

# CURSO: LA MEJORA GENÓMICA EN PORCINO

## Capítulo 2 - La selección genómica

*Joan Estany & Romi Pena*

*Departamento de Ciencia Animal, Universidad de Lleida – Agrotecnio Center, Lleida.*

En el segundo capítulo del curso se presenta cómo se integra la variación en la secuencia de ADN en la selección de los futuros reproductores.

### 1.1. La selección

El proceso de selección genética implica elegir a los individuos que tienen mejor valor genético para los caracteres de interés de entre todos los candidatos a ser seleccionados. No resulta evidente conocer el valor genético de la mayor parte de los caracteres que son de interés en porcino, ya que se trata de caracteres complejos, cuyo fenotipo es el resultado del efecto de muchos genes, por lo general desconocidos y con un efecto individual pequeño sobre la variabilidad del carácter, y del efecto del ambiente. Por ello, el valor genético de un individuo para un carácter determinado no puede deducirse todavía directamente del genotipo sino que debe predecirse a partir de observaciones relacionadas. Se dispone de tres fuentes de datos para la predicción del valor genético de un cerdo (**Figura 1**): (1) los fenotipos, esto es, los registros de uno o más caracteres; (2) la genealogía, esto es, las relaciones de parentesco entre individuos; y (3) los marcadores moleculares, esto es, variaciones individuales en la secuencia del ADN, bien sea una sola mutación causal o un panel de muchos marcadores.

Tradicionalmente la predicción del valor genético en porcino se ha basado en los datos fenotípicos del individuo a valorar, tanto del carácter a mejorar (selección directa) como de otros caracteres correlacionados genéticamente con él (selección indirecta), y en los datos fenotípicos de sus parientes (selección genealógica). Así, por ejemplo, el valor genético de un cerdo para el índice de conversión, se predice teniendo en cuenta su propia conversión, la conversión de todos sus parientes, así como los datos de ganancia diaria de peso, tanto de él como de sus parientes, puesto que ambos caracteres están correlacionados negativamente. Cada una de estas observaciones puede haber sido registrada en un ambiente distinto y por

tanto no son directamente comparables. No es lo mismo tener un buen índice de conversión con una dieta alta o baja en proteína, o si se trata de un macho o una hembra. El BLUP (que responde a las siglas en inglés de mejor predictor lineal insesgado) es el procedimiento estadístico que mejor pondera linealmente los datos propios y los de los parientes, sea de uno o más caracteres, a la vez que simultáneamente corrige cada uno de ellos por el efecto del ambiente donde se han tomado (por ejemplo, granja, dieta o sexo). A lo largo de estos últimos treinta años, el BLUP se ha convertido en la herramienta clave de la valoración genética y muchos de los cambios ocurridos en estos años en los programas de mejora del porcino, tanto en estructura como en tamaño o posicionamiento en el mercado, tienen en él su origen. Un buen ejemplo son las líneas maternas prolíficas, las cuales no hubieran podido desarrollarse sin el avance que supuso el BLUP en la mejora del tamaño de camada.

## 1.2. Genes y marcadores genéticos

El primer marcador que se comercializó en porcino fue el gen del halotano (HAL-1843®), en 1991, cuya mutación – un cambio de una timina por una citosina en la posición 1843 del gen del receptor de la rianodina - permite distinguir los cerdos libres de los portadores de la variante que provoca sensibilidad a la hipertermia maligna y mala calidad de la carne. El éxito de esta primera aportación de la tecnología del DNA a la mejora porcina se debió a que la hipertermia maligna – de hecho un SNP, ver capítulo 1- depende de un solo gen. No obstante, tal como hemos visto, este no es el caso de la mayoría de los caracteres de interés económico, también llamados cuantitativos, que dependen de muchos genes. Durante los siguientes diez años se diseñaron toda una serie de experimentos con el objetivo de identificar mutaciones causales en estos genes, llamados también QTL (del inglés, genes con efecto en caracteres cuantitativos), o, en su defecto, de marcadores que estuvieran estrechamente ligados a ellos. El método seguido para identificar estos QTL fue y sigue siendo la detección de asociación entre variaciones en la secuencia del DNA y los fenotipos de interés.

Quizás el diseño más productivo en este sentido ha sido el análisis de la variación de la secuencia del DNA en genes con una función biológica relevante sobre un carácter objetivo, los llamados genes candidato. El primer QTL descubierto en cerdos mediante este procedimiento se localizó en gen del receptor de estrógenos (Rothschild y Plastow, 1999), con el que, según el tipo genético, pueden llegar a explicarse diferencias de hasta un lechón por parto. Siguiendo esta metodología, nuestro grupo ha probado que un SNP localizado en la región del promotor del gen esteroil-CoA desaturasa afecta el contenido de ácidos grasos monoinsaturados en la grasa de cerdo (Estany y col., 2014; **Figura 2**). Una limitación de este enfoque es que la búsqueda de marcadores se restringe a genes ya conocidos. Una alternativa más general es ampliar el escrutinio de marcadores a todo el genoma o al menos, como aproximación, a los SNP incluidos en un chip de genotipado (capítulo 1). Obviamente, cuantos más marcadores

tenga el chip, más denso será el escrutinio y más acotada la identificación de marcadores. A este procedimiento de búsqueda de marcadores se le conoce como GWAS (del inglés, estudios de asociación del genoma completo). En la **Figura 3** se presenta un ejemplo de GWAS (Ros-Freixedes y col., 2016). Los análisis GWAS han servido para confirmar que la mayoría de QTL tienen efectos pequeños y que, para una línea genética en particular, son muy escasos los QTL con un efecto relevante que segregan a frecuencias intermedias.

### 1.3. La selección con marcadores

La selección apoyada en unos pocos marcadores bien conocidos no es eficaz, ya que los QTLs descubiertos hasta la fecha sólo explican una pequeña fracción de la variación genética. De hecho, es esperable que la variante favorable de un gen con un efecto importante sobre un carácter largamente seleccionado esté ya fijada. Es por ello, que en la práctica la selección nunca se ha hecho exclusivamente por marcadores sino en el mejor de los casos con marcadores. Así, el genotipo de genes y/o marcadores se ha incluido como un factor más en los procedimientos BLUP de valoración genética, lo que ha dado lugar al denominado BLUP asistido por marcadores. Sin embargo, debido a la falta de marcadores con efecto consistente, la selección asistida con marcadores ha tenido un escaso impacto comercial y en general se ha limitado a asistir el BLUP con algunos marcadores conocidos (Dekkers, 2004; van der Steen y col., 2005). Alguna empresa, no obstante, procuró capitalizar este tipo de selección, en especial reivindicando la utilización de marcadores propios, sobre unos 30 por carácter, en la evaluación genética de sus líneas. Ya más recientemente, con la llegada de los SNP, se abrió la posibilidad de realizar genotipados a mayor escala y, con ello, el desarrollo de los primeros paneles con más de mil marcadores (Van Eenennaam y col., 2014).

### 1.4. La selección genómica

La comercialización de un chip de genotipado (capítulo 1) propició un salto tecnológico clave para el uso comercial de la genómica, pues puso a disposición de todas las empresas un panel de 60.000 SNP a un precio razonable. A diferencia de los marcadores anteriores, estos SNP son anónimos, es decir, no se corresponden con ningún QTL ni se les atribuye función biológica alguna. Lo que pretende la selección genómica es aprovechar esta variación anónima para mejorar la predicción del valor genético. ¿Cómo conseguirlo si se desconoce el efecto de estos SNP? Para ello se parte de la hipótesis de que muchos SNP bien distribuidos por todo el genoma pueden captar la mayor parte de la variación genética, ya que al menos habrá uno (aunque él mismo no tenga efecto como tal) asociado a cada uno de los QTL que sí que afectan el carácter objetivo. Se denomina valor genómico a la suma del efecto de todos los alelos de todos los marcadores que componen el genotipo de un individuo determinado (**Figura**

4). Es importante notar que en el cálculo del valor genómico, a diferencia de lo que sucede con la selección asistida con marcadores, se tiene en cuenta el efecto de todos los marcadores independientemente de su significación estadística.

La asignación de un valor a cada uno de los 60.000 alelos es una tarea compleja, pues hay más efectos a estimar que registros (cerdos con fenotipo) disponibles. Para ello se han desarrollado nuevos procedimientos estadísticos, que básicamente difieren en la distribución que asumen para el efecto de los marcadores. En el caso que ésta sea normal (esto es, que la mayoría de marcadores tiene un efecto similar), el valor genómico coincide con el valor genético BLUP en la que la matriz que recoge las relaciones de parentesco entre individuos, calculada a partir del pedigrí, se ha substituido por otra de genómica, en la que el parecido genético entre dos individuos se mide en función del número de alelos que comparten. Por eso, a este procedimiento se le conoce como *G-BLUP*. La ventaja de la matriz genómica (y por ende del *G-BLUP*) es que, con un buen número de marcadores bien distribuidos por todo el genoma, afina más el parentesco que la genealógica, ya que opera con la proporción real de genoma compartido entre individuos y no con la esperada, que es lo que se obtiene con el pedigrí. Así, por ejemplo, según el parentesco genealógico, los cuatro hermanos de la **Figura 4** son igual de parientes entre sí, mientras que con el parentesco genómico, que implícitamente considera la segregación dentro de familia, se tiene en cuenta que H1, por ejemplo, se parece más a H2 que a H4. Esto permite que el *G-BLUP* distinga entre hermanos que todavía carecen de registros propios, de tal manera que siguiendo con el ejemplo de la Figura 4, si tuviéramos que seleccionar a uno de los cuatro hermanos anteriores, con el *G-BLUP* elegiríamos a H3, que es el que presenta un valor genómico superior; por el contrario, con el BLUP convencional, no podríamos distinguir entre ellos y la elección tendría que hacerse al azar, por lo que sólo en un 25% de las ocasiones elegiríamos a H3, que es el mejor.

En principio, la selección genómica obviaría la necesidad de tomar datos fenotípicos, los cuales sólo se registrarían en una primera generación de referencia para poder estimar el efecto de los marcadores. El problema es que este efecto va variando conforme la población se reproduce, puesto que marcadores y QTL no están completamente ligados: cuanto más alejados están uno de otro, más probable es que durante la meiosis se produzca un suceso de recombinación entre ambos y, por tanto, que la asociación se rompa. Esto obliga a reestimar periódicamente las asociaciones QTL-marcador, lo que significa crear cada cierto tiempo una nueva generación de referencia. Esta forma de proceder, en la que se alternarían generaciones de referencia y generaciones de selección, introduce un cambio demasiado profundo de asumir para los programas de mejora, pues implica desaprovechar los animales no genotipados, que pueden ser muchos, y los registros fenotípicos que se sigan tomando. Como una solución de compromiso, surge la necesidad de encontrar una alternativa que combine el valor genómico *G-BLUP* con el valor genético BLUP convencional. Afortunadamente, esta alternativa existe, es fácil de aplicar y se llama single-step GBLUP (GBLUP en una etapa) o simplemente SS-BLUP.

La ventaja del SS-BLUP como procedimiento estadístico es que produce directamente un valor genómico final sin necesidad de tener que calcular previamente los valores G-BLUP y BLUP por separado para luego combinarlos. En buena medida el SS-BLUP es el responsable de que la selección genómica (así entendida) se haya empezado a implantar en porcino.

El valor genómico SS-BLUP es más preciso que el valor genético BLUP convencional porque, mediante la información de los marcadores, las relaciones de parentesco entre individuos se ajustan mejor a las reales, lo que a su vez permite ponderar mejor los datos de los parientes. El gran mérito del SS-BLUP es que integra en un sólo procedimiento las tres fuentes informativas (fenotipos, genealogía, y marcadores) que sirven para predecir el potencial genético de un individuo (Figura 1). Además, el SS-BLUP, por su similitud con el BLUP convencional, se adapta muy bien a las esquemas y rutinas de selección, de tal manera que el valor genómico, igual que sucede con el valor genético BLUP convencional, se actualiza en cada nuevo ciclo de evaluación con los nuevos datos que se hayan producido desde el ciclo anterior, sean del tipo que sean, y se calcula para todos los animales, estén genotipados o no. En la práctica, ello ha llevado a que actualmente en porcino la información genómica sirva fundamentalmente para refinar las relaciones genéticas entre animales, y con ello la precisión con la que estos se evalúan. En este contexto, la estimación de los efectos de los marcadores ha dejado de ser propiamente un objetivo y por tanto más que de generaciones de referencia y de selección, lo que se impone es definir la estrategia de fenotipado y genotipado con mejor relación coste-beneficio.

Los avances que hemos descrito en estos dos primeros capítulos están cambiando la mejora porcina (Knol y col., 2016). Tres de ellos han resultado clave: 1) se han identificado (aunque su función sea desconocida) muchos marcadores (miles de SNP) distribuidos a lo largo del genoma; 2) que se pueden genotipar de forma coste-eficiente (chips de genotipado); y, 3) que se dispone del cuerpo metodológico (SS-BLUP) que resuelve cómo integrar su variación con los datos fenotípicos y la genealogía. Debido a ello, la selección genómica se está utilizando cada vez más, lo que, al igual que pasó con el BLUP en su tiempo, está provocando un cambio en la organización de los programas de mejora (capítulo 3). Sin embargo, puesto que la selección genómica requiere inversiones importantes, su aplicación no siempre resulta evidente, particularmente en porcino (Ibáñez-Escriche y col., 2014), donde, en términos relativos, los núcleos de selección no son grandes, los objetivos diversos y el intervalo generacional corto. Es por ello que, aparte de intereses comerciales y de competencia entre compañías, la influencia que la genómica ejerza en los programas de mejora del porcino dependerá, al menos a corto plazo, de la posibilidad de disponer de una estrategia de genotipado (y fenotipado) que sea económicamente sostenible.

## Referencias

Dekkers JCM. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science* 82 (E. Suppl.):E313–E328

Estany, J., R Ros-Freixedes, M Tor, RN Pena. 2014. A functional variant in the stearyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. *PLoS ONE* 9: e86177.

Ibáñez-Escriche N, S Forni, JL Noguera, L Varona, 2014. Genomic information in pig breeding: Science meets industry. *Livestock Science* 166: 94–100

Knol EF, B Nielsen, PW Knap. 2016. Genomic selection in commercial pig breeding. *Animal Frontiers* 6(1): 15-22

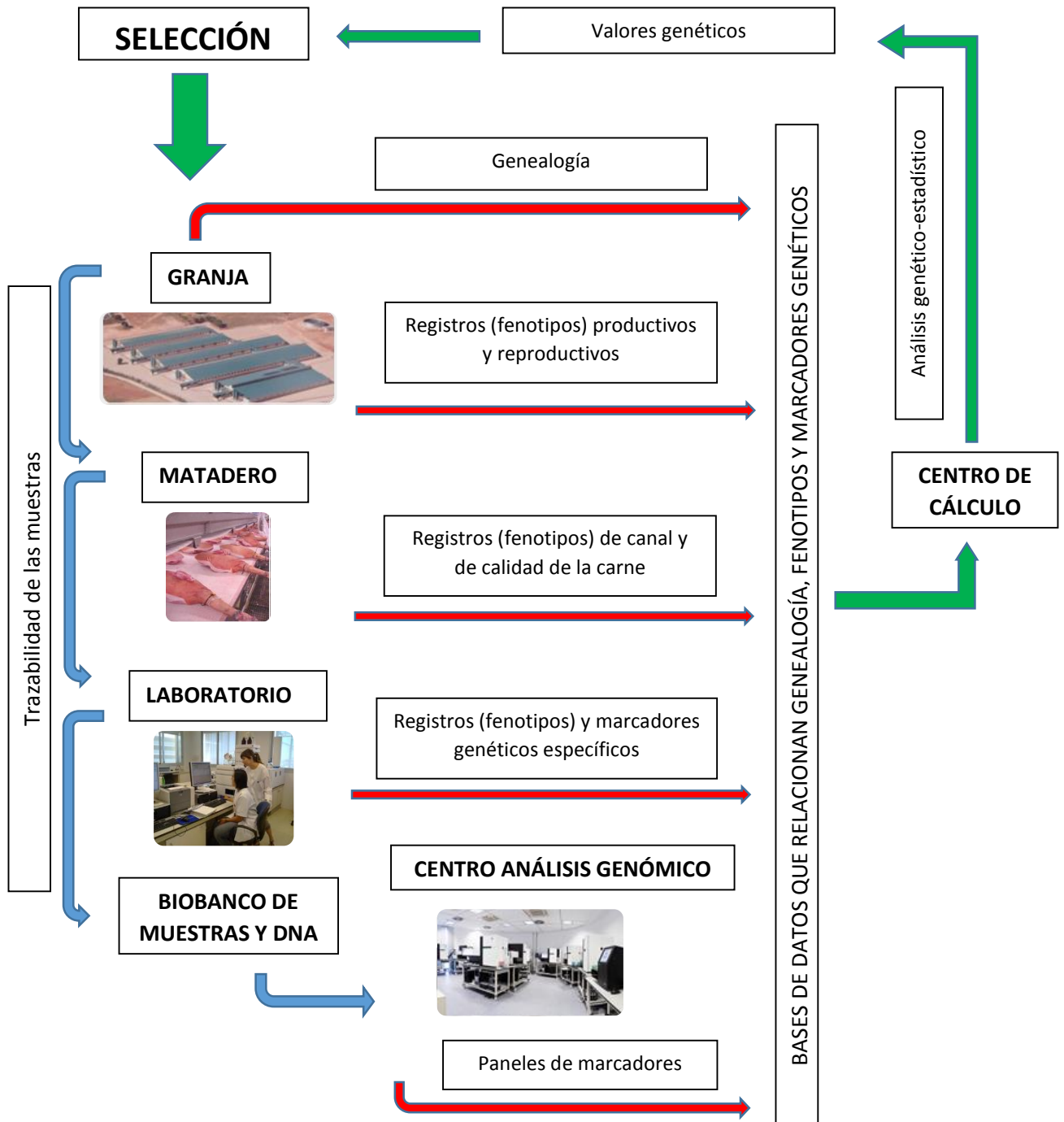
Ros-Freixedes R, S Gol, RN Pena, M Tor, N Ibáñez-Escriche, JCM Dekkers, J Estany. 2016. Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. *PLoS ONE* 11: e0152496

Rothschild MF, GS Plastow. 1999. Advances in pig genomics and industry applications. *AgBiotechNet*, 10:1–8.

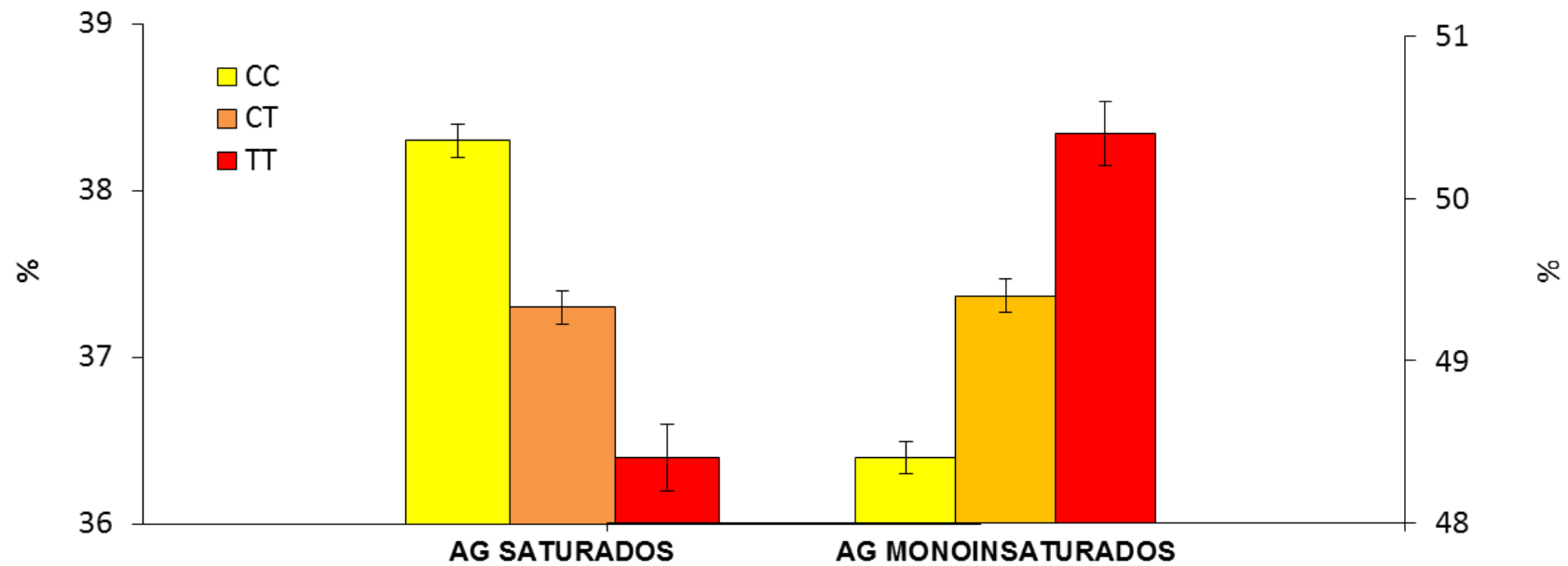
Van der Steen HAM, GFW Prall, GS Plastow. 2005. Application of genomics to the pork industry. *Journal of Animal Science* 83 (E. Suppl.):E1–E8

Van Eenennaam AL, KA Weigel, AE Young, MA Cleveland, JCM Dekkers. 2014. *Annual Review of Animal Biosciences* 2:105–39

**Figura 1. Flujo de información en el proceso de selección.** Los nuevos reproductores se eligen entre aquellos candidatos que presentan un mejor valor genético. La predicción del valor genético de un cerdo se realiza ponderando tres fuentes de información: (1) los registros fenotípicos de los caracteres de interés comercial; (2) la genealogía, esto es, las relaciones de parentesco entre individuos recogidas en su pedigrí; y (3) los marcadores moleculares.

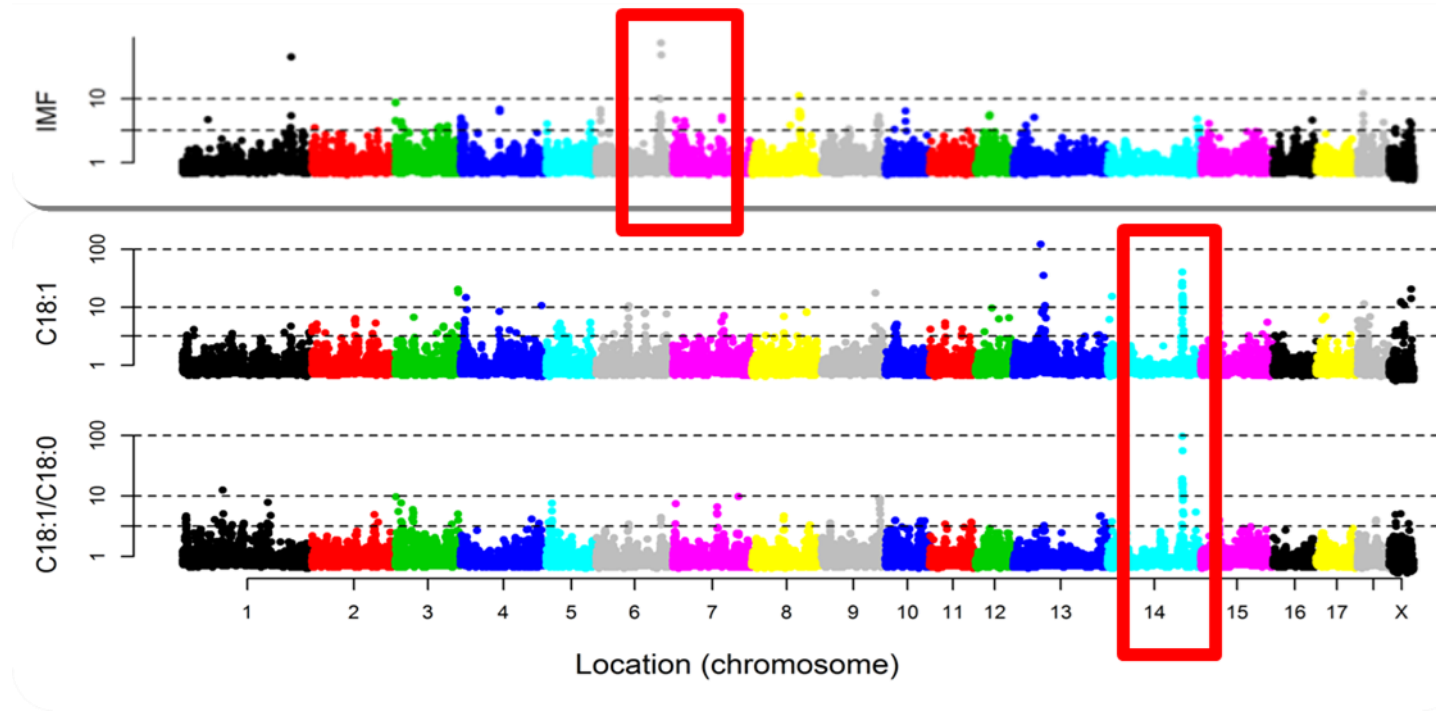


**Figura 2. Efecto del marcador SCD sobre la composición de la grasa en cerdo.** Hay dos alelos posibles para este marcador: C y T. Según los alelos que hayan recibido de su padre y de su madre, los cerdos pueden tener genotipo CC, CT o TT. Los cerdos TT presentan un 1% más de ácidos grasos monoinsaturados (y un 1% menos de ácidos grasos saturados) que los cerdos CT y éstos a su vez un 1% más de ácidos grasos monoinsaturados (y un 1% menos de ácidos grasos saturados) que los CC en la grasa intramuscular del músculo glúteo medio. Los resultados han sido obtenidos en una muestra de 1.087 cerdos Duroc.

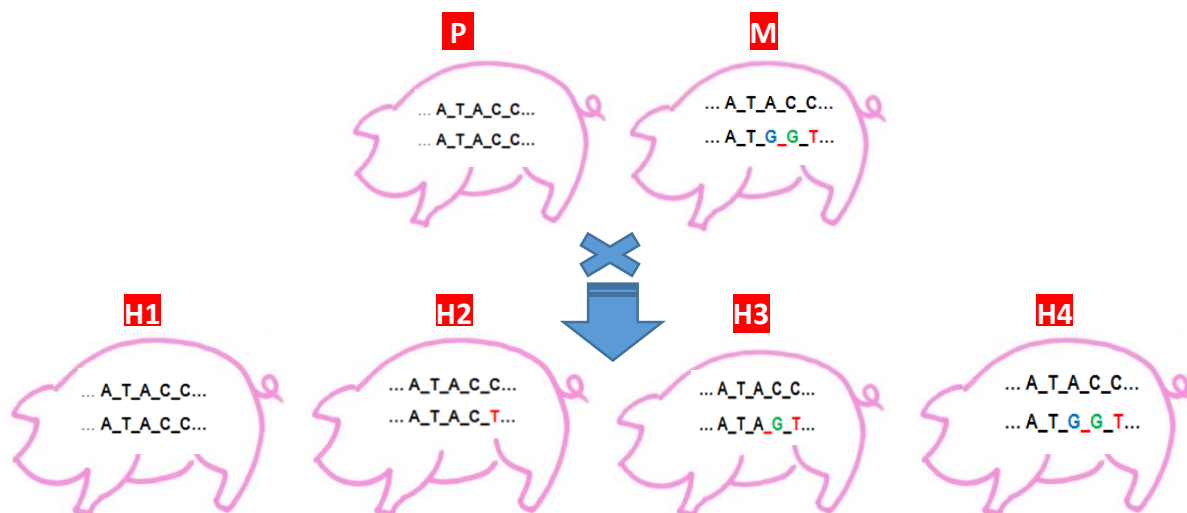




**Figura 3. Estudio de asociación genómica (GWAS) sobre contenido (IMF: grasa intramuscular, %) y composición de la grasa (C18:1: ácido oleico, %; C18:1/C18:0: relación entre contenido en ácido oleico y ácido esteárico) en una línea de cerdos Duroc.** En el eje horizontal se indica la localización genómica y en el eje vertical una medida de la asociación. Cada punto representa un SNP. En recuadro, dos regiones con evidencias de asociación. La región del cromosoma 14 se corresponde con el marcador SCD y la del cromosoma 6 con otro marcador conocido, el del receptor de la leptina.



**Figura 4. Predicción del valor genómico de una familia de hermanos a partir de un panel de cinco polimorfismos (SNP1 a SNP5).** Cada alelo de cada marcador tiene asignado un valor. El valor genómico resulta de la suma del efecto aditivo de los dos alelos de todos los marcadores del genotipo definido por los cinco SNP. El valor genómico se aprovecha de la segregación alélica dentro de familia para diferenciar entre hermanos antes de que estos dispongan de registros propios. H1 se parece más a H2 que a H4, mientras que H3, de acuerdo con su valor genómico, es el mejor.



	SNP1	SNP2	SNP3	SNP4	SNP5	
<b>EFEECTO ESTIMADO</b>	A= +4	T= +1	A= +5	C= - 2	T= +10	
	G= -4	C= -1	G= -5	G= +2	C= -10	
<b>CERDO</b>	<b>GENOTIPO</b>					<b>VALOR GENÓMICO</b>
P	AA	TT	AA	CC	CC	-4
M	AA	TT	AG	CG	CT	10
H1	AA	TT	AA	CC	CC	-4
H2	AA	TT	AA	CC	CT	16
H3	AA	TT	AA	CG	CT	20
H4	AA	TT	AG	CG	CT	10