

Estimación de la heredabilidad usando información de pedigree y genómica en caracteres de composición de carcasa y calidad de carne

*Tesis presentada para optar al título de Magister de la Universidad de Buenos Aires,  
Área Biometría y Mejoramiento*

**Belcy Karine Angarita Barajas**  
Zootecnista - Universidad Industrial de Santander - 2011

Lugar de trabajo: Cátedra Mejoramiento genético animal. Facultad de Agronomía.  
Universidad de Buenos Aires



Escuela para Graduados Ing. Agr. Alberto Soriano  
Facultad de Agronomía – Universidad de Buenos Aires

## COMITÉ CONSEJERO

Director de Tesis

**Natalia Soledad Forneris**

Ing. Agr. (Universidad de Buenos Aires)

Doctora (Universidad de Buenos Aires)

Co-director de Tesis

**Rodolfo Juan Carlos Cantet**

Ing. Agr. (Universidad de Buenos Aires)

MSc (Montana State University)

MSc (University of Illinois)

PhD (Iowa State University)

Consejero de Estudios

**Juan Pedro Steibel**

Ing. Agr. (Universidad Nacional de la Pampa)

MSc (Universidad de Buenos Aires)

PhD (Michigan State University)

## JURADO DE TESIS

JURADO

Daniel Omar Maizón

Med. Vet. (Universidad de Buenos Aires, Argentina)

MSc. (Universidad de Buenos Aires, Argentina)

Ph.D. (Cornell University, E.E.U.U)

JURADO

**Luis Varona Aguado**

Lic. Vet. (Universidad de Zaragoza, España)

MSc. (Universidad de Zaragoza, España)

Doctor (Universidad de Zaragoza, España)

JURADO

**Fernando Sebastián Baldi Rey**

Ing. Agr. (Universidad de la República, Uruguay)

MSc. (Universidade Estadual Paulista, Brasil)

Doutor (Universidade Estadual Paulista, Brasil)

Fecha de defensa de la tesis: 29 de junio de 2021

*A mis padres: Orlando Angarita y Blanca Barajas, quienes me han dado ejemplo de trabajar con pasión, entrega y perseverancia.*

*A mis hermanos: Newman, Ronald y Angello, quienes cuidan y confortan mis pasos.*

## AGRADECIMIENTOS

*Agradezco a Dios y a la Virgen María por fortalecer mi espíritu.*

*A mi comité consejero: Dr. Forneris, Dr. Cantet y Dr. Steibel. Les agradezco por ser mis maestros y trasmitirme su sabiduría, por creer y confiar en mí, por el tiempo que han dedicado orientándome y enseñándome en mi proceso de formación, por darme libertad para avanzar paso a paso.*

*Nati, gracias por guiarme, motivarme y aconsejarme, su incondicional apoyo me ayudó a vencer temores, a trabajar perseverantemente y dar lo mejor de mí, delineando mi propio ritmo de aprendizaje.*

*Dr. Cantet, mi profundo agradecimiento por creer en mí y permitirme formar parte del grupo MGA. Gracias por darme la oportunidad de recorrer nuevos caminos, que me han mostrado mis fortalezas ocultas, dejando huella y transformándome como persona y como profesional.*

*Dr. Juan Pedro, gracias por inspirarme y motivarme a ser mejor, por desafiarme continuamente para vencer las barreras de mis propios temores. Gracias por darme la oportunidad de aprender metódicamente.*

*A los jurados, quienes dedicaron su tiempo en la lectura de esta tesis.*

*A mis amigos de la Cátedra de Mejoramiento Genético Animal de la FAUBA: Caro, Dani, Matías, Esteban, Gastón, Sebas, Andrés, Martín, por su apoyo incondicional y gratos momentos.*

*A los miembros, profesores y compañeros de la Escuela para Graduados, quienes contribuyeron en mi proceso de formación, y me hicieron pasar gratas experiencias.*

*A mi familia por su amor incondicional y apoyo constante, reconfortándome a conseguir y perseverar para lograr mis objetivos.*

*A mis amigos Astrid Villamil y Jorge Herrera, por alentarme y estimularme en todo momento.*

*A Hercilia Rey, por apoyarme para que diera mi primer paso en el inicio de esta meta.*

*A todas aquellas personas que me han apoyado.*

*Declaro que el material incluido en esta tesis es, a mi mejor saber y entender, original producto de mi propio trabajo (salvo en la medida en que se identifique explícitamente las contribuciones de otros), y que este material no lo he presentado, en forma parcial o total, como una tesis en esta u otra institución.*

Belcy Karine Angarita Barajas

## ÍNDICE GENERAL

|  | <b>página</b> |
|--|---------------|
| DEDICATORIA.....   | iii           |
| AGRADECIMIENTOS.....   | iv            |
| DECLARACIÓN.....   | v             |
| ÍNDICE GENERAL.....  | vi            |
| ÍNDICE DE TABLAS.....  | viii          |
| ÍNDICE DE FIGURAS.....   | ix            |
| ABREVIATURAS.....  | x             |
| RESUMEN.....   | xi            |
| ABSTRACT.....  | xii           |
| <br>   |               |
| <b>CAPÍTULO 1: <i>Introducción</i></b> .....   | <b>1</b>      |
| <br>   |               |
| <b>CAPÍTULO 2: <i>Materiales y métodos</i></b> .....                                     | <b>5</b>      |
| 2.1. Animales.....   | 6             |
| 2.2. Fenotipos.....  | 6             |
| 2.3. Genotipado y edición de datos genómicos.....  | 6             |
| 2.4. Estimación de las matrices de relaciones de parentesco.....                         | 7             |
| 2.4.1. Matriz de relaciones aditivas.....  | 8             |
| 2.4.2. Matriz de relaciones genómicas IBS.....   | 8             |
| 2.4.3. Matriz de relaciones genómicas IBD.....   | 8             |
| 2.5. Estimación de los componentes de varianza y heredabilidad de los<br>caracteres..... | 9             |
| 2.5.1. Modelo estadístico.....   | 9             |
| 2.5.2. Estimación de la heredabilidad.....   | 10            |
| 2.5.3. Test de hipótesis para la varianza genética aditiva de los modelos.....           | 10            |
| 2.6. Estimación de la habilidad predictiva.....  | 10            |
| 2.6.1. Validación cruzada de 5 iteraciones ( <i>5-fold cross-validation</i> ).....       | 11            |
| 2.6.2. Test de comparación de medias de las habilidades predictivas.....                 | 11            |

|   |    |
|---|----|
| <b>CAPÍTULO 3: <i>Resultados</i></b> .....  | 12 |
| 3.1. Análisis descriptivo de las matrices de relaciones de parentesco estimadas.....                                | 13 |
| 3.2. Componentes de varianza estimados en caracteres de composición de carcasa y calidad de la carne de cerdos..... | 16 |
| 3.3. Heredabilidad estimada en caracteres de composición de carcasa y calidad de la carne de cerdos.....            | 18 |
| 3.4. Habilidad predictiva de los modelos.....   | 20 |
| <br>  |    |
| <b>CAPÍTULO 4: <i>Discusión</i></b> .....   | 22 |
| 4.1. Análisis descriptivo de las matrices de relaciones de parentesco estimadas.....                                | 23 |
| 4.2. Componentes de varianza estimados en caracteres de composición de carcasa y calidad de la carne de cerdos..... | 24 |
| 4.3. Heredabilidad estimada en caracteres de composición de carcasa y calidad de la carne de cerdos.....            | 26 |
| 4.4. Habilidad predictiva de los modelos.....   | 27 |
| <br>  |    |
| <b>CAPÍTULO 5: <i>Conclusiones</i></b> .....  | 30 |
| <br>  |    |
| <b>BIBLIOGRAFÍA</b> .....   | 32 |

## ÍNDICE DE TABLAS

| <b>TABLA</b>  | <b>página</b> |
|---|---------------|
| 2.1. Número de observaciones, media fenotípica y desvío estándar de los fenotipos analizados.....   | 7             |
| 3.1. Distribución de los elementos diagonales y fuera de la diagonal de las matrices de relaciones de parentesco estimadas.....   | 13            |
| 3.2. Coeficientes de regresión (arriba de la diagonal), varianza (diagonal), correlaciones (debajo de la diagonal) de los elementos fuera de la diagonal de las matrices de relaciones de parentesco estimadas..... | 15            |
| 3.3. Estimadores de componentes de varianza usando tres modelos lineales mixtos en caracteres de composición de carcasa y calidad de la carne de cerdos (Error estándar).....                                       | 17            |
| 3.4. Heredabilidad estimada usando tres modelos lineales mixtos en caracteres de composición de carcasa y calidad de la carne de cerdos (Error estándar).....   | 19            |
| 3.5. Habilidad predictiva ( $\hat{r}_{y,\hat{y}}$ ) de los diferentes modelos estimada por validación cruzada (Error estándar).....   | 21            |



## ÍNDICE DE FIGURAS

| <b>FIGURA</b>   | <b>página</b> |
|---|---------------|
| 3.1. Autovalores de las matrices de relaciones de parentesco.....                       | 14            |
| 3.2. Regresión de los elementos fuera de la diagonal de las matrices de parentesco..... | 15            |

## ABREVIATURAS

|              |  |
|--------------|--|
| <b>ADN</b>   | Ácido desoxirribonucleico  |
| <b>BLUP</b>  | Predicción lineal insesgada de mínima varianza (del inglés <i>Best Linear Unbiased Prediction</i> )                    |
| <b>BV</b>    | Valor de Cría (del inglés <i>Breeding value</i> )  |
| <b>GBLUP</b> | Metodología BLUP de predicción del valor de cría genómico (del inglés, <i>genomic best linear unbiased predictor</i> ) |
| <b>HMM</b>   | Modelos Ocultos de Markov (del inglés <i>Hidden markov models</i> )  |
| <b>IBD</b>   | Identidad por Descendencia (del inglés <i>Identity by descent</i> )  |
| <b>IBDP</b>  | proporción del genoma idéntico por descendencia  |
| <b>IBS</b>   | Identidad por estado (del inglés <i>Identity by state</i> )  |
| <b>LD</b>    | Desequilibrio de Ligamiento (del inglés <i>Linkage disequilibrium</i> )  |
| <b>LE</b>    | Equilibrio de ligamiento (del inglés, <i>Linkage equilibrium</i> )   |
| <b>LRT</b>   | Test del cociente de verosimilitud (del inglés, <i>Likelihood ratio test</i> )   |
| <b>MAF</b>   | Alelo en menor proporción (del inglés, <i>Minor allele frequency</i> )   |
| <b>QTL</b>   | Locus de un carácter cuantitativo (del inglés <i>quantitative trait locus</i> )  |
| <b>REML</b>  | Máxima verosimilitud residual (del inglés, <i>Restricted maximum likelihood</i> )                                      |
| <b>SD</b>    | Desvío Estándar  |
| <b>SG</b>    | Selección genómica   |
| <b>SNP</b>   | Polimorfismos genéticos de un solo nucleótido (del inglés <i>Single nucleotide polymorphysm</i> )                      |

## RESUMEN

### *Estimación de la heredabilidad usando información de pedigree y genómica en caracteres de composición de carcasa y calidad de carne*

La información de un panel de marcadores moleculares, condicional o no en la información del pedigree, permite estimar la proporción realizada del genoma compartido idéntico por descendencia (IBD) entre un par de individuos. Dicha medida ha demostrado refinar los parentescos y aumentar la exactitud de la predicción de los valores de cría, respecto del modelo tradicional de pedigree. Asimismo, los parentescos genómicos pueden reemplazar los de pedigree en la estimación de la heredabilidad ( $\hat{h}^2$ ) y los componentes de varianza genética. Sin embargo, en poblaciones reales no hay antecedentes suficientes que aseguren que el parecido genómico, estimado con algoritmos IBD, logre mayor  $\hat{h}^2$  y/o mayor habilidad predictiva ( $\hat{r}_{y,\hat{y}}$ ). El objetivo de este trabajo fue evaluar el impacto de usar diferentes matrices de parentesco, que ponderen diferencialmente la información del pedigree y genómica, ajustando un modelo lineal mixto, en la estimación de los componentes de varianza y  $\hat{h}^2$  para caracteres de composición de carcasa y calidad de la carne, medidos en una población experimental de cerdos. Los modelos comparados fueron: BLUP convencional ( $\mathbf{A}_{22}$ ); GBLUP con relaciones genómicas basadas en los enfoques de identidad por estado (IBS) ( $\mathbf{G}_{IBS}$ ) y de IBD ( $\mathbf{G}_{IBD}$ ). El desempeño de los modelos se evaluó mediante validación cruzada. Las estimaciones de la varianza aditiva y  $\hat{h}^2$  fueron ligeramente mayores con el modelo  $\mathbf{G}_{IBD}$  que con el modelo  $\mathbf{G}_{IBS}$  para la mayoría de los caracteres, aunque las diferencias no tuvieron significancia estadística. Los modelos genómicos tuvieron mejor  $\hat{r}_{y,\hat{y}}$  que el modelo convencional de pedigree. La  $\hat{r}_{y,\hat{y}}$  fue mayor con el modelo  $\mathbf{G}_{IBS}$  que el modelo  $\mathbf{G}_{IBD}$ , sugiriendo que la matriz  $\mathbf{G}_{IBS}$  recuperaría mayor variabilidad en las relaciones de parentesco. Nuestros resultados revelaron que el uso de estimaciones más precisas del parentesco genómico, como las de algoritmos basados en IBD, no se traducen necesariamente en estimaciones más precisa de  $\hat{h}^2$  y/o de la  $\hat{r}_{y,\hat{y}}$ .

**Palabras clave:** heredabilidad, identidad por descendencia, habilidad predictiva, métodos de estimación, relaciones de parentesco genómicas, selección genómica.

## ABSTRACT

### *Estimation of heritability using pedigree and genomic information in meat quality and carcass traits*

The availability of molecular markers allows to estimate the realized proportion of the genome shared identical-by-descent (IBD) between two individuals. The result is an increase in the accuracy of both the estimates of relationships and the prediction of breeding values, when compared to the traditional pedigree model. Also, the IBD notion has been used to obtain more refine estimates of heritability ( $\hat{h}^2$ ) and genetic variance components. However, there is insufficient evidence to ensure that using the estimated genomic resemblance among relatives with IBD algorithms (instead of using an identity-by-state (IBS) approach) achieves greater  $\hat{h}^2$  and increases predictive ability ( $\hat{r}_{y,\hat{y}}$ ). The goal of this study was to assess the impact of using different relationship matrices, which differentially weighted information on the pedigree and on genomic markers. In doing that, we fitted mixed linear models to estimate variance components and  $\hat{h}^2$  for meat quality and carcass traits. Data were from an experimental pig population. The models compared were three: *conventional* BLUP ( $A_{22}$ ); GBLUP *with genomic relationships*, either based on the notion of IBS ( $G_{IBS}$ ) or on the notion of IBD ( $G_{IBD}$ ). The performance of the models was assessed through cross-validation. Estimates of additive variance and  $\hat{h}^2$  were slightly higher with the  $G_{IBD}$  model than with the  $G_{IBS}$  model for most traits, although these differences were not show significant difference. The genomic models had better  $\hat{r}_{y,\hat{y}}$  than the conventional pedigree model. Predictive ability was greater with the  $G_{IBS}$  model than with the  $G_{IBD}$  model, suggesting that the  $G_{IBS}$  matrix might recover greater variability in relationships. Our results suggested that the use of more accurate estimates of genomic relationships, such as those of IBD-based algorithms, does not necessarily translate into more accurate estimates of  $\hat{h}^2$  and/ or  $\hat{r}_{y,\hat{y}}$ .

**Key words:** heritability, identity by descent, predictive ability, estimation methods, genomic relationships, genomic selection.

## **Capítulo 1: *Introducción***

# Capítulo 1

## Introducción.

En los programas de selección genética, la predicción del mérito genético o valor de cría (BV, *Breeding Value*, en inglés), también conocida como evaluación genética, permite la identificación de los mejores animales, con el objeto de que sean utilizados como padres de la siguiente generación (Van Vleck, 1993). En el modelo infinitesimal de herencia poligénica (Falconer y Mackay, 1996), los BVs son controlados por un número muy grande de genes que influyen sobre un carácter cuantitativo, o bien de QTLs (*quantitative trait loci*, en inglés). Tradicionalmente, los BVs son predichos a partir de la información de los registros fenotípicos propios y/o de individuos emparentados. Dado que los BVs de individuos emparentados no son independientes, los individuos se “prestan” información a través de las relaciones aditivas de parentesco (Wright, 1922). Dichas relaciones se basan en la información del pedigree y constituyen el valor esperado de la proporción de alelos idénticos por descendencia (IBD, *identity by descent*, en inglés) compartidos por un par de individuos en un locus, mostrado al azar en el genoma (Malécot, 1948).

La disponibilidad de una tecnología de evaluación de genotipos (“genotipado”) con una alta densidad de marcadores moleculares asociados a polimorfismos de un solo nucleótido o SNPs (*single nucleotide polymorphism*, en inglés) ha permitido predecir el BV mediante una metodología conocida como *selección genómica* (SG; Meuwissen *et al.*, 2001). El supuesto de la SG es que, dada una elevada densidad de marcadores distribuidos a lo largo del genoma, cada QTL se asociaría con al menos un SNP en las cercanías y que, al ajustar un efecto para cada marcador, se capturaría información acerca de los QTLs (Meuwissen *et al.*, 2001). Dicha asociación se conoce como desequilibrio gamético o de ligamiento (LD), y consiste en la ausencia de independencia estadística entre las frecuencias de los alelos en dos o más loci (o segmentos de ADN) en el mismo, o en distinto, cromosoma. Un modelo de SG utilizado actualmente es el BLUP genómico o GBLUP (*genomic best linear unbiased predictor*, en inglés; VanRaden, 2008). En términos generales, este modelo relaciona los fenotipos con los genotipos de los marcadores de forma lineal, calcula una matriz de relaciones de parentesco genómicas entre los animales basada en todos los SNPs y produce un encogimiento o *shrinkage* homogéneo de los efectos de los marcadores. A diferencia de las relaciones aditivas, las relaciones de parentesco genómicas son valores realizados causados por la segregación y recombinación aleatoria más que el valor esperado por pedigree. De ese modo, capturan la información que está presente en el residuo de segregación mendeliana del modelo aditivo para el BV (Bulmer, 1985; Goddard, 2008; Cantet *et al.*, 2017). Cada padre pasa la mitad de sus genes a cada miembro de su prole, transmitiendo mitades distintas a cada hijo, dado los mencionados procesos de segregación y recombinación aleatoria de segmentos cromosómicos durante la meiosis.

En cualquier caso, las relaciones genómicas tal como se calculan actualmente (VanRaden, 2008) se sustentan en el enfoque de identidad por estado (IBS, *identity by state*, en inglés) dado que se emplea solamente marcadores y se ignora el pedigree, para inferir la proporción de genoma compartido entre individuos. Ahora bien, por razones de costo y logística, en poblaciones animales no es posible genotipar a todos los individuos de una población. Aquellos con genotipo evaluado son, en general, machos de alta difusión; también machos jóvenes promisorios y, posiblemente, hembras

potenciales candidatas a la selección (por ejemplo, madres para transferencia embrionaria, Legarra *et al.*, 2009). Coexisten, entonces, individuos genotipados e individuos no genotipados. En SG, la metodología conocida como *Single Step* permite calcular las relaciones entre animales genotipados y no genotipados teniendo en cuenta las relaciones genómicas entre animales genotipados y la genealogía completa (Aguilar *et al.*, 2014). De este modo, la precisión de las estimaciones de las relaciones entre individuos genotipados no sólo impactará en los BVs de estos animales sino en los de todo el pedigree.

Recientemente se propusieron diversos algoritmos, fundamentalmente en el área de la genética humana, para refinar el cálculo de las relaciones genómicas empleando simultáneamente la información de los SNPs y del pedigree en lugar de utilizar una u otra fuente de manera aislada. La idea subyacente es estimar las relaciones genómicas entre individuos genotipados a través de la proporción *realizada* del genoma compartido IBD (IBDP). Esto se logra calculando las probabilidades *a posteriori* de los estados de identidad en cada posición del genoma condicionales tanto en el pedigree como en la secuencia del panel de marcadores. Los métodos tradicionales de estimación de las probabilidades de IBD se basan en modelos Markov “ocultos” (HMM, *Hidden Markov Models*, en inglés; Rabiner, 1989) que incorporan dependencia estadística entre los estados de IBD para loci marcadores vecinos, pero asumen equilibrio gamético (LE) (Elston y Stewart, 1971; Lander y Green, 1987). Sin embargo, este supuesto no se cumple en un panel denso de SNPs, donde puede existir un extenso LD entre los marcadores. Han y Abney (2011) propusieron un enfoque alternativo para inferir las probabilidades de IBD en pedigrees humanos de gran tamaño con muchos individuos sin genotipar utilizando una densidad elevada de SNPs. El enfoque de estos autores modela eficientemente en términos computacionales la relación entre un par de individuos genotipados teniendo en cuenta el LD generado por ligamiento, así como aquel producido por las relaciones ancestrales que van más allá del pedigree conocido.

La incorporación del enfoque de IBD extendido al genoma completo ha demostrado aumentar la exactitud tanto de las estimaciones de las relaciones genómicas como de la predicción de los BVs (Cantet *et al.*, 2017; Forneris *et al.*, 2016; García-Baccino *et al.*, 2016). Asimismo, dicho enfoque ha sido utilizado para calcular estimaciones refinadas de la heredabilidad ( $h^2$ ) y la partición genómica de la varianza genética (Visscher, 2009; Visscher *et al.*, 2006). La estimación de  $h^2$  tiene un rol crucial porque determina el nivel de precisión con la cual el valor genético puede ser predicho desde la información fenotípica, y por lo tanto delimita los esquemas de mejoramiento animal (Visscher, 2009; Visscher *et al.*, 2014). Sin embargo, aún no es claro si el parecido genómico estimado con algoritmos más precisos (como los basados en el enfoque de IBD genómico) lograrán mayor  $h^2$  y/o mayor habilidad predictiva que sus correspondientes basados en IBS, cuando se emplean datos de poblaciones reales. Este proyecto propone estudiar el impacto de utilizar diferentes matrices de relaciones de parentesco, que consideran la información del pedigree y/o genómica entre individuos, ajustando un modelo lineal mixto de predicción, en la estimación de los componentes de varianza y la  $h^2$ , para caracteres de composición de carcasa y calidad de la carne medidos en una población experimental de cerdos. Adicionalmente, el segundo objetivo es evaluar el desempeño de los modelos propuestos estimando su habilidad predictiva mediante validación cruzada. Esto permitirá investigar si el uso de una estimación más precisa del parentesco genómico, como es el uso de algoritmos que se sustentan en el enfoque de IBD, necesariamente se traducirá en una estimación más precisa de la heredabilidad y/o de la habilidad predictiva.

El documento está organizado en cinco capítulos siendo el primero la presente introducción. El Capítulo 2 describe la metodología utilizada para el la estimación de las matrices de relaciones de parentesco, los modelos lineales mixtos ajustados para estimar los componentes de varianza y la heredabilidad de los caracteres analizados y el procedimiento de validación cruzada para estimar la habilidad predictiva de cada modelo. Los resultados obtenidos se presentan en el Capítulo 3. El Capítulo 4 contiene la discusión de los resultados y el 5 desarrolla las conclusiones generales de la tesis.



## **Capítulo 2. *Materiales y Métodos***

## Capítulo 2

### Materiales y Métodos

#### 2.1. Animales.

La base de datos utilizada para realizar el análisis proviene de una población experimental de cerdos Duroc x Pietrain distribuidos en tres generaciones ( $F_0$ ,  $F_1$  y  $F_2$ ) creada en Michigan State University Swine Teaching and Research Center, East Lansing, Michigan (Casiró *et al.*, 2017; Edwards *et al.*, 2008). Brevemente, la generación  $F_0$  consistió en 4 machos de la raza Duroc cruzados con 16 hembras de la raza Pietrain mediante inseminación artificial, de estos apareamientos se produjeron los individuos  $F_1$ , de los cuales se seleccionaron 51 hembras y 6 machos como padres de la siguiente generación. La generación  $F_2$  reunió 960 lechones de ambos sexos.

#### 2.2. Fenotipos.

Se dispuso de una base de datos con registros correspondientes a 38 caracteres de composición de carcasa y calidad de la carne evaluados en aproximadamente 948 animales pertenecientes a la generación  $F_2$  (Casiró *et al.*, 2017; Edwards *et al.*, 2008). Para los análisis propuestos en la presente investigación y del total de caracteres registrados, se eligieron 20 con al menos 900 individuos con dato fenotípico. En la Tabla 2.1 se describen el número de observaciones, media fenotípica y desvío estándar de los fenotipos utilizados en los análisis de esta investigación.

#### 2.3. Genotipado y edición de datos genómicos.

El procedimiento de genotipado y edición fue reportado por Casiró *et al.* (2017). Brevemente, el genotipado se realizó con dos paneles de SNP de diferentes densidades. Todos los animales de la generación  $F_0$  y  $F_1$  y 336  $F_2$  fueron genotipados utilizando el chip Illumina PorcineSNP60 beadchip, que contenía aproximadamente 62,000 SNP. Además, 612 animales de la generación  $F_2$  fueron genotipados a baja densidad usando el panel GeneSeek Genomic Profiler – Porcine LD (GGP-Porcine LD, GeneSeek a Neogen Company, Lincoln, NE), que contenía aproximadamente 8,836 marcadores. El procedimiento para la edición de los genotipos consistió en eliminar SNP con genotipos faltantes en todos los animales, inconsistencias mendelianas, aquellos SNP a los que les faltaba más del 10% de los genotipos y cuya frecuencia alélica en el alelo menos frecuente fuera inferior a 1% ( $MAF < 0.01$ ). Después del proceso de edición, se llevó a cabo la imputación de genotipos faltantes en el chip de 60K y aquellos no anotados en el chip de baja densidad usando el programa FImpute (Sargolzaei *et al.*, 2014). Luego de filtrar animales por inconsistencias mendelianas y porcentaje de SNPs anotados (*call rate*), quedaron finalmente para los análisis posteriores un total de 1015 animales con genotipos para 42,087 SNP.

**Tabla 2.1.** Número de observaciones, media fenotípica y desvío estándar de los 20 fenotipos analizados.

| <b>Fenotipo</b>                            | <b>Abreviatura</b> | <b>n</b> | <b>Media</b> | <b>SD</b> |
|--|--------------------|----------|--------------|-----------|
| <i>Medidas de carcasa</i>                  |                    |          |              |           |
| Rendimiento en canal (%)                   | RCptg              | 931      | 73.013       | 2.113     |
| pH 24 horas post-sacrificio                | pH24               | 910      | 5.513        | 0.139     |
| Longitud de la canal (cm)                  | Carlng             | 930      | 78.718       | 2.515     |
| Grasa dorsal última vertebra lumbar (mm)   | Glumb              | 929      | 22.249       | 6.221     |
| Grasa dorsal decima costilla (mm)          | G10c               | 931      | 19.899       | 6.396     |
| Area de lomo (cm <sup>2</sup> )            | ALm                | 931      | 37.083       | 4.814     |
| <i>Cortes primarios</i>                    |                    |          |              |           |
| Peso de la panceta (kg)                    | Panckg             | 930      | 5.029        | 0.673     |
| Peso de la paleta - <i>Boston</i> (kg)     | Bostkg             | 930      | 3.904        | 0.560     |
| Peso de Jamón (kg)                         | Jamkg              | 930      | 9.632        | 0.769     |
| Peso de Lomo (kg)                          | Lomkg              | 930      | 8.287        | 0.830     |
| Peso de paleta - <i>Picnic</i> (kg)        | Picnkg             | 930      | 3.716        | 0.570     |
| Peso del costillar del pecho (kg)          | Costkg             | 927      | 1.528        | 0.198     |
| <i>Calidad de carne</i>                    |                    |          |              |           |
| Grasa (%)                                  | Grptg              | 919      | 3.179        | 1.396     |
| Humedad (%)                                | Humptg             | 919      | 73.946       | 1.536     |
| Proteína (%)                               | Protptg            | 918      | 23.438       | 1.129     |
| <i>Análisis de Laboratorio</i>             |                    |          |              |           |
| Rendimiento a la cocción (%)               | RCoc               | 921      | 77.264       | 2.829     |
| Pérdida de agua por goteo (%)              | PAptg              | 929      | 1.844        | 1.174     |
| Resistencia al corte (Warner-Bratzler, kg) | SF                 | 920      | 3.208        | 0.687     |
| <i>Análisis de paneles sensoriales</i>     |                    |          |              |           |
| Jugosidad (1 a 8)                          | Jug                | 925      | 5.234        | 0.590     |
| Terneza (1 a 8)                            | Tern               | 925      | 5.554        | 0.614     |

SD desviación estándar.

#### 2.4. Estimación de las matrices de relaciones de parentesco.

Las matrices de relaciones de parentesco fueron estimadas mediante tres métodos, el primero sólo considera la información del pedigree para calcular la matriz de relaciones de parentesco aditivas entre los animales genotipados de la población experimental de cerdos; el segundo método utiliza solo la información genómica para estimar una matriz de relaciones de parentesco entre los animales genotipados basada en el enfoque de identidad por estado (IBS); el tercer método hace uso de la información genómica condicionada en el pedigree y calcula una matriz de relaciones genómicas entre animales, mediante un algoritmo que se basa en el enfoque de identidad por descendencia (IBD).

### 2.4.1. Matriz $A$ de relaciones aditivas.

El pedigree de los 1015 animales genotipados contenía 1037 individuos (519 machos y 518 hembras), distribuidos en tres generaciones ( $F_0 = 20$ ,  $F_1 = 57$ ,  $F_2 = 960$ ). Sólo 22 animales de la generación  $F_2$  no poseían genotipo. Para construir la matriz  $A$  de relaciones de parentesco, se calcularon las relaciones genéticas aditivas entre los individuos del pedigree (Wright, 1934). Los elementos de  $A$  fueron calculados con la función `kinship` disponible en el paquete `kinship2` (Sinnwell *et al.*, 2014) del software R (*R v 4.0.2; Foundation for Statistical Computing, Viena, Austria*). Luego,  $A$  fue particionada en animales genotipados y animales no genotipados como:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

La submatriz  $A_{22}$  contiene las relaciones aditivas entre los individuos con genotipos y es la que finalmente se utilizó para realizar los análisis que se describen más adelante.

### 2.4.2. Matriz $G_{IBS}$ de relaciones genómicas IBS.

Las relaciones genómicas fueron calculadas según la metodología propuesta por VanRaden (2008), que utiliza la información de los marcadores moleculares sin considerar la genealogía. Para estimar dichas relaciones, los genotipos se expresaron como conteo de alelos y se almacenaron en una matriz  $M$ , con dimensión  $n$  (número de individuos, 1015)  $\times$   $m$  (número de marcadores SNP, 42,087).  $M$  se estandarizó restando primero el conteo alélico esperado de cada SNP a cada elemento para luego dividir por la raíz cuadrada de la varianza del conteo de alelos. Una vez que se obtuvo  $M$ ,  $G_{IBS}$  se calculó multiplicando la matriz de marcadores estandarizada  $Z$  por su transpuesta, siendo la fórmula de cálculo igual a  $G_{IBS} = ZZ'$ , donde  $Z_{ij} = \frac{M_{ij} - 2p_j}{\sqrt{m[2p_j(1-p_j)]}}$ ;  $i$  es el  $i$ -ésimo animal,  $m$  el número de marcadores y  $p_j$  la frecuencia del marcador  $j$ . El procedimiento de estandarización permite que  $G_{IBS}$  sea análoga a la matriz  $A$  de relaciones aditivas.

### 2.4.3. Matriz $G_{IBD}$ de relaciones genómicas IBD.

Los elementos de esta matriz se computaron con el algoritmo “IBDL” v3.13 (Han y Abney, 2011, 2013). El algoritmo ajusta un modelo oculto de Markov (HMM), para estimar las probabilidades *a posteriori* de los 9 estados condensados de IBD ( $\{\check{\Delta}_{1t}, \check{\Delta}_{2t}, \dots, \check{\Delta}_{9t}\}$ ) en el  $t$ -ésimo SNP ( $t = 1:nsnp$ ;  $nsnp = 42.087$ ), entre un par de animales genotipados, condicional a la información de todo el panel de SNPs y del pedigree. Para ello, utiliza como probabilidades *a priori* las calculadas exclusivamente con información del pedigree, tal como se describe en Jacquard (1970) ( $\{\Delta_1, \Delta_2, \dots, \Delta_9\}$ ). Como resultado, se obtiene un coeficiente de coancestría empírico (o realizado) por cromosoma y otro genómico, que surge como un promedio de los anteriores ponderado por el largo de cada cromosoma según la siguiente expresión:

$$\check{\theta}_{ij}(t) = \check{\Delta}_{1t} + \frac{1}{2}(\Delta_{3t} + \Delta_{5t} + \Delta_{7t}) + \frac{1}{4}\Delta_{8t}$$

Donde  $\check{\theta}_{ij}(t)$  es el coeficiente de coancestría realizado entre un par de individuos  $i$  y  $j$  en el SNP  $t$ . Así, el coeficiente de relación aditiva realizado en el SNP  $t$  es igual a

$2\check{\theta}_{ij}(t)$ , la coancestría promedio para el cromosoma  $k$  con  $nsnp_k$  marcadores es igual a  $\check{\theta}_{ij}(k) = \sum_{t=1}^{nsnp_k} \check{\theta}_{ij}(t)/nsnp_k$  y para un genoma de largo  $L$  y cromosomas de largo  $l_k$  ( $k = 1:nchr$ ;  $nchr = 18$ ) la actual relación  $\check{R}_{ij}$ , es la siguiente (Guo, 1995):

$$\check{R}_{ij} = \frac{1}{L} \sum_{k=1}^{nchr} 2\check{\theta}_{ij}(k) \times l_k$$

De las 5 opciones que ofrece el programa para calcular la coancestría empírica, se utilizó la opción IBDLD-RR, que tiene en cuenta el desequilibrio de ligamiento (LD) entre SNPs, al condicionar las probabilidades de IBD en un locus en los 20 SNPs anteriores mediante una regresión “ridge”. Los elementos de la matriz  $\mathbf{G}_{IBD}$  se obtuvieron multiplicando la coancestría genómica por dos. Finalmente, la función `nearPD` del paquete `Matrix` en R fue utilizada para calcular la matriz positiva definida más cercana a la  $\mathbf{G}_{IBD}$  original (Cheng y Higham, 1998; Higham, 2002), los elementos de esta matriz fueron empleados para los análisis estadísticos.

## 2.5. Estimación de componentes de varianza y heredabilidad de los caracteres.

La estimación de los componentes de varianza y la heredabilidad en los caracteres de composición de la carcasa y calidad de carne en cerdos, se realizó ajustando un modelo lineal mixto unicaracter, mediante la función `gblup` del paquete `gwaR` implementado en R (<https://github.com/steibelj/gwaR>) tal como detallan Casiró *et al.* (2017).

### 2.5.1. Modelo estadístico.

El modelo lineal mixto empleado para estimar los parámetros de covarianza en los 20 fenotipos de composición de carcasa y calidad de carne, es esencialmente un modelo unicaracter donde el fenotipo es función de los efectos fijos más un efecto aditivo

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad [1]$$

En [1],  $\mathbf{y}$  es el vector  $n \times 1$  del fenotipo observado,  $\mathbf{X}$  es la matriz de incidencia  $n \times p$  relacionando los registros al vector de efectos fijos  $\boldsymbol{\beta}$  de orden  $p$ , los cuales incluyen el sexo (2 niveles) y el grupo de sacrificio (33 niveles) como variables clasificatorias y el peso de la carcasa como covariable,  $\mathbf{Z}$  es la matriz de incidencia que relaciona a los fenotipos en  $\mathbf{y}$  con el vector aleatorio de efectos genéticos aditivos  $\mathbf{u}$  ( $q \times 1$ ), y  $\mathbf{e}$  ( $n \times 1$ ) es el vector de errores aleatorios distribuidos de modo tal que  $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$ , donde  $\mathbf{I}$  es la matriz identidad y  $\sigma_e^2$  es la varianza del error. Se asumió una distribución de los valores de cría igual a  $\mathbf{u} \sim N(0, \mathbf{G}\sigma_u^2)$ , donde  $\mathbf{G}$  es alguna de las matrices de relaciones de parentesco  $\mathbf{A}_{22}$ ,  $\mathbf{G}_{IBS}$  o  $\mathbf{G}_{IBD}$  y  $\sigma_u^2$  es la varianza genética aditiva. Las variables respuesta fueron los caracteres de composición de la carcasa y calidad de la carne (Tabla 2.1). La ecuación del modelo [1] fue ajustada con cada matriz de parentesco en cada uno de los caracteres. Consecuentemente, para diferenciar los tres modelos, se los denominó  $\mathbf{A}_{22}$ ,  $\mathbf{G}_{IBS}$  y  $\mathbf{G}_{IBD}$ , respectivamente. Los componentes de varianza de cada modelo fueron estimados mediante máxima verosimilitud residual (REML, Patterson and Thompson, 1971) empleando el algoritmo de segundas derivadas Newton-Raphson.

### 2.5.2. Estimación de la heredabilidad.

El parámetro heredabilidad ( $h^2$ ) fue estimado como el cociente entre la varianza genética aditiva ( $\sigma_u^2$ ) y la varianza fenotípica ( $\sigma_p^2$ ), del modo siguiente:

$$\hat{h}^2 = \frac{\sigma_u^2}{\sigma_p^2} = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2}$$

Los escalares  $\hat{\sigma}_u^2$  y  $\hat{\sigma}_e^2$  son los estimadores de la varianza genética aditiva y la del error respectivamente, calculados para cada uno de los tres modelos y para cada caracter.

### 2.5.3. Test de hipótesis para la varianza genética aditiva de los modelos.

Teniendo en cuenta el modelo lineal mixto de la ecuación [1], el parámetro de interés es la varianza genética aditiva ( $\hat{\sigma}_u^2$ ), la cual fue estimada ajustando los modelos  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$ . Como el parámetro de interés en testear es solamente uno, el valor del verdadero parámetro puede estar en los límites del espacio paramétrico (el límite inferior es 0). En tal caso, Self y Liang (1987) propusieron una aproximación de la distribución del test del Cociente de Verosimilitud (Likelihood Ratio Test, LRT), en la cual, bajo hipótesis nula ( $H_0$ ) de cero varianza, la distribución asintótica del estadístico LRT es una mezcla 50:50 chi-cuadrado con 0 y 1 grados de libertad ( $\frac{1}{2}\chi^2_0 : \frac{1}{2}\chi^2_1$ ). Asumiendo esta distribución, el valor de probabilidad para el estadístico LRT obtenido ( $P$  – *valor*) puede ser calculado dividiendo a la mitad el correspondiente  $P$  – *valor* de una distribución  $\chi^2_{gl}$  (Stram y Lee, 1994; Visscher, 2006).

Con el objeto de realizar la prueba de hipótesis con los modelos propuestos, se testeó para pares de modelos la diferencia entre los estimadores del componente de varianza  $\hat{\sigma}_u^2$ . Siendo las comparaciones  $G_{IBS}$  vs  $A_{22}$ ,  $G_{IBD}$  vs  $A_{22}$  y  $G_{IBD}$  vs  $G_{IBS}$  en cada caracter. Por tanto, las hipótesis evaluadas fueron:  $H_0: \hat{\sigma}_{u_i}^2 - \hat{\sigma}_{u_j}^2 = 0$  contra la alternativa  $H_1: \hat{\sigma}_{u_i}^2 - \hat{\sigma}_{u_j}^2 \neq 0$  y el estadístico fue calculado de la siguiente manera:

$$\left| \frac{\hat{\sigma}_{u_i}^2}{SE_i} - \frac{\hat{\sigma}_{u_j}^2}{SE_j} \right| \sim \chi^2_{0,5gl}$$

Donde,  $\hat{\sigma}_{u_i}^2$  y  $\hat{\sigma}_{u_j}^2$  son los estimadores REML de los componentes de varianza genética aditiva de cada modelo y  $SE_i$ ,  $SE_j$  sus respectivos errores estándar de estimación. Se empleó un nivel de significancia del 5%.

### 2.6. Estimación de la habilidad predictiva.

Para evaluar y comparar los resultados de los modelos propuestos,  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$ , se estimó su habilidad predictiva como la correlación entre el fenotipo observado y el predicho a través de una validación cruzada con 5 iteraciones (en inglés, *5-fold cross-validation*, Jia, 2017).

### 2.6.1. Validación cruzada de 5 iteraciones (5-fold cross-validation).

El procedimiento de validación cruzada de 5 iteraciones fue programado en un código en lenguaje R. El mismo consistió en particionar aleatoriamente los datos para cada carácter en 5 subconjuntos. En cada iteración, se ocultaron los fenotipos de los animales de uno de los subconjuntos, denominado conjunto de datos de validación (“test set”, en inglés) y se utilizaron los fenotipos de los  $k - 1$  subconjuntos restantes, cuatro en este caso, para “entrenar” el modelo, llamado conjunto de datos de entrenamiento (“training set”, en inglés). Como resultado de cada iteración, se obtuvo una solución para los efectos fijos y aleatorios del modelo, y con ellos se contruyó la predicción del fenotipo para los animales del “test set”. Una vez finalizadas las 5 iteraciones, se calculó el coeficiente de correlación entre el fenotipo observado (i.e., el observado en la base de datos original) y el fenotipo predicho para cada animal ( $\hat{r}_{y,\hat{y}}$ ). El procedimiento se replicó 300 veces. Se descartaron aquellas réplicas que, durante el proceso de estimación de los componentes de varianza, no alcanzaron convergencia en la estimación de  $\hat{r}_{y,\hat{y}}$ . Para los análisis posteriores, solo se consideraron los estimadores de  $\hat{r}_{y,\hat{y}}$ , correspondientes a 213 réplicas que alcanzaron el criterio de convergencia con los 3 modelos ajustados. Se emplearon los paquetes `lm` y `lsmeans` para calcular la media estimada de  $r_{y,\hat{y}}$  y su error estándar (SE) en cada carácter, mediante el ajuste del modelo lineal  $\hat{r}_{(y,\hat{y})ij} = \tau_i + \varepsilon_{ij}$ , donde la variable respuesta  $\hat{r}_{(y,\hat{y})ij}$  es la  $j$  – ésima correlación estimada ( $j = 1, \dots, 213$ ) con la  $i$  – ésima matriz de parentesco ( $i = 1, 2, 3, \mathbf{A}_{22}, \mathbf{G}_{IBS}$  o  $\mathbf{G}_{IBD}$ ) y  $\varepsilon_{ij}$  es el término de error aleatorio.

### 2.6.2. Test de comparación de medias de las habilidades predictivas.

El método de Tukey fue el procedimiento de uso para la comparación múltiple de las medias de  $\hat{r}_{(y,\hat{y})ij}$  con los tres modelos ajustados en cada carácter, mediante el uso del paquete `multcomp` disponible en R (Hothorn *et al.*, 2008). Se empleó un nivel de significancia del 5%.

## **Capítulo 3. *Resultados***



## Capítulo 3

### Resultados

#### 3.1. Análisis descriptivo de las matrices de relaciones de parentesco estimadas.

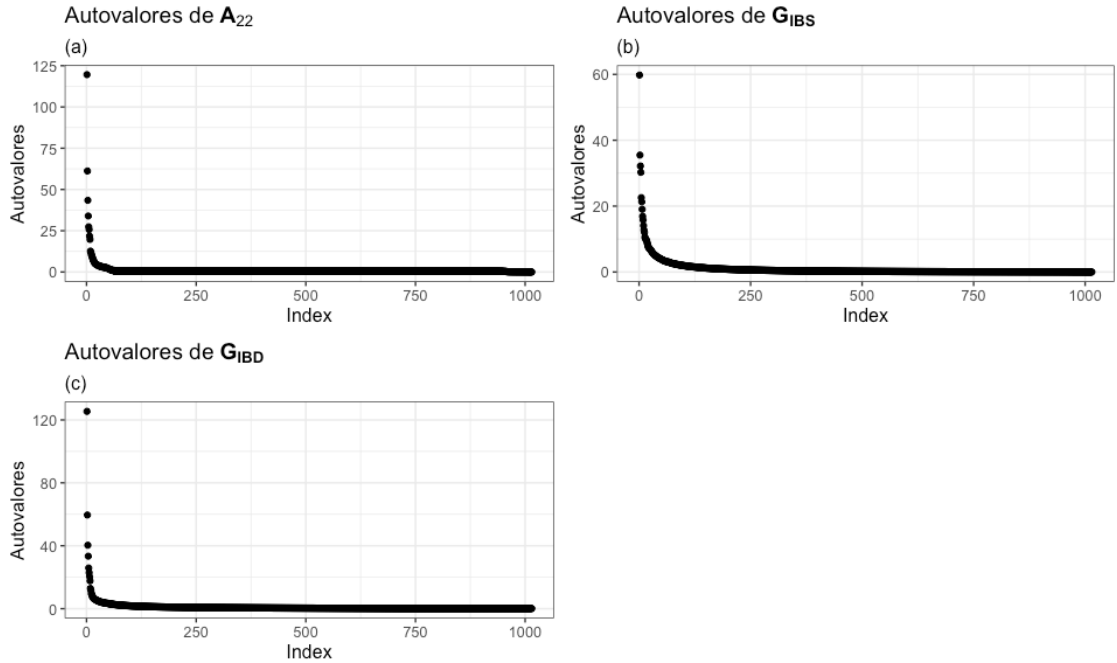
En la Tabla 3.1 se presenta los estadísticos descriptivos de la distribución de los elementos diagonales y no diagonales calculados con 3 generaciones de animales ( $n = 1015$ ) en cada matriz de parentesco. Hubo individuos consanguíneos en la población. Las medias estimadas de los coeficientes diagonales, en orden creciente, fueron 0,94 con  $G_{IBS}$ , 1,00 con  $A_{22}$  y 1,06 con  $G_{IBD}$ , teniendo mayor desviación estándar los coeficientes estimados con la información genómica, respecto de aquellos obtenidos con información de pedigree. El rango de los elementos diagonales estimados en  $A_{22}$  fue 1,00 a 1,25, y entre 1,02 a 1,36 con  $G_{IBD}$ , mientras que con  $G_{IBS}$  el rango fue mayor siendo el valor mínimo 0,61 y el valor máximo 1,42. Las medias estimadas de los elementos fuera de la diagonal fueron 0,112,  $-0,001$  y 0,12 para  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$  respectivamente. La desviación estándar fue similar para las tres matrices. Sin embargo, la distribución de sus elementos presentó algunas diferencias. En primer lugar, los máximos alcanzados con matrices genómicas fueron mayores que con  $A_{22}$  (0,92 vs. 0,75). En segundo lugar, el mínimo con  $G_{IBS}$  ( $-0,51$ ) fue más negativo que con las otras matrices  $-0,08$  y 0, para  $G_{IBD}$  y  $A_{22}$ , respectivamente. En el caso particular de  $G_{IBD}$ , los valores negativos surgieron al aplicar la función `nearPD` de R, programada para hacer la matriz inversible, aún cuando los valores previos tenían un mínimo igual a 0. Finalmente, la distribución de los elementos en  $G_{IBD}$  fue más simétrica que con las otras matrices, al coincidir la media y la mediana.

**Tabla 3.1.** Distribución de los elementos diagonales y fuera de la diagonal de las matrices de relaciones de parentesco estimadas.

|                     | Diagonal |           |           | Fuera de la diagonal |           |           |
|---------------------|----------|-----------|-----------|----------------------|-----------|-----------|
|                     | $A_{22}$ | $G_{IBS}$ | $G_{IBD}$ | $A_{22}$             | $G_{IBS}$ | $G_{IBD}$ |
| <b>Media</b>        | 1,001    | 0,944     | 1,065     | 0,112                | $-0,001$  | 0,119     |
| <b>SD</b>           | 0,016    | 0,079     | 0,035     | 0,102                | 0,101     | 0,099     |
| <b>Valor Mínimo</b> | 1,000    | 0,610     | 1,016     | 0,000                | $-0,510$  | $-0,086$  |
| <b>Quantil 25%</b>  | 1,000    | 0,901     | 1,041     | 0,062                | $-0,071$  | 0,051     |
| <b>Mediana</b>      | 1,000    | 0,939     | 1,056     | 0,062                | $-0,020$  | 0,101     |
| <b>Quantil 75%</b>  | 1,000    | 0,980     | 1,079     | 0,125                | 0,051     | 0,171     |
| <b>Valor Máximo</b> | 1,250    | 1,425     | 1,364     | 0,750                | 0,917     | 0,920     |

$A_{22}$  matriz de relaciones Aditivas.  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado.  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia. **SD** desviación estándar.

El patrón de los autovalores de cada matriz de relación de parentesco estimada se contrasta en la Figura 3.1. El rango de autovalores fue similar para las matrices  $A_{22}$  (0,01 - 119,59) y  $G_{IBD}$  (0,12 - 125,42), mientras que para  $G_{IBS}$  el rango fue significativamente menor (0 - 59,78). Cabe mencionar que  $G_{IBS}$  no fue ponderada por  $A_{22}$  en su construcción para lograr la no singularidad.



**Figura 3.1.** Autovalores de las matrices de relaciones de parentesco. (a)  $A_{22}$  matriz de relaciones Aditivas. (b)  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado. (c)  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia.

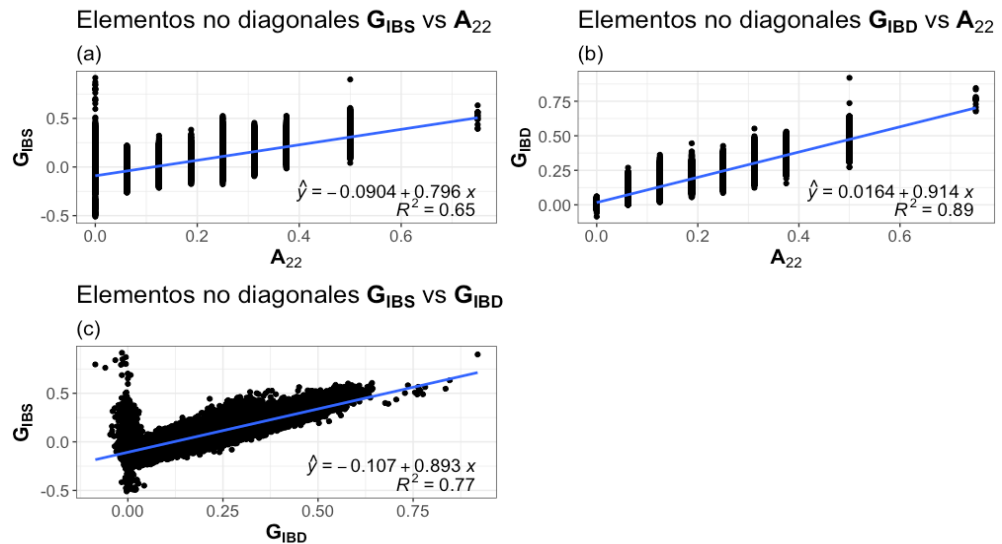
La relación entre los elementos fuera de la diagonal de las matrices de parentesco estimadas se muestra en la Tabla 3.2 y Figura 3.2. El coeficiente de correlación ( $\hat{r}$ ) entre los elementos no diagonales fue más alto entre las relaciones estimadas con  $A_{22}$  y  $G_{IBD}$  (0,94) que  $\hat{r}$  entre  $A_{22}$  y  $G_{IBS}$  (0,808) o entre  $G_{IBS}$  y  $G_{IBD}$  (0,878). Por su parte, los coeficientes de regresión estimados  $\hat{\beta}$  de los elementos no diagonales de  $G_{IBS}$  y  $G_{IBD}$  sobre  $A_{22}$  fueron menores a 1 ( $\hat{\beta}$  igual a 0,796 y 0,914 para  $G_{IBS}$  y  $G_{IBD}$ , respectivamente). Lo mismo ocurrió con la regresión de  $G_{IBS}$  en  $G_{IBD}$  ( $\hat{\beta} = 0,892$ ). En particular, la regresión de  $G_{IBD}$  en  $A_{22}$  menor a 1 indicaría que las relaciones realizadas entre los animales fueron menores, en promedio, a su valor esperado únicamente por pedigree.

**Tabla 3.2.** Coeficientes de regresión (arriba de la diagonal), varianza (diagonal), correlaciones (debajo de la diagonal) de los elementos fuera de la diagonal de las matrices de relaciones de parentesco estimadas.

|           | $A_{22}$      | $G_{IBD}$     | $G_{IBS}$     |
|-----------|---------------|---------------|---------------|
| $A_{22}$  | <b>0,0105</b> | 0,9144        | 0,7957        |
| $G_{IBD}$ | 0,9441        | <b>0,0098</b> | 0,8925        |
| $G_{IBS}$ | 0,8084        | 0,8782        | <b>0,0102</b> |

Los coeficientes de regresión  $\hat{\beta}$  son de la variable en la columna  $j$  sobre la variable en la fila  $i$ . Coeficiente de correlación de Pearson ( $\hat{r}$ ).  $A_{22}$  matriz de relaciones Aditivas.  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado.  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia.

La Figura 3.2 ilustra mejor cómo la matriz genómica basada en el enfoque de IBS recupera relaciones diferentes de 0 para aquellas parejas de animales no relacionadas por pedigré. En efecto, para valores de  $A_{22}$  o  $G_{IBD}$  iguales a 0, correspondientes a individuos no relacionados,  $G_{IBS}$  tomó valores entre  $-0,51$  y  $0,9$ . Esto indicaría la habilidad de esta matriz de capturar parentescos más allá del pedigré conocido, o de detectar errores de pedigré.



**Figura 3.2.** Regresión de los elementos fuera de la diagonal de las matrices de parentesco. (a) Regresión de matriz  $G_{IBS}$  sobre matriz  $A_{22}$ . (b) Regresión de matriz  $G_{IBD}$  sobre la matriz  $A_{22}$ . (c) Regresión de matriz  $G_{IBS}$  sobre matriz  $G_{IBD}$ .

### 3.2. Componentes de varianza estimados en caracteres de composición de carcasa y calidad de la carne de cerdos.

En la Tabla 3.3 se reportan los componentes de varianza estimados en caracteres de composición de carcasa y calidad de la carne con cada modelo. Los 20 caracteres se clasificaron dentro de 5 grupos, los cuales son ‘medidas de carcasa’, ‘cortes primarios’, ‘calidad de carne’, ‘análisis de laboratorio’ y ‘análisis de paneles sensoriales’. En términos generales, las estimaciones de la varianza genética aditiva ( $\hat{\sigma}_u^2$ ) obtenidos en los cinco grupos de caracteres fueron mayores con los modelos que usaron matrices de relaciones genómicas (modelos  $G_{IBS}$  y  $G_{IBD}$ ) en comparación con el modelo que utiliza solo pedigree ( $A_{22}$ ). Además, los errores estándar de  $\hat{\sigma}_u^2$  fueron de menor magnitud empleando los modelos  $G_{IBS}$  y/o  $G_{IBD}$ , que los estimados con el modelo  $A_{22}$ . Este resultado estuvo también acompañado por menores valores de la varianza residual estimada ( $\hat{\sigma}_e^2$ ) con los modelos genómicos que con el modelo convencional de pedigree. El test de hipótesis mostró que la diferencia entre estimadores de  $\sigma_u^2$  obtenidos con los modelos  $G_{IBS}$  y/o  $G_{IBD}$  versus el obtenido con el modelo  $A_{22}$ , fueron mayores en todos los caracteres del grupo ‘calidad de carne’ (Grptg, Humptg, Protptg), cuatro caracteres del grupo ‘medidas de carcasa’ (Carlng, Glumb, G10c, ALm) y dos del grupo ‘cortes primarios’ (Jamkg, Lomkg), con  $P$ -valores entre un rango de 0,016 a 0.04. Aun cuando en cuatro caracteres, relacionados con los grupos ‘cortes primarios’ (Bostkg, Costkg) y ‘análisis de laboratorio’ (RCoc, PAptg), el parámetro  $\hat{\sigma}_u^2$  obtenido fue mayor para el modelo  $A_{22}$  que para los modelos  $G_{IBS}$  y/o  $G_{IBD}$ , no se encontró diferencia estadística ( $P > 0,05$ ) entre estos estimadores del parámetro.

Por otra parte, al contrastar los resultados de los estimadores de los componentes de varianza entre los modelos  $G_{IBS}$  vs  $G_{IBD}$ , es de notar que, en 18 de los 20 caracteres el modelo  $G_{IBD}$  estimó valores de  $\hat{\sigma}_u^2$  levemente mayores y con errores estándar de mayor magnitud que los estimados con el modelo  $G_{IBS}$ . Además, la varianza estimada del error  $\sigma_e^2$  fue menor en el modelo  $G_{IBD}$  respecto del modelo  $G_{IBS}$ . En referencia al carácter *ph 24 horas post-sacrificio* (ph24), los estimadores de los componentes de varianza fueron iguales con ambos modelos ( $\hat{\sigma}_u^2 = 0.003$  y  $\hat{\sigma}_e^2 = 0.013$ ), mientras que para el carácter *Terneza* (Tern), la varianza genética aditiva estimada fue levemente mayor con el modelo  $G_{IBS}$  ( $\hat{\sigma}_u^2 = 0.101$ ,  $SD = 0.02$ ) comparado con la estimación obtenida con el modelo  $G_{IBD}$  ( $\hat{\sigma}_u^2 = 0.096$ ,  $SD = 0.02$ ). Al evaluar las diferencias observadas entre los estimadores de  $\sigma_u^2$  para los modelos genómicos (modelo  $G_{IBS}$  vs modelo  $G_{IBD}$ ) mediante pruebas de hipótesis, no se encontraron diferencias significativas ( $P > 0,05$ ) para todos los caracteres analizados.

**Tabla 3.3.** Estimadores de componentes de varianza usando tres modelos lineales mixtos en caracteres de composición de carcasa y calidad de la carne de cerdos (Error estándar).

| Caracter            | Modelo $A_{22}$    |                    | Modelo $G_{IBS}$   |                    | Modelo $G_{IBD}$   |                    |
|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                     | $\hat{\sigma}_u^2$ | $\hat{\sigma}_e^2$ | $\hat{\sigma}_u^2$ | $\hat{\sigma}_e^2$ | $\hat{\sigma}_u^2$ | $\hat{\sigma}_e^2$ |
| RCptg <sup>1</sup>  | 0,488 (0,23)       | 2,410 (0,18)       | 0,782 (0,17)       | 2,101 (0,13)       | 0,869 (0,20)       | 2,091 (0,15)       |
| pH24                | 0,002 (0,001)      | 0,013 (0,001)      | 0,003 (0,001)      | 0,013 (0,001)      | 0,003 (0,001)      | 0,013 (0,01)       |
| Carlng              | 1,564 (0,51)       | 2,503 (0,31)       | 1,839 (0,30)*      | 2,083 (0,14)       | 2,089 (0,33)*      | 1,931 (0,17)       |
| Glumb               | 7,503 (2,90)       | 21,75 (2,0)        | 11,696 (2,1)*      | 17,53 (1,17)       | 12,27 (2,29)*      | 17,22 (1,38)       |
| G10c                | 2,889 (1,41)       | 16,25 (1,16)       | 7,445 (1,38)*      | 12,33 (0,81)       | 8,453 (1,57)*      | 11,80 (0,95)       |
| ALm                 | 3,949 (1,43)       | 9,389 (0,94)       | 5,094 (0,92)*      | 7,965 (0,53)       | 5,797 (1,04)*      | 7,481 (0,61)       |
| Panckg <sup>2</sup> | 0,018 (0,009)      | 0,105 (0,007)      | 0,024 (0,006)      | 0,10 (0,005)       | 0,028 (0,01)       | 0,097 (0,01)       |
| Bostkg              | 0,031 (0,01)       | 0,047 (0,006)      | 0,018 (0,004)      | 0,055 (0,003)      | 0,02 (0,005)       | 0,05 (0,003)       |
| Jamkg               | 0,062 (0,02)       | 0,137 (0,014)      | 0,096 (0,01)*      | 0,102 (0,007)      | 0,103 (0,02)*      | 0,098 (0,01)       |
| Lomkg               | 0,038 (0,018)      | 0,208 (0,015)      | 0,075 (0,02)*      | 0,174 (0,011)      | 0,088 (0,02)*      | 0,168 (0,01)       |
| Picnkg              | 0,010 (0,005)      | 0,068 (0,004)      | 0,011 (0,003)      | 0,066 (0,003)      | 0,012 (0,004)      | 0,06 (0,004)       |
| Costkg              | 0,008 (0,002)      | 0,008 (0,001)      | 0,006 (0,001)      | 0,01 (0,001)       | 0,006 (0,001)      | 0,01 (0,001)       |
| Grptg <sup>3</sup>  | 0,736 (0,23)       | 1,079 (0,14)       | 0,944 (0,13)*      | 0,805 (0,06)       | 1,107 (0,15)*      | 0,690 (0,07)       |
| Humptg              | 0,618 (0,21)       | 1,320 (0,14)       | 0,723 (0,13)*      | 1,164 (0,07)       | 0,843 (0,15)*      | 1,074 (0,08)       |
| Protptg             | 0,231 (0,11)       | 0,914 (0,07)       | 0,449 (0,08)*      | 0,732 (0,04)       | 0,533 (0,10)*      | 0,670 (0,05)       |
| RCoc <sup>4</sup>   | 2,713 (0,93)       | 5,174 (0,58)       | 2,399 (0,50)       | 5,260 (0,33)       | 2,641 (0,56)       | 5,026 (0,38)       |
| PAptg               | 0,503 (0,15)       | 0,695 (0,09)       | 0,306 (0,06)       | 0,790 (0,04)       | 0,391 (0,08)       | 0,735 (0,05)       |
| SF                  | 0,097 (0,04)       | 0,303 (0,03)       | 0,116 (0,025)      | 0,288 (0,01)       | 0,120 (0,03)       | 0,281 (0,02)       |
| Jug <sup>5</sup>    | 0,020 (0,01)       | 0,305 (0,02)       | 0,022 (0,01)       | 0,302 (0,01)       | 0,034 (0,02)       | 0,293 (0,02)       |
| Tern                | 0,086 (0,03)       | 0,264 (0,02)       | 0,101 (0,02)       | 0,250 (0,015)      | 0,096 (0,02)       | 0,250 (0,02)       |

$A_{22}$  matriz de relaciones Aditivas.  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado.  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia.  $\hat{\sigma}_u^2$  varianza genética aditiva.  $\hat{\sigma}_e^2$  varianza del error. <sup>1</sup>Caracteres de medidas de carcasa. <sup>2</sup>Caracteres de cortes primarios. <sup>3</sup>Caracteres de calidad de carne. <sup>4</sup>Caracteres de análisis de laboratorio. <sup>5</sup>Caracteres de análisis de paneles sensoriales. \* Diferencia estadística significativa entre modelo  $G_{IBS}$  vs  $A_{22}$ , y  $G_{IBD}$  vs  $A_{22}$  ( $\alpha = 0,05$ ;  $H_0: \hat{\sigma}_{uG}^2 - \hat{\sigma}_{uA}^2 = 0$ ).

### 3.3. Heredabilidad estimada en caracteres de composición de carcasa y calidad de la carne de cerdos.

Los estimadores de heredabilidad ( $\hat{h}^2$ ) ajustando cada modelo en caracteres de composición de carcasa y calidad de la carne de cerdos se reportan en la Tabla 3.4. Para los cuatro grupos de caracteres ‘medidas de carcasa’, ‘cortes primarios’, ‘calidad de carne’ y ‘análisis de laboratorio’ las heredabilidades estimadas fueron entre moderadas y altas, mientras que en el grupo ‘análisis de paneles sensoriales’ fueron entre bajas y moderadas. Asimismo, del conjunto de caracteres analizados, en dieciséis de ellos la proporción de la varianza genética estimada fue mayor con los modelos que emplean información genómica (modelos  $G_{IBS}$  y/o  $G_{IBD}$ ) que el modelo basado en la información de pedigree (modelo  $A_{22}$ ). En cambio, solo en cuatro caracteres los valores de  $\hat{h}^2$  obtenidos con el modelo  $A_{22}$  fueron mayores a aquellos encontrados con los modelos  $G_{IBS}$  y/o  $G_{IBD}$ . Estos resultados siguieron la misma tendencia que los reportados para la varianza aditiva en la Tabla 3.3.

En particular, los caracteres dentro de los grupos ‘medidas de carcasa’, ‘calidad de carne’ y ‘análisis de paneles sensoriales’ (Grupo 1, 3 y 5 en Tabla 3.4.), la diferencia en el valor de  $\hat{h}^2$  entre  $A_{22}$  y  $G_{IBS}$  y/o  $G_{IBD}$  es substancial. En estos grupos, el rango de  $\hat{h}^2$  ajustando  $A_{22}$  varió entre 0,15 a 0,40, en el caso del modelo  $G_{IBS}$  entre 0,06 a 0,54, y con  $G_{IBD}$  entre 0,10 a 0,61. Además, se puede observar que los estimadores del error estándar asociados al parámetro fueron de menor magnitud en los modelos genómicos que con el modelo convencional de pedigree. Por ejemplo, en el carácter *grasa dorsal medido en la décima costilla* (G10c), el valor de heredabilidad estimado en  $A_{22}$  fue igual a 0,15, a diferencia de  $G_{IBS}$  y  $G_{IBD}$  con los cuales se estimaron valores de 0,37 y 0,42, respectivamente. Similar resultado fue obtenido con el carácter *porcentaje de proteína* (Pptg), para el cual los valores de  $\hat{h}^2$  fueron 0,20 en  $A_{22}$ , 0,38 en  $G_{IBS}$  y 0,43 en  $G_{IBD}$ . En los dos caracteres anteriores, con los modelos genómicos el incremento de  $\hat{h}^2$  representa hasta 2 veces más el valor estimado de  $A_{22}$ .

Los grupos de caracteres ‘cortes primarios’ y ‘análisis de laboratorio’ (Grupo 2 y 4 Tabla 3.4) presentaron resultados contrastantes: en cinco de los nueve caracteres, la diferencia en las estimaciones de heredabilidad se mantuvo a favor de los modelos genómicos respecto del tradicional, aunque en menor magnitud. En cambio, en los caracteres *peso de la paleta* y *peso del costillar* (Bostkg, Costkg) pertenecientes al grupo de ‘cortes primarios’, así como en *rendimiento en cocción* y *porcentaje de pérdida de agua por goteo* (RCoc, PAptg) del grupo ‘análisis de laboratorio’, la diferencia fue a favor de la matriz de relaciones aditivas, siendo  $\hat{h}^2$  con  $A_{22}$  superior a las estimadas por los modelos  $G_{IBS}$  y/o  $G_{IBD}$ , pero con un error estándar de mayor magnitud. Por ejemplo, para el carácter Costkg, la  $\hat{h}^2$  fue igual a 0,50, 0,37 y 0,41 con  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$  y los errores estándar fueron 0,13, 0,05 y 0,06 respectivamente.

Respecto a la comparación de los estimadores de la heredabilidad entre modelos genómicos ( $G_{IBS}$  vs  $G_{IBD}$ ), los resultados no mostraron diferencias de magnitud, siendo la matriz de relaciones genómicas IBD (modelo  $G_{IBD}$ ) la que estimó, tanto  $h^2$  como errores estándar ligeramente mayores, respecto de las estimaciones para  $G_{IBS}$ , exceptuando los caracteres *ph 24 horas post-sacrificio* (ph24) y *Terneza* (Ter), para los cuales el resultado fue similar entre  $G_{IBS}$  y  $G_{IBD}$ .

**Tabla 3.4.** Heredabilidad estimada usando tres modelos lineales mixtos en caracteres de composición de carcasa y calidad de la carne de cerdos (Error estándar).

| Caracter            | Heredabilidad |              |              |
|---------------------|---------------|--------------|--------------|
|                     | $A_{22}$      | $G_{IBS}$    | $G_{IBD}$    |
| RCptg <sup>1</sup>  | 0,168 (0,07)  | 0,271 (0,05) | 0,293 (0,06) |
| pH24                | 0,164 (0,07)  | 0,189 (0,04) | 0,187 (0,05) |
| CarIng              | 0,384 (0,11)  | 0,468 (0,05) | 0,519 (0,06) |
| Glumb               | 0,256 (0,10)  | 0,400 (0,05) | 0,416 (0,06) |
| G10c                | 0,151 (0,07)  | 0,376 (0,05) | 0,417 (0,06) |
| ALm                 | 0,296 (0,09)  | 0,390 (0,05) | 0,436 (0,06) |
| Panckg <sup>2</sup> | 0,150 (0,07)  | 0,194 (0,04) | 0,223 (0,05) |
| Bostkg              | 0,396 (0,11)  | 0,244 (0,05) | 0,278 (0,06) |
| Jamkg               | 0,313 (0,10)  | 0,484 (0,05) | 0,513 (0,06) |
| Lomkg               | 0,154 (0,07)  | 0,300 (0,05) | 0,344 (0,06) |
| Picnkg              | 0,128 (0,06)  | 0,151 (0,04) | 0,160 (0,05) |
| Costkg              | 0,506 (0,13)  | 0,373 (0,05) | 0,413 (0,06) |
| Grptg <sup>3</sup>  | 0,405 (0,12)  | 0,539 (0,05) | 0,615 (0,06) |
| Humptg              | 0,319 (0,10)  | 0,383 (0,05) | 0,439 (0,06) |
| Protptg             | 0,202 (0,1)   | 0,380 (0,05) | 0,433 (0,06) |
| RCoc <sup>4</sup>   | 0,344 (0,10)  | 0,313 (0,05) | 0,344 (0,06) |
| PAptg               | 0,419 (0,12)  | 0,279 (0,05) | 0,347 (0,06) |
| SF                  | 0,242 (0,10)  | 0,287 (0,05) | 0,298 (0,06) |
| Jug <sup>5</sup>    | 0,062 (0,05)  | 0,069 (0,03) | 0,105 (0,04) |
| Tern                | 0,246 (0,09)  | 0,287 (0,05) | 0,278 (0,06) |

$A_{22}$  matriz de relaciones Aditivas.  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado.  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia. <sup>1</sup>Caracteres de medidas de carcasa. <sup>2</sup>Caracteres de cortes primarios. <sup>3</sup>Caracteres de calidad de carne. <sup>4</sup>Caracteres de análisis de laboratorio. <sup>5</sup>Caracteres de análisis de paneles sensoriales.

### 3.4. Habilidad predictiva de los modelos.

El coeficiente de correlación de Pearson entre el fenotipo observado y el fenotipo predicho ( $\hat{r}_{y,\hat{y}}$ ), fue la medida de prueba usada para evaluar la habilidad predictiva de cada modelo mediante validación cruzada. Los resultados se muestran en la Tabla 3.5. Como resultado global, sin considerar las matrices de parentesco empleadas en cada modelo, la habilidad predictiva estimada fue alta en el grupo de caracteres ‘cortes primarios’ (0,80 – 0,87), moderada en los caracteres ‘medida de carcasa’ (0,46 – 0,7), ‘calidad de carne’ (0,37 – 0,61) y ‘análisis de laboratorio’ (0,34 – 0,53), y baja en los caracteres de ‘análisis de paneles sensoriales’ (0,21 – 0,40). Por otra parte, en términos de selección de modelos, la habilidad predictiva de los modelos que incluyen la información genómica (modelos  $G_{IBS}$  y/o  $G_{IBD}$ ) fueron significativamente superiores ( $P < 0,05$ ) comparados con el modelo basado en la información de pedigree (modelo  $A_{22}$ ). El resultado anterior se obtuvo en el 90% del conjunto de caracteres (18 de 20), los cuales pertenecen a los grupos ‘medidas de carcasa’, ‘calidad de carne’, ‘análisis de laboratorio’ y ‘análisis de paneles sensoriales’. El incremento en la habilidad predictiva obtenido por usar matrices genómicas (modelo  $G_{IBS}$  vs modelo  $G_{IBD}$ ) respecto de la matriz de pedigree (modelo  $A_{22}$ ), fue más notorio en el grupo de caracteres ‘calidad de carne’, el cual estuvo entre un rango de 14% a 37%, y de menor magnitud en el grupo ‘cortes primarios’ donde el rango varía entre 0,6% y 4%. En los caracteres *peso de la paleta* (Bostkg, Picnkg) integrados en el grupo de ‘cortes primarios’ no se encontraron diferencias significativas ( $P > 0,05$ ) en  $\hat{r}_{y,\hat{y}}$  entre los modelos. Ahora bien, comparando los resultados entre modelos genómicos ( $G_{IBS}$  vs  $G_{IBD}$ ), se puede detallar que en 14 caracteres, la habilidad predictiva medida como el coeficiente  $\hat{r}_{y,\hat{y}}$  fue significativamente mejor ( $P < 0,05$ ) con el modelo  $G_{IBS}$ , comparada con la  $\hat{r}_{y,\hat{y}}$  del modelo  $G_{IBD}$ , mostrando que en este conjunto de caracteres el incremento al ajustar el modelo con la matriz  $G_{IBS}$  respecto del modelo con la matriz  $G_{IBD}$  puede variar en un rango de 0,1% a 5,5%. Adicionalmente, en 4 caracteres (Jamkg, Lomkg, Humptg, Jug) no hubo diferencias significativas entre las  $\hat{r}_{y,\hat{y}}$  de los modelos genómicos.

En resumen, empleando la información fenotípica, genealógica y de marcadores de la población experimental de cerdos analizada en esta investigación, los resultados en su conjunto mostraron, en primer lugar, que, en aquellos caracteres donde los modelos genómicos estimaron un componente de varianza aditivo estadísticamente significativo (Tabla 3.3) y mayor heredabilidad (Tabla 3.4), también mostraron en la validación una mejor habilidad predictiva (Tabla 3.5), en comparación con el modelo que usó la matriz de relaciones aditivas. En segundo lugar, al evaluar el desempeño entre los modelos incluyendo las matrices genómicas ( $G_{IBD}$  vs  $G_{IBS}$ ), los caracteres que obtuvieron estimaciones ligeramente superiores de la varianza aditiva (estadísticamente no significativas) y de heredabilidad con  $G_{IBD}$ , no mostraron una habilidad predictiva significativamente superior que con  $G_{IBS}$ . Tercero, en aquellos caracteres donde los estimadores de  $\sigma_u^2$  y de  $h^2$  fueron mayores con  $A_{22}$  que con  $G_{IBS}$  o  $G_{IBD}$ , tanto los estimadores de  $\sigma_u^2$  como los de  $r_{y,\hat{y}}$  no fueron estadísticamente diferentes entre los tres modelos, de acuerdo con la prueba de hipótesis entre varianzas y la prueba de comparación de medias. Finalmente, en caracteres donde la heredabilidad estimada por los modelos es baja, la habilidad predictiva fue significativamente mejor con los modelos genómicos que con el modelo de pedigree.



**Tabla 3.5.** Habilidad predictiva ( $\hat{r}_{y,\hat{y}}$ ) de los diferentes modelos estimada por validación cruzada (Error estándar).

| Caracter            | Habilidad predictiva ( $\hat{r}_{y,\hat{y}}$ ) |                             |                              |
|---------------------|--|-----------------------------|------------------------------|
|                     | Modelo $A_{22}$                                | Modelo $G_{IBS}$            | Modelo $G_{IBD}$             |
| RCptg <sup>1</sup>  | 0,5980 (0,002) <sup>c</sup>                    | 0,6406 (0,002) <sup>a</sup> | 0,6235 (0,002) <sup>b</sup>  |
| pH24                | 0,4634 (0,001) <sup>c</sup>                    | 0,4801 (0,001) <sup>a</sup> | 0,4701 (0,001) <sup>b</sup>  |
| Carlng              | 0,6538 (0,003) <sup>c</sup>                    | 0,7286 (0,003) <sup>a</sup> | 0,7132 (0,003) <sup>b</sup>  |
| Glumb               | 0,5407 (0,002) <sup>c</sup>                    | 0,6192 (0,002) <sup>a</sup> | 0,5932 (0,002) <sup>b</sup>  |
| G10c                | 0,7299 (0,002) <sup>c</sup>                    | 0,7687 (0,002) <sup>a</sup> | 0,7556 (0,002) <sup>b</sup>  |
| ALm                 | 0,6643 (0,002) <sup>c</sup>                    | 0,7141 (0,002) <sup>a</sup> | 0,7061 (0,002) <sup>b</sup>  |
| Panckg <sup>2</sup> | 0,852 (0,0001) <sup>c</sup>                    | 0,857 (0,0001) <sup>a</sup> | 0,856 (0,0001) <sup>b</sup>  |
| Bostkg              | 0,8845 (0,002) <sup>a</sup>                    | 0,8788 (0,002) <sup>a</sup> | 0,8851 (0,002) <sup>a</sup>  |
| Jamkg               | 0,8276 (0,002) <sup>b</sup>                    | 0,8642 (0,002) <sup>a</sup> | 0,8592 (0,002) <sup>a</sup>  |
| Lomkg               | 0,7980 (0,002) <sup>b</sup>                    | 0,8234 (0,002) <sup>a</sup> | 0,8181 (0,002) <sup>a</sup>  |
| Picnkg              | 0,8721 (0,001) <sup>a</sup>                    | 0,8723 (0,001) <sup>a</sup> | 0,8737 (0,0001) <sup>a</sup> |
| Costkg              | 0,808 (0,0001) <sup>c</sup>                    | 0,825 (0,0001) <sup>a</sup> | 0,819 (0,0001) <sup>b</sup>  |
| Grptg <sup>3</sup>  | 0,4456 (0,001) <sup>c</sup>                    | 0,6111 (0,001) <sup>a</sup> | 0,5981 (0,001) <sup>b</sup>  |
| Humptg              | 0,5032 (0,002) <sup>b</sup>                    | 0,5733 (0,002) <sup>a</sup> | 0,5710 (0,002) <sup>a</sup>  |
| Protptg             | 0,3787 (0,001) <sup>c</sup>                    | 0,4878 (0,001) <sup>a</sup> | 0,4746 (0,001) <sup>b</sup>  |
| RCoc <sup>4</sup>   | 0,3473 (0,001) <sup>c</sup>                    | 0,4046 (0,001) <sup>a</sup> | 0,3954 (0,001) <sup>b</sup>  |
| PAptg               | 0,5028 (0,001) <sup>c</sup>                    | 0,5309 (0,001) <sup>a</sup> | 0,5234 (0,001) <sup>b</sup>  |
| SF                  | 0,4361 (0,001) <sup>c</sup>                    | 0,4747 (0,001) <sup>a</sup> | 0,4657 (0,001) <sup>b</sup>  |
| Jug <sup>5</sup>    | 0,2158 (0,001) <sup>b</sup>                    | 0,2240 (0,001) <sup>a</sup> | 0,2264 (0,001) <sup>a</sup>  |
| Tern                | 0,3548 (0,001) <sup>c</sup>                    | 0,4044 (0,001) <sup>a</sup> | 0,3830 (0,001) <sup>b</sup>  |

$A_{22}$  matriz de relaciones Aditivas.  $G_{IBS}$  matriz de relaciones genómicas de identidad por estado.  $G_{IBD}$  matriz de relaciones genómicas de identidad por descendencia. <sup>1</sup>Caracteres de medidas de carcasa. <sup>2</sup>Caracteres de cortes primarios. <sup>3</sup>Caracteres de calidad de carne. <sup>4</sup>Caracteres de análisis de laboratorio. <sup>5</sup>Caracteres de análisis de paneles sensoriales. Tukey múltiple comparación, valores con distintas letras (a, b, c) son significativamente diferentes ( $\alpha = 0,05$ ).

## **Capítulo 4. *Discusión***

## Capítulo 4

### Discusión

En este trabajo se evaluaron tres modelos animales en términos de estimación de los componentes de varianza, heredabilidad y habilidad predictiva, en caracteres de composición de carcasa y calidad de la carne registrados en una población experimental de cerdos de la generación F<sub>2</sub>. Esencialmente, los tres modelos animales difieren en la especificación de la matriz de varianzas y covarianzas entre los valores de cría. Las diferencias se dieron en el modo de estimar las relaciones de parentesco, ya sea: 1) utilizando información únicamente de pedigree (modelo  $A_{22}$ ), 2) utilizando información únicamente de marcadores moleculares y basándose en el enfoque de identidad por estado (modelo  $G_{IBS}$ ), o 3) condicionando la información de genómica en la del pedigree, basándose en el enfoque de identidad por descendencia (modelo  $G_{IBD}$ ). En este capítulo, se discutirán los principales resultados obtenidos, siguiendo el orden de su presentación en el capítulo anterior. Primero se hará referencia al análisis descriptivo de las matrices de relaciones de parentesco estimadas, segundo a los componentes de varianza estimados con cada modelo en caracteres de composición de carcasa y calidad de la carne de cerdos, tercero a la heredabilidad estimada en cada carácter y finalmente a la habilidad predictiva de los modelos ajustados.

#### 4.1. Análisis descriptivo de las matrices de relaciones de parentesco estimadas.

Los estadísticos de la distribución de los elementos diagonales y no diagonales de las diferentes matrices de parentesco evaluadas ( $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$ ) mostraron mayor variación en las matrices genómicas ( $G_{IBS}$  y  $G_{IBD}$ ) que en la matriz de relaciones aditivas basada en el pedigree ( $A_{22}$ ). Este resultado era esperable, ya que las relaciones genómicas estiman la proporción ‘realizada’ del genoma que dos individuos comparten, en torno a los valores esperados presentes en la matriz de pedigree (Goddard *et al.*, 2011). A su vez, el rango de variación fue mayor con  $G_{IBS}$  que con  $G_{IBD}$ . En el caso de la matriz  $G_{IBD}$ , los valores realizados del parentesco se obtienen a partir de probabilidades de IBD condicionales en la información genómica y de pedigree (Han y Abney, 2011, 2013); como probabilidades condicionales al pedigree, se acota el rango de variación. En esta investigación la matriz  $G_{IBS}$  fue calculada mediante el primer método presentado en VanRaden (2008), donde el denominador escala la matriz para que sea similar a  $A$ . Esta formulación puede generar algunos elementos fuera de la diagonal negativos, elementos diagonales menores que 1 o que la matriz sea semidefinida positiva con autovalores mayores o iguales a 0, como fue observado en nuestros resultados de la Tabla 3.1 y Figura 3.1. Tales elementos negativos indican que algunos individuos están molecularmente menos correlacionados respecto de lo esperado si la población estuviera en equilibrio Hardy-Weinberg (Yu *et al.*, 2017; Toro *et al.*, 2002).  $G_{IBS}$  es positiva semidefinida y puede ser singular si el número de loci es limitado o dos individuos tienen idénticos genotipos en todos los marcadores (Forni *et al.*, 2011; VanRaden, 2008). Estos resultados van en línea con lo reportado en otros estudios, en los que, empleando datos reales y/o simulados, se obtuvieron estimadores de relaciones de parentesco genómicas con mayor variabilidad cuando éstos se calcularon empleando el enfoque de IBS antes que el de IBD (Solaymani *et al.*, 2020; Forneris *et al.*, 2016; García-Baccino *et al.*, 2016).

Los coeficientes de regresión de  $G_{IBS}$  o  $G_{IBD}$  sobre  $A_{22}$  mostraron que  $G_{IBS}$  arrojó, en promedio, relaciones genómicas más lejanas que lo esperado por pedigree, probablemente, por ser más sensibles a las frecuencias alélicas de la muestra. Asimismo, el coeficiente de correlación fue más alto entre elementos de  $G_{IBD}$  y  $A_{22}$ , que entre  $G_{IBS}$  y  $A_{22}$ , debido a que, por construcción, la primera matriz está condicionada en la información del pedigree, mientras que la segunda no. El mismo patrón en los coeficientes de regresión y correlación fue observado por Meuwissen *et al.* (2014) al comparar matrices de parentesco genómico en una población de cerdos más numerosa.

#### 4.2. Componentes de varianza estimados en caracteres de composición de carcasa y calidad de la carne de cerdos.

El ajuste de los modelos  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$  permitió estimar la varianza aditiva y residual en 20 caracteres. En el 80% de ellos, la varianza genética aditiva estimada con la información genómica ( $G_{IBS}$ ) o con información genómica y pedigree ( $G_{IBD}$ ) fue de mayor magnitud y más precisa (menor error estándar) respecto del modelo  $A_{22}$ . Si bien existen antecedentes de estimaciones de componentes de varianza en este tipo de caracteres en cerdos, aun no se han reportado estimaciones que incorporen el enfoque de IBD genómica. Los modelos ajustados utilizan información del pedigree, o bien IBS o una combinación de ambas (Willson *et al.*, 2020; Iqbal *et al.*, 2019; Khanal *et al.*, 2019; Wijk *et al.*, 2005). Por su parte, existen, en otros caracteres, estudios de estimaciones de componentes de varianza que comparen matrices genómicas IBS vs IBD (Clark *et al.*, 2013; Hickey *et al.*, 2012; Luan *et al.*, 2012). Estos estudios coinciden en el cálculo de  $G_{IBS}$ , pero difieren en el modo de calcular  $G_{IBD}$  respecto de las metodologías propuestas en esta tesis; para calcular  $G_{IBD}$  utilizan, por ejemplo, similitud de haplotipos (Hickey *et al.*, 2012) o análisis de ligamiento (Luan *et al.*, 2012). Particularmente, en Hickey *et al.* (2012) los datos fueron simulados, mientras que en Luan *et al.* (2012) y Clark *et al.* (2013), se estimaron con los tres modelos los componentes de varianza en caracteres de producción de leche en bovinos y “espesor del ojo de bife” en ovinos, respectivamente. En sus estudios, los citados autores encontraron que los modelos genómicos recuperaban más varianza que el modelo  $A_{22}$ , explicando más varianza el modelo  $G_{IBD}$  propuesto por ellos que los modelos  $G_{IBS}$  y  $A_{22}$ , resultados que son similares a los obtenidos en la presente investigación. Sin embargo, a diferencia de nuestro trabajo, en dichos estudios no se realizó una prueba de hipótesis que confirmara la significancia del estimador de la varianza aditiva entre  $G_{IBS}$  y/o  $G_{IBD}$  versus  $A_{22}$ , o entre  $G_{IBS}$  versus  $G_{IBD}$ .

Un resultado importante de esta tesis ha sido la realización de una prueba de hipótesis de las diferencias entre los valores estimados de la varianza aditiva con cada uno de los modelos evaluados. La diferencia entre los modelos genómicos vs  $A_{22}$  fue, en su mayoría, no significativa. Esto puede deberse a que, en los tres métodos utilizados, el estimador capturó el *mismo tipo* de variabilidad genética. Esta última es cuantificada por los componentes de varianza, los cuales son funciones lineales del parecido genético entre individuos dentro de la población, estimado por las relaciones aditivas entre dichos individuos (Cockerham, 1954; Kempthorne, 1955). Sin embargo, para cierta relación esperada existe variabilidad en la relación ‘realizada’, debido a procesos estocásticos que ocurren durante la transmisión del ADN, como son la segregación mendeliana, la recombinación, la consanguinidad, el ligamiento físico entre

genes causales y el desequilibrio gamético o LD (Hill y Weir, 2011, 2012). En poblaciones finitas, Weir y Cockerham (1968, 1969) estudiaron los efectos del ligamiento sobre la probabilidad de IBD ( $P(\text{IBD})$ ) para dos genes ligados. Invocando un resultado de J.B.S. Haldane en 1949, estos autores describieron una función del *desequilibrio de identidad* (ID) que resulta de observar que la  $P(\text{IBD})$  conjunta entre dos loci ( $P_C(\text{IBD})$ ) es estrictamente mayor que el producto de las  $P(\text{IBD})$  en cada uno de ambos loci ( $P_D(\text{IBD})$ ); representando el símbolo ‘ $\equiv$ ’ al evento en que los genes causales de los individuos 1 y 2 sean IBD, la expresión formal de la mencionada desigualdad es la siguiente:

$$P(X_1 \equiv Y_1, X_2 \equiv Y_2) > P(X_1 \equiv Y_1)P(X_2 \equiv Y_2)$$

Esta expresión indica que la  $P_C(\text{IBD})$  entre dos loci ligados es siempre mayor que la  $P_D(\text{IBD})$  entre loci independientes. Consecuentemente, las relaciones de parentesco calculadas por *segmentos* entre sitios ligados diferenciarán entre parentescos con la misma relación  $A_{XY}$  para loci independientes dado que poseen más *información*. Es así como las relaciones abuelo(a) – nieto(a), tío(a) – sobrino(a) y medios hermanos, todas las cuales son iguales a  $A_{XY} = 0.25$  en ausencia de consanguinidad, son diferenciables para distintos valores de la tasa de recombinación (Thompson, 1988) o de la probabilidad de integridad gamética (Saavedra *et al.*, 2020). Guo (1995) utilizó el enfoque de la  $P_C(\text{IBD})$  como proveniente de una distribución *continua* (Donnelly, 1983) para definir la relación de parentesco como la proporción de genoma compartido IBD entre dos individuos emparentados (IBDP). Para asociar esta definición con el proceso genealógico y el genético de transmisión del genoma entre generaciones, Cantet (2020, comunicación personal) atribuye a la  $P_D(\text{IBD})$  discreta (en un punto o loci cualquiera del genoma) de ser la medida esperada del parecido genético que cuantifica al proceso ancestral; mientras que la  $P_C(\text{IBD})$  cuantifica al proceso genético o ID.

Para lograr diferencias sustanciales entre estimaciones de las relaciones de parentesco con información genómica, y consecuentemente, entre los componentes de varianza, es necesario emplear estadísticos cuyo valor esperado reflejen los dos procesos de transmisión arriba mencionados y no sólo el ancestral. Csurös (2013) observó que con marcadores bialélicos como lo son los SNP, sólo es posible estimar funciones lineales del conjunto de nueve (9)  $P_D(\text{IBD})$  definidas por Jacquard (1972), tales como el coeficiente de coancestría de Malecot (1948) o el de consanguinidad de uno o de ambos individuos en la relación. La mayoría de las comparaciones que existen en la literatura de métodos que estiman relaciones genómicas, por ejemplo, los de García-Baccino *et al* (2016), tienen como valor esperado al proceso ancestral pero no el ID. Asimismo, en esta tesis, los dos estadísticos utilizados para estimar relaciones de parentesco genómicas pueden considerarse estimadores puntuales de funciones lineales de las  $P_D(\text{IBD})$ , ya que estiman los valores realizados del coeficiente de coancestría de Malecot. Por lo tanto, los valores esperados de las diferencias entre modelos, i.e.  $E(\mathbf{G}_{IBS} - \mathbf{A}_{22})$ ,  $E(\mathbf{G}_{IBD} - \mathbf{A}_{22})$  y/o  $E(\mathbf{G}_{IBS} - \mathbf{G}_{IBD})$ , son iguales a cero y la variabilidad observada es sinónimo de error de estimación *pero no de ID*. Entonces, en los nueve caracteres donde se encontraron diferencias entre los modelos  $\mathbf{G}_{IBS}$  y/o  $\mathbf{G}_{IBD}$  versus el modelo  $\mathbf{A}_{22}$ , la variabilidad podría atribuirse a error de estimación en las relaciones de parentesco *pero no al ID*. En cuanto a los otros caracteres, la varianza del error de estimación fue similar entre métodos, y por lo tanto no se encontró ninguna diferencia entre los modelos evaluados. Por lo tanto, para estimar las  $P_C(\text{IBD})$  sería necesario dejar de lado estimadores de relaciones de parentesco en un sitio específico del genoma

asumiendo independencia (herencia puntual) y considerar la *herencia en segmentos cromosómicos*, estimando el largo de los segmentos IBD mediante la detección de los eventos de recombinación.

### 4.3. Heredabilidad estimada en caracteres de composición de carcasa y calidad de la carne de cerdos.

La estimación de la heredabilidad ( $\hat{h}^2$ ) de un caracter es importante porque determina la precisión con la cual el valor genético puede ser predicho desde la información fenotípica, delimitando los esquemas de mejoramiento animal (Visscher *et al.*, 2006). El conjunto de caracteres analizados mostró heredabilidades variadas entre bajas y altas, y en el 80% de los caracteres hubo un incremento y mejor precisión de estimación con los modelos  $G_{IBS}$  y  $G_{IBD}$  que con el modelo  $A_{22}$ . Los estimadores de la heredabilidad de este conjunto de datos ya han sido reportados por Casiró *et al.* (2017) ajustando solamente el modelo  $G_{IBS}$ . En otros estudios con poblaciones de cerdos puros o cruza, la heredabilidad de caracteres de composición de carcasa y calidad de la carne ha sido estimada empleando  $A_{22}$  y/o  $G_{IBS}$  (Willson *et al.*, 2020; Iqbal *et al.*, 2019; Miar *et al.*, 2014; Wijk *et al.*, 2005), o combinando en una matriz de parentesco la información del pedigree y genómica IBS (Khanal *et al.*, 2019). En dichos estudios se obtuvieron estimadores de similar magnitud al obtenido por nosotros con el modelo  $A_{22}$  y el modelo  $G_{IBS}$ . Sin embargo, a la fecha no hay estudios que empleen el modelo  $G_{IBD}$ , lo cual dificulta la comparación con otros estudios.

Nuestros estimadores de heredabilidad para cada caracter siguieron el mismo patrón que los estimadores de los componentes de varianza en cada modelo ajustado. Hayes y Goddard (2008) mostraron que utilizando un número suficiente de marcadores la heredabilidad estimada puede ser más exacta que empleando la información de pedigree, porque la información genómica no está sujeta a los errores del pedigree. En contraste, en aquellos caracteres donde el modelo  $A_{22}$  estima mayor heredabilidad que los modelos genómicos, pareciera existir un problema de “heredabilidad perdida” (en inglés “missing heritability”), el cual posiblemente se deba al imperfecto LD entre marcadores y QTL (de los Campos *et al.*, 2013, 2015). Sin embargo, teniendo en cuenta que no se encontró significancia estadística entre los estimadores de la varianza aditiva para  $A_{22}$ ,  $G_{IBS}$  y/o  $G_{IBD}$ , este resultado podría deberse a la variabilidad observada en el error de estimación de las matrices genómicas. Comparando las matrices  $G_{IBS}$  y  $G_{IBD}$  ésta última provee un estimador más exacto de las relaciones de parentesco que la primera (i.e., la mejor estimación de  $\hat{h}^2$  fue con  $G_{IBD}$  que con  $G_{IBS}$ ). Sin embargo, nuestros resultados mostraron similares valores de  $\hat{h}^2$  con ambos modelos, resultado congruente con lo mencionado en la sección anterior, dado que a pesar que el método IBD es más eficiente que el IBS, las estimaciones de las  $P(\text{IBD})$  son para un “punto” (en inglés “single point”), es decir, condicionan en la información genotípica de *un sólo locus* asumiéndose que todos ellos segregan de manera independiente. Nuestros resultados son similares a los obtenidos por Clark *et al.* (2013), quienes con datos reales de una población ovina reportaron estimadores de heredabilidad similares entre modelos  $G_{IBS}$  y  $G_{IBD}$ .

La imputación de genotipos podría llevar a errores en las estimaciones de los parentescos genómicos y consecuentemente en las estimaciones de la heredabilidad. Sin embargo, en este estudio, se utilizó una base de datos de tres generaciones de animales genotipados en la cual sólo una parte de los animales de la última generación (el 64.5%)

fueron imputados de un chip de 9K a uno de 60K. En un estudio previo en la misma población, Gualdrón Duarte *et al.* (2013) reportaron una exactitud de la imputación de 0,99 al utilizar el chip comercial de 9K en los animales F<sub>2</sub>. La condición que aseguró esta alta exactitud fue el hecho de contar con genotipos F<sub>0</sub> y F<sub>1</sub> genotipados a alta densidad, siendo estos abuelos y padres, respectivamente, de los animales F<sub>2</sub>.

#### 4.4. Habilidad predictiva de los modelos.

Uno de los objetivos de este trabajo fue evaluar el desempeño de los diferentes modelos en términos de su habilidad predictiva mediante la metodología de validación cruzada, ya que se trata del fin último de las valoraciones genéticas y genómicas en mejoramiento animal. (Erbe *et al.*, 2010; Legarra *et al.*, 2008; Luan *et al.*, 2009). Nuestros resultados mostraron que la habilidad predictiva de los modelos que usan la información genómica (modelos  $G_{IBS}$  y/o  $G_{IBD}$ ) es significativamente mayor que el modelo convencional de pedigree ( $A_{22}$ ) en los caracteres evaluados. Este resultado es similar a los obtenidos mediante simulación estocástica (Hickey *et al.*, 2012; Vela-Avitúa *et al.*, 2015; Forneris *et al.*, 2016) o datos reales (Meuwissen *et al.*, 2014), no obstante, el modo en que estos autores calculan la matriz  $G_{IBD}$  es distinta a la propuesta en esta tesis. Así mismo, Daetwyler *et al.* (2012) en caracteres de calidad de carne y carcasa medidos en una población ovina, calcularon la exactitud de predicción como una correlación entre BVs predichos y fenotipos observados o ajustados, y obtuvieron mejores resultados con el modelo IBS que con el modelo de pedigree, aunque ellos calcularon la matriz de relaciones IBS mediante otra metodología. Por otra parte, Iqbal *et al.* (2019) ajustaron el modelo  $G_{IBS}$  en caracteres de peso de la canal y grasa dorsal para tres razas de cerdos y obtuvieron mediante validación cruzada estimaciones de la exactitud calculada como correlación entre el BV genómico y el fenotipo individual entre 0,09 y 0,27 para peso de la canal y entre 0,30 y 0,46 para grasa dorsal. En esta tesis el peso de la canal es considerado un efecto fijo del modelo, existiendo un grupo de caracteres definido como ‘medidas de carcasa’ para los cuales la habilidad predictiva se estimó dentro de un rango de 0,48 a 0,76 con  $G_{IBS}$ .

En la presente investigación la habilidad predictiva del 70% de los caracteres fue significativamente mayor con el modelo  $G_{IBS}$  que con el modelo  $G_{IBD}$ , mientras que en los demás caracteres tuvieron el mismo desempeño. Una limitación de la habilidad predictiva, como estadístico de selección de modelos, es que sus resultados dependen del archivo de datos empleados en el cálculo, por lo que no permite generalizar resultados. La predicción real aplicada por los mejoradores empleando BLUP es condicional a un modelo de *referencia* y el ajuste es independiente de los datos recogidos en un año u otro. En tal sentido los estudios de simulación o con datos reales han reportado resultados diversos al comparar las dos metodologías, debido a dicha dependencia. Por ejemplo, resultados en línea con los nuestros fueron encontrados por Meuwissen *et al.* (2014), donde el  $G_{IBS}$  fue significativamente mejor que  $G_{IBD}$ . Asimismo, en la simulación de Vela-Avitúa *et al.* (2015) el modelo IBS fue más exacto que el modelo IBD, aunque el primero fue más sensible a la reducción en la densidad de marcadores. Debe aclararse que el IBD calculado por dichos autores no tiene en cuenta el desequilibrio gamético. En tanto Hickey *et al.* (2012) y Clark *et al.* (2013) no encontraron diferencia entre los métodos IBS e IBD. Por su parte en el trabajo de Forneris *et al.* (2016), las diferencias encontradas en precisión entre los estimadores de las relaciones genómicas se diluyeron cuando se compararon las exactitudes de predicción de los BV entre métodos.

La mejor performance de los modelos genómicos ( $G_{IBS}$  y/o  $G_{IBD}$ ) sobre el modelo de pedigree ( $A_{22}$ ), es probablemente debida al hecho que ellos capturan mejor el término de segregación mendeliana (Daetwyler *et al.*, 2013), particularmente a través de los colaterales (medio hermanos, hermanos enteros) del animal. Además, de los Campos *et al.* (2013) encontraron que la efectividad de los modelos genómicos depende de cuán bien las relaciones genómicas derivadas de los marcadores describen la relación genética ‘realizada’ en el loci causal no observado. De acuerdo con Habier *et al.* (2013) la información del pedigree, cosegregación y el LD poblacional contribuyen a la habilidad predictiva de los modelos genómicos. La información del pedigree es capturada directamente por los modelos  $A_{22}$  y  $G_{IBD}$ , mientras  $G_{IBS}$  usa los genotipos marcadores que son heredados. La cosegregación es una desviación de la segregación independiente de alelos en la misma gameta de loci ligados, que puede ser capturada por las relaciones genómicas IBS o IBD, especialmente con colaterales. Es un supuesto importante en selección genómica que los marcadores deban estar en suficiente LD con el QTL (Meuwissen *et al.*, 2001). El nivel de LD entre marcadores y QTL o entre marcadores, puede ser cuantificado como el cuadrado del coeficiente de correlación entre dos loci ( $r^2$ ). Este parámetro es el calculado comúnmente, porque da una medida directa de la proporción de la varianza en un locus que puede ser predicha desde el marcador (Hill, 1981). En la población analizada se observó una lenta caída del LD sobre la distancia, encontrándose un  $r^2$  promedio igual a 0.28 dentro de 1 megabase (resultados no mostrados). Los resultados de simulaciones realizadas por Calus *et al.* (2008) y Meuwissen *et al.* (2001) sugieren que un nivel de  $r^2$  mayor de 0.20 es suficiente para lograr altas exactitudes. En tal escenario, nuestros resultados sugieren que la información del LD ayudó a obtener altas habilidades predictivas. Por otra parte, la habilidad predictiva es influenciada por factores como el tamaño del “training set” y “test set”, las relaciones de parentesco entre individuos del “training set” y “test set”, la heredabilidad del carácter y la densidad de marcadores (Karaman *et al.*, 2016; Clark *et al.*, 2013; Hickey *et al.*, 2012; Hayes *et al.*, 2009; Habier *et al.*, 2007). Se pueden calcular correlaciones elevadas con un “training set” grande y un “test set” pequeño, siendo la subdivisión de 5 *folds* un buen compromiso al usar el 20% de los datos en el “test set” (Erbe *et al.*, 2010). Varios estudios han resaltado la importancia de las relaciones sobre la predicción de los modelos genómicos mostrando que, a medida que la relaciones entre individuos del “training set” y “test set” son distantes o no hay relación de parentesco, la exactitud de la predicción es reducida (Clark *et al.*, 2013; Clark *et al.*, 2012; Habier *et al.*, 2010; Habier *et al.*, 2007). En nuestro estudio, la elección de los animales del “training set” y “test set” es aleatoria, y no tiene en cuenta la relación de parentesco. Los datos usados para la validación cruzada son de los individuos de la generación  $F_2$ , en la cual el rango de las relaciones de parentesco entre individuos está entre 0 y 0,75,  $-0,276$  y  $0,9$ ,  $-0,26$  y  $0,92$  para  $A_{22}$ ,  $G_{IBS}$  y  $G_{IBD}$  respectivamente. Al definir aleatoriamente el subconjunto de animales para la validación cruzada, la habilidad predictiva de los caracteres fue alta. Se han reportado resultados similares cuando se investigó la predicción genómica de caracteres de carcasa y calidad de carne en ovejas (Brito *et al.*, 2017; Daetwyler *et al.*, 2012). La heredabilidad de los caracteres también es crucial, dado que distintos estudios mostraron que a mayor heredabilidad del carácter, mayor es la exactitud de predicción (Vela-Avitúa *et al.*, 2015; Calus *et al.*, 2008). En esta investigación, se encontró que caracteres en heredabilidad moderada (0,25-0,50) la habilidad predictiva está en un rango de 0,46 a 0,88, en tanto que caracteres de baja heredabilidad (menor a 0,25) la habilidad predictiva varió entre 0,22 y 0,48.



Finalmente, la mejor habilidad predictiva obtenida con  $G_{IBS}$  respecto de  $G_{IBD}$  parecería no deberse a una mayor heredabilidad estimada, ni tampoco a una mayor correlación con los valores esperados por pedigree. Pareciera que la razón es la de lograr una mayor variación en los parentescos genómicos involucrados en la ecuación de regresión del valor de cría de un animal a una matriz más densa (más relaciones distintas de cero al no condicionar en el pedigree), lo que implicaría, para un animal, un número mayor de fuentes de información y más variables entre sí, con la consecuente ganancia en precisión. Es importante destacar que los resultados de la validación cruzada, como de cualquier otro criterio de selección de modelos, son condicionales a los datos empleados en el análisis. Esta no es la situación que se utiliza en la evaluación genética empleando modelos mixtos y BLUP donde el modelo de análisis es considerado de referencia y no suele ser modificado en el tiempo. En estas circunstancias, la robustez y consistencia de las predicciones entre años son esenciales para la credibilidad del criador y la predicción BLUP con matrices IBS, e incluso el pedigree (Legarra *et al.*, 2009) ha mostrado ser inestable entre años debido a la información incompleta para calcular la matriz de relaciones genómicas. En este sentido, modelos recientes como la regresión ancestral (Cantet *et al.*, 2017), sobre la base de IBD calculado por segmentos brindando una descripción dinámica de la herencia entre generaciones en individuos diploides, parecería ser una alternativa promisoría para la generación de una matriz IBD que produzca predicciones del valor de cría que maximicen la respuesta a la selección y sean consistentes en el tiempo.

## **Capítulo 5. Conclusiones**

## Capítulo 5

### Conclusiones

Empleando la información fenotípica, genealógica y genómica de la población experimental de cerdos analizada en esta tesis, se pudo concluir, en primer lugar, que, al comparar los modelos genómicos con el modelo convencional, en aquellos caracteres donde los primeros estimaron un componente de varianza aditiva significativamente mayor y una mayor heredabilidad, también mostraron una mejor habilidad predictiva mediante validación cruzada. En segundo lugar, al comparar el desempeño entre modelos genómicos ( $G_{IBD}$  vs  $G_{IBS}$ ), las estimaciones ligeramente superiores de la varianza aditiva (estadísticamente no significativas) y de heredabilidad obtenidas con  $G_{IBD}$  no redundaron en una habilidad predictiva significativamente superior. Tercero, en caracteres donde la varianza aditiva estimada fue mayor con el modelo  $A_{22}$  que con los modelos genómicos, la diferencia no resultó significativa y la exactitud obtenida por validación fue similar entre modelos. Finalmente, en caracteres donde la heredabilidad estimada por las distintas matrices fue baja, la habilidad predictiva fue significativamente superior con los modelos genómicos que con el modelo de pedigree. Nuestros resultados sugieren que, al haber utilizado matrices genómicas con un enfoque discreto y no continuo para la estimación del parentesco genómico, las ganancias en habilidad predictiva observadas con el modelo  $G_{IBS}$  respecto a del modelo  $G_{IBD}$  surgen principalmente de capturar una mayor variabilidad entre animales en la matriz de coeficientes (i.e SD de la relaciones genómicas) de la covarianza entre los valores de cría y no en una mayor estimación de la heredabilidad.

## BIBLIOGRAFÍA.

- Aguilar, I., Misztal, I., Tsuruta, S., Legarra, A. y Wang, H. 2014. PREGSF90 – POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 Programs. 10th World Congress of Genetics Applied to Livestock Production, Vancouver, Canada, poster 680.
- Brito, L. F., Clarke, S. M., McEwan, J. C., Miller, S. P., Pickering, N. K., Bain, W. E., Dodds, K. G., Sargolzaei, M. y Schenkel, F. S. 2017. Prediction of genomic breeding values for growth, carcass and meat quality traits in a multi-breed sheep population using a HD SNP chip. *BMC Genetics*, 18:7. <https://doi.org/10.1186/s12863-017-0476-8>
- Bulmer, M. G. 1985. *The mathematical theory of quantitative genetics*. Oxford: Clarendon Press, NY, USA.
- Calus, M. P. L., Meuwissen, T. H. E., De Roos, A. P. W. y Veerkamp, R. F. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics*, 178:553–561. <https://doi.org/10.1534/genetics.107.080838>
- Cantet, R. J. C., García-Baccino, C. A., Rogberg-Muñoz, A., Forneris, N. S. y Munilla, S. 2017. Beyond genomic selection: The animal model strikes back (one generation)!. *J. Anim. Breed. Genet.* 134: 224–231. <https://doi.org/10.1111/jbg.12271>
- Casiró, S., Velez-Irizarry, D., Ernst, C. W., Raney, N. E., Bates, R. O., Charles, M. G. y Steibel, J. P. 2017. Genome-Wide association study in an F2 duroc x pietrain resource population for economically important meat quality and carcass traits. *J. Anim. Sci.* 95: 545–558. <https://doi.org/10.2527/jas2016.1003>
- Cheng, S. H. y Higham, N. J. 1998. A modified cholesky algorithm based on a symmetric indefinite factorization. *SIAM J. on Matrix Anal.*, 19:1097–1110.
- Clark, S A, Kinghorn, B. P. y van der Werf, J. H. J. 2013. Comparisons of Identical by State and Identical by Descent Relationship Matrices Derived from SNP Markers in Genomic Evaluation. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 20:261–265.
- Clark, Samuel A., Hickey, J. M., Daetwyler, H. D. y van der Werf, J. H. J. 2012. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet. Sel. Evol.* 44: 4. <https://doi.org/10.1186/1297-9686-44-4>
- Cockerham, C. C. 1954. An Extension of the Concept of Partitioning Hereditary Variance for Analysis of Covariances among Relatives when Epistasis is Present. *Genetics*. 39:799–882. <https://doi.org/10.1111/medu.12876>
- Csurös, M. 2013. Non-identifiability of identity coefficients at biallelic loci. *Theor. Popul. Bio.* 92: 22–29. <https://doi.org/10.1016/j.tpb.2013.11.001>
- Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de los Campos, G., y Hickey, J. M. 2013. Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics*. 193:347-365. <https://doi.org/10.1534/genetics.112.147983>

- Daetwyler, H. D., Swan, A. A., Van Der Werf, J. H. y Hayes, B. J. 2012. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. *Genet. Sel. Evol.* 44:1–11. <https://doi.org/10.1186/1297-9686-44-33>
- de los Campos, G., Sorensen, D. y Gianola, D. 2015. Genomic Heritability: What Is It?. *PLoS Genet.* 11:(5).e1005048. <https://doi.org/10.1371/journal.pgen.1005048>
- de los Campos, G., Vazquez, A. I., Fernando, R., Klimentidis, Y. C. y Sorensen, D. 2013. Prediction of Complex Human Traits Using the Genomic Best Linear Unbiased Predictor. *PLoS Genet.* 9(7). e1003608. <https://doi.org/10.1371/journal.pgen.1003608>
- Donnelly, K. P. 1983. The probability that related individuals share some section of genome identical by descent. *Theoret. Pop. Biol.* 23:34–63. [https://doi.org/10.1016/0040-5809\(83\)90004-7](https://doi.org/10.1016/0040-5809(83)90004-7)
- Edwards, D. B., Ernst, C. W., Raney, N. E., Doumit, M. E., Hoge, M. D. y Bates, R. O. 2008. Quantitative trait locus mapping in an F2 Duroc x Pietrain resource population: II. Carcass and meat quality traits. *J. Anim. Sci.* 86:254–266. <https://doi.org/10.2527/jas.2006-626>
- Elston, R. C., Stewart, J. 1971. A general model for the genetic analysis of pedigree data. *21:523–542.*
- Erbe, M., Pimentel, E. C. G., Sharifi, A. R. y Simianer, H. 2010. Assessment of cross-validation strategies for genomic prediction in cattle. 9th World Congress of Genetics Applied to Livestock Production, Edited by German Society for Animal Science. German Society for Animal Science, Leipzig, Germany. p. 129–132.
- Falconer, D.S., Mackay, T.F.C. 1996 Introduction to quantitative genetics. Longman New York.
- Forneris, N. S., Steibel, J. P., Legarra, A., Vitezica, Z. G., Bates, R. O., Ernst, C. W., Basso, A. L. y Cantet, R. J. C. 2016. A comparison of methods to estimate genomic relationships using pedigree and markers in livestock populations. *J. Anim. Breed. Genet.* 133:452–462. <https://doi.org/10.1111/jbgs.12217>
- Forni, S., Aguilar, I. y Misztal, I. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genet. Sel. Evol.* 43:1. <https://doi.org/10.1186/1297-9686-43-1>
- García-Baccino, C. A., Munilla, S., Legarra, A., Vitezica, Z. G., Forneris, N. S., Bates, R. O., Ernst, C. W., Raney, N. E., Steibel, J. P. y Cantet, R. J. C. 2016. Estimates of the actual relationship between half-sibs in a pig population. *J. Anim. Breed. Genet.* 134:109–118. <https://doi.org/10.1111/jbgs.12236>
- Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica.* 136:245–257. <https://doi.org/10.1007/s10709-008-9308-0>
- Goddard, M. E., Hayes, B. J. y Meuwissen, T. H. E. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J. Anim. Breed. Genet.* 128:409–421. <https://doi.org/10.1111/j.1439-0388.2011.00964.x>

- Gualdrón Duarte, J. L., Bates, R. O., Ernst, C. W., Raney, N. E., Cantet, R. J. C., y Steibel, J. P. 2013. Genotype imputation accuracy in a F2 pig population using high density and low density SNP panels. *BMC Genetics*, 14. <https://doi.org/10.1186/1471-2156-14-38>
- Guo, S. W. 1995). Proportion of genome shared identical by descent by relatives: Concept, computation, and applications. *Am. J. Hum. Genet.* 56:1468–1476.
- Habier, D., Fernando, R. L. y Dekkers, J. C. M. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. 177:2389–2397. <https://doi.org/10.1534/genetics.107.081190>
- Habier, David, Tetens, J., Seefried, F. R., Lichtner, P. y Thaller, G. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* 42:5. <https://doi.org/10.1186/1297-9686-42-5>
- Habier, D., Fernando, R. L., y Garrick, D. J. 2013. Genomic BLUP decoded: A look into the black box of genomic prediction. *Genetics*. 194:597–607. <https://doi.org/10.1534/genetics.113.152207>
- Han, L. y Abney, M. 2011. Identity by descent estimation with dense genome-wide genotype data. *Genet. Epidemiol.* 35: 557–567. <https://doi.org/10.1002/gepi.20606>.Identity
- Han, L. y Abney, M. 2013. Using identity by descent estimation with dense genotype data to detect positive selection. *Europ. J. Hum. Genet.* 21(2): 205–211. <https://doi.org/10.1038/ejhg.2012.148>
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J. y Goddard, M. E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92:433–443. <https://doi.org/10.3168/jds.2008-1646>
- Hayes, B. J. y Goddard, M. E. 2008. Technical note: Prediction of breeding values using marker-derived relationship matrices. *J. Anim. Sci.* 86:2089–2092. <https://doi.org/10.2527/jas.2007-0733>
- Hickey, J. M., Kinghorn, B. P., Tier, B., Clark, S. A., van der Werf, J. H. J. y Gorjanc, G. 2012. Genomic evaluations using similarity between haplotypes. *J. Anim. Breed. Genet.* 130:259–269. <https://doi.org/10.1111/jbg.12020>
- Higham, N. J. 2002. Computing the nearest correlation matrix - A problem from finance. *IMA J. Num. Anal.* 22:329–343. <https://doi.org/10.1093/imanum/22.3.329>
- Hill, W. G. y Weir, B. S. 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Res.* 93: 47–64. <https://doi.org/10.1017/S0016672310000480>
- Hill, W. G. y Weir, B. S. 2012. Variation in actual relationship among descendants of inbred individuals. *Genet. Res.* 94:267–274. <https://doi.org/10.1017/S0016672312000468>
- Hill, William G. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet. Res.* 38:209–216. <https://doi.org/10.1017/S0016672300020553>
- Hothorn, T., Bretz, F., y Westfall, P. 2008. Simultaneous Inference in General

- Parametric Models. Technical Report Number 019. *Biom. J.* 50(3), 346–363. <http://www.stat.uni-muenchen.de>
- Iqbal, A., Choi, T. J., Kim, Y. S., Lee, Y. M., Zahangir Alam, M., Jung, J. H., Choe, H. S. y Kim, J. J. 2019. Comparison of genomic predictions for carcass and reproduction traits in Berkshire, Duroc and Yorkshire populations in Korea. *Asian-Australas. J. Anim. Sci.* 32:11.1657–1663. <https://doi.org/10.5713/ajas.18.0672>
- Jacquard, A. 1974. The genetic structure of populations. Springer-Verlag, New York.
- Jia, Z. 2017. Controlling the overfitting of heritability in genomic selection through cross validation. *Sci. Rep.s*, 7:13678. <https://doi.org/10.1038/s41598-017-14070-z>
- Karaman, E., Cheng, H., Firat, M. Z., Garrick, D. J., y Fernando, R. L. 2016. An upper bound for accuracy of prediction using GBLUP. *PLoS ONE*. 11: 1–18. <https://doi.org/10.1371/journal.pone.0161054>
- Kempthorne, O. 1955. The correlations between relatives in random mating Populations. *Proceedings of the Royal Society of London. Series B*. 143:102–113.
- Khanal, P., Maltecca, C., Schwab, C., Gray, K. y Tiezzi, F. 2019. Genetic parameters of meat quality, carcass composition, and growth traits in commercial swine. *J. Anim. Sci.* 97:3669–3683. <https://doi.org/10.1093/jas/skz247>
- Lander, E. S. y Green, P. 1987. Construction of multilocus genetic linkage maps in humans. *Proc. Nat. Acad. Sci.* 84:2363–2367. <https://doi.org/10.1073/pnas.84.8.2363>
- Legarra, A., Aguilar, I. y Misztal, I. 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92:4656–4663. <https://doi.org/10.3168/jds.2009-2061>
- Legarra, Andrés, Robert-Granié, C., Manfredi, E. y Elsen, J. M. 2008. Performance of genomic selection in mice. *Genetics*. 180:611–618. <https://doi.org/10.1534/genetics.108.088575>
- Luan, T., Woolliams, J. A., Degård, J., Dolezal, M., Roman-Ponce, S. I., Bagnato, A. y Meuwissen, T. H. E. 2012. The importance of identity-by-state information for the accuracy of genomic selection. *Genet. Sel. Evol.* 44:28. <https://doi.org/10.1186/1297-9686-44-28>
- Luan, T., Woolliams, J. A., Lien, S., Kent, M., Svendsen, M. y Meuwissen, T. H. E. 2009. The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. *Genetics*. 183:1119–1126. <https://doi.org/10.1534/genetics.109.107391>
- Malécot, G. 1948. Les mathématiques de l'hérédité. Masson et Cie, Paris.
- Meuwissen, T. H.E., Hayes, B. J. y Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157:1819–1829.
- Meuwissen, T., Hayes, B. y Goddard, M. 2013. Accelerating improvement of livestock with genomic selection. *Annu. Rev. of Anim. Biosci.* 1: 221–237. <https://doi.org/10.1146/annurev-animal-031412-103705>

- Meuwissen, Theo H.E., Odegard, J., Andersen-Ranberg, I. y Grindflek, E. 2014. On the distance of genetic relationships and the accuracy of genomic prediction in pig breeding. *Genet. Sel. Evol.* 46:49. <https://doi.org/10.1186/1297-9686-46-49>
- Miar, Y., Plastow, G. S., Moore, S. S., Manafiazar, G., Charagu, P., Kemp, R. A., van Haandel, B., Huisman, A. E., Zhang, C. Y., McKay, R. M., Bruce, H. L. y Wang, Z. 2014. Genetic and phenotypic parameters for carcass and meat quality traits in commercial crossbred pigs. *J. Anim. Sci.* 92: 2869–2884. <https://doi.org/10.2527/jas.2014-7685>
- Patterson, H. D. y Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika.* 58: 545–554. <https://doi.org/10.1093/biomet/58.3.545>
- Rabiner, L.R. 1989. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE.* 77:257-286.
- Saavedra, R. R., Zubiri-Gaitán, A., y Cantet, R. J. C. 2020. El desequilibrio de identidad en las relaciones de parentesco bajo recombinación. *Congreso de la Asoc. Argentina de Producción Animal-RAPA XXXX.* Vol. x,y.
- Sargolzaei, M., Chesnais, J. P. y Schenkel, F. S. 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics.* 1:478. <https://doi.org/10.1186/1471-2164-15-478>
- Self, S. G. Y Liang, K.Y. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.* 82:605–610.
- Sinnwell, J. P., Therneau, T. M., y Schaid, D. J. 2014. The kinship2 R package for pedigree data. *Human Heredity.* 78(2), 91–93. <https://doi.org/10.1159/000363105>
- Solaymani, S., Ayatollahi Mehrgardi, A., Esmailzadeh, A., Tusell, L. y Momen, M. 2020. Performance of pedigree and various forms of marker-derived relationship coefficients in genomic prediction and their correlations. *J. Anim. Breed. Genet.* 13:423–437. <https://doi.org/10.1111/jbg.12467>
- Stram, D. O. y Lee, J. W. 1994. Variance components testing in the longitudinal mixed effects model. *Biometrics.* 50:1171–1177.
- Thompson, E. A. 1988. Two-locus and three-locus gene identity by descent in pedigrees. *IMA J. Math. Appl. Med. Biol.* 5(4):261–279. <https://doi.org/10.1093/imammb/5.4.261>
- Thompson, Elizabeth A. 2013. Identity by descent: Variation in meiosis, across genomes, and in populations. *Genetics.* 194:2. 301–326. <https://doi.org/10.1534/genetics.112.148825>
- Toro, M., Barragán, C., Óvilo, C., Rodrigañez, J., Rodríguez, C. y Silió, L. 2002. Estimation of coancestry in Iberian pigs using molecular markers. *Conservation Genetics.* 3:309–320. <https://doi.org/10.1023/A:1019921131171>
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Van Vleck, L. D. 1993. Selection Index and Introduction to Mixed Model Methods.



- Vela-Avitúa, S., Meuwissen, T. H., Luan, T. y Ødegård, J. 2015. Accuracy of genomic selection for a sib-evaluated trait using identity-by-state and identity-by-descent relationships. *Genet. Sel. Evol.* 47:9. <https://doi.org/10.1186/s12711-014-0084-2>
- Visscher, P. M. 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Res. Hum. Genet.* 9:490–495. <https://doi.org/10.1375/183242706778024928>
- Visscher, P. M., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Zhu, G., Cornes, B. K., Montgomery, G. W., y Martin, N. G. 2006. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* 2:e41. 0316–0325. <https://doi.org/10.1371/journal.pgen.0020041>
- Visscher, P. M. 2009. Whole genome approaches to quantitative genetics. *Genetica.* 136: 351–358. <https://doi.org/10.1007/s10709-008-9301-7>
- Visscher, P. M., Hemani, G., Vinkhuyzen, A. A. E., Chen, G. B., Lee, S. H., Wray, N. R., Goddard, M. E., y Yang, J. 2014. Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. *PLoS Genet.* 10(4). <https://doi.org/10.1371/journal.pgen.1004269>
- Weir, B. S. y Cockerham, C. C. 1968. Pedigree mating with two linked loci. *Genetics.* 61:923–924.
- Weir, B. S. y Cockerham, C. C. 1969. Group inbreeding with two linked loci. *Genetics.* 63:711–742.
- Wijk, H. J. Van, Arts, D. J. G., Matthews, J. O., Webster, M., Ducro, B. J. y Knol, E. F. 2005. Genetic parameters for carcass composition and pork quality estimated in a commercial production chain. *J. Anim. Sci.* 83:324–333.
- Willson, H. e., Rojas de Oliveira, H., Schinkel, A. P., Grossi, D. y Brito, L. F. 2020. Estimation of genetic parameters for pork quality, novel carcass, primal-cut and growth traits in Duroc pigs. *Animals.* 10: 779.
- Wright, S. 1934. The method of path coefficients. *The Annals of Mathematical Statistics.* 5(3), 161–215. <https://doi.org/10.1214/aoms/1177732676>
- Wright S. 1922. Coefficients of inbreeding and relationship. *Amer. Nat.* 56:330-338.
- Yu, H., Spangler, M. L., Lewis, R. M. y Morota, G. 2017. Genomic relatedness strengthens genetic connectedness across management units. *G3(Bethesda).* 7:3543–3556. <https://doi.org/10.1534/g3.117.300151>