

# Nuevas técnicas moleculares en genética porcina: selección genómica



**LUIS VARONA**

*Área de Genética Cuantitativa y Mejora Animal. Facultad de Veterinaria. Universidad de Zaragoza.*

*Instituto Agroalimentario de Aragón (IA2). Universidad de Zaragoza.*

## RESUMEN

La valoración genética de reproductores es un instrumento clave para la implementación de los procedimientos de mejora genética porcina. Tradicionalmente, los candidatos a la reproducción han sido seleccionados a partir de la valoración genética obtenida mediante BLUP, que utiliza la información genealógica y fenotípica. Sin embargo, en los últimos años, el desarrollo de la genética molecular ha facilitado el desarrollo de los procedimientos de selección genómica. Esta utiliza la información proporcionada por un gran número de polimorfismos de nucleótido simple —SNP—, que son genotipados mediante un dispositivo comercial. Hoy en día, la selección genómica es aplicada por la mayoría de las empresas de mejora genética porcina con la aproximación “single-step”, que permite aprovechar adicionalmente la información fenotípica y genealógica de los individuos no genotipados. Sin embargo es muy plausible que, en los próximos años, se introduzcan nuevas innovaciones tanto metodológicas como a partir de las grandes cantidades de nueva información que

van a generar tecnologías “ómicas” como la microbiómica, la metabolómica o la transcriciónómica, y el desarrollo de la ganadería de precisión.

La estructura de la mejora genética porcina del cerdo blanco está basada en el cruzamiento industrial de una cerda híbrida con un macho finalizador. La cerda híbrida se genera habitualmente a partir del cruce de individuos de dos poblaciones seleccionadas de manera muy intensa para el incremento del rendimiento reproductivo y productivo. Por otra parte, el macho finalizador puede proceder de una única línea o del cruce de dos líneas seleccionadas para crecimiento y/o eficiencia alimentaria, calidad de la canal y, en ocasiones, calidad de la carne.

La producción de carne de porcino en función de un esquema de cruzamiento exige que las empresas de selección mantengan poblaciones en pureza para abastecer mediante el esquema piramidal a una base productiva muy amplia. En este sentido, la estrategia básica de la mejora genética en porcino consiste en la aplicación de una intensa presión de selección en las poblaciones de la cúpula del esquema para que la respuesta a la selección se transmita con posterioridad a los estratos de multiplicación y producción.

El progreso genético de las poblaciones de selección se garantiza gracias a la correcta elección de los reproductores a

partir de su valoración genética para los criterios de selección. Desde los años 80 del siglo pasado, el procedimiento habitual para la valoración genética de reproductores ha sido el mejor predictor lineal insesgado o BLUP —Best Linear Unbiased Predictor— (Henderson, 1975). El BLUP aprovecha de manera óptima la información fenotípica del propio candidato y, adicionalmente, la información procedente de los individuos emparentados con él. Además, permite realizar una corrección por los efectos ambientales que pudiesen afectar a la expresión del carácter, como el sexo, la época de parto, la edad, etcétera.

La valoración genética mediante BLUP utiliza exclusivamente la información procedente de la genealogía y de los fenotipos para los caracteres de interés. Sin embargo, la genética molecular ha sufrido una grandísima evolución desde el descubrimiento del código genético (Watson y Crick, 1953). El planteamiento inicial para la utilización de la información molecular fue la selección asistida por genes y marcadores (Dekkers, 2004), que postulaba que era necesario conocer los polimorfismos genéticos directamente asociados con la variabilidad genética (genes), o bien polimorfismos o marcadores moleculares neutros que se encuentren en desequilibrio de ligamiento con estos genes (marcadores).

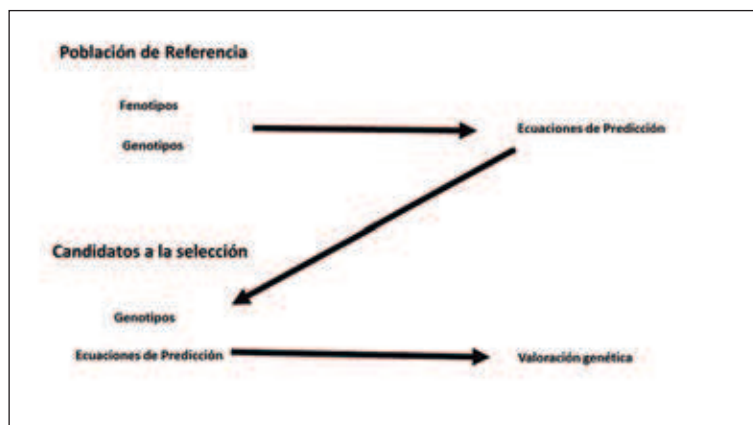
Desde este punto de partida, a partir de la última década del siglo pasado se desarrollaron numerosos protocolos experimentales con el objetivo de identificar los polimorfismos genéticos asociados a los caracteres de interés. El resultado de estos estudios localizó una gran cantidad de regiones del genoma asociadas a estos caracteres (QTL: *Quantitative Trait Loci*). De hecho, en julio de 2019, la base de datos [www.animalgenome.org/QTLdb](http://www.animalgenome.org/QTLdb) identifica en la especie porcina hasta 29,045 QTL para 678 caracteres diferentes. Lamentablemente, la mayoría de estos QTL no ha sido posible asociarlos a un polimorfismo concreto y, salvo notables excepciones como, entre otros, el RYR1 (receptor de la rianodina 1), el IGF2 (factor de crecimiento asociado a la insulina), el RN (rendimiento Napole), el MC4R (receptor 4 de la melanocortina) o la CAST (calpastatina), no han dado lugar a polimorfismos

de aplicación directa en los procedimientos de valoración genética de los reproductores.

Sin embargo, la utilización de la información molecular en la mejora genética porcina ha dado un vuelco radical gracias al desarrollo del mapa del genoma porcino (Groenen et al., 2012) y la posibilidad del desarrollo de *chips* de genotipado masivo de marcadores SNP (*Single Nucleotide Polymorphism*). Los marcadores SNP son polimorfismos que afectan a un único nucleótido, muy frecuentes en el genoma, y repartidos homogéneamente a lo largo del mismo. Estos dispositivos permiten disponer a un precio razonable (decenas de euros) del genotipo de un individuo para un número muy grande de marcadores SNP (decenas o centenares de miles).

La utilización de esta información genómica masiva con fines de selección fue sugerida de manera teórica por Meuwissen et al., (2001). El planteamiento inicial (Figura 1) implicaba disponer de una población de referencia suficientemente grande de la que se dispondría tanto el genotipado de los individuos como los fenotipos para los caracteres de interés. En esta población se debe realizar la calibración de las ecuaciones de predicción que incluirán las asociaciones de los marcadores SNP con cada uno de los caracteres de interés. Posteriormente, esta ecuación de predicción se puede utilizar para predecir el mérito genético de los futuros reproductores a partir de su información molecular, exclusivamente. Este planteamiento teórico se enfrentó con una importante complejidad estadística, ya que necesita estimar un número muy grande de efectos asociados a los marcadores moleculares SNP a partir de un número limitado de datos fenotípicos. Desde el punto de vista estadístico, se trata de un problema de *large p, small n*, que ha sido abordado mediante el desarrollo de un numeroso grupo de procedimientos (Gianola, 2013). Estos procedimientos difieren entre sí en las asunciones que hacen acerca de la distribución de los efectos asociados a los marcadores SNP.

La formulación original de la selección genómica puede proporcionar ventajas en cuanto a la precisión de la valoración genética de reproductores, pero implica la definición de una población de referencia y exige un replanteamiento radical de la organización de los esquemas de mejora por parte de las empresas de selección. En porcino, la mayoría de las empresas tienen una larga trayectoria en cuanto a la recogida sistemática de datos fenotípicos y genealógicos, que de haberse implantado este planteamiento original hubiera podido ser desperdiciada. Sin embargo, a finales de la década pasada, se desarrolló una propuesta alternativa



**Figura 1.** Esquema de la implementación de la selección genómica.



Figura 2. Parentesco Genómico vs. Parentesco Genealógico.

que se basaba en la construcción de una matriz de relaciones genómica ( $G$ ), definida por VanRaden (2008). Esta matriz de relaciones se calcula a partir de los genotipos de los SNP, y puede ser utilizada en la valoración genética mediante BLUP. Este BLUP genómico o GBLUP se diferencia de la valoración tradicional por BLUP en que utiliza la matriz de relaciones genómica ( $G$ ) en lugar de la matriz de relaciones calculada a partir de la genealogía ( $A$ ), tal y como postula BLUP.

La diferencia entre las dos matrices de relaciones de parentesco se ilustra en la Figura 2. En ella se presenta una familia de dos progenitores (A y B) y tres hermanos completos (C, D y E). Los animales tienen un genoma simplificado de 4 cromosomas sin recombinación (8 segmentos genómicos). Se puede observar que el parentesco genómico y genealógico entre padres e hijos siempre es 0.5, mientras que entre hermanos completos pueden existir importantes diferencias. El parentesco genealógico entre ellos es siempre 0.5, mientras que el parentesco genómico depende del proceso de segregación y puede hacer que algunos hermanos completos tengan un parentesco mayor que 0.5, mientras que otras parejas tengan parentescos inferiores a 0.5. Así, en el ejemplo, el parentesco genómico entre C y D es 0.75, ya que comparten 6 de los 8 segmentos del genoma, mientras que los parentescos genómicos entre C y E y entre D y E son solamente de 0.375 al compartir 3 de los 8 segmentos. La clave de esta diferencia es que la matriz de parentesco genealógica ( $A$ ) calcula el parentesco esperado entre los individuos, mientras que la matriz de parentesco genómico ( $G$ ) aproxima el parentesco realizado, como el porcentaje del genoma que es idéntico por descendencia; es decir, que procede de un mismo ancestro común. Además, una de las mayores ventajas de GBLUP consiste en que mediante un procedimiento denominado “single-step” (Misztal et al., 2009; Aguilar et al., 2011), es factible integrar la información fenotípica y genealógica acumulada históricamente por las empresas de selección con la información procedente del genotipado de SNPs. Este método permite valorar a

los reproductores genotipados y no genotipados y se ha convertido en el procedimiento habitual de las empresas de genética porcina para elegir a los futuros reproductores. Pese a todo, las aplicaciones de la información molecular a la mejora genética porcina no han hecho sino comenzar. En estos momentos existen numerosas líneas de investigación que probablemente darán frutos en los próximos años o décadas. Por una parte, debe tenerse en cuenta que una particularidad de la producción porcina es que el producto final a comercializar es un individuo procedente de un cruce a tres o cuatro vías. Hoy en día, numerosas empresas de selección utilizan la información de los animales cruzados para la predicción de los valores genéticos de los individuos de las poblaciones puras. En este sentido, la selección genómica ofrece una alternativa a la selección recíproca recurrente (Comstock et al., 1949), que planteaba la selección de los candidatos a la reproducción en función de sus descendientes en la población cruzada y que, pese a sus propiedades teóricas, no ha sido implementada de manera regular ya que exige disponer de datos de descendientes cruzados de los individuos de las poblaciones puras e implica un retraso en el progreso genético. La selección genómica permite calibrar los efectos de los marcadores SNP en la población cruzada (Kinghorn et al., 2010) y evaluar a los animales de las poblaciones en pureza a partir de los resultados de esta calibración. Además, las estrategias de selección se han basado hasta el momento en el componente genético aditivo, pero la información molecular permite predecir el rendimiento futuro de los individuos a partir de los efectos aditivos, de dominancia y epistáticos (Varona et al., 2018). Por lo tanto, la clasificación de los individuos ya no será estática, sino que dependerá de que con qué individuo (o con qué población) se van a cruzar.

Desde otro punto de vista, la mayoría de los procedimientos actualmente utilizados en la implementación de la selección genómica no incluyen información adicional que permita atribuir mayor o menor relevancia a distintas regiones del genoma. Por lo tanto presuponen que, a priori, todo el genoma colabora con la misma intensidad en la determinación de la variabilidad genética del carácter. Sin embargo, en los últimos años se están desarrollando procedimientos que pueden utilizar información a priori alternativa. Por ejemplo, es posible usar la información procedente de análisis de expresión génica o la información relativa a la contribución de los genes localizados en cada región del genoma a las rutas bioquímicas o metabólicas relacionadas con el carácter de interés (MacLeod et al., 2016).

Por otra parte, y al margen de la información puramente genómica, las aportaciones de las tecnologías “ómicas” a la mejora genética del porcino no han hecho sino comenzar



ya que es plausible que, en los próximos años o décadas, se desarrollen procedimientos que permitan utilizar la información generada por otras tecnologías “ómicas”, como la microbiómica, la transcriptómica o la metabolómica, que aportan información, respectivamente, del genoma de la flora microbiana de los individuos, de los niveles de expresión génica en los tejidos de interés y de la cuantificación de los niveles de numerosos metabolitos que pueden actuar como caracteres correlacionados de los objetivos de selección. Además, la ganadería de precisión va a generar una cantidad ingente de nuevos datos (“fenómica”) mediante la utilización de sensores, cámaras o dispositivos conectados a internet, que también podrán ser utilizados en la valoración genética de reproductores. Toda esta cantidad de nuevos datos seguramente va a introducir los procedimientos de valoración genética de reproductores dentro del ámbito de *big data* y permitirá incluir nuevos caracteres en los objetivos y criterios de selección en los esquemas de mejora genética porcina, probablemente relacionados con la salud animal y humana, la resiliencia o la huella de carbono, entre otros. ■

#### BIBLIOGRAFÍA

- Aguilar I, Misztal I, Johnson D, Legarra A, Tsuruta S, Lawlor T, 2010. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743-52.
- Comstock RE, Robinson HF, Harvey PH, 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. *Agron J.* 41:360-7.
- Dekkers JC, 2004. Commercial application of marker and gene assisted selection in livestock: strategies and lessons. *J. Anim. Sci.* 82:313-28.
- Gianola D, 2013. Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* 194:573-96.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, et al, 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491: 393-8.
- Henderson CR, 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics* 31:423-47.
- Kinghorn BP, Hickey JM, Van der Werf JHJ. 2010. Reciprocal recurrent genomic selection for total merit in crossbred individuals. In: Proc. 9<sup>th</sup> WCGALP. Leipzig, Germany.
- MacLeod IM, Bowman PJ, Vander Jaqt CJ, Haile-Mariam M, Kemper KE, Chamberlain AJ, Schrooten C, Hayes BJ, Goddard ME, 2016. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* 17: 144.
- Meuwissen THE, Hayes BJ, Goddard ME, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-29.
- VanRaden PM, 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91: 4414-23.
- Varona L, Legarra A, Toro MA, Vitezica ZG, 2018. Non-additive effects in genomic selection. *Front. Genet.* 9: 78.
- Watson JD, Crick FHC, 1953. A Structure of Deoxyribose Nucleic Acid. *Nature* 171:737-8.