



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

MÁSTER EN MEJORA GENÉTICA ANIMAL Y BIOTECNOLOGÍA DE LA
REPRODUCCIÓN

**ESTUDIO DE ASOCIACIÓN DEL GENOMA COMPLETO PARA EL
CRECIMIENTO Y CARACTERÍSTICAS DE LA CANAL EN LA POBLACIÓN
DE VACUNO DE CARNE RUBIA GALLEGA.**

Proyecto Final de Máster
Carlos Mirokys Then Rodríguez

Universitat Politècnica de València
Junio 2020

Director: Dr. Luis Varona Aguado

ACKNOWLEDGMENT

Me gustaría agradecer a Luis Varona por la oportunidad de realizar este trabajo de Fin de Máster con él y por el tiempo dedicado en este trabajo.

Quería agradecer a mis compañeros Manuel, Oscar y Ramon por el apoyo brindado. A Estefany Espinal por el apoyo incondicional que me ha ofrecido en el transcurso del tiempo y, Por último, a mi familia y a mis amigos por estar conmigo siempre.

ABSTRACT:

Single-step genomic selection analyses were performed using a data set of five traits that include 288,960 records for birth weight (BW), 75,631 for weaning weight (WW), 76,868 for cold carcass weight (CCW), 76,859 for conformation using the SEUROP scale (CON) and 76,723 for fatness (FAT). The pedigree included 417,903 individual-sire-dam entries and 713 of them were genotyped with the Affymetrix Axiom Bovine Chip. After a standard filtering, genotypes for 45,279 SNP markers were available. From the results of the Single-step genomic analyses, we used the equivalence between the GBLUP and SNPBLUP models to achieve the solutions for SNP effects and they were used to calculate the percentage of variance explained by genomic regions from 0.5 to 3 Mb and between 10 to 60 SNP. The results allow us to identify genomic regions of interest in chromosomes 1, 2, 3, 5, 6, 8, 9, 11, 15, 16, 19, 20, 21 and 23. It is noteworthy that some of them have a pleiotropic effect such as those located in BTA1 (14-15mb) -BW, WW, CON and FAT-, BTA1 (131-132mb) -BW, FAT, CCW-, BTA2 (90-92mb) - WW, CCW, CON and FAT-, BTA2 (126-127mb) -BW, WW, CCW, CON and FAT-, BTA3 (32mb) -BW, WW, CCW, CON and FAT-, BTA6 (36 -38mb) - BW, CCW, FAT-, BTA16 (24-25mb) –BW, WW, CCW, CON and FAT-, BTA19 (55-56mb) - BW, WW, CCW, CON and FAT-, BTA20 (20-22mb) - BW, WW, CCW, CON and FAT-, BTA21 (56-57mb) -BW, WW CCW, CON y FAT- and BTA23 (39mb) -BW, WW, CCW, CON and FAT. Among the genes located in These regions are worth highlighting the following: MRPS6 and NCK1 in BTA1, MSTN and SESN2 in BTA2, KNCNA3 in BTA 3, DDIT3 in BTA5, LCORL and NCAPG in BTA6, BPNT1 in BTA16, ACOX1 in BTA 19, PLK2 in BTA20, SLC2A4 in BTA21 and RNF144B in BTA23.

KeyWords: Beef cattle, Single-step GBLUP, SNP, Candidate genes, GWAS

RESUMEN:

Se ha realizado un análisis mediante ssGBLUP –single-step Genomic Selection BLUP- utilizando cinco bases de datos que incluían 288,960 registros de pesos al nacimiento (BW), 75,631 para Peso al destete (WW), 76,868 para peso de la canal fría (CCW), 76,859 para conformación usando la escala SUEROP (CONF) y 76,732 para engrasamiento (FAT) en la raza bovina Rubia Gallega. Además, se utilizó una genealogía que incluyó 417,903 entradas de individuo-padre-madre. Además, se usaron los genotipados de 713 individuos con el Affymetrix Axion Bovine Chip. Después de un filtrado estándar se utilizaron 45,279 SNPs marcadores. A partir de los resultados del ssGBLUP se han obtenido las soluciones de los efectos SNP a través de la equivalencia entre los modelos GBLUP y SNP-BLUP. Estas soluciones se han utilizado posteriormente para calcular el porcentaje de la varianza genética explicada por regiones definidas en función del tamaño del genoma (0.5 a 3 Mb) o el número de marcadores SNP (10 a 60). Los resultados nos han permitido identificar regiones genómicas de interés en los cromosomas 1, 2, 3, 5, 6, 8, 9, 11, 15, 16, 19, 20, 21 y 23. Es destacable que algunas de ellas tienen un efecto pleiotrópico como las localizadas en BTA1 (14-15mb) -BW, WW, CON y FAT-, BTA1(131-132mb) -BW, FAT y CCW- , BTA2 (90-92mb) -WW, CCW, CON y FAT- , BTA2(126-127mb) -BW, WW, CCW, CON y FAT-, BTA3(32mb) -BW, WW, CCW, CON y FAT- , BTA6 (36-38mb) -BW, CCW y FAT-, BTA16(24-25mb) –BW, WW, CCW, CON y FAT- , BTA19 (55-56mb)- BW, WW, CCW, CON y FAT-, BTA20 (20-22mb) – BW, WW, CCW, CON y FAT-, BTA21(56-57mb) -BW, WW, CCW, CON y FAT-, BTA23 (39mb) -BW, WW, CCW, CON y FAT-. Entre los genes localizados en estas regiones merece la pena destacar los siguientes: MRPS6 y NCK1 en BTA1, MSTN y SESN2 en BTA2, KNCNA3 en BTA 3, DDIT3 en BTA5, LCORL y NCAPG en BTA6, BPNT1 en BTA16, ACOX1 en BTA 19, PLK2 en BTA20, SLC2A4 en BTA21 y el RNF144B en BTA23.

Palabras claves: Bovino de Carne, Single-step GBLUP, SNP, Genes candidatos, GWAS

RESUM

S'ha realitzat una anàlisi per mitjà de ssGBLUP -single-step Genomic Selection BLUP- utilitzant cinc bases de dades que incloïen 288,960 registres de pesos al naixement (BW), 75,631 per a Pes al deslletament (WW), 76,868 per a pes de la canal freda (CCW), 76,859 per a conformació usant l'escala SUEROP (CONF) i 76,732 per a greixatge (FAT) en la raça bovina Rossa Gallega. A més, es va utilitzar una genealogia que va incloure 417,903 entrades d'individu-pare-mare. A més, es van usar els genotipatges de 713 individus amb l'Affymetrix Axion Bovine Xip. Després d'un filtrat estàndard es van utilitzar 45,279 SNPs marcadors. A partir dels resultats del ssGBLUP s'han obtingut les solucions dels efectes SNP a través de l'equivalència entre els models GBLUP i SNP-BLUP. Estes solucions s'han utilitzat posteriorment per a calcular el percentatge de la varianza genètica explicada per regions definides en funció de la grandària del genoma (0.5 a 3 Mb) o el nombre de marcadors SNP (10 a 60). Els resultats ens han permès identificar regions genòmiques d'interés en els cromosomes 1, 2, 3, 5, 6, 8, 9, 11, 15, 16, 19, 20, 21 i 23. És destacable que algunes d'elles tenen un efecte pleiotrópic com les localitzades en BTA1 (14-15mb) -BW, WW, AMB i FAT-, BTA1 (131-132mb) -BW, FAT i CCW-, BTA2 (90-92mb) -WW, CCW, AMB i FAT-, BTA2 (126-127mb) -BW, WW, CCW, CON i FAT-, BTA3 (32mb) -BW, WW, CCW, AMB i FAT-, BTA6 (36-38mb) -BW, CCW i FAT-, BTA16 (24-25mb) -BW, WW, CCW, AMB i FAT-, BTA19 (55-56mb) -BW, WW, CCW, AMB i FAT-, BTA20 (20- 22mb) -BW, WW, CCW, AMB i FAT-, BTA21 (56-57mb) -BW, WW, CCW, AMB i FAT-, BTA23 (39mb) -BW, WW, CCW, AMB i FAT-. Entre els gens localitzats en estes regions val la pena destacar els següents: MRPS6 i NCK1 en BTA1, MSTN i SESN2 en BTA2, KNCNA3 en BTA 3, DDIT3 en BTA5, LCORL i NCAPG en BTA6, BPNT1 en BTA16, ACOX1 en BTA 19, PLK2 en BTA20, SLC2A4 en BTA21 i el RNF144B en BTA23.

Paraules claus: Boví de Carn, Single-step GBLUP, SNP, Gens candidats, GWAS

ÍNDICE GENERAL

ACKNOWLEDGMENT	ii
ABSTRACT:	iii
RESUMEN:	iv
RESUM	v
.....	v
ÍNDICE GENERAL	vi
LISTAS DE CUADROS	vii
LISTA DE FIGURAS	viii
1. INTRODUCCIÓN.....	1
1.1. La Rubia Gallega	1
1.2. Valoración Genética de Reproductores.....	5
1.3. Asociación de genoma completo (GWAS).....	9
2. OBJETIVOS	12
3. MATERIAL Y MÉTODOS	13
3.1. Material fenotípico y genotípico.....	13
3.2. Control de calidad y filtrado de muestras y SNPs.....	15
3.3. Análisis de asociación del genoma completo (GWAS).....	15
4. RESULTADOS	19
4.1. Estimación de Componentes de Varianza y Covarianza.....	19
4.2. Peso al nacimiento (BW).....	20
4.3. Peso al destete: Efecto directo (WWD).....	21
4.4. Peso al destete: Efecto Materno (WWM).....	23
4.5. Peso de la Canal Fría (CCW).....	25
4.6. Conformación (CON).....	27
4.7. Engrasamiento (FAT).....	29
5. DISCUSION.....	32
6. CONCLUSIONES	43
7. REFERENCIAS	44
8. APÉNDICE I.	56
9. APÉNDICE II.	62

LISTAS DE CUADROS

Tabla 1.1. Datos productivos medios de la población Rubia Gallega.....	2
Tabla 3.1 Descripción de la codificación de los datos de Conformación.....	13
Tabla 3.2. Descripción de la codificación de los datos de Engrasamiento.	14
Tabla 3.3. Media y Desviación Típica de los caracteres analizados.	14
Tabla 3.4 Efectos sistemáticos.....	16
Tabla 4.1 Estimadores (y error estándar) de máxima verosimilitud restringida de los componentes de varianza y covarianza.	19
Tabla 4.2 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Peso al Nacimiento (BW).....	21
Tabla 4.3 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el efecto directo del Peso a los 210 días (WWD).	23
Tabla 4.4 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el efecto materno del Peso a los 210 días (WWM).....	25
Tabla 4.5 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el Peso de la Canal Fría (CCW).	27
Tabla 4.6 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Conformación (CON).....	29
Tabla 4.7 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Engrasamiento (FAT).	30
Tabla 5.1. Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y para Peso al Nacimiento.....	33
Tabla 5.2. Regiones del genoma que mostraron un resultando estandarizado por encima del percentil 95% en al menos cuatro caracteres.....	35

LISTA DE FIGURAS

Figura 1.1. Escala de valoración SEUROP del carácter CONFORMACIÓN.....	3
Figura 1.2. Escala de valoración SEUROP del carácter ENGRASAMIENTO DE LA CANAL.	3
Figura 4.1 Barrido Genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Peso al Nacimiento (BW).	20
Figura 4.2. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Peso al Nacimiento (BW)...	20
Figura 4.3. Barrido Genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el efecto directo del Peso a los 210 días (WWD).....	22
Figura 4.4. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el efecto directo del Peso a los 210 días (WWD).	22
Figura 4.5. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el efecto materno del Peso a los 210 días (WWM).	24
Figura 4.6. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el efecto materno del Peso a los 210 días (WWM).	24
Figura 4.7. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el Peso de la Canal Fría (CCW).	26
Figura 4.8. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el Peso de la Canal Fría (CCW).	26
Figura 4.9. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Conformación (CON).....	28

Figura 4.10 Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Conformación (CON).	28
Figura 4.11. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Engrasamiento (FAT)..	29
Figura 4.12. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Engrasamiento (FAT).	30
Figura 5.1. Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso al Nacimiento.	33
Figura 5.2 Correlaciones entre los resultados estandarizados del porcentaje de la varianza explicada para cada marcador SNP y entre Peso al Nacimiento (BW), efecto directo en el Peso a los 210 días (WWD), efecto materno en el Peso a los 210 días (WWM), Peso de la Canal Fría (CCW), Conformación (CON) y Engrasamiento (FAT).	34

1. INTRODUCCIÓN

1.1. La Rubia Gallega

La población bovina Rubia Gallega representa el biotipo racial más característico del ganado bovino de carne de la región noroeste de España y forma parte del paisaje de la Comunidad Autónoma de Galicia (FEAGAS, 2018). La estructura de producción de la raza Rubia Gallega, se basa, fundamentalmente, en pequeños rebaños familiares en sistema de pastoreo que comercializan terneros de 7 a 10 meses de edad. Los terneros de la raza Rubia Gallega proporcionan canales ligeras y suficientemente engrasadas (Carballo et al., 2005), que son muy apreciadas en el mercado regional y nacional, y constituyen una parte sustancial de la base genética de la Indicación Geográfica Protegida “Ternera Gallega” desde 1989.

El libro genealógico y el plan de mejora genética de la población se encuentran gestionados por la Asociación de Criadores de Raza Rubia Gallega (ACRUGA). Según esta asociación, en el año 2018 existían unos 1.900 criadores, con una mayor concentración en la Comunidad Autónoma de Galicia. El censo de reproductores constaba de 27,698 hembras y 474 machos (<https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/razasganaderas/razas/catalogo/autoctona-fomento/bovino/rubia-gallega>), que la sitúa a como la de mayor censo de entre las poblaciones autóctonas de ganado vacuno de carne.

El actual programa de mejora genética de la población (https://acruga.com/archivos/docs/PROGRAMA_MEJORA.pdf) se inició en 1990. Los datos productivos medio de la raza, según indica el propio plan de mejora se presentan en la Tabla 1.1.

Actualmente, los objetivos de mejora de la población son:

- Mantener la vocación productiva de la raza y su capacidad adaptativa a las condiciones de producción y manejo.
- Incrementar/homogeneizar la producción cárnica de la población
- Evitar cambios en la precocidad sexual.
- Mejorar la esfera reproductiva y reducir la dificultad al parto
- Mantener/mejorar la funcionalidad y armonía morfológica de los animales.

Tabla 1.1. Datos productivos medios de la población Rubia Gallega.

	Machos	Hembras
Peso Nacimiento	43 kg.	40 kg.
Peso al Destete (210 días)	300 kg.	270 kg.
Velocidad de Crecimiento en Lactación	1,200 kg. /d	1,100 kg. /d
Edad al sacrificio	8-10 meses	
Peso vivo al sacrificio	300-400 Kg	
Edad al primer parto	27 meses	
Intervalo entre partos	409 días	
Intervalo parto – 1ª cubrición	78 días	
Intervalo parto – concepción	118 días	
Duración de la gestación	291 días	
Facilidad de parto (% de hembras que paren solas)	85%	

El control de rendimientos implementado por ACRUGA consiste en la obtención de registros en el animal vivo y en el matadero. En el animal vivo se registran los siguientes caracteres:

- Peso al nacimiento y dificultad al parto.
- Peso al destete a los 210 días.
- Consumo de alimento (en estación de testaje).
- Caracteres morfológicos funcionales.
- Caracteres reproductivos.

En el matadero se obtienen la siguiente información:

- Peso de la Canal.
- Conformación.
- Engrasamiento de la canal.

Estos dos últimos caracteres se registrar según la escala de valoración SEUROP que se muestra en las figuras 1.1 y 1.2

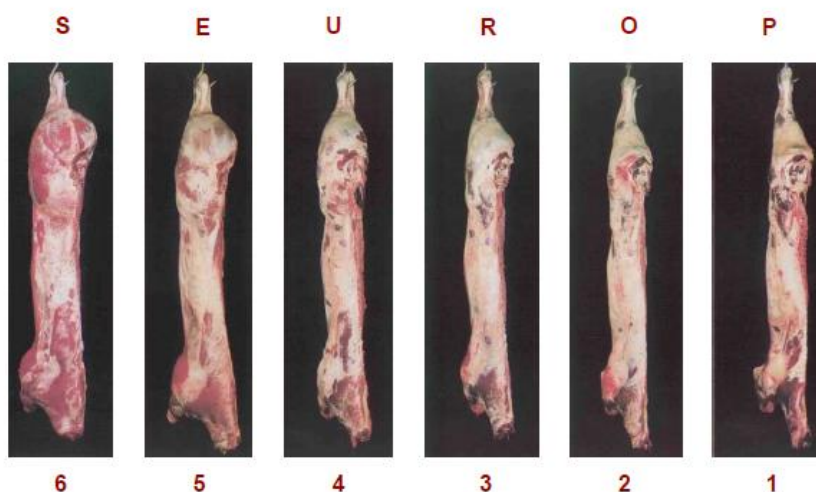


Figura 1.1. Escala de valoración SEUROP del carácter CONFORMACIÓN.

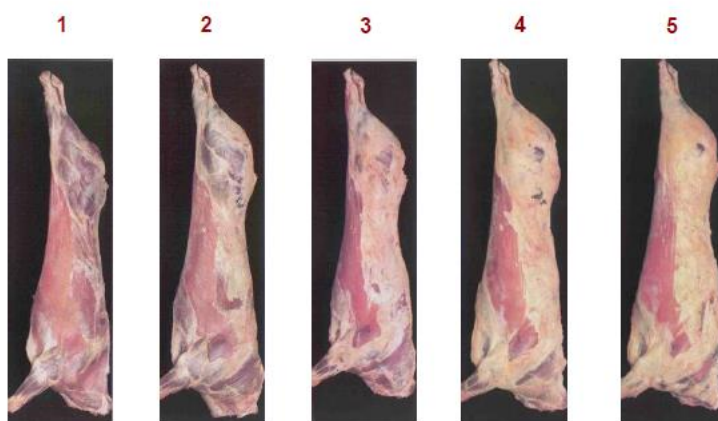


Figura 1.2. Escala de valoración SEUROP del carácter ENGRASAMIENTO DE LA CANAL.

La valoración genética de los caracteres se realiza mediante el mejor predictor lineal insesgado –BLUP- (Henderson, 1984) utilizando los siguientes modelos:

- ✓ $BW_{ijklm} = \mu + S_i + E_{Mj} + REA_k + EBV_{BWl} + e_{ijklm}$
- ✓ $WW_{ijklmn} = \mu + E_i + S_j + E_{Mk} + REA_l + EBV_{WWm} + EBV_{EMN} + e_{ijklmn}$
- ✓ $CCW_{ijkl} = \mu + E_i + S_j + REA_k + Mat_l + EBV_{CCWm} + e_{ijklm}$
- ✓ $CON_{ijkl} = \mu + E_i + S_j + REA_k + Mat_l + EBV_{CONm} + e_{ijklm}$
- ✓ $FAT_{ijkl} = \mu + E_i + S_j + REA_k + Mat_l + EBV_{FATm} + e_{ijklm}$

Donde BW el peso al nacimiento (Birth Weigth), WW es el peso al destete a los 210 días (Weaning Weigth), CCW es el peso de la canal fría (Cold Carcass Weight), CON es la medida de la conformación (Conformation) y FAT es la medida de engrasamiento (Fatness). Además, los EBV son los valores genéticos aditivos, S es el sexo, E es la edad, E_M es la edad de la madre, REA es el efecto asociado al rebaño-año-estación y MAT es el matadero. Finalmente, e son los residuos.

La implementación de este esquema de selección se basa en la valoración genética de los animales, que se obtiene mediante índice compuestos de selección (ICO). Estos índices se calculan a partir de los resultados de la valoración genética mediante BLUP para caracteres registrados en vivo (peso al nacimiento (BW) y peso al destete -210 días (WW) aproximadamente-) y en matadero (peso de la canal fría, y valoraciones SEUROP para conformación, engrasamiento y color). En concreto se utilizan dos índices compuestos de selección (ICO vida -IVC- e ICO carne -ICC-). El ICO Vida se calcula a partir de las valoraciones genéticas para peso al nacimiento (BW), el efecto materno para peso del destete (WWM), la conformación (CON) y el engrasamiento (FAT).

$$IVC = 0,2 EBV_{BW} + 0,2 EBV_{WWM} + 0,2 EBV_{CON} + 0,4 EBV_{FAT}$$

Y es posteriormente estandarizado con media y varianza 100. El cálculo del ICO Carne (Cruzamiento) utiliza el peso al nacimiento (con valoración negativa), el peso de la canal (CCW), el peso al destete (WW) y la conformación.

$$ICC = 0,2 EBV_{BW} + 0,2 EBV_{WW} + 0,2 EBV_{CCW} + 0,4 EBV_{CON}$$

También estandarizado con media 100 y varianza 100

1.2. Valoración Genética de Reproductores.

La valoración genética de reproductores es un elemento fundamental para el desarrollo de los esquemas de mejora genética de las poblaciones ganaderas. La mayoría de los caracteres de interés en ganadería son caracteres cuantitativos (Falconer y MacKay, 1996), es decir, muestran una variación continua. La valoración genética de los caracteres cuantitativos se asume que su variabilidad genética depende de un número quasi-infinito de genes independientes y con efecto muy pequeño. Este modelo teórico se denomina modelo infinitesimal y fue desarrollado por Fisher (1918), y ha supuesto la base metodológica para el desarrollo de procedimientos como el índice de selección (Hazel y Lush, 1942) y el Mejor Predictor Linear Insesgado o BLUP -Best Linear Unbiased Predictor- (Henderson, 1984). Desde los años 80 del siglo pasado, el BLUP se ha establecido como la herramienta habitual para la valoración genética de reproductores, y utiliza como fuentes de información tanto los fenotipos como el parentesco entre los individuos.

Sin embargo, el desarrollo de la genética molecular ha permitido, a partir de la década de los 90, la posibilidad de utilizar esta información para la valoración genética de reproductores. Los primeros intentos de utilizar la información proporcionada por la genética molecular se basaban en el concepto de Selección Asistida por Marcadores (Marker Assisted Selection –MAS-, Dekkers, 2004), que se introdujo por primera vez a principios del siglo XX (Sax, 1923). La MAS proponía el uso de la información molecular en la valoración genética como un proceso en varios pasos. En primer lugar, se requería la detección de uno o varios QTL (*Quantitative Trait Loci*), que se definen como regiones del genoma asociadas con la variabilidad genética de un carácter cuantitativo. Posteriormente, se requería identificar la mutación causal en el gen responsable de la variabilidad o de un polimorfismo en desequilibrio de ligamiento con este y, finalmente, se necesita el genotipado del polimorfismo seleccionado que se incluye como un efecto sistemático en los modelos de valoración genética mediante BLUP.

Sin embargo, la variabilidad genética de los caracteres cuantitativos está regulada por numerosos genes o mutaciones, que han dado lugar a la

detección de números QTL (www.animalgenome.org/QTLdatabase), pero la gran mayoría de ellos únicamente explican una pequeña proporción de su varianza genética (VanRaden et al., 2009; Riedelsheimer et al, 2012) y, por regla general, no se ha conseguido identificar el polimorfismo responsable de la variabilidad genética, ya que la localización y estima de los efectos asociados a polimorfismos que expliquen un porcentaje muy pequeño de la variabilidad genética exigiría un número de datos muy grande (Falconer y Mackay, 1996).

A principios de este siglo, Meuwissen et al. (2001) propusieron una alternativa a la MAS denominado Selección Genómica, que no tiene la intención de localizar QTL asociados a los caracteres de interés, sino que se pretende utilizar todos los marcadores disponibles para la predicción de los valores mejorantes de los candidatos a la selección. La clave del procedimiento se basa en la disponibilidad de un número suficientemente grande de marcadores para poder cubrir la totalidad del genoma con la suficiente densidad para capturar la inmensa mayoría de la varianza genética aditiva de los caracteres, gracias al desequilibrio de ligamiento con las mutaciones causales. En este sentido, la disponibilidad de dispositivos de genotipado masivo de SNPs (*Single Nucleotide Polymorphisms*) tras la secuenciación de los genomas de las especies ganaderas (The Bovine Genome Sequencing and Analysis Consortium, 2009; Groenen et al., 2012) ha permitido la implementación práctica de la selección genómica en la mayoría de las poblaciones ganaderas (Meuwissen et al., 2016).

En el planteamiento original de la selección genómica por Meuwissen et al. (2001), su implementación exigía la definición de una población de referencia suficientemente grande, cuyos miembros deberían ser fenotipados para los caracteres de interés y genotipados para un número suficientemente grande de marcadores moleculares, como los SNP. La información procedente del genotipado y fenotipado de la población de referencia se utiliza para el cálculo de una función predictiva que se compone de los efectos aditivos asociados a todos los marcadores SNP. Esta función predictiva se puede utilizar posteriormente para predecir los valores mejorantes de los candidatos a la selección o GEBV (*Genomic Breeding Values*), incluso antes de producir un fenotipo propio. Por este motivo, el atractivo de este planteamiento de la selección genómica es evidente para aquellos caracteres donde la evaluación

genética mediante BLUP proporciona poca precisión debido a la dificultad de obtener información fenotípica en los candidatos a la selección (calidad de la carne, calidad de la canal, longevidad), en caracteres con una heredabilidad baja, como los caracteres reproductivos, o en caracteres que solo se expresan en un sexo como la producción de leche o los caracteres reproductivos asociados a la madre (Garrick, 2010). Además, esta valoración genética puede obtenerse cuando los candidatos a la selección son muy jóvenes, y, por lo tanto, permite reducir el intervalo generacional de una manera muy importante. En concreto, esta particularidad ha permitido la implantación de la selección genómica de manera masiva en el ganado vacuno de leche (Hayes et al., 2009).

Sin embargo, la aplicación práctica de este planteamiento presenta dificultades de ámbito estadístico, ya que implica la estimación de un número de efectos asociados a los marcadores moleculares SNP que, habitualmente, supera a la cantidad de datos fenotípicos disponibles. Este fenómeno corresponde con un problema estadístico de “large p, small n” y ha sido resuelto mediante el desarrollo de un gran número de procedimientos de regularización o de asunciones a priori de la distribución de probabilidad de los efectos asociados a los marcadores SNP. Así, se han descrito métodos que asumen que los efectos asociados a los marcadores siguen una distribución normal (Random Regression BLUP, Meuwissen et al., 2001), una t de Student (Bayes A, Meuwissen et al., 2001), una mezcla de una distribución t y una masa de probabilidad (Bayes B, Meuwissen et al., 2001 o Bayes Cπ, Habier et al., 2011) o una distribución doble exponencial (Bayesian Lasso, De los Campos et al., 2009), entre otras. En la literatura científica existe controversia en cuanto al rendimiento de todos estos procedimientos, aunque existe consenso en que los procedimientos basados en la distribución normal ofrecen una precisión similar a métodos más complejos cuando la variabilidad del carácter está gobernada por un número gran de loci que se comportan de manera aditiva (De los Campos et al., 2013).

Además, una ventaja adicional de la utilización de un procedimiento de regularización Gaussiana (Random Regression BLUP) es que es equivalente a un modelo definido en términos de los valores genéticos aditivos de los individuos. El resultado de esta parametrización da lugar a un modelo de evaluación equivalente al BLUP, que se denomina GBLUP o BLUP genómico

(Habier et al., 2007). La principal diferencia entre el GBLUP y el BLUP calculado a partir de la genealogía consiste en que la matriz de parentesco numerador (**A**), que se requiere para la construcción de la matriz de ecuaciones de modelo mixto se sustituye por una matriz de parentesco genómico (**G**), que fue desarrollada por VanRaden (2008). Este planteamiento de la selección genómica no busca calcular los efectos aditivos asociados a los marcadores SNP, sino que utiliza la información molecular proporcionada por los chips de genotipado para refinar el parentesco de los individuos. De esta manera, el parentesco genealógico que se refleja en la matriz **A** se puede entender como el parentesco esperado entre los individuos, mientras que el parentesco genómico corresponde a una estimación del parentesco realizado. Además, una ventaja adicional de este planteamiento de la selección genómica es que se han desarrollado procedimientos estadísticos que permiten integrar de manera conjunta el parentesco genómico y el parentesco genealógico para aquellos individuos de los que no se dispone de información molecular mediante la definición de una nueva matriz de relaciones **H** (Legarra et al., 2009; Christensen y Lund, 2010) en un procedimiento denominado como "single-step genomic selection" o single-step GBLUP -ssGBLUP- (Misztal et al., 2009; Aguilar et al., 2010). Este procedimiento se ha convertido en la estrategia de selección en poblaciones con un censo reducido y con una capacidad de inversión en genotipados limitada, como la Rubia Gallega, ya que permite integrar toda la información fenotípica y genealogía registrada históricamente por el programa de mejora.

La implementación de la selección genómica mediante GBLUP o ssGBLUP asigna, a priori, la misma probabilidad de contribuir a la variabilidad genética del carácter a todos los marcadores SNP o regiones del genoma. Sin embargo, y pese a que la mayoría de los caracteres de interés en las poblaciones ganaderas, la variabilidad genética no se distribuye de manera homogénea a lo largo de todo el genoma (Hayes et al., 2010). En este sentido, se han desarrollado varios procedimientos que permiten asignar pesos diferentes a cada región del genoma como el denominado weighted-GBLUP (Tiezzi y Maltecca, 2015; Tiezzi et al., 2018). Este procedimiento utiliza una versión modificada de la matriz de relaciones genómicas que pondera la información proporcionada por cada SNP

a partir de un análisis de asociación de genoma completo (GWAS – Genome Wide Association).

1.3. Asociación de genoma completo (GWAS).

Desde los inicios del desarrollo de la genética se han realizado esfuerzos para localizar los genes o las regiones del genoma (QTL, *Quantitative Trait Loci*) responsables de la variabilidad de los caracteres cuantitativos (Sax, 1923). Antiguamente, únicamente se disponía de la información procedente de polimorfismos observables directamente, o bien a través de pruebas bioquímicas. Más adelante, en la década de los 80, se generalizó el uso de un grupo de marcadores muy polimórficos denominados micro y mini satélites que, junto al desarrollo de procedimientos estadísticos (Lander y Botstein, 1989; Haley y Knott, 1992) y a la ejecución de numerosos protocolos experimentales en todas las especies ganaderas, dieron lugar a la detección de una gran cantidad de regiones asociadas a caracteres de interés (www.animalgenome.org/QTLdatabase). Sin embargo, la precisión que proporcionaban los marcadores microsatélites era pequeña ya que, habitualmente, se localizaban muy espaciados entre sí.

A partir de la década de los 2000, se desarrollaron los proyectos de secuenciación de los genomas de las especies domésticas (The Bovine Genome Sequencing and Analysis Consortium, 2009; Groenen et al., 2012) que facilitaron el desarrollo de dispositivos de genotipado masivo de marcadores SNP (*Single Nucleotide Polymorphism*). La disponibilidad de estos dispositivos ha permitido el desarrollo de la selección genómica, descrita en el capítulo anterior, y de los procedimientos de asociación de genoma completo o GWAS -*Genome Wide Association*- (Georges, 2007). El planteamiento inicial de los procedimientos de asociación genoma completo consistió en la implementación de un estudio de asociación estadística para cada uno de los marcadores SNP disponibles en una población, que proporcionan un valor p o de probabilidad de obtener un estadístico de mayor magnitud como consecuencia del azar bajo la hipótesis nula de no asociación entre el marcador SNP y el carácter cuantitativo. De esta manera, una vez realizado el análisis se dispone de numerosos valores p que

pueden representarse de forma logarítmica a lo largo del genoma en un tipo de gráfico denominado *Manhattan plot*.

La identificación de regiones del genoma asociada con la variabilidad del carácter cuantitativo se realiza mediante la localización de los marcadores cuyo p-value sea inferior a un determinado umbral que se calcula teniendo en cuenta que se realizan numerosos tests. Los procedimientos más habituales para la corrección por testaje múltiple son la corrección de Bonferroni (Bonferroni, 1936), o la tasa de falso descubrimiento o False Discovery Rate – FDR- (Benjamini y Hochberg, 1995).

En mejora genética animal, los modelos para la realización de los análisis de asociación de genoma completo (GWAS) consideran habitualmente la presencia de efectos ambientales sistemáticos (sexo, edad, etc) y aleatorios (Rebaño-Año-Estación, Camada) y efectos genéticos poligénicos que se definen como distribuidos de manera normal multivariante con una matriz de relaciones construida a partir de la genealogía (**A**) o de los marcadores SNP (**G**). De esta manera, los efectos asociados a los marcadores SNP son identificados corriendo otros posibles efectos que pueden enmascarar los resultados. Además, el modelo de selección genómica mediante GBLUP o ssGBLUP se puede parametrizar en términos de los marcadores SNP. Así, Wang et al. (2012) ha desarrollado un procedimiento que permite obtener los valores p asociados a cada uno de los marcadores mediante una transformación de las predicciones de los valores genéticos aditivos.

Sin embargo, la utilización de los valores p es controvertida, ya que los valores p obtenidos por SNP contiguos, o muy cercanos entre sí, suelen tener una magnitud similar, como consecuencia del desequilibrio de ligamiento entre los marcadores SNP. Además, las regiones con un grado de desequilibrio de ligamiento alto entre los marcadores diluyen el grado de asociación con los caracteres analizados y presentan valores p más pequeños, incluso aunque estén implicadas en un porcentaje relativamente grande de la variabilidad genética aditiva del carácter. Para solventar este problema, algunos autores (Fernando et al., 2017; Legarra y Varona, 2018) han propuesto utilizar el

porcentaje de la variabilidad genética capturada por un grupo de marcadores consecutivos en lugar de los valores p .

2. OBJETIVOS

El objetivo principal de este trabajo consiste en la localización de las regiones genómicas asociadas con la variabilidad genética en los caracteres utilizados en los índices de selección de la población Rubia Gallega: Peso al Nacimiento, Peso al Destete, Peso de la Canal Fría, Conformación y Engrasamiento.

De este objetivo principal, se desprende dos subobjetivos:

1. La identificación de posibles genes candidatos que puedan contener las mutaciones causales que expliquen un porcentaje importante de la variabilidad genética de los caracteres de interés en la población Rubia Gallega.
2. El estudio del grado de similitud entre la arquitectura genética de los caracteres de interés en función de la caracterización de regiones pleiotrópicas entre ellos.

3. MATERIAL Y MÉTODOS

3.1. Material fenotípico y genotípico.

Los datos utilizados para este estudio fueron datos genealógicos, fenotípicos y de genotipado proporcionados por ACRUGA (Asociación Nacional de Criadores de Ganado Vacuno Selecto de Raza Rubia Gallega). Los ficheros de datos incluyeron datos fenotípicos de 288,960 registros para peso al nacimiento (BW), 75,631 de peso a los 210 días (WW), 76,868 pesos de la canal fría (CCW), 76,859 datos de Conformación utilizando la escala SEUROP (CON), 76,723 de la valoración de engrasamiento en el matadero (FAT).

La clasificación de las canales en matadero se regula según el Real Decreto 225/08. Para conformación, la valoración incluye las categorías S (superior), E (excelente), U (muy buena), R (buena), O (menos buena) y P (mediocre). Cada una de las categorías se divide en tres subcategorías (Superior +, Central = e Inferior -). La codificación de los fenotipos se describe en la Tabla 3.1.

Tabla 3.1 Descripción de la codificación de los datos de Conformación.

CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	FENOTIPO
S (superior)	Todos los perfiles extremadamente convexos: desarrollo muscular excepcional con dobles músculos (tipo: "culón")	+	6.000
		=	5.666
		-	5.333
E (excelente)	Todos los perfiles de convexos a superconvexos: desarrollo muscular excepcional.	+	5.000
		=	4.666
		-	4.333
U (muy buena)	Perfiles convexos en conjunto: fuerte desarrollo muscular	+	4.000
		=	3.666
		-	3.333
R (buena)	Perfiles rectilíneos a cóncavos: desarrollo muscular medio	+	3.000
		=	2.666
		-	2.333
O (menos buena)	Perfiles rectilíneos a cóncavos: desarrollo muscular medio	+	2.000
		=	1.666
		-	1.333
P (mediocre)	Todos los perfiles de cóncavos a muy cóncavos: escaso desarrollo muscular	+	1.000
		=	0.666
		-	0.333

Para el grado de engrasamiento, las clases descritas en el Real Decreto 225/08 son: 1 (no graso), 2 (poco cubierto), 3 (cubierto), 4 (graso) y 5 (muy graso). Como en el caso anterior, se definen tres subcategorías (Superior +, Central = e Inferior -). La codificación de los fenotipos se describe en la Tabla 3.2.

Tabla 3.2. Descripción de la codificación de los datos de Engrasamiento.

CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	FENOTIPO
5 (muy graso)	Toda la canal cubierta de grasa, acúmulos importantes de grasa en el interior de la cavidad torácica.	+	5.000
		=	4.666
		-	4.333
4 (graso)	Músculos cubiertos de grasa, pero aun parcialmente visibles a nivel de cadera y de la paletilla, algunos acúmulos pronunciados de grasa en el interior de la cavidad torácica.	+	4.000
		=	3.666
		-	3.333
3 (cubierto)	Músculos, excepto cadera y paletilla, casi siempre cubiertos, escasos acúmulos de grasa en el interior de la cavidad torácica.	+	3.000
		=	2.666
		-	2.333
2 (poco cubierto)	Ligera cobertura de grasa, músculos casi siempre aparentes.	+	2.000
		=	1.666
		-	1.333
1 (no graso)	Cobertura de grasa inexistente o muy débil	+	1.000
		=	0.666
		-	0.333

Las medias y desviaciones típicas de los datos disponibles se presentan en la Tabla 3.3.

Tabla 3.3. Media y Desviación Típica de los caracteres analizados.

	Media	Desv. Típica.
Peso al Nacimiento	42.72	7.46
Peso a los 210 días	282.12	46.98
Peso de la Canal Fría	222.85	40.44
Engrasamiento	2.27	0.45
Conformación	3.91	0.66

La genealogía incluyó 417,903 entradas con la información del individuo, su padre y su madre. Adicionalmente, se dispuso del genotipado de 723 individuos con el Axiom ® Bovine Genotyping Array. Los individuos genotipados fueron seleccionados en función de la precisión de su valoración genética en la evaluación de reproductores del año 2018 dentro del marco del plan de mejora de la población Rubia Gallega.

3.2. Control de calidad y filtrado de muestras y SNPs.

Los controles de calidad sobre las muestras genotipadas se realizaron usando la plataforma Axiom Analysis Suite v5.0.1. (ThermoFisher Scientific, 2020). Los criterios utilizados para excluir las muestras con baja calidad de genotipado fueron: señal de contraste en los locus no polimórficos (DQC)<0.89, eficiencia del genotipado (genotipos faltantes por individuo o QC) <96%, promedio de la eficiencia del genotipado por placa <98.5% y porcentajes de muestras que superan DQC y QC <95%.

El resto de los controles de calidad sobre los SNPs se realizaron con el software PLINK v1.9 (Chang et al., 2015). En primer lugar, se fijó que el número de genotipos faltantes por individuos debía ser menor del 95%. Adicionalmente, se excluyeron los marcadores SNP que no cumplieran los siguientes criterios: genotipos faltantes por SNP menor del 5% y frecuencia del alelo minoritario (MAF) mayor del 0.01. Además, se seleccionaron exclusivamente los marcadores localizados en los cromosomas autosómicos. El resultado final se seleccionaron 45,279 marcadores SNPs de 713 individuos.

3.3. Análisis de asociación del genoma completo (GWAS).

En primer lugar, se realizó una estimación de componentes de la varianza mediante máxima verosimilitud restringida -REML- (Patterson y Thompson, 1971) utilizando el paquete airemlf90 (Miszta et al., 2002). Para los caracteres BW, CCW, CON y FAT, el modelo asumido fue:

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

donde **y** es el vector de observaciones fenotípicas, **b** es el vector de efectos sistemáticos, **u** es el vector de efectos genéticos aditivos, **p** es el vector efectos aleatorios asociados al rebaño-año-estación y **e** es el vector de residuos. **X**, **Z** y

W son las matrices de incidencia que relacionan los fenotipos con los efectos sistemáticos y genéticos aditivos, respectivamente.

Para el peso a los 210 días (WW), el modelo incluyó también efectos maternos, tanto genéticos como ambientales. Por lo tanto, el modelo de análisis fue:

$$y = Xb + Zu + Tm + Kp + Wp + e,$$

donde **m** son los efectos genéticos aditivos maternos y **p** son los efectos permanentes maternos. **T** y **K** son las correspondientes matrices de incidencia.

Los efectos sistemáticos incluidos en los modelos de evaluación se describen en la siguiente tabla.

Tabla 3.4 Efectos sistemáticos.

	Edad	Edad de la madre	Sexo	Matadero
Peso al Nacimiento		X	X	
Peso a los 210 días	X	X	X	
Peso de la Canal Fría	X	X	X	X
Engrasamiento	X	X	X	X
Conformación	X	X	X	X

En BW, CCW, FAT y CON, las varianzas de los efectos aleatorios son:

$$\text{var} \begin{bmatrix} u \\ p \\ e \end{bmatrix} = \begin{bmatrix} H\sigma_a^2 & 0 & 0 \\ 0 & I\sigma_p^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

y en WW fue:

$$\text{var} \begin{bmatrix} u \\ m \\ d \\ p \\ e \end{bmatrix} = \begin{bmatrix} H\sigma_a^2 & H\sigma_{am} & 0 & 0 & 0 \\ H\sigma_{am} & H\sigma_m^2 & 0 & 0 & 0 \\ 0 & 0 & I\sigma_d^2 & 0 & 0 \\ 0 & 0 & 0 & I\sigma_p^2 & 0 \\ 0 & 0 & 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

Donde σ_a^2 es la varianza genética aditiva, σ_m^2 es la varianza genética materna, σ_{am} es la covarianza entre efectos genéticos directos y maternos, σ_d^2 es la varianza ambiental permanente materna, σ_p^2 es la varianza permanente asociada al rebaño-año-estación y σ_e^2 es la varianza residual. **H** es la matriz que combina la de parentesco numerador con la matriz genómica según fue descrita por Aguilar et al (2010) e **I** es la matriz de identidad. La inversa de la matriz **H** es

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

donde **A** es la matriz de parentesco numerador para todos los individuos, **A₂₂** es la matriz de parentesco numerador para animales genotipados, y **G** es la matriz de parentesco genómico que se calculó según lo descrito por VanRaden (2008) como:

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{\sum_{i=1}^N 2p_i(1-p_i)}$$

donde **Z** es una matriz de contenido de genes ajustada para las frecuencias alélicas de cada SNP (p_i) y N es el número de SNP (VanRaden, 2008).

El análisis de asociación de genoma completo se ha realizado mediante el programa postGSf90 (Aguilar et al.,2010). En primer lugar, el procedimiento obtiene las estimaciones de los efectos de los SNP (**g**) (Wang et al., 2012) a partir del siguiente procedimiento.

$$\hat{\mathbf{g}} = \mathbf{cov}(\mathbf{g}, \mathbf{u}')[\mathbf{var}(\mathbf{u})]\hat{\mathbf{u}}$$

$$\hat{\mathbf{g}} = \frac{\mathbf{Z}'\mathbf{G}^{-1}\hat{\mathbf{u}}}{\sum_{i=1}^N 2p_i(1-p_i)}$$

$$\hat{\mathbf{g}} = \mathbf{Z}'(\mathbf{Z}'\mathbf{Z})^{-1}\hat{\mathbf{u}}$$

Los efectos de los SNP pueden proporcionar una estima con mucho ruido estadístico como consecuencia del desequilibrio de ligamiento entre marcadores. Por este motivo, se ha utilizado la opción *windows_variance* del programa postGSf90 para calcular el porcentaje de la varianza genética aditiva asociada a una región del genoma determinada por un número de marcadores SNP o una distancia de mapa en pares de bases. La determinación de la

amplitud del segmento no es un proceso sencillo, por este motivo se han utilizado hasta 12 definiciones de segmentos del genoma. 6 de ellas se basan en el número de marcadores SNP (10, 20, 30, 40, 50 y 60 marcadores SNP) y otras 6 en función de la cantidad de pares de bases (0.5, 1, 1.5, 2, 2.5 y 3 Megabases). El porcentaje de varianza explicada se calculó en ventanas deslizantes centrada en cada uno de los SNP incluidos en el análisis. Finalmente, los resultados de los 12 procedimientos se estandarizaron con media 0 y desviación típica 1 y fueron promediados en un resultado final que se utilizó para identificar las regiones del genoma asociadas con un mayor porcentaje de varianza genética.

Por último, se seleccionaron las regiones del genoma que explicaban un mayor porcentaje de la varianza genética aditiva y se utilizó la herramienta BiomartTool (www.emsembl.org) que contiene la información de la última versión del genoma bovino (*ARS-UCD1.2*) para identificar los genes presentes en las regiones del genoma seleccionadas.

4. RESULTADOS

4.1. Estimación de Componentes de Varianza y Covarianza.

En primer lugar, se realizó la estimación de componentes de varianza y covarianza con el programa airemlf90 (Misztal et al.,2002) mediante el procedimiento single-step GBLUP. En el análisis, el valor medio de la diagonal de la matriz genómica fue de 1.025, con un rango comprendido entre 0.906 y 1.352. La correlación entre los elementos de la diagonal de la matriz G con los elementos de la matriz A fue de 0.349, mientras que la correlación entre los elementos no diagonales fue de 0.828. Los resultados de la estimación de componentes de varianza y covarianza mediante máxima verosimilitud restringida se presentan en la tabla 4.1.

Tabla 4.1 Estimadores (y error estándar) de máxima verosimilitud restringida de los componentes de varianza y covarianza.

	BW	WW	CCW	CON	FAT
σ_a^2	13.231 ± (0.128)	781.68 ± (52.69)	722.98± (17.22)	0.224 ± (0.061)	0.075 (0.027)
σ_m^2	-	279.75± (22.59)	-	-	-
σ_{am}	-	-321.77± (31.47)	-	-	-
σ_p^2	12.202± (0.218)	374.80± (7.87)	184.35± (4.21)	0.034 ± (0.011)	0.012 ± (0.005)
σ_e^2	24.692 ± (0.147)	683.84± (27.07)	220.40± (10.98)	0.145 ± (0.041)	0.103 ± (0.020)
h_a^2	0.264	0.372	0.641	0.556	0.395
h_m^2	-	0.133	-	-	-
r_{am}	-	-0.687	-	-	-

σ_a^2 = Varianza genética aditiva directa, σ_m^2 = Varianza genética aditiva materna
 σ_{am} = covarianza genética aditiva materna, σ_p^2 = Varianza permanente,
 σ_e^2 =Varianza del residual, h_a^2 = heredabilidad directa, h_m^2 =heredabilidad materna,
 r_{am} =Correlación genética aditiva materna.

4.2. Peso al nacimiento (BW).

Los resultados del promedio estandarizado de los 12 procedimientos para el barrido genómico se presentan en la figura 4.1. Los resultados detallados de los barridos genómicos específicos para los segmentos de 10 a 60 marcadores SNP y para las regiones de 0.5 a 3 Mb se presentan en el apéndice 1.1.

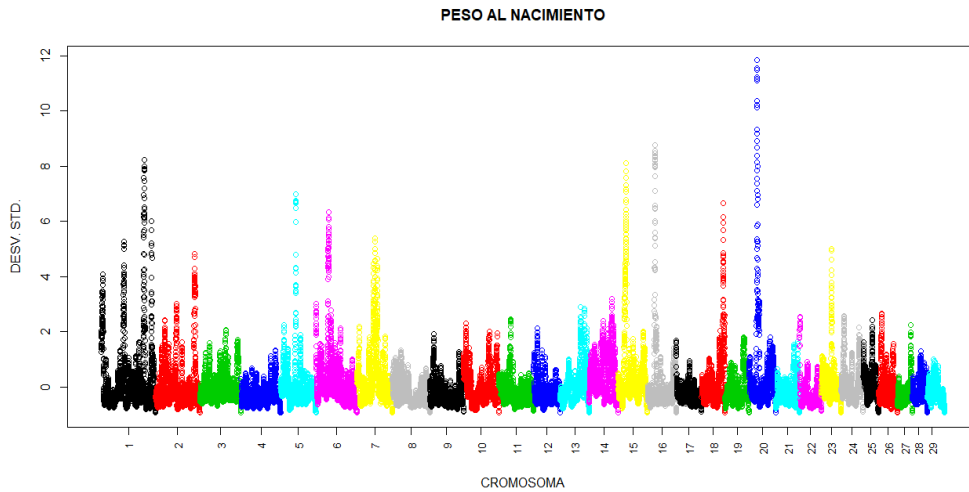


Figura 4.1 Barrido Genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Peso al Nacimiento (BW).

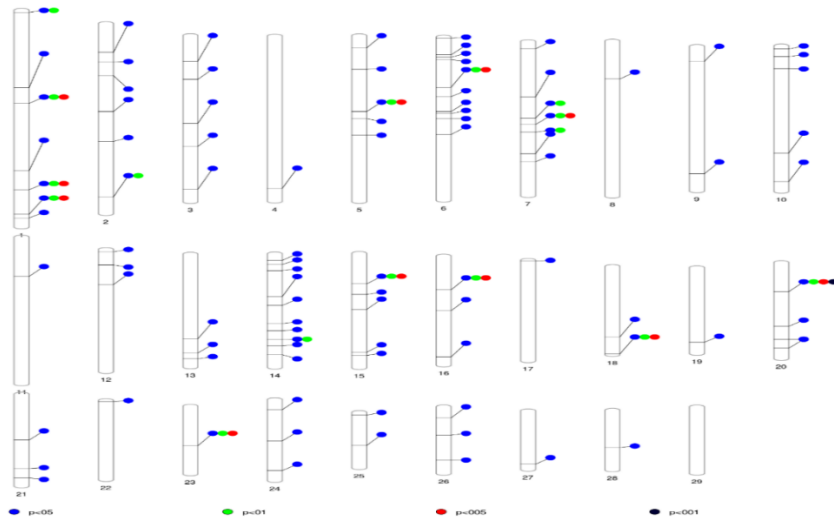


Figura 4.2. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Peso al Nacimiento (BW).

Finalmente, los genes localización en las regiones el genoma cuyo promedio estandarizado supera el percentil 99.5% se presentan en la Tabla 4.2.

Tabla 4.2 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Peso al Nacimiento (BW).

Cromosoma	Inicio	Fin	Genes
1	68353295	68457522	ROPN1
1	13172159 9	132347489	IL20RB, NCK1, SLC35G2.
1	14985650 0	150008106	KCNJ6
5	56237069	56541109	LRP1, SNORA62, NAB2, NEMP1, MYO1A, TAC3, ZBTB39, GPR182, RDH16
6	36708862	38689886	SPP1, MEPE, IBSP, LAP3, MED28, FAM184B, NCAPG, DCAF16, LCORL
7	60957222	61093739	ARHGEF37, PPARGC1B
15	24150048	24216893	-
16	24810362	25689363	DUSP10
18	65063534	65231440	ZSCAN4
20	20538223	22016148	RAB3C, GAPT, PLK2, ACTBL2
23	30595850	30735804	ZNF165

4.3. Peso al destete: Efecto directo (WWD).

Los resultados de la asociación de genoma completo con el promedio estandarizado de los 12 procedimientos para el efecto directo del peso al destete se presentan en la figura 4.3. Los resultados específicos de cada análisis se presentan en la Apéndice 1.2.

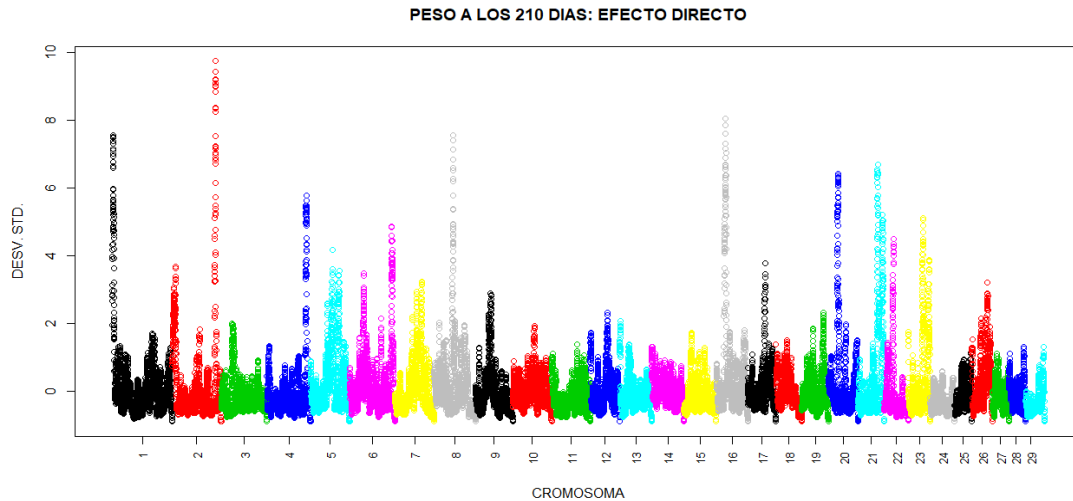


Figura 4.3. Barrido Genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el efecto directo del Peso a los 210 días (WWD).

La identificación de las regiones genómicas que mostraron un resultado de varianza explicada por encima del percentil 95 para el efecto directo del peso al deteste (WW) se presentan en la Figura 4.4.

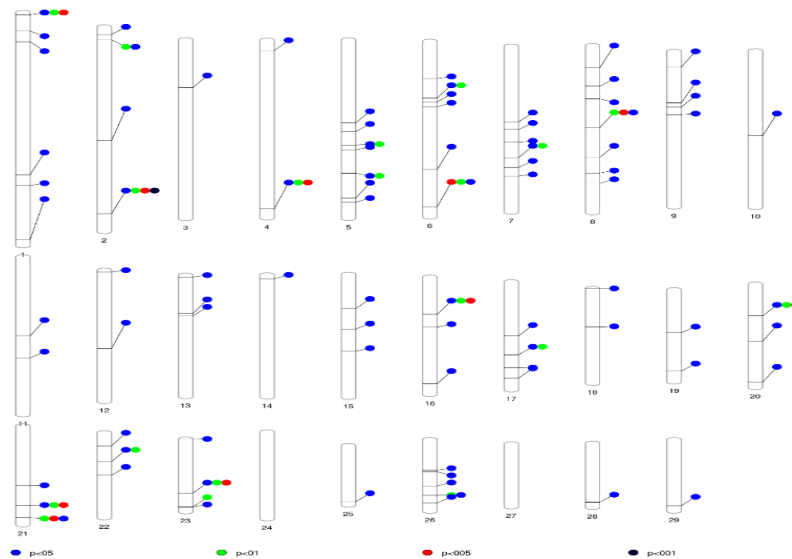


Figura 4.4. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el efecto directo del Peso a los 210 días (WWD).

En la tabla 4.3 se muestra una lista completa de los genes localizados en las regiones del genoma que superan el percentil 99.5%.

Tabla 4.3 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el efecto directo del Peso a los 210 días (WWD).

Cromosoma	Inicio	Fin	Genes
1	1143137	1889540	MRPS6, ATP5PO, ITSN1
2	126126854	127480295	SLC9A1, TENT5B, TRNP1, KDF1, NUDC, NR0B2, GPATCH3, GPN2, SFN, ZDHHC18, PIGV, ARID1A, RPS6KA1, DHDDS, LIN28A, ZNF683, CRYBG2, CD52, UBXN11, SH3BGRL3, CEP85, CATSPER4, CNKSR1, ZNF593, FAM110D, PDIK1L, TRIM63, SLC30A2, EXTL1, PAFAH2, STMN1, PAQR7, AUNIP, MTFR1L, SELENON, MAN1C1.
4	114081367	114436171	PRKAG2
6	112382578	112413487	-
8	55225038	55807984	TLE4
16	25738442	25944271	-
20	25733134	25944271	-
21	57156944	57550317	SLC24A4, RIN3, LGMN, GOLGA5
21	65063551	66951310	EVL, DEGS2, YY1, SLC25A29, SCL25A47, WAR1, WDR25, BEGAIN, DLK1, RTL1, SNORD112, SNORD113-1, SNORD113-2, SNORD113-4, SNORD 113-6, SNORD 113-7, SNORD 113-8, SNORD 113-9.
23	39066459	23	-

4.4. Peso al destete: Efecto Materno (WWM).

El resultado estandarizado de la varianza explicada por los 12 procedimientos de definición de las regiones del genoma para el efecto materno del peso al destete se muestra en la figura 4.5. Los resultados específicos para cada uno de los procedimientos se encuentran en el apéndice 1.3.

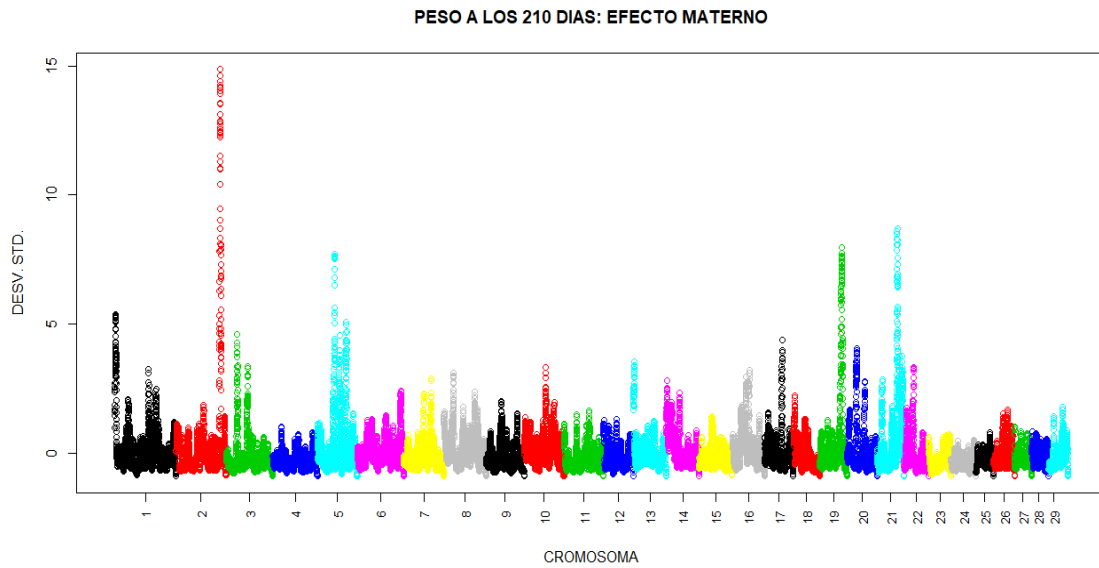


Figura 4.5. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el efecto materno del Peso a los 210 días (WWM).

Las regiones del genoma cuyo resultado supero el percentil 95% se presentan en la figura 4.6.

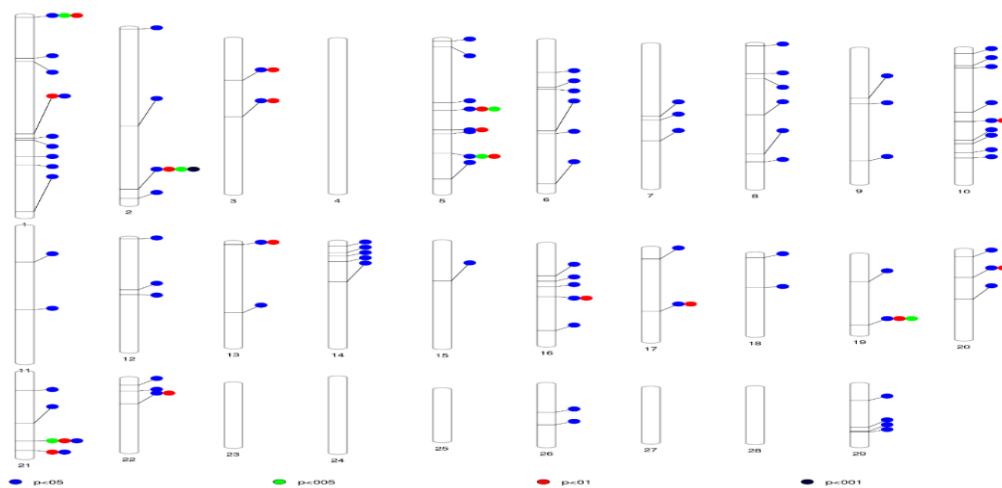


Figura 4.6. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el efecto materno del Peso a los 210 días (WWM).

Finalmente, los genes localizados en las regiones cuyo promedio estandarizado para el peso a los 210 días se encontró por encima del percentil 99.5% se presentan en la tabla 4.4

Tabla 4.4 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el efecto materno del Peso a los 210 días (WWM).

Cromosoma	Inicio	Fin	Genes
1	1513277	1562560	-
2	125826843	127592315	WASF2, GPR3, CD164L2, MAP3K6, SYTL1, TMEM222, WDTC1, SLC9A1, TENT5B, TRNP1, KDF1, NUDC, NR0B2, GPATCH3, GPN2, SFN, ZDHHC18, PIGV, ARID1A, RPS6KA1, DHDDS, LIN28A, ZNF683, CRYBG2, CD52, UBXN11, SH3BGRL3, CEP85, CATSPER4, CNKSR1, ZNF593, FAM110D, PDIK1L, TRIM63, SLC30A2, EXTL1, PAFAH2, STMN1, PAQR7, AUNIP, MTFR1L, SELENON
5	127592315	56552682	SLC26A10, ARHGEF25, DTX3, PIP4K2C, KIF5A, DCTN2, MBD6, DDIT3, MARS1, ARHGAP9, GLI1, INHBE, INHBC, R3HDM2, STAC3, NDUFA4L2, SHMT2, NXPH4, LRP1, SNORA62, STAT6, NAB2, NEMP1, MYO1A, TAC3, ZBTB39, GPR182, RDH16, SDR9C7
5	89545151	89793252	-
19	56104218	56559004	GRB2, SLC25A19, MIF4GD, MRPS7, GGA3, NUP85, SUMO2, JPT1, NT5C, ARMC7, SLC16A5, KCTD2, ATP5PD, MRPL58, CDR2L, HID1, OTOP3, OTOP2, USH1G, FADS6, FDXR, GRIN2C.
21	56679149	57559563	TRIP11, ATXN3, NDUFB1, CPSF2, SLC24A4, LGMN, GOLGA5.

4.5. Peso de la Canal Fría (CCW).

En primer lugar, en la figura 4.7 se presenta el barrido genómico del resultado estandarizado de la varianza explicada por los segmentos del genoma para el

peso de la canal fría (CCW). Los resultados detallados se encuentran en el apéndice 1.4

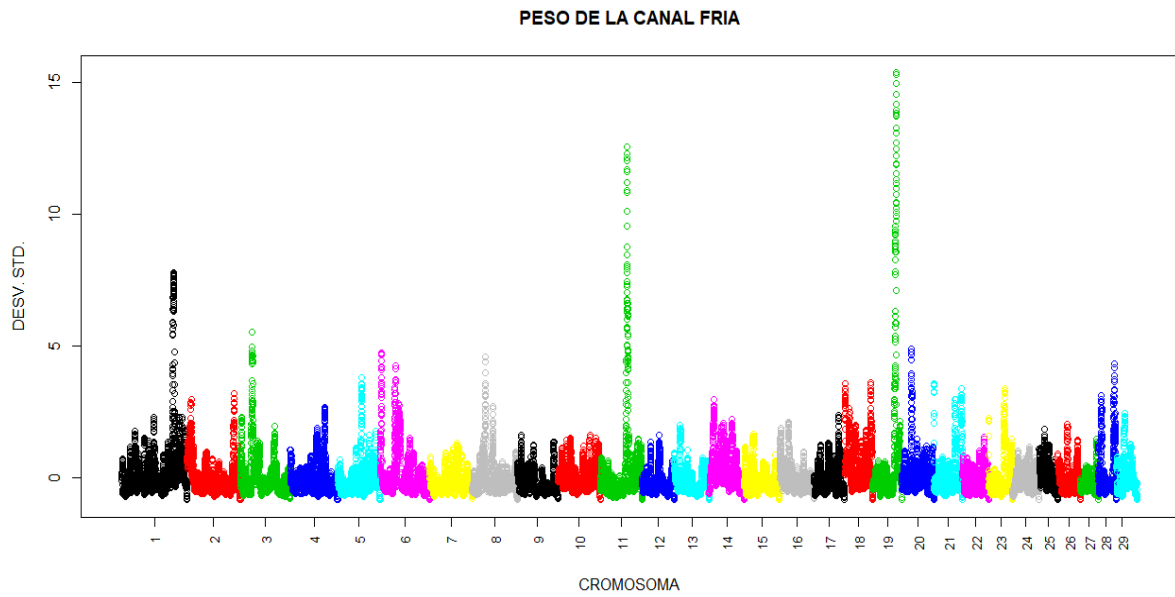


Figura 4.7. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para el Peso de la Canal Fría (CCW).

Las regiones genómicas de que muestran una relevancia por encima de los percentiles 95% (azul), 99% (verde), 99.5% (rojo) y 99.9% (negro) para el carácter Peso de la Canal Fría se presentan en la figura 4.8.

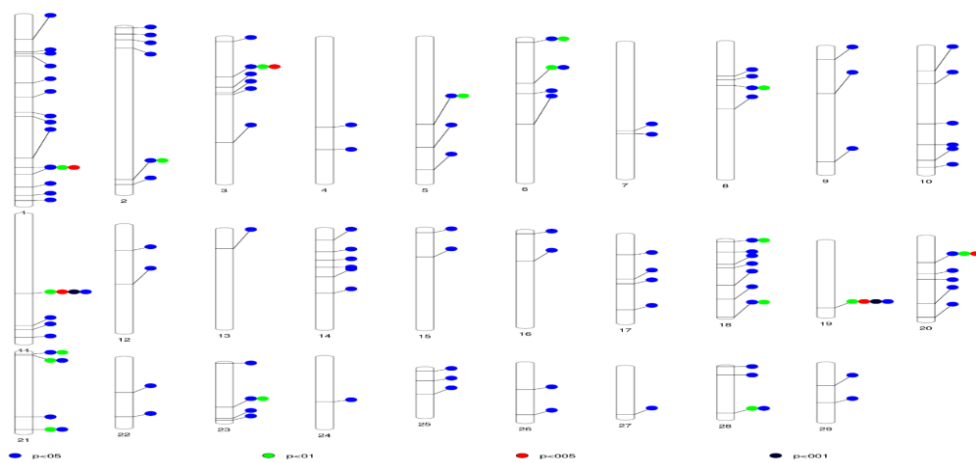


Figura 4.8. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para el Peso de la Canal Fría (CCW).

Finalmente, los genes situados en las regiones del genoma con un porcentaje de la varianza explicada por encima del percentil 99.5% se presentan en la tabla 4.5.

Tabla 4.5 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para el Peso de la Canal Fría (CCW).

Cromosoma	Inicio	Fin	Genes
1	131721599	132437217	IL20RB, NCK1, SLC35G2.
3	32479004	32788575	CD53, KCNA3
11	65386179	68938216	ETAA1, C1D, WDR92, PNO1, PPP3R1, CNRIP1, PLEK, FBXO48, APLF, PROKR1, ARHGAP25, BMP10, GKN3P, GNK2, GNK1, ANTXR1, 7SK, GFPT1, NFU1, AAK1, ANXA4, GMCL1, SNRNP27, MXD1, ASPRV1, PCBP1, TIA1, PCYOX1, SNRPG, EHD3, CAPN14, GALNT14.
19	55569727	56555839	EXOC7, ZACN, GALR2, SRP68, EVPL, CDK3, TEN1, ACOX1, FBF1, MRPL38, TRIM65, TRIM47, WBP2, UNC13D, UNK, H3F3B, GALK1, ITGB4, SAP30BP, RECQL5, SMIM6, SMIM5, LLGL2, TSEN54, CASKIN2, TMEM94, GRB2, SLC25A19, MIF4GD, MRPS7, GGA3, NUP85, SUMO2, JPT1, NT5C, ARMC7, SLC16A5, KCTD2, ATP5PD, MRPL58, CDR2L, HID1, OTOP3, OTOP2, USH1G, FADS6, FDXR, GRIN2C.
20	21930618	21930618	-

4.6. Conformación (CON).

Los resultados del barrido genómico para el carácter Conformación (CON) y de las regiones del genoma que explican un mayor porcentaje de la varianza genética se muestran en las figuras 4.9 y 4.10 y en la tabla 4. 6. Los resultados detallados para todos los métodos de definición de las regiones genómicas se presentan en el apéndice 1.5

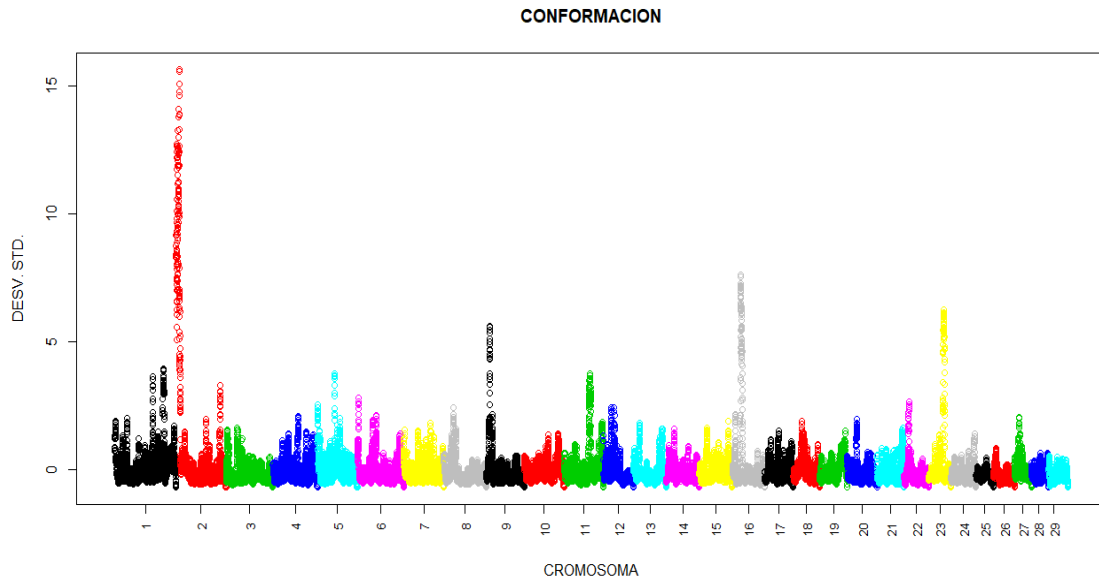


Figura 4.9. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Conformación (CON).

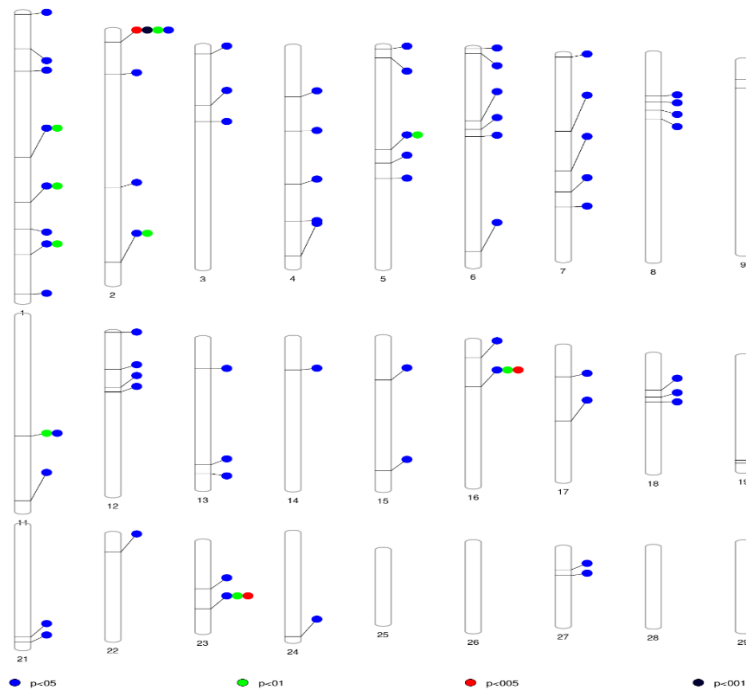


Figura 4.10 Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Conformación (CON).

A partir de las regiones genómicas cuyo promedio ajustado se ha encontrado por encima del percentil del 99.5% se muestran en la Tabla 4.6

Tabla 4.6 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Conformación (CON).

Cromosoma	Inicio	Fin	Genes
2	189886	11792974	LGSN, OCA2, HERC2, NIPA1, NIPA2, CYFIP1, TUBGCP5, CCDC115, IMP4, PTPN18, AMER3, ARHGEF4, FAM168B, PLEKHB2, HS6ST1, UGGT1, SAP130, AMMECR1L, POLR2D, WDR33, SFT2D3, LIMS2,
9	10328193	10711235	OGFRL1
16	25185581	25933132	DUSP10
23	38810995	39312609	RNF144B

4.7. Engrasamiento (FAT).

Finalmente, en las figuras 4.11, 4.12 y la tabla 4.7 se presentan los resultados para el carácter Engrasamiento (FAT). Los resultados detallados para cada uno de las 12 definiciones de las regiones genómicas se presentan en el apéndice 1.6.

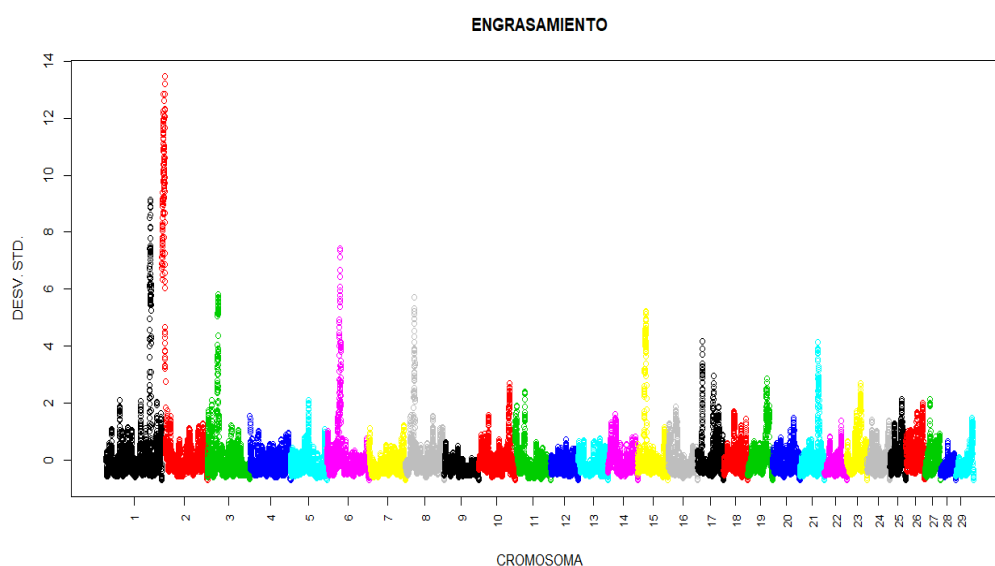


Figura 4.11. Barrido genómico del resultado estandarizado de la varianza genética explicada en torno a cada SNP para Engrasamiento (FAT).

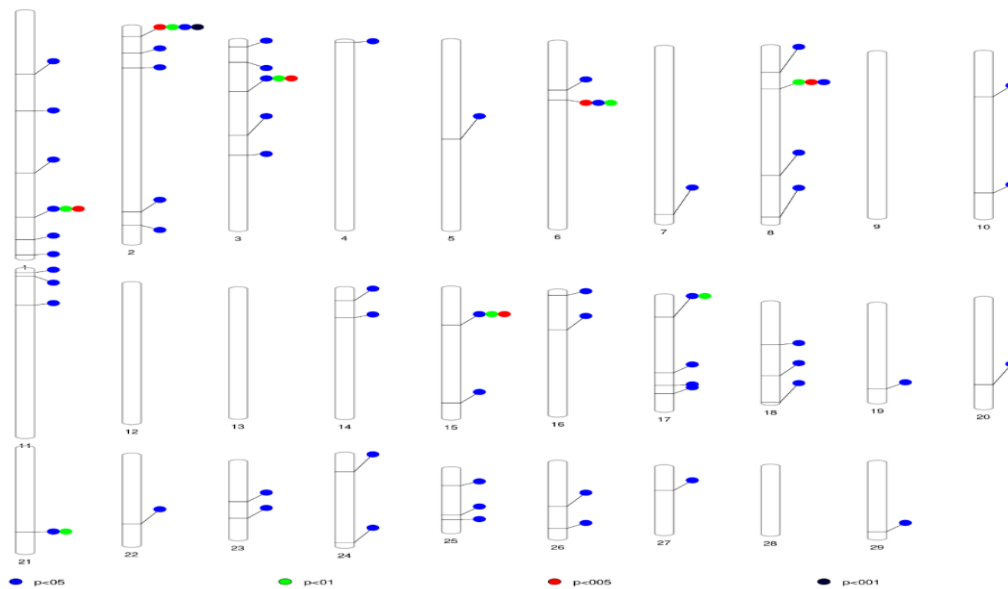


Figura 4.12. Regiones genómicas con un resultado estandarizado del porcentaje de varianza explicada situado por encima de los percentiles 95 (azul), 99 (verde), 99.5 (rojo) y 99.9 % (negro) para Engrasamiento (FAT).

Tabla 4.7 Genes localizados en las regiones de genoma cuyo promedio estandarizado de la varianza explicada superó el percentil 99.5% para Engrasamiento (FAT).

Cromosoma	Inicio	Fin	Genes
1	131423039	132497074	IL20RB, NCK1, SLC35G2
2	189886	9845870	LGSN, OCA2, HERC2, NIPA1, NIPA2, CYFIP1, TUBGCP5, CCDC115, IMP4, PTPN18, AMER3, ARHGEF4, PLEKHB2, HS6ST1, UGGT1, SAP130, AMMECR1L, POLR2D, WDR33, SFT2D3, LIMS2, GPR17, MYO7B, IWS1, PROC, MAP3K2, ERCC3, CYP27C1, BIN1, NAB1, NEMP2, MFSD6, INPP1, HIBCH, MSTN, PMS1, ORMDL1, OSGEPL1, ANKAR, ASNSD1, SLC40A1, WDR75, COL5A2, COL3A1, GULP1, 7SK, TFPI, CALCRL, ZSWIM2, FAM171B, ITGAV, ZC3H15.

3	32788575	32907274	KCNA2
6	36829725	37501365	MEPE, IBSP, LAP3, MED28, FAM184B, NCAPG, DCAF16, LCORL
8	26729617	26971008	SH3GL2, CNTLN
15	24156753	24171575	

5. DISCUSION

Los resultados de la estimación de componentes de la varianza confirmaron que existe un porcentaje muy alto de la variabilidad fenotípica que es explicada por causas genéticas aditivas. En concreto, las estimaciones de las heredabilidades oscilaron entre 0.264 para BW hasta 0.641 para CCW. Estas estimaciones están dentro del rango de las obtenidas para otras poblaciones (Simm, 1998; Eriksson et al., 2003; Hickey et al., 2007; Altarriba et al. 2009; Savoia et al., 2019; Srivastava et al., 2019). Además, en WW, el modelo de análisis utilizado por ACRUGA incluye los efectos genéticos maternos además de los efectos aditivos. El resultado de la estimación por máxima verosimilitud proporcionó un estimador de la varianza genética materna (y de su heredabilidad) similar a los registrados en la literatura (Robinson, 1996; Meyer, 1997). En este mismo carácter (WW), la correlación entre efectos genéticos directos y maternos fue alta y negativa (-0.687), como es habitual es los estudios de peso al destete en las poblaciones de vacuno de carne (Robinson, 1996; Meyer, 1997; Varona et al., 2015).

Los resultados del ssGWAS mostraron que esta importante variabilidad genética aditiva no se distribuye de manera homogénea a lo largo del genoma, sino que hay regiones del mismo asociadas a una mayor cantidad. En este sentido, si se calcula una curva de Lorenz como la que se presenta en la Figura 5.1 para BW, se observa cómo los 20,000 marcadores SNP con menor efecto en la varianza genética aditiva explican menos del 20% de su magnitud, independientemente del procedimiento para la determinación de las regiones genómicas.

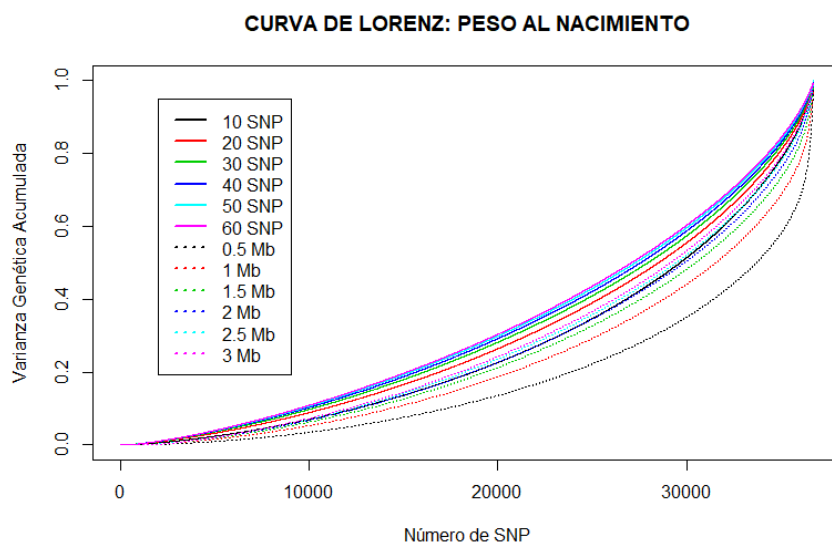


Figura 5.1. Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso al Nacimiento.

En este mismo sentido, en la Tabla 5.1, se presenta el cálculo de coeficiente de desigualdad de Gini (Gini, 1912), que, para BW, osciló entre 0.346, con regiones definidas por 60 marcadores SNP, y 0.638, para regiones definidas por 0.5 Mb. Como era esperable, a medida que las regiones se definieron con un tamaño mayor, el coeficiente de desigualdad se redujo.

Tabla 5.1. Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y para Peso al Nacimiento.

	Gini Index	10%	20%	50%
10 SNP	0.461	1.42%	4.21%	19.28%
20 SNP	0.406	1.71%	5.05%	22.41%
30 SNP	0.381	1.94%	5.59%	23.95%
40 SNP	0.365	2.17%	6.00%	25.00%
50 SNP	0.354	2.39%	6.33%	25.74%
60 SNP	0.346	2.55%	6.61%	26.30%
0.5 Mb	0.638	0.23%	0.64%	8.29%
1 Mb	0.542	0.44%	1.56%	14.13%
1.5 Mb	0.495	0.72%	2.59%	17.07%
2 Mb	0.470	1.00%	3.33%	18.68%
2.5 Mb	0.453	1.24%	3.95%	19.85%
3 Mb	0.441	1.45 %	4.40%	20.70%

En la misma tabla, puede observarse como el 10% de la variabilidad genética aditiva se explica por un rango de entre el 0.23% y el 2.55% de los marcadores SNP, mientras que el 50% de variación genética aditiva fue explicada por entre el 8.29 % y el 26.30 % de los marcadores SNP. Los resultados para los demás caracteres fueron similares y se presentan en el apéndice II.

Los modelos de análisis que se utilizaron en el procedimiento de ssGWAS fueron univariantes, aunque existe evidencias de que los caracteres de crecimiento y de calidad de la canal, como los utilizados la valoración genética están correlacionados entre sí (Altarriba et al., 2009; Savoia et al., 2019; Srivastava et al., 2019). Esta correlación genética entre los caracteres se pone de manifiesto si estudiamos la correlación entre los estimadores de la varianza explicada por los SNP estandarizados para los 6 caracteres analizados que se presentan en la Figura 5.2.

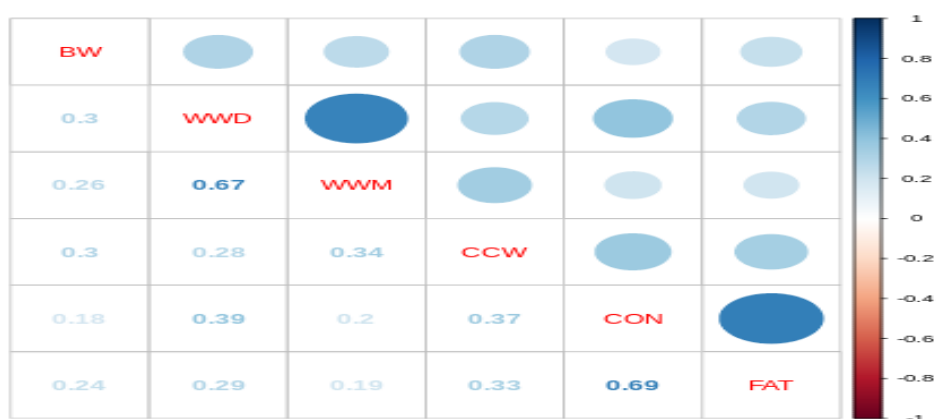


Figura 5.2 Correlaciones entre los resultados estandarizados del porcentaje de la varianza explicada para cada marcador SNP y entre Peso al Nacimiento (BW), efecto directo en el Peso a los 210 días (WWD), efecto materno en el Peso a los 210 días (WWM), Peso de la Canal Fría (CCW), Conformación (CON) y Engrasamiento (FAT).

Como se puede observar, todas las correlaciones entre los caracteres son positivas y ponen de manifiesto la presencia de pleiotropía entre los 6 caracteres analizados. Las correlaciones más elevadas aparecen entre WWD y WWM y entre CON y FAT, y son coherentes con la correlación genética entre efectos directo y materno en el peso al destete (Robinson, 1996; Meyer, 1997; Varona

et al., 2015) y entre con la correlación genética previamente descrita entre caracteres de conformación y engrasamiento (Altarriba et al., 2009, Savoia et al., 2019; Srivastava et al., 2019). Además, se confirman los resultados de Purfield et al. (2020), que mostraron un alto grado de pleiotropía entre estos grupos de caracteres.

La presencia de pleiotropía es confirmada cuando se analizan las regiones del genoma que explican un porcentaje de varianza relevante para los caracteres analizados. Así, en la tabla 5.2 se presentan 12 regiones genómicas cuyo estimador estandarizado de la varianza genética explicada se situó por encima del percentil 95 en al menos 4 de los caracteres.

Tabla 5.2. Regiones del genoma que mostraron un resultado estandarizado por encima del percentil 95% en al menos cuatro caracteres.

BTA	Inicio	Fin	BW	WWD	WWM	CCW	CON	FAT
1	1492106	1562560	**	****	***	-	*	*
1	131326810	132456625	***	-	-	***	**	***
2	904065	9202511	-	**	*	**	****	****
2	126126854	127514835	***	****	****	**	**	*
3	32148583	32829586	*	*	***	***	*	***
5	55865567	56541588	***	*	***	*	**	*
6	36518232	38689886	***	**	*	***	*	***
16	24983313	25934724	****	****	*	**	***	*
19	55820166	56556421	*	*	***	****	*	*
20	20764625	22126788	****	***	**	***	*	*
21	56758194	57579759	*	***	****	**	*	**
23	39066459	39091834	***	***	-	**	***	*

El símbolo **** identifica a las regiones con unos resultados por encima del percentil 99.9%, y los símbolos ***, **, * señalan a las regiones con unos resultados por encima del 99.5%, 99% y 95%, respectivamente.

Entre estas regiones del genoma hay dos, localizadas en los cromosomas BTA2 y BTA6 que ya han sido destacadas por Purfield et al. (2020) como susceptibles de albergar genes relacionados con caracteres de crecimiento y de calidad de la

canal en numerosos estudios. La primera región se localiza entre los pares de bases 904065 y 9202511 y contiene en su interior al gen de la *miostatina* (MTSN). El gen *MSTN* está asociado con el fenotipo de doble músculo en vacuno (Grobet et al., 1997) y los resultados de este estudio muestran una fuerte asociación con la variabilidad genética de los caracteres CON y FAT, confirmando resultados previamente obtenidos en la literatura (Martin et al., 2019; Purfield et al., 2020). Pese a que la política oficial de la población Rubia Gallega implica la selección en contra del carácter “culón” debido a su asociación con la dificultad en el parto (Billinge et al. 2005), se sabe que alguna de las mutaciones encontradas en el gen MTSN (Dunner et al., 2003) está segregando en la población, ya que en estudios previos se encontró una fuerte señal de selección asociada a la misma (González-Rodríguez et al., 2017). Sin embargo, la enorme longitud de la región puede indicar que existe otros genes asociados con la variabilidad genética de los caracteres analizados. En este sentido, Gutiérrez-Gil et al. (2015) sugieren también como posibles candidatos a los siguientes genes: *SLC40A1* (*solute carrier family 40, member 1*), *COL5A2* (*collagen, type V, alpha 2*), *COL3A1* (*collagen, type III, alpha 1*), *CALCRL* (*calcitonin receptor-like*) y *ITGAV* (*integrin, alpha V*).

La segunda de las regiones destacadas por Purfield et al. (2020), se localiza en el cromosoma BTA6 entre los pares de bases 36518232 y 38689886. En ella se localizan dos genes que han sido propuestos como candidatos en numerosas publicaciones: *LCORL* (*Ligand dependent nuclear receptor corepressor like*) y *NCAPG* (*Condensin complex subunit 3*). El gen *LCORL* codifica un correpresor del receptor nuclear dependiente de ligando, y se ha detectado que afecta a la estatura en humanos (Horikosi et al., 2013), caballos (Metzger et al., 2013) y vacas (Pryce et al., 2011). El gen *NCAPG* codifica un gen que actúa en la condensación de los cromosomas y, junto con el gen *LCORL*, se ha asociado al crecimiento fetal y el peso de la canal en vacuno (Olsen et al 2005, Eberlein et al, 2009, Setoguchi et al., 2009; Saatchi et al., 2014), en la deposición grasa (Setoguchi et al., 2009, Weikard, et al., 2010, Setoguchi et al., 2011, Lindholm-Perry et al., 2011, Hoshiba, et al., 2013), el consumo de alimento (Snelling et al., 2010; Snelling et al., 2011, Lindholm-Perry et al., 2011, Bongiorno et al., 2012, Nishimura et al., 2012) y a la calidad de la canal y la carne (Wang et al., 2020;

Pegolo et al., 2020). Además, se trata de una región del genoma que ha sido identificada en numerosas ocasiones como sujeto de un proceso de selección reciente (Gutiérrez-Gil et al., 2015; González-Rodríguez et al., 2017).

Al margen de estas dos regiones del genoma para las que existe un consenso generalizado en la literatura se han identificado otras regiones del genoma con resultados interesantes. En el cromosoma BTA1 se identifican dos regiones de especial interés, la primera de ellas se localiza entre las pares de bases 1492106 y 1562560 y afecta a BW, WWD, WWM y, en menor medida a CON y FAT. Entre los genes localizados dentro de la región o en su cercanía destaca el gen *MRPS6* (*Mitochondrial Ribosomal Protein S6*), que codifica una proteína mitocondrial implicada en el metabolismo energético y que cuya expresión genética ha sido asociada diferencialmente con vacas en distintos estados fisiológicos (Valour et al., 2014). La segunda región genómica con efectos pleiotrópicos dentro del BTA1 se localiza entre las posiciones 131326810 y 132456625, y afecta de manera importante a BW, CCW, CON y FAT. Entre los genes localizados dentro de la región destaca el *NCK1* (*NCK Adaptor Protein 1*) que está relacionado a la obesidad en humanos, ya que modula el desarrollo del tejido adiposo blanco WAT (*white adipose tissue*) través de la remodelación de preadipocitos dependiente de PDGFR α en humanos y ratones (Haider et al., 2018). Además, ha sido asociado con tolerancia a la glucosa en ratones (Latreille et al., 2011), y, en vacuno de carne, ha sido identificado como un posible gen candidato para caracteres de crecimiento por Seabury et al. (2017).

Además de la región que alberga el gen *MTSN*, en el BTA2 se localizó otra zona entre los pares de bases 126126854 y 127514835, que fue especialmente relevante en los caracteres BW, WWD y WWM. En ella se localiza el gen *SESN2* (*Sestrin 2*), que está implicado en la regulación de las funciones mitocondriales (Kovaleva et al., 2020) y regula el metabolismo energético (Ben-Sahra et al., 2013). En vacuno, regula de la producción de caseína en las células mamarias (Luo et al., 2018) y ha sido asociado con la producción y composición de la leche (Li et al., 2019). Estos resultados concuerdan con el relevante porcentaje de la variabilidad genética asociada a esta región del genoma en los efectos maternos (WWM) y en los pesos al nacimiento (BW) y al destete (WWD).

En el cromosoma BTA3 se localizó una región genómica situada entre los pares de bases 32148583 y 32829586, donde se localiza el gen *KCNA3* (*Potassium Voltage-Gated Channel Subfamily A Member 3*), según estudio realizado por Li et al, (2017) se asoció al marmóreo de la carne en ganado nativo coreano.

La región situada en el cromosoma BTA5, entre los pares de bases 55865567 y 56541588, se encuentra el gen *DDIT3* (*DNA damage-inducible transcript 3*), que codifica un miembro de la familia de factores de transcripción CCAAT / proteína de unión al potenciador (C / EBP). Varios miembros de esta familia participan en la adipogénesis como el C/ EBP α que es un regulador importante de la adipogénesis (Rosen et al.,2002). Los miembros de la familia C/EBP se expresan en estadios tempranos de la adipogénesis (Darlington et al., 1998). C/EPB β Y C/EBP δ son considerados factores pro-adipogénicos, mientras que las proteínas (C / EBP γ y C / EBP) homóloga del DDIT3 parecen ser supresores en el programa adipogénico (Rosen y MacDougald, 2006). Además, el gen DDIT3 ha sido también asociado a la producción y composición de la leche (Li et al., 2019) y concuerda con el efecto relevante que tiene esta región en WWM y BW.

En el cromosoma BTA16, y entre las pares de bases 24983313 y 25934724, se ha detectado una región del genoma que explica un porcentaje importante de la variabilidad genética para BW, WWD y CON. En ella se localiza el gen *BPNT1* (*3'(2')*, *5'-Bisphosphate Nucleotidase 1*), que ha sido asociado con el contenido de grasa intramuscular (Hay y Roberts, 2018) y con la edad a la pubertad (Melo et al., 2019).

El gen *ACOX1* (*Acyl-CoA Oxidase 1*) codifica la primera enzima de la vía de oxidación beta de los ácidos grasos, que juega un papel importante en el metabolismo de los ácidos grasos y la deposición de las grasas (Jiao et al.,2011) y se localiza en una región genómica del cromosoma BTA19 (55820166-56556421 bp) que explica una parte importante de la varianza genética en WWM y CCW. Esta región fue asociada con la variabilidad genética en la ingesta residual en una población Hereford (Saatchi et al., 2014)

En el cromosoma 20 y entre los pares de bases 20764625 y 22126788 se ha localizado un segmento del genoma que se relaciona en mayor medida con a la variabilidad genética aditiva para BW, WWD y CCW. Entre los genes localizados

dentro del segmento o en su vecindad destaca la presencia del gen *PLK2* (*Serine/threonine-protein kinase PLK2*), que codifica una kinasa que tiene un papel en la división celular (Archambault y Glover, 2009) y que ha sido asociado al crecimiento de ratones en sus primeras fases de vida (Ma et al., 2003).

En el cromosoma 21 (BTA21) y entre los pares de bases 56758194 y 57579759 se localizó una región con un efecto importante en WWD y WWM. Esta región contiene al gen *SLC24A4* (*Solute Carrier Family 24 Member 4*), que codifica un miembro de la familia de proteínas de intercambio de sodio o calcio dependientes de potasio y que se ha asociado a al intervalo de parto y los días abiertos en el ganado Holstein lechero canadiense (Nayeri et al., 2016) y, más recientemente, ha sido propuesto por Yan et al. (2019) como gen candidato para la conformación mamaria, lo que concuerda con su relación con el efecto materno al destete y con su asociación con la variabilidad genética al destete.

Finalmente, en el BTA23 y entre las pares de bases 39066459 y 39091834 se localiza una región del genoma que explica parte de la variabilidad genética para BW, WW y CON y que ha sido asociada con QTL para peso de la canal (McClure et al., 2010) y ganancia media diaria (Nkrumah et al., 2007). Entre los genes presentes en es su vecindad puede ser destacado el *RNF144B* (*Ring Finger Protein 144B*). que ha sido asociado con susceptibilidad a la tuberculosis (Raphaka et al., 2017) y que interviene con la respuesta inmunológica de los individuos (Baek et al. 2016).

Al margen de estas regiones con efectos pleiotrópicos, algunas regiones han mostrado un efecto muy destacado en la regulación de la variabilidad genética de alguno de los caracteres analizados. Así, para BW existe una región con un efecto importante en el BTA15 entre los pares de bases 24150048 y 24216893, que no contiene ningún gen en su interior pero que se encuentra cerca del gene *DRD2* (*Dopamine Receptor D2*), que es un receptor de la dopamina que, a su vez, está implicada en la regulación de la hormona del crecimiento. Para WW, es destacable una región del BTA8, entre las posiciones 55225038 y 55807984, que contiene el gen *TLE4* (*Transducin-like enhancer protein 4*) y que ha sido identificada por Ma et al. (2019) como una de las zonas del genoma que han

sufrido cambios relevantes como consecuencia de la selección en la población Holstein.

En el peso de la canal fría (CCW) existe una región del BTA11. situada entre los pares de bases 65386179 y 68938216, que ya fue detectada por González-Rodríguez et al. (2017) como susceptible de haber sufrido una señal de selección reciente, y que alberga entre otros el gen *BMP10* (*Bone Morphogenetic Protein 10*), que induce a la formación de hueso, cartílago y tejido conjuntivo (Chen et al., 2014). Además, en su cercanía se localiza dos de los genes que codifican calpaínas que están implicadas en procesos de proteólisis y han sido asociadas con la terneza de la carne (Kemp et al., 2010).

Por último, para CON y FAT se pueden destacar adicionalmente dos regiones del genoma situadas en los cromosomas 9 (10328193-10711235 bp) y 8 (26729617-26971008 bp), respectivamente. Esta última región contiene al gen SH3GL2 que está relacionado con el metabolismo lipídico (Taneera et al., 2012) y que ha sido identificado por Lim et al. (2016) como un gen candidato a haber sufrido selección reciente en ganado coreano.

Los resultados obtenidos en este trabajo nos ofrecen una visión inicial de la heterogeneidad en la contribución en la variabilidad genética aditiva a lo largo del genoma, y puede permitir, en un futuro, la modificación de los procedimientos de valoración genómica en la población Rubia Gallega. En estos momentos, la implantación de la valoración genómica se encuentra en las fases iniciales y utiliza un procedimiento GBLUP a partir del cálculo de la matriz genómica de VanRaden (2008). Sin embargo, este procedimiento asume que, a priori, todos los marcadores SNP incluidos en el análisis están relacionados de la misma manera con la variabilidad genética del carácter a estudiar. Sin embargo, como se ha observado en los resultados de este y otros trabajos (Bolormaa et al., 2014; Zhang et al., 2016; Fang et al 2019), esta asunción dista mucho de la realidad y existen algunas regiones genómicas mucho más implicadas que otras en la regulación de los caracteres de interés. En este sentido, Tiezzi y Maltecca (2015) propusieron la utilización de un procedimiento llamado "Weighted GBLUP" que pondera la información procedente de cada marcador SNP en la generación de la matriz de relaciones genómicas. Sin embargo, esta ponderación es

dependiente del carácter y, por lo tanto, dificulta la implementación multicarácter de la valoración genómica. En el caso de caracteres correlacionados, como los que se incluyen en el índice de selección de la población Rubia Gallega, la valoración multicarácter puede proporcionar un incremento de precisión relevante, sobre todo para aquellos caracteres que son valorados *post-mortem* (CCW, CON y FAT). En este sentido, en futuros estudios se deberá valorar la implementación de procedimientos que permitan la ponderación de los marcadores SNP para cada carácter y que sean compatibles con la implementación multicarácter con los propuestos recientemente por Karaman et al. (2018, 2020).

6. CONCLUSIONES

Los resultados obtenidos en este trabajo nos han permitido extraer las siguientes conclusiones:

1. Se ha detectado que la influencia de las zonas del genoma en la variabilidad genética aditiva de los caracteres de interés en el plan de mejora de la población Rubia Gallega es heterogénea a lo largo del genoma.
2. Se ha confirmado que las correlaciones genéticas entre los caracteres de crecimiento y características de la canal se ocasionan por los efectos pleiotrópicos asociados a algunas regiones del genoma.
3. Se han detectado doce regiones del genoma distribuidas a lo largo de los cromosomas 1, 2, 3, 5, 6, 16, 19, 20, 21 y 23 que afectan al menos a 4 de los 6 caracteres analizados.
4. La observación detallada de las regiones con mayor influencia en la variabilidad genética aditiva ha permitido sugerir numerosos genes candidatos que deben ser confirmados en futuros estudios.

7. REFERENCIAS

- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93:743-752. doi:<https://doi.org/10.3168/jds.2009-2730>
- Altarriba, J., Yagüe, G., Moreno, C., & Varona, L. (2009). Exploring the possibilities of genetic improvement from traceability data: An example in the Pirenaica beef cattle. *Livestocks Science*, 125:115-120. doi:<https://doi.org/10.1016/j.livsci.2009.03.013>
- Archambault, V., & Glover, D. M. (2009). Polo-like kinases: conservation and divergence in their functions and regulation. *Nature Reviews Biología Celular Molecular volumen*, 10:265-275. doi:<https://doi.org/10.1038/nrm2653>
- Baek, S. H., Huang, B., & Chang, H. W. (2016). RNF144b is a negative regulator in TLR2-mediated NF- κ B activation. *The Journal of Immunology*, 196(1):132-134.
- Bellinge, R. H., Liberles, D. A., Laschi, P. A., O'Brien, P. A., & Tay, G. K. (2005). Myostatin and its implications on animal breeding: a review. *Animal Genetics*, 36:1-6. doi:<https://doi.org/10.1111/j.1365-2052.2004.01229.x>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of Royal Statistical Society, Series B*, 57(1):289-300. doi:<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Ben-Sahra, I., Dirat, B., Laurent, K., Puissant, A., Auberger, P., & Bost, F. (2013). Sestrin2 integrates Akt and mTOR signaling to protect cells against energetic stress-induced death. *Journal of Biological Chemistry*, 20:611-619.

- Bolormmaa, S., Pryce, J. E., Reverter, A., Zhang, Y., Barendse, W., Kemper, K., . . . Goddard, M. E. (2014). A multi-trait, Meta-analysis for Detecting Pleiotropic Polymorphism for Statute, Fatness and Reproduction in Beef cattle. *Plos Genetics*, 10(3):e1004198. doi:10.1371/journal.pgen.1004198
- Bonferroni, C. E. (1936). Teoria Statistica della classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* .
- Carballo, J. A., Olieta, B., Moreno, T., Sanchez, L., & Monserrat, L. (2005). Categorizacion de las canales de Ternero Producidas en Galicia. *Archivos de Zootecnia*, 53: 119-128.
- Chang, C. C., Chow , C. C., Tellier , L. C., Vattikut, S., Purcell, S. M., & Lee , J. J. (2015). Second-generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *Gigascience*, 4:1-16. doi:10.1186/s13742-015-0047-8
- Chen, H., Shi, S., Acosta, L., Li, W., Lu, J., Bao, S., . . . Shou, W. (2004). BMP10 is essential for maintaining cardiac growth during murine cardiogenesis. *Development*, 84(2):248-256. doi:10.1242/dev.01094
- Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, 42:2. doi:https://doi.org/10.1186/1297-9686-42-2
- Darlington, G. J., Ross, S. E., & MacDougald, O. A. (1998). The Role of C/EBP Genes in Adipocyte Differentiation. *Journal of Biological Chemistry*, 273:30057–30060. doi:doi: 10.1074/jbc.273.46.30057
- De los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. (2013). Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics*, 193 (2): 327-345. doi:https://doi.org/10.1534/genetics.112.143313
- De los Campos, G., Naya, H., Gianola, D., Grossa, J., Legarra, A., Manfredi, E., . . . Cotes, J. M. (2009). Predicting Quantitative Traits With Regression Models for Dense Molecular Markers and Pedigree. *Genetics*, 182(1): 375-385. doi:https://doi.org/10.1534/genetics.109.101501

- Dekkers, J. M. (2004). Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science*, 82(13):E313-E328. doi:https://doi.org/10.2527/2004.8213_supplE313x
- Dunner, S., Miranda, M. E., Amigues, Y., Cañón, J., George, M., Hanset, R., & Ménissier, F. (2003). Haplotype diversity of the myostatin gene among beef cattle breeds. *Genetics Selections Evolution*, 35:103-118.
- Falconer, D. S., & Mackey, T. C. (1996). *Introduction to Quantitative Genetics*, 4th ed. Essex, United Kingdom: Longman Group Ltd.
- Fang, Z. H., & Pausch, H. (2019). Multi-Trait meta-analyses reveal 25 quantitative trait loci for economical important traits in Brown Swiss cattle. *BMC Genomics*, 20:695. doi:<https://doi.org/10.1186/s12864-019-6066-6>
- Feagas. (2018). AGRUGA, 50 años de Evolucion . *Federacion Española de Asociaciones de Ganado selceto(feagas)*, 18-27.
- Fernando, R., Toosi, A., Wolc, A., Garrick, D., & Dekkers, J. (2017). Application of Whole-Genome Prediction Methods for Genome-Wide Association Studies: A Bayesian Approach. *Journal of Agriculture, Biological and Environmental Statistics*, 22:172-193. doi:<https://doi.org/10.1007/s13253-017-0277-6>
- Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans R Soc Edin*, 52:399-433.
- Garrick, D. J. (2010). The Nature, Scope And Impact Of Some WholeGenome Analyses In Beef Cattle. *Proc 9th World Congress on Genentics Applied to Livestocks Production*, (pág. 23). Liepzig, Germany.
- Georges, M. (2007). Mapping, Fine Mapping, and Molecular Dissection of Quantitative Trait Loci in Domestic Animals. *Annual Review of Genomics and Human Genetics*, 8:131-162. doi:<https://doi.org/10.1146/annurev.genom.8.080706.092408>
- Gini, C. (1912). *Variabilità e Mutabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*. C. Cuppini.

- González-Rodríguez, A., Munilla, A., Mourensan, E. F., Cañas-Álvarez, J. J., Baró, J. A., Díaz, C., . . . Varona, L. (2017). Genomic differentiation between Asturiana de los Valles, Avileña-Negra Ibérica, Bruna dels Pirineus, Morucha, Pirenaica, Retinta and Rubia Gallega cattle breeds. *Animal*, 11:1667:1679. doi: <https://doi.org/10.1017/S1751731117000398>
- Groenen, M. M., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., & Rothschild, M. F. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 491:393-398. doi:<https://doi.org/10.1038/nature11622>
- Gutiérrez-Gil, B., Arranz, J. J., & Wiener, P. (2015). An interpretive review of selective sweep studies in *Bos taurus* cattle populations: identification of unique and shared selection signals across breeds. *Frontiers in Genetics*, 6,167. doi:<https://doi.org/10.3389/fgene.2015.00167>
- Habier, D., Fernando, R. L., & Dekkers, J. M. (2007). The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics*, 177(4):2389-2397. doi:<https://doi.org/10.1534/genetics.107.081190>
- Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*, 12:186. doi:doi: 10.1186/1471-2105-12-186
- Haley, C. S., & Knott, S. A. (1992). A Simple Regression Method for Mapping Quantitative Trait Loci in Line Crosses Using Flanking Markers. *Heredity The Genetical Society of Great Britain*, 69:315-324. doi:DOI: 10.1038/hdy.1992.131
- Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Quareshi, A., & Hunter, D. J. (2008). A Genome-Wide Association Study Identifies Novel Alleles Associated with Hair Color and Skin Pigmentation. *Plos Genetics*, 4(5). doi: doi: 10.1371/journal.pgen.1000074
- Hay, E., & Roberts, A. (2018). Genome-wide association study for carcass traits in a composite beef cattle breed. *Livestock Science*, 213:35-43. doi:<https://doi.org/10.1016/j.livsci.2018.04.018>

- Hayes , B. J., Pryce, J., Chamberlain , A. J., Bowman, P. J., & Goddard, M. E. (2010). Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits. *Plos Genetics*, 6(9): e1001139. doi:doi: 10.1371/journal.pgen.1001139
- Hayes, B. J., Bowman, P. J., Chamberlan, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, 92:433-443.
- Hazel, L. N., & Lush, J. L. (1942). THE EFFICIENCY OF THREE METHODS OF SELECTION. *Journal of Heredity*, 33(11):393-399. doi:https://doi.org/10.1093/oxfordjournals.jhered.a105102
- Henderson , C. R. (1984). *Aplications of Lineal Models in animal breeding*. Guelph, Canada: University of Guelph.
- Hickey, J. M., Keane, M. G., Cromie, D. A., & Veerkamp, R. F. (2007). Genetic parameters for EUROP carcass traits within different groups of cattle in Ireland. *Journal of Animal Science*, 85:314-321. doi:https://doi.org/10.2527/jas.2006-263
- Karaman, E., Lund, M. S., & Su, G. (2020). Multi-trait single-step genomic prediction accounting for heterogeneous (co)variances over the genome. *Heredity*, 124(2):274-287. doi:https://doi.org/10.1038/s41437-019-0273-4
- Karaman, E., Lund, M. S., Anche, M. T., Janss, L., & Su, G. (2018). Genomic Prediction Using Multi-trait Weighted GBLUP Accounting for Heterogeneous Variances and Covariances Across the Genome. *G3(Bethesda)*, 8(11):3549-3558.
- Kemp, C. M., Sensky, P. L., Bardsley, R. G., Buttery, P. J., & Parr, T. (2010). Tenderness – An enzymatic view. *Meat Science*, 84(2):248-256. doi:https://doi.org/10.1016/j.meatsci.2009.06.008
- Koch, R. M., & Cundiff, L. V. (1982). Heredabilities and genetics, enviromental and phenotypic correlations of carcass traits in a populations of diverse biological types and their implications in selection progrms. *Journal of America Science*, 55:1319-1329.

- Kovaleva, I. E., Tokarchuk, A. V., Zheltukhin, A. O., Dalina, A. A., Safronov, G. G., Evstafieva, A. G., & Budanov, A. V. (2020). Mitochondrial localization of SESN2. *Plos One*, 15(4): e0226862. doi:<https://doi.org/10.1371/journal.pone.0226862>
- Lader, E. S., & Boststein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121:185-199. Obtenido de <https://www.genetics.org/content/genetics/121/1/185.full.pdf>
- Latreille, M., Laberge, M. K., Bourret, G., Yamani, L., & Laros, L. (2011). Deletion of Nck1 attenuates hepatic ER stress signaling and improves glucose tolerance and insulin signaling in liver of obese mice. *Journal Physiol Endocrinol Metabolism*, 300:423-434. doi:<https://doi.org/10.1152/ajpendo.00088.2010>
- Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, 92(9): 4656-4663. doi:<https://doi.org/10.3168/jds.2009-2061>
- Legarra, A., Ricard, A., & Varona, L. (2018). GWAS by GBLUP: Single and Multimarker EMMAX and Bayes Factors, with an Example in Detection of a Major Gene for Horse Gait. *G3: Genes, Genomes, Genetics*, 8:2301-2308. doi:<https://doi.org/10.1534/g3.118.200336>
- Li, X., Kim, S. W., Do, K. T., Ha, Y. K., Lee, Y. M., Yoon, S. H., & Kim, K. S. (2010). Analyses of porcine public SNPs in coding-gene regions by re-sequencing and phenotypic association studies. *Molecular Biology Reports*, 38:3805–3820. doi:<https://doi.org/10.1007/s11033-010-0496-1>
- Li, Y., Gao, Y., Kim, Y. S., Iqbal, A., & Kim, J. J. (2017). A whole genome association study to detect additive and dominant single nucleotide polymorphisms for growth and carcass traits in Korean native cattle, Hanwoo. *Asia-Australasian of Animal Journal*, 30(1):8-19. doi:[doi:10.5713/ajas.16.0170](https://doi.org/10.5713/ajas.16.0170)
- Li, Y., Han, B., Lui, L., Zhao, F., Liang, W., Jian, J., . . . Sun, D. (2019). Genetic association of DDIT 3, RPL 23A, SESN 2 and NR 4A1 genes with milk yield and composition in dairy cattle. *Animal Genetics*, 50:123-135.

- Lim, D., Strucken, E. M., Choi, B. H., Chai, H. H., Cho, Y. M., Jang, G. W., . . . Lee, S. H. (2016). Genomic Footprints in Selected and Unselected Beef Cattle Breeds in Korea. *Plos One*, 11(3):e0151324. doi:10.1371/journal.pone.0151324
- Luo, C., Zhao, S., Zhang, M., Gao, Y., Wang, J., Haning, M. D., & Zheng, Z. (2018). SESN2 negatively regulates cell proliferation and casein synthesis by inhibition the amino acid-mediated mTORC1 pathway in cow mammary epithelial cells. *Scientific Reports*, 8:3912. doi:https://doi.org/10.1038/s41598-018-22208-w
- Ma, L., Sonstegard, T. S., Cole, J. B., VanTassell, C. P., Wiggans, G. R., Crooker, B. A., . . . Da, Y. (2019). Genome changes due to artificial selection in U.S. Holstein cattle. *BMC Genomics*, 20:128. doi:https://doi.org/10.1186/s12864-019-5459-x
- Ma, S., Charron, J., & Erikson, R. L. (2003). Role of Plk2 (Snk) in Mouse Development and Cell Proliferation. *Molecular and Cellular Biology*, 23:6936-6943. doi: 10.1128/MCB.23.19.6936-6943.2003
- Maiorana, A. M., Lourenco, D. L., Tsurata, S., Toro Opina, A. M., Bovino Stafuzza, N., Masuda, Y., & de Vasconcelo Silva, J. A. (2018). Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *Plos One*, 13(8). doi:doi:10.1371/journal.pone.0200694
- Martin , P., Taussant, S., Vinet, A., Krauss, D., Maupetit, D., & Renard, G. (2019). Genetic parameters and genome-wide association study regarding feed efficiency and slaughter traits in Charolais cows. *Journal of Animal Science*, 97:3864-3698. doi:https://doi.org/10.1093/jas/skz240
- Mayer, K. (1997). Estimates of genetic parameters for weaning weight of beef cattle accounting for weaning weight of beef cattle accounting for direct-maternal enviromental covariance. *Livestock Production Science*, 52:187-199.
- McClure, M. C., Morsci, N. S., Schnabel, R. D., Kim, J. W., Yao, P., Rolf, M. M., . . . Taylor, J. F. (2010). A genome scan for quantitative trait loci influencing

carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Animal Genetics*, 41:597-607. doi: <https://doi.org/10.1111/j.1365-2052.2010.02063.x>

Melo, T. P., Fortes, M. S., Fernandes, G., Albuquerque, L. G., & Carneiro, R. (2019). Multi-breed validation study unravelled genomic regions associated with puberty traits segregating across tropically adapted breeds. *Journal of Animal Genetics*, 97(7):3027-3033. doi:<https://doi.org/10.1093/jas/skz121>

Meuwissen, T. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, 157(4):1819-1829.

Meuwissen, T., Hayes, B. J., & Goddard, M. E. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers*, 6:614. doi:<https://doi.org/10.2527/af.2016-0002>

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, B., & Lee, D. H. (19-23 de August de 2002). BLUPF90 AND RELATED PROGRAMS(BGF90). *In Proceeding of the 7th World congress on Genetics Applied to Livestocks Production*.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, B., & Lee, D. H. (19-23 de August de 2002). BLUPF90 AND RELATED PROGRAMS (BGF90). *In Proceeding of the 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002, Montpellier, France*.

Nayeri, S., Sargolzaei, M., Abo-Ismael, M. K., May, N., Miller, S. P., Schenkel, F., & Stohard, P. (2016). Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genetics*, 17:75. doi:<https://doi.org/10.1186/s12863-016-0386-1>

Nkrumah, J. D., Sherman, E. L., Li, C., Marques, E., Crews, D. H., Bartusiak, R., . . . Moore, S. S. (2007). Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *Journal of Animal Science*, 85:3 170-3181. doi:<https://doi.org/10.2527/jas.2007-0234>

- PATTERSON, H. D., & THOMPSON, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545-554. doi:<https://doi.org/10.1093/biomet/58.3.545>
- Pegolo, S., Cecchinato, A., Savoia, S., Di Stasio, S., Pauciullo, A., Brugiapaglia, A., . . . Albera, G. (2020). Genome-wide association and pathway analysis of carcass and meat quality traits in Piemontese young bulls. *Animal*, 14:243-252. doi:<https://doi.org/10.1017/S1751731119001812>
- Peng, M., Han, J., Li, L., & Ma, H. (2016). Suppression of fat deposition in broiler chickens by (-)-hydroxycitric acid supplementation: A proteomics perspective. *Scientific Reports*, 6:32580. doi:<https://doi.org/10.1038/srep32580>
- Pulfield, D. C., Evans, R. D., & Berry, D. P. (2019). Reaffirmation of known major genes and the identification of novel candidate genes associated with carcass-related metrics based on whole genome sequence within a large multi-breed cattle population. *BMC Genomics*, 20:720. doi:<https://doi.org/10.1186/s12864-019-6071-9>
- Raphaka, K., Matika, O., Sánchez-Molano, E., Mrode, R., Coffey, M. P., Riggio, V., . . . Banos, G. (2017). Genomic regions underlying susceptibility to bovine tuberculosis in Holstein-Friesian cattle. *BMC Genetics*, 18:27.
- Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., . . . Melchinger, A. E. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *NATURE GENETICS*, 44:217-220. doi:<https://doi.org/10.1038/ng.1033>
- Robinson, D. L. (1996). Estimation and interpretation of direct and maternal genetic parameters for weights of Australian Angus cattle. *Livestock Production Science*, 45:1-11. doi:[https://doi.org/10.1016/0301-6226\(95\)00083-6](https://doi.org/10.1016/0301-6226(95)00083-6)
- Rosen, E. D., & MacDougald, O. A. (2006). Adipocyte Differentiation From the Inside Out. *Nature Reviews Molecular Cell Biology*, 7:885-896. doi:DOI: 10.1038 / nrm2066

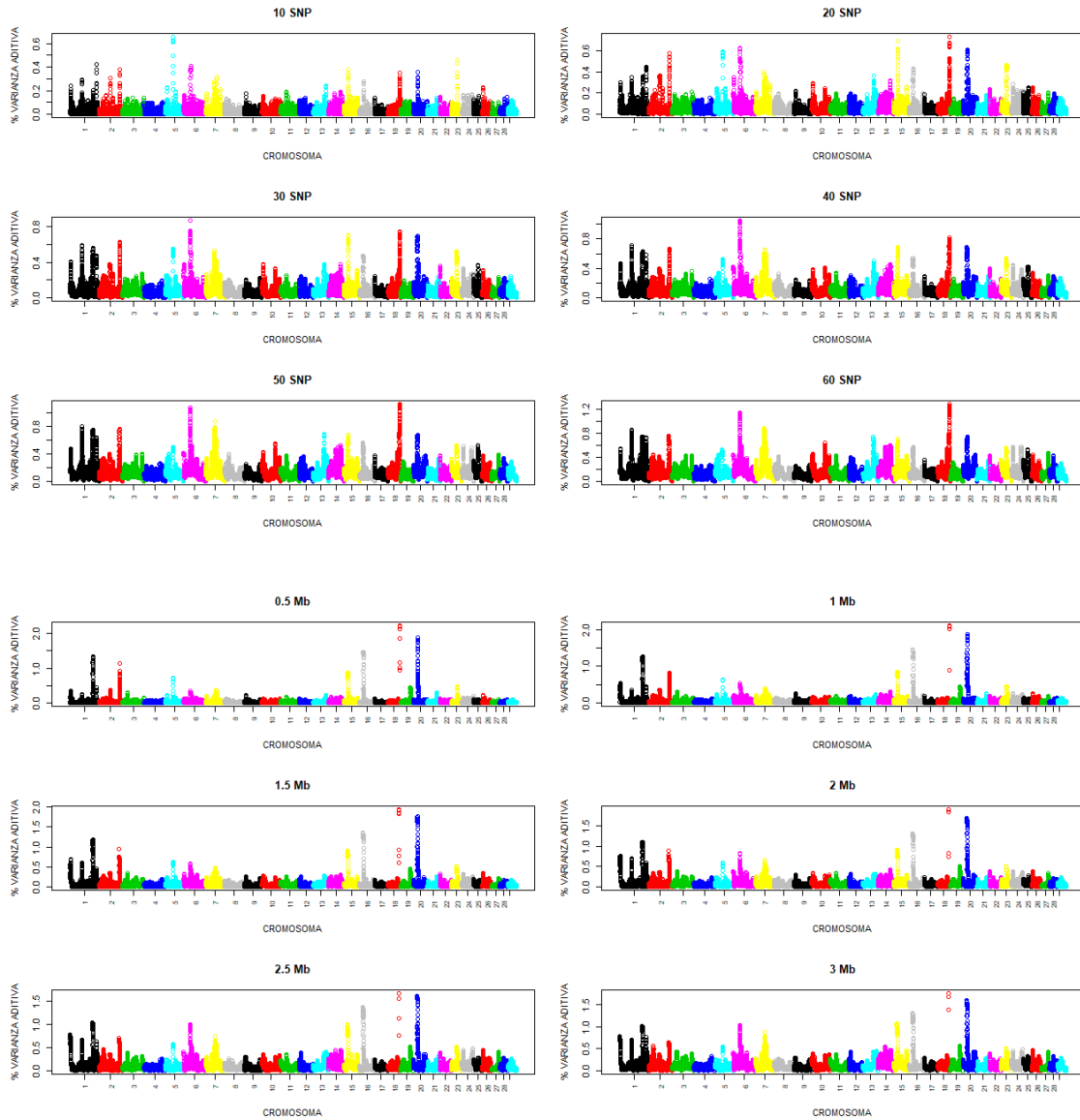
- Rosen, E. D., Hsu, C. H., Wang, X., Sakai, S., Freeman, M. W., Gonzalez, F. J., & Spiegelman, B. M. (2002). C/EBPalpha Induces Adipogenesis Through PPARgamma: A Unified Pathway. *Genes & Development*, 16(1):22-26. doi: 10.1101/gad.948702
- Saatchi, M., Beever,, J. E., Decker, J. E., Faulkner, D. B., Freetly, H. C., Hansen, S. L., . . . Taylor, J. F. (2014). QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. *BMC Genomics*, 15:1004.
- Saatchi, M., Schnabel, R. D., Taylor, J. F., & Garrick, D. J. (2014). Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics*, 15:442. doi:https://doi.org/10.1186/1471-2164-15-442
- Satih, S., Chalabi, N., Rabiau, N., Bosviel, R., Fontana , L., Bignon, Y. J., & Bernard-Gallon, D. J. (2010). Gene Expression Profiling of Breast Cancer Cell Lines in Response to Soy Isoflavones Using a Pangenomic Microarray Approach. *OMICS: Journal of integrative Biology*, 14(3):231-238. doi:10.1089/omi.2009.0124
- Savoia, S., Albera, A., Brugiapaglia, A., Di Stasio, L., Cecchinatto, A., & Brittante, G. (2019). Heritability and genetic correlations of carcass and meat quality traits in Piemontese young bulls. *Meat Science*, 156:111-117. doi:https://doi.org/10.1016/j.meatsci.2019.05.024
- Sax, K. (1923). The association of Size Differences with Seed-Coat Pattern and Pigmentation in *Phaseolus Vulgaris*. *Genetics*, 8:552-560.
- Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., . . . Taylor, J. F. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics*, 18:386. doi:https://doi.org/10.1186/s12864-017-3754-y
- Simm, G. (1998). *Genetics improvement of cattle and sheep*. Ipswich, United Kingdom: Farming press.
- Srivastava, S., López, B., Heras-Saldana, S. L., Park, J. E., Shin, D. H., Chai, H. H., . . . Lim, D. (2019). Estimation of Genetic Parameters by Single-Trait

- and Multi-Trait Models for Carcass Traits in Hanwoo Cattle. *Animals*, 9:12. doi:<https://doi.org/10.3390/ani9121061>
- Taneera, J., Lang, S., Sharma, A., Fadista, J., Zhou, Y., Ahlqvist, E., . . . Groop, L. (2012). A Systems Genetics Approach Identifies Genes and Pathways for Type 2 Diabetes in Human Islets. *Cell Metabolism*, 16(1):122-134. doi:<https://doi.org/10.1016/j.cmet.2012.06.006>
- The Bovine Genome Sequencing and Analysis Consorti, Elvik, C. G., Tellam, R. L., & Worley, K. C. (2009). The Genome Sequence of Taurine Cattle: A Window to ruminant biology and evolution . *Science* , 324:522-528.
- Tiezzi, F., & Maltecca , C. (2015). Accounting for trait architecture in genomic predictions of US Holstein cattle using a weighted realized relationship matrix. *Genetics Selection Evolution*, 47:24. doi:<https://doi.org/10.1186/s12711-015-0100-1>
- Tiezzi, F., Arceo, M. E., Cole, J. B., & Maltecca, C. (2018). Including gene networks to predict calving difficulty in Holstein, Brown Swiss and Jersey cattle. *BMC Genetics*, 19(1):20. doi:<https://doi.org/10.1186/s12863-018-0606-y>
- Valour, S. A., Degrelle, A. A., Ponter, C., Giraud-Delville, C., Campion, E., Guayader-Joly, C., . . . Grimard, B. (2014). Energy and lipid metabolism gene expression of D18 embryos in dairy cows is related to dam physiological status. *Physiological Genomics*, 46:39-56. doi:<https://doi.org/10.1152/physiolgenomics.00091.2013>
- VanRaden , P. M., Van Tassell, C. P., Wiggans , G. R., Sonstegard, T. S., Schanabel, R. D., & Taylor, J. F. (2012). Invited Review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science*, 92:16-24. doi:<https://doi.org/10.3168/jds.2008-1514>
- VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11):4414-4423. doi:<https://doi.org/10.3168/jds.2007-0980>
- Varona, L., Moreno, C., & Altarriba, J. (2009). A model with heterogeneous thresholds for subjective traits: Fat cover and conformation score in the

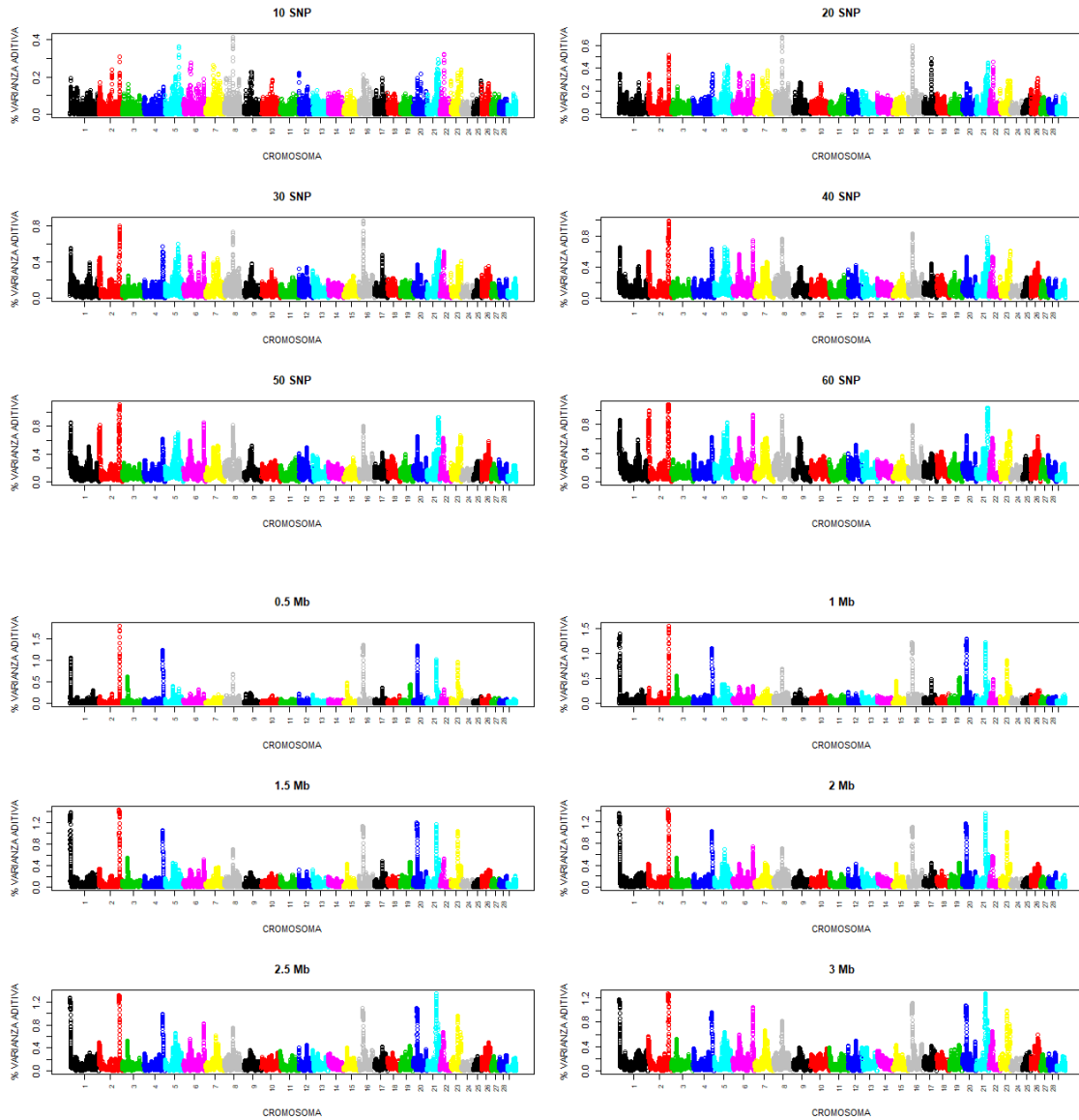
- Pirenaica beef cattle. *Journal of Animal science*, 87(4):1210-1217.
doi:<https://doi.org/10.2527/jas.2008-1361>
- Varona, L., Munilla, S., Casellas, J., Moreno, C., & Altarriba, J. (2015). Consequences of paternally inherited effects on the genetic evaluation of maternal effects on the genetics evaluation of maternal effects. *Genetics Selection Evolution*, 47:63. doi:<https://doi.org/10.1186/s12711-015-0141-5>
- Wang, H., Misztal, I., Legarra, A., & Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, 94(2):73-83. doi:<https://doi.org/10.1017/S0016672312000274>
- Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., . . . Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics*, 21:38. doi:<https://doi.org/10.1186/s12864-019-6273-1>
- Yan, Z., Wang, Z., Zhang, Q., Yue, S., Jiang, Y., & Shi, K. (2019). Identification of whole-genome significant single nucleotide polymorphisms in candidate genes associated with body conformation traits in Chinese Holstein cattle. *Animal Genetics*, 41:141-142. doi:10.1111/age.12865
- Zhang, F., Wang, Y., Mubiibi, R., Chen, L., Vinsky, M., Plastow, G., . . . Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. *BMC Genomics*, 21:36. doi:<https://doi.org/10.1186/s12864-019-6362-1>
- Zhang, Q., Guldbrechtsen, B., Thomasen, J. R., Lund, M. S., & Shana, G. (2016). Genome-wide association study for longevity with Whole-genome sequencing in 3 cattle breeds. *Journal of Dairy Science*, 99(9):7289-7298. doi:<https://doi.org/10.3168/jds.2015-10697>

8. APÉNDICE I.

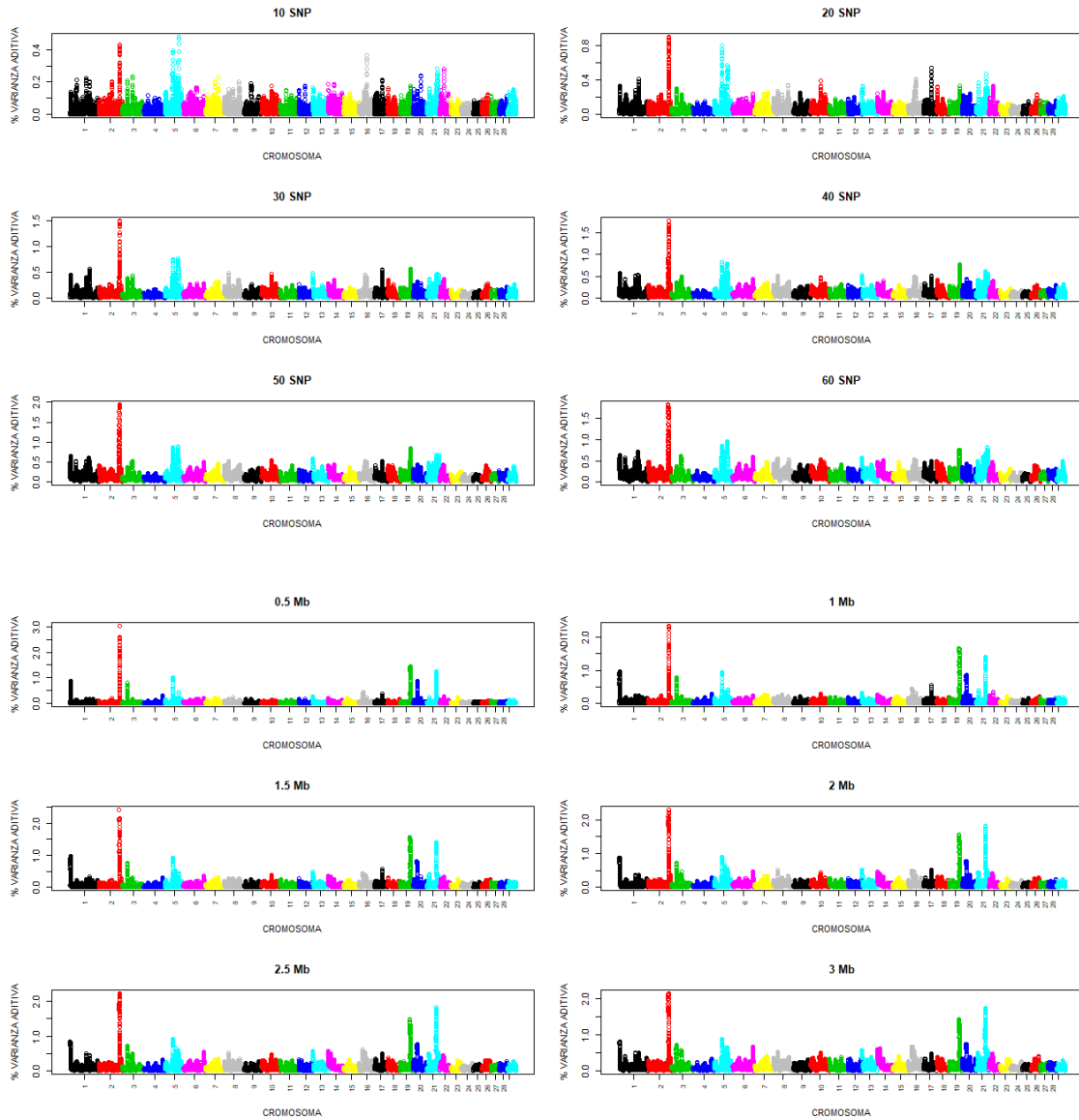
Apéndice1.1 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para Peso al Nacimiento (BW).



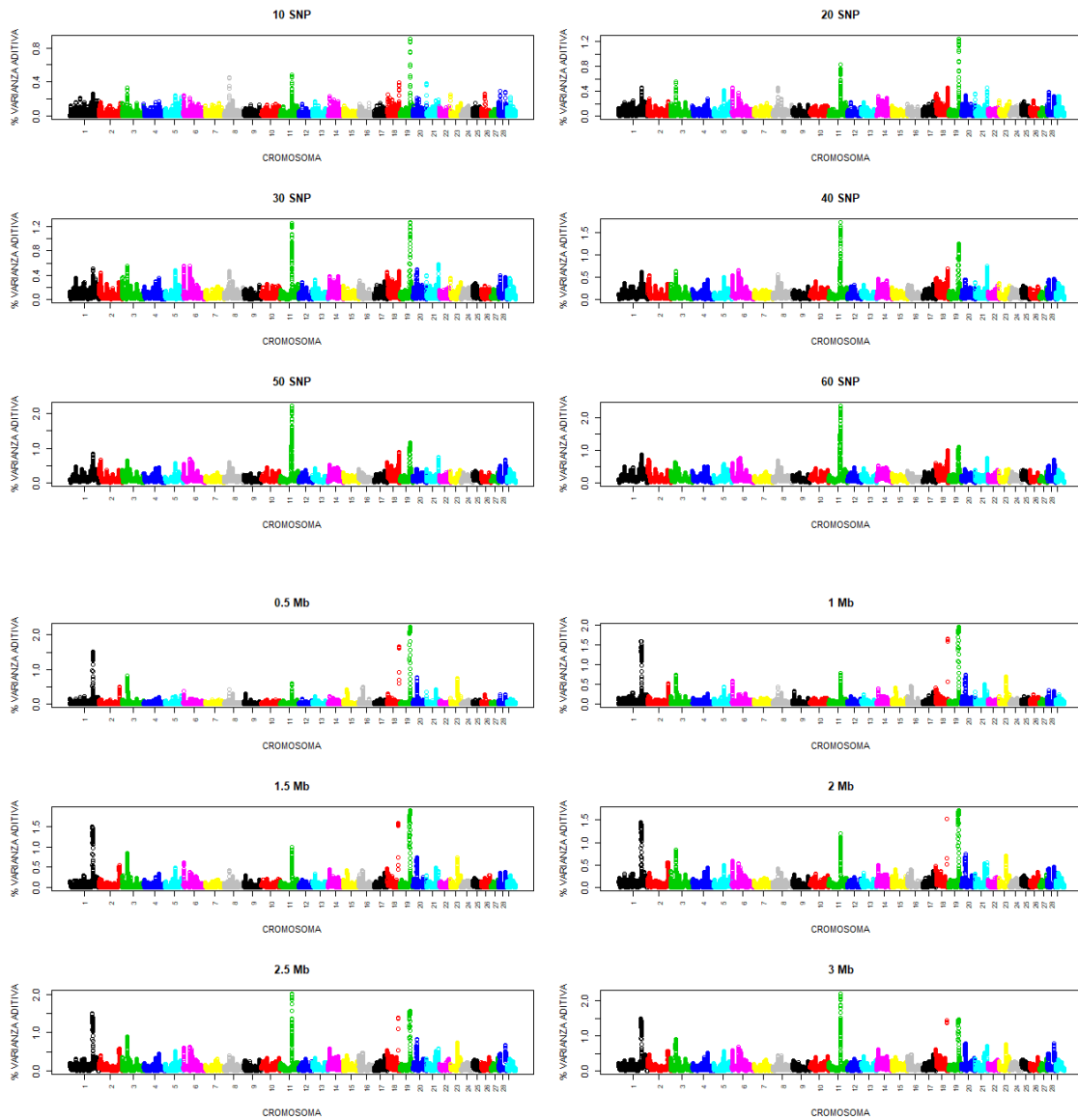
Apéndice 1.2 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para Peso a los 210 días Efector Directo (WWD).



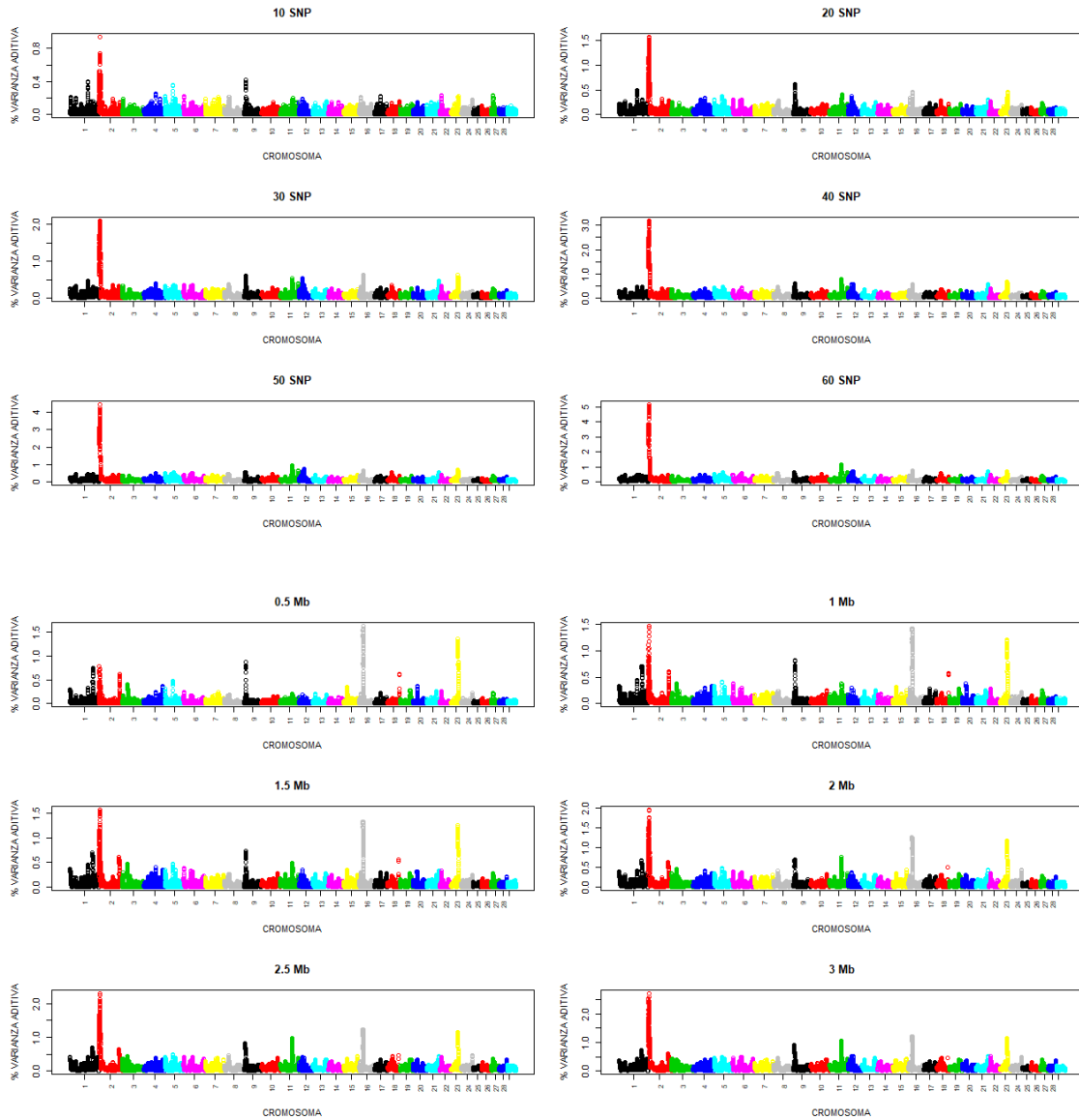
Apéndice 1.3 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para Peso a los 210 días Efector Materno (WWM).



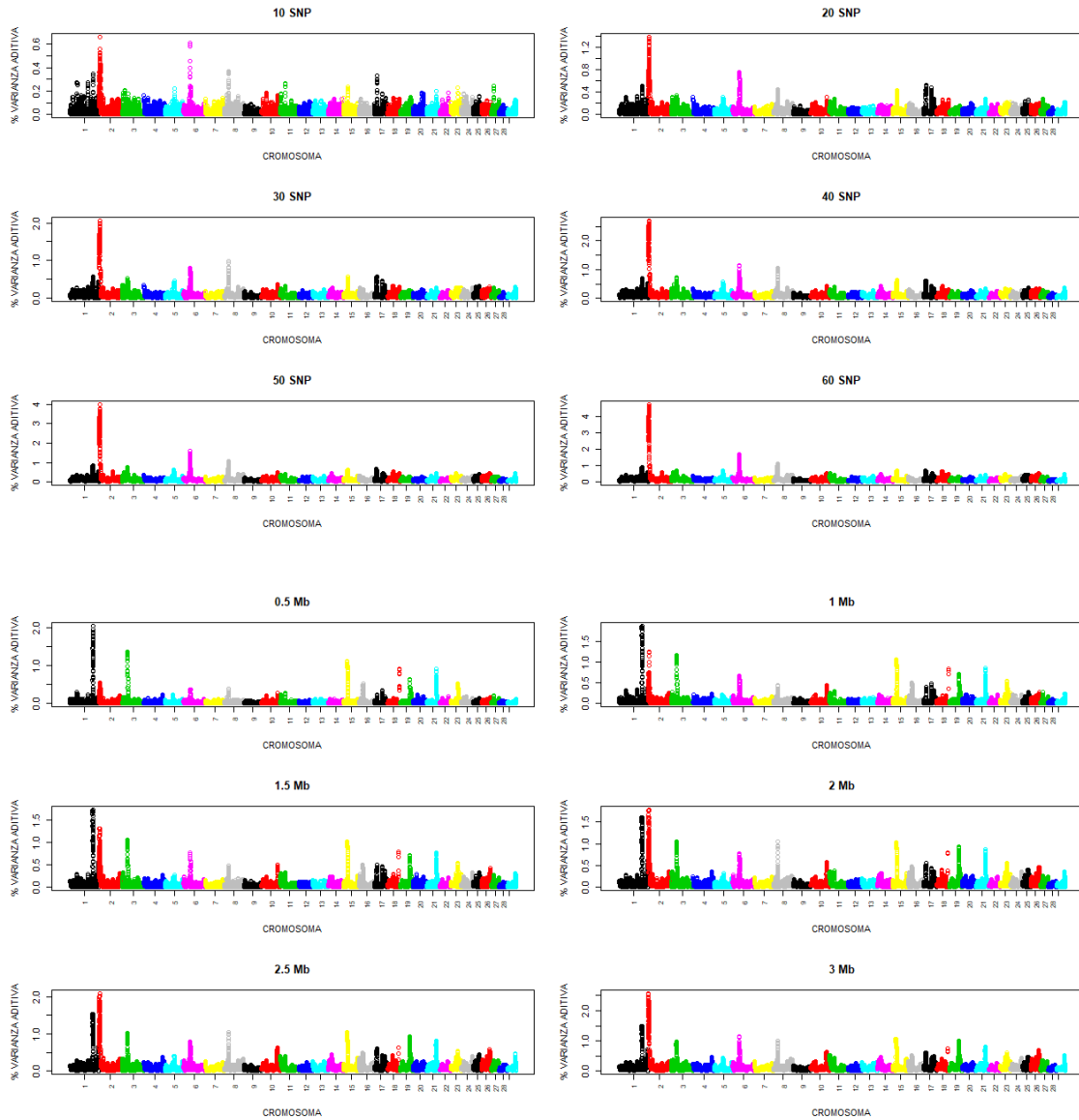
Apéndice 1.4 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para Peso de la Canal Fría (CCW).



Apéndice 1.5 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para la Conformación (CON).

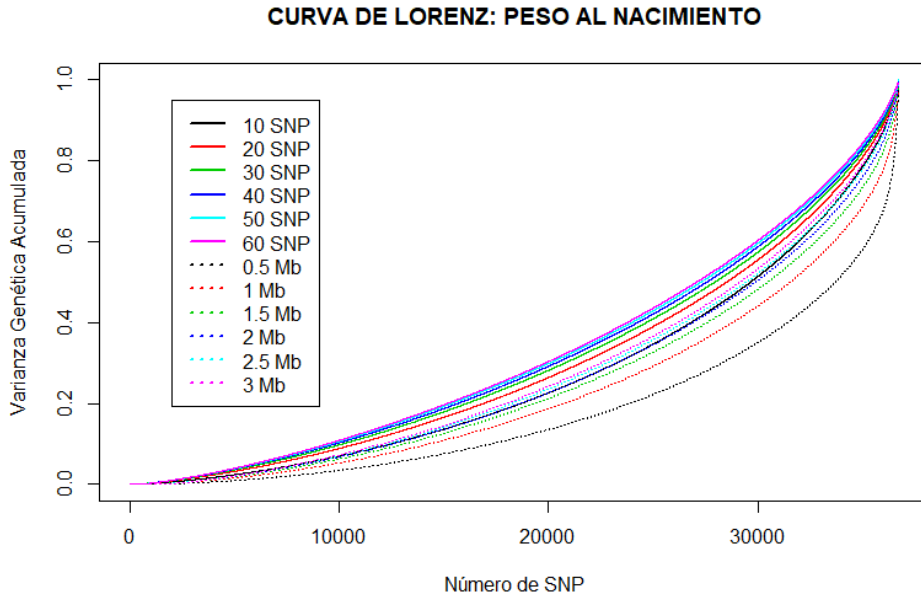


Apéndice 1.6 Barridos Genómicos de los resultados de la varianza genética explicada por las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) explicada en torno a cada SNP para el Engrasamiento.

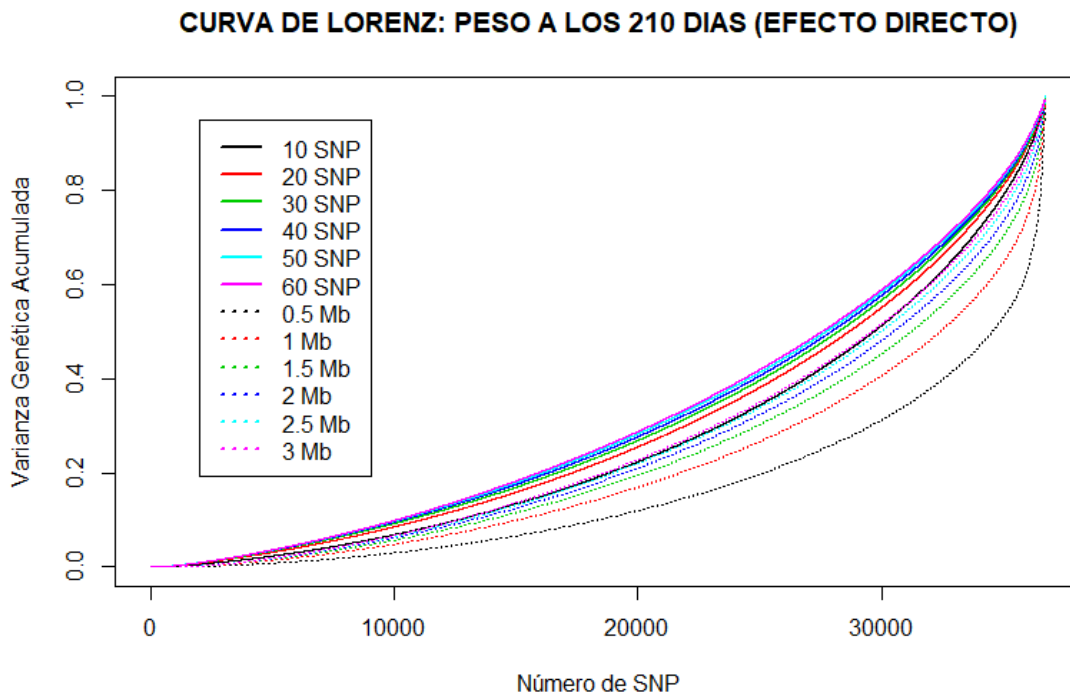


9. APÉNDICE II.

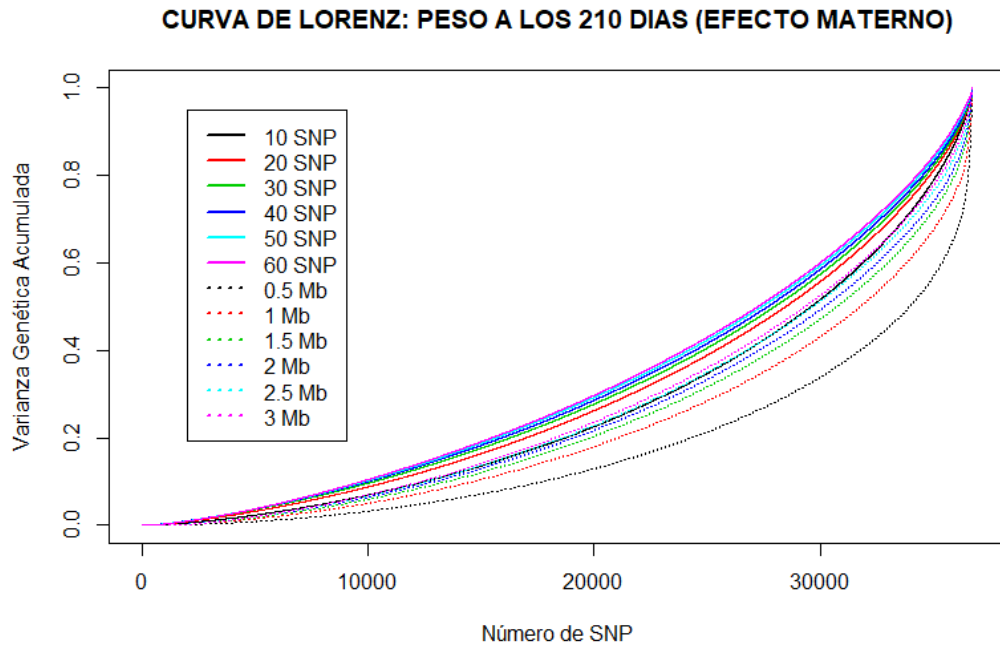
Apéndice 2.1 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso al Nacimiento.



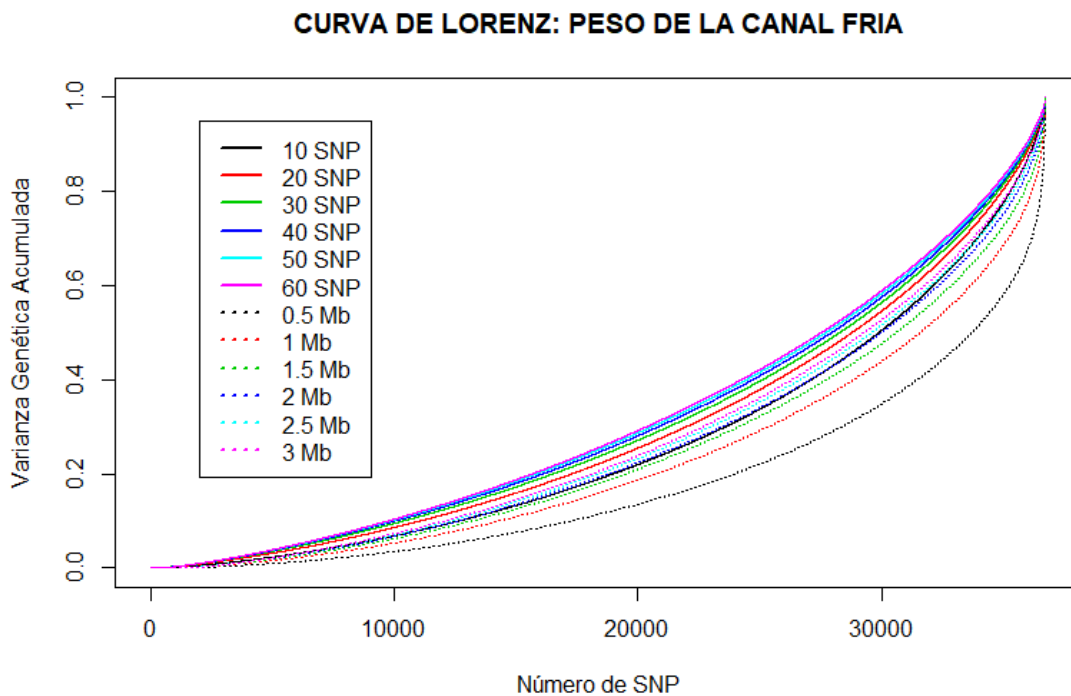
Apéndice 2.2 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso a los 210 días Efecto Directo.



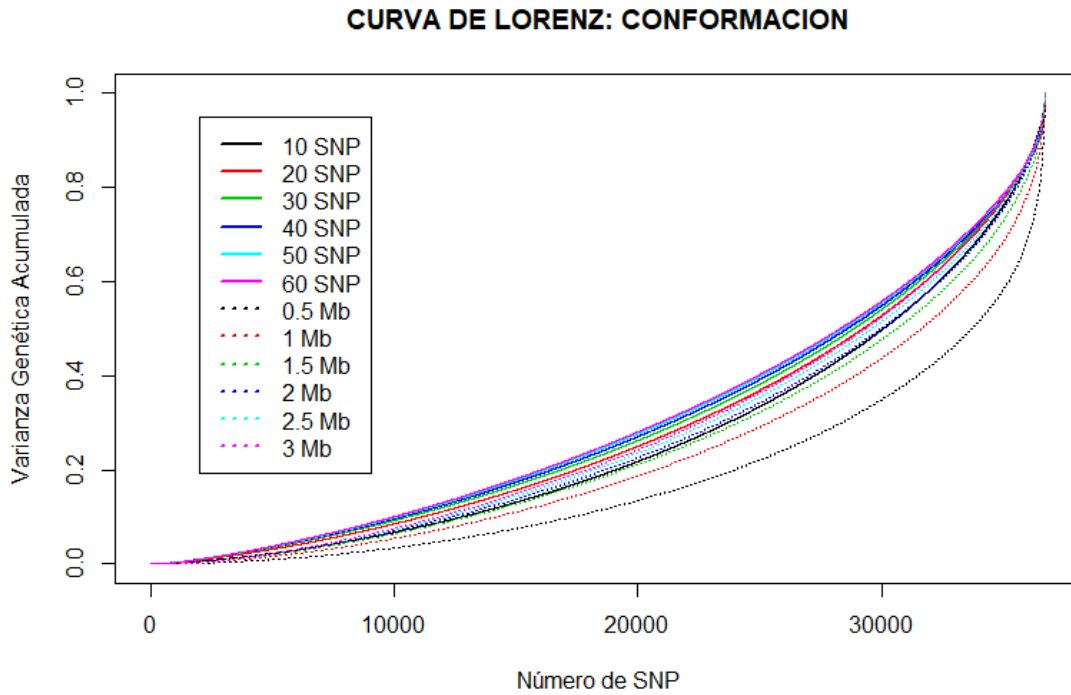
Apéndice 2.3 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso a los 210 días Efecto Materno.



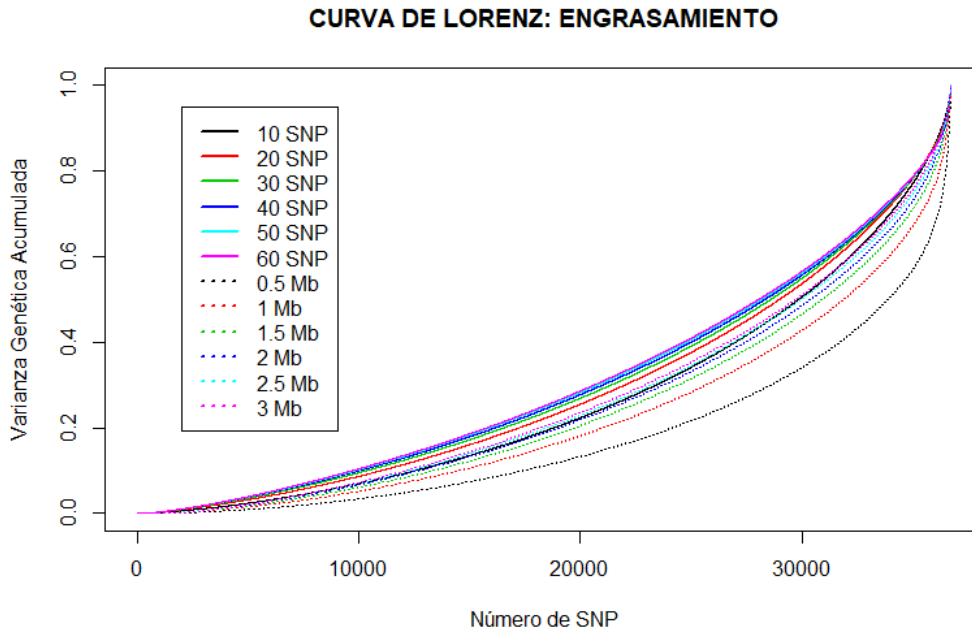
Apéndice 2.4 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Peso de la Canal Fría.



Apéndice 2.5 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Conformación.



Apéndice 2.6 Curva de Lorenz entre la varianza genética explicada y el número de SNP para Engrasamiento.



Apéndice 2.7 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y para Peso al Nacimiento

	Gini Index	10%	20%	50%
10 SNP	0.461	1.42%	4.21%	19.28%
20 SNP	0.406	1.71%	5.05%	22.41%
30 SNP	0.381	1.94%	5.59%	23.95%
40 SNP	0.365	2.17%	6.00%	25.00%
50 SNP	0.354	2.39%	6.33%	25.74%
60 SNP	0.346	2.55%	6.61%	26.30%
0.5 Mb	0.638	0.23%	0.64%	8.29%
1 Mb	0.542	0.44%	1.56%	14.13%
1.5 Mb	0.495	0.72%	2.59%	17.07%
2 Mb	0.470	1.00%	3.33%	18.68%
2.5 Mb	0.453	1.24%	3.95%	19.85%
3 Mb	0.441	1.45%	4.40%	20.70%

Apéndice 2.8 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y para Peso a los 210 días (Efecto Directo).

	Gini Index	10%	20%	50%
10 SNP	0.462	1.57%	4.38%	19.22%
20 SNP	0.414	1.86%	5.13%	21.99%
30 SNP	0.395	2.00%	5.51%	23.19%
40 SNP	0.383	2.12%	5.74%	23.96%
50 SNP	0.375	2.21%	5.96%	24.55%
60 SNP	0.368	2.31%	6.17%	25.00%
0.5 Mb	0.674	0.23%	0.53%	5.94%
1 Mb	0.575	0.41%	1.17%	11.79%
1.5 Mb	0.526	0.62%	2.00%	15.03%
2 Mb	0.497	0.82%	2.68%	17.01%
2.5 Mb	0.477	1.08%	3.38%	18.43%
3 Mb	0.463	1.29%	3.88%	19.46%

Apéndice 2.9 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y Peso a los 210 días (Efecto Materno).

	Gini Index	10%	20%	50%
10 SNP	0.460	1.50%	4.30%	19.42%
20 SNP	0.407	1.73%	5.13%	22.50%
30 SNP	0.385	1.91%	5.61%	23.80%
40 SNP	0.371	2.13%	5.98%	24.67%
50 SNP	0.361	2.37%	6.35%	25.38%
60 SNP	0.353	2.59%	6.66%	25.84%
0.5 Mb	0.651	0.17%	0.51%	7.55%
1 Mb	0.553	0.32%	1.24%	13.50%
1.5 Mb	0.510	0.52%	2.14%	16.33%
2 Mb	0.486	0.67%	2.64%	17.83%
2.5 Mb	0.466	0.92%	3.36%	19.23%
3 Mb	0.453	1.16%	3.86%	20.16%

Apéndice 2.10 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y Peso de la Canal Fría.

	Gini Index	10%	20%	50%
10 SNP	0.471	1.34%	3.96%	18.65%
20 SNP	0.418	1.64%	4.88%	21.66%
30 SNP	0.394	1.83%	5.35%	23.04%
40 SNP	0.380	1.98%	5.67%	23.95%
50 SNP	0.371	2.09%	5.91%	24.59%
60 SNP	0.363	2.20%	6.14%	25.09%
0.5 Mb	0.639	0.14%	0.51%	8.25%
1 Mb	0.543	0.32%	1.53%	14.05%
1.5 Mb	0.501	0.56%	2.39%	16.69%
2 Mb	0.475	0.81%	3.13%	18.38%
2.5 Mb	0.460	1.01%	3.61%	19.48%
3 Mb	0.448	1.22%	4.04%	20.29%

Apéndice 2.11 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y la Conformación.

	Gini Index	10%	20%	50%
10 SNP	0.478	1.06%	3.46%	18.12%
20 SNP	0.434	1.00%	3.78%	20.41%
30 SNP	0.417	0.90%	3.89%	21.42%
40 SNP	0.406	0.83%	3.99%	22.15%
50 SNP	0.399	0.76%	4.04%	22.60%
60 SNP	0.394	0.71%	4.10%	22.96%
0.5 Mb	0.639	0.22%	0.69%	8.21%
1 Mb	0.542	0.47%	1.73%	13.90%
1.5 Mb	0.498	0.68%	2.58%	16.65%
2 Mb	0.472	0.86%	3.20%	18.38%
2.5 Mb	0.457	0.94%	3.60%	19.44%
3 Mb	0.446	1.01%	3.93%	20.17%

Apéndice 2.12 Coeficiente de Desigualdad de Gini y porcentaje de marcadores SNP que explican el 10, 20 y 50% de la varianza aditiva con las 12 definiciones (10, 20, 30, 40, 50, 60 SNP o 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 Mb) de las regiones genómicas y el Engrasamiento.

	Gini Index	10%	20%	50%
10 SNP	0.468	1.16%	3.74 %	18.71%
20 SNP	0.425	1.03%	4.03 %	21.09%
30 SNP	0.407	0.90%	4.01%	22.17%
40 SNP	0.396	0.82%	4.06%	22.88%
50 SNP	0.390	0.78%	4.08%	23.29%
60 SNP	0.385	0.76%	4.12 %	23.65%
0.5 Mb	0.645	0.22%	0.65%	7.77%
1 Mb	0.553	0.43%	1.45%	13.20%
1.5 Mb	0.511	0.64%	2.12%	15.85%
2 Mb	0.489	0.79%	2.59%	17.26%
2.5 Mb	0.471	0.90%	2.97%	18.43%
3 Mb	0.461	0.99%	3.27%	19.12%