1007128>
- Spitschak, M., & Hoeflich, A. (2018). Potential Functions of IGFBP-2 for Ovarian Folliculogenesis and Steroidogenesis. *Frontiers in endocrinology*, 9, 119. <https://doi.org/10.3389/fendo.2018.00119>
- Toro, M., Barragán, C., Óvilo, C., Rodrigañez, J., Rodríguez, C., & Silió, L. (2002). Estimation of coancestry in Iberian pigs using molecular markers. *Conservation Genetics*, 3(3), 309-320. <https://doi.org/10.1023/a:1019921131171>
- Toro, M. A., Silió, L., Rodrigañez, J., & Dobao, M. T. (1988). Inbreeding and family index selection for prolificacy in pigs. *Animal Science*, 46(1), 79-85. <https://doi.org/10.1017/s0003356100003135>
- van der Weyden, L., Arends, M. J., Chausiaux, O. E., Ellis, P. J., Lange, U. C., Surani, M. A., Affara, N., Murakami, Y., Adams, D. J., & Bradley, A. (2006). Loss of TSLC1 causes male infertility due to a defect at the spermatid stage of spermatogenesis. *Molecular and cellular biology*, 26(9), 3595–3609. <https://doi.org/10.1128/MCB.26.9.3595-3609.2006>
- Van Rechem, C., Boulay, G., & Leprince, D. (2009). HIC1 interacts with a specific subunit of SWI/SNF complexes, ARID1A/BAF250A. *Biochemical and biophysical research communications*, 385(4), 586–590. <https://doi.org/10.1016/j.bbrc.2009.05.115>
- Velázquez, M. M., Salvetti, N. R., Amweg, A. N., Díaz, P. U., Matiller, V., & Ortega, H. H. (2013). Changes in the expression of Heat Shock Proteins in ovaries from bovines with cystic ovarian disease induced by ACTH. *Research in veterinary science*, 95(3), 1059–1067. <https://doi.org/10.1016/j.rvsc.2013.07.012>
- Wang, K., Kang, Z., Jiang, E., Yan, H., Zhu, H., Liu, J., Qu, L., Lan, X., & Pan, C. (2020). Genetic effects of DSCAML1 identified in genome-wide association study revealing strong associations with litter size and semen quality in goat (*Capra hircus*). *Theriogenology*, 146, 20–25. <https://doi.org/10.1016/j.theriogenology.2020.01.079>
- Wang, T., Zhang, M., Jiang, Z., & Seli, E. (2017). Mitochondrial dysfunction and ovarian aging. *American journal of reproductive immunology* (New York, N.Y. : 1989), 77(5). <https://doi.org/10.1111/aji.12651>
- Wang, X., Liu, X., Deng, D., Yu, M., & Li, X. (2016). Genetic determinants of pig birth weight variability. *BMC genetics*, 17 Suppl 1(Suppl 1), 15. <https://doi.org/10.1186/s12863-015-0309-6>

- Wright, S. (1949). The Genetical Structure Of Populations. *Annals Of Eugenics*, 15(1), 323-354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Xu, M., Che, L., Yang, Z., Zhang, P., Shi, J., Li, J., Lin, Y., Fang, Z., Che, L., Feng, B., Wu, D., & Xu, S. (2017). Proteomic Analysis of Fetal Ovaries Reveals That Primordial Follicle Formation and Transition Are Differentially Regulated. *BioMed research international*, 2017, 6972030. <https://doi.org/10.1155/2017/6972030>
- Yang, W., Lee, Y. H., Jones, A. E., Woolnough, J. L., Zhou, D., Dai, Q., Wu, Q., Giles, K. E., Townes, T. M., & Wang, H. (2014). The histone H2A deubiquitinase Usp16 regulates embryonic stem cell gene expression and lineage commitment. *Nature communications*, 5, 3818. <https://doi.org/10.1038/ncomms4818>
- Yao, G. D., Shi, S. L., Song, W. Y., Jin, H. X., Peng, Z. F., Yang, H. Y., Wang, E. Y., & Sun, Y. P. (2015). Role of PFAH1B1 in human spermatogenesis, fertilization and early embryonic development. *Reproductive biomedicine online*, 31(5), 613–624. <https://doi.org/10.1016/j.rbmo.2015.07.010>
- Yao, Y. W., Shi, Y., Jia, Z. F., Jiang, Y. H., Gu, Z., Wang, J., Aljofan, M., & Sun, Z. G. (2011). PTOV1 is associated with UCH-L1 and in response to estrogen stimuli during the mouse oocyte development. *Histochemistry and cell biology*, 136(2), 205–215. <https://doi.org/10.1007/s00418-011-0825-z>
- Zeng, S., Ulbrich, S. E., & Bauersachs, S. (2019). Spatial organization of endometrial gene expression at the onset of embryo attachment in pigs. *BMC genomics*, 20(1), 895. <https://doi.org/10.1186/s12864-019-6264-2>
- Zhang, Z., Chen, Z., Ye, S., He, Y., Huang, S., Yuan, X., Chen, Z., Zhang, H., & Li, J. (2019). Genome-Wide Association Study for Reproductive Traits in a Duroc Pig Population. *Animals : an open access journal from MDPI*, 9(10), 732. <https://doi.org/10.3390/ani9100732>
- Zhong, X. , Sun, H. , Zhang, Z. , Zhao, Q. , Olasege, B. , Qiu-meng, L. , YUE, Y., MA, P., Zhang, X., Wang, Q. & Pan, Y. (2020). Genome-wide detection of selective signatures in a Jinhua pig population. *Journal of Integrative Agriculture*. 19. 1314-1322. [https://doi.org/10.1016/S2095-3119\(19\)62833-9](https://doi.org/10.1016/S2095-3119(19)62833-9).