

## **EVALUACIÓN GENÓMICA DE LA RAZA HOLSTEIN EN ESPAÑA**

### **1. Cálculo del valor genómico directo (VGD)**

Los valores genómicos directos se calculan mediante el método GBLUP o BLUP genómico. Este método es un BLUP que usa la matriz de relaciones genómicas como matriz de relaciones entre los SNP. El BLUP genómico asume que todos los SNP tienen la misma varianza y que ésta sigue una distribución normal. La matriz genómica o matriz de relaciones entre genotipos llamada *G*, es una matriz de similitud entre genotipos.

El uso de *G* implica que individuos con genotipos muy parecidos tendrán predicciones genómicas similares.

La valoración genómica es una valoración uni-carácter, donde el modelo lineal que se utiliza para estimar los valores genómicos directos (VGD) de cada carácter se escribe de la siguiente forma:

$$y = 1\mu + Zg + e$$

donde:

*y*: Pruebas derregresadas (PDR) de los animales de la población de referencia en un carácter determinado calculadas a partir de las pruebas MACE que publica Interbull (Jairath *et al.*, 1998).

$\mu$ : Media de la población.

*g*: Vector del efecto genético.

*Z*: Matriz de diseño que relaciona los valores genéticos de los animales con los datos fenotípicos (en este caso, las PDR).

Se asume que el efecto genético (*g*) y el efecto ambiental (*e*) siguen distribuciones normales:

$$g \sim N \left( 0, G\sigma_g^2 \right) \quad y \quad e \sim N \left( 0, D\sigma_e^2 \right)$$

donde:

*G*: Matriz de relaciones genómicas que se definió siguiendo el método 2 de VanRaden (VanRaden, 2008), donde los marcadores están ponderados cada uno por su varianza esperada:

$$G = \frac{MM'}{\sum 2p_j (1 - p_j)}$$

siendo:

*M*: Matriz de SNP (n animales x m SNP).

$p_j$ : Frecuencia alélica en el SNP<sub>*i*</sub>.

$\sigma_g^2$ : Varianza genética aditiva.

*D*: Matriz diagonal con los elementos  $d_{ii} = 1/\omega$ , siendo  $\omega$  un peso para la PDR<sub>*i*</sub>:

$$\omega = \frac{r_{PDR}^2}{(1 - r_{PDR}^2)} \quad y \quad r_{PDR}^2 \quad \text{la fiabilidad de las pruebas derregresadas}$$

$\sigma_e^2$ : Varianza residual

Se utiliza  $\omega$ , para tomar en cuenta las diferencias en fiabilidad entre las pruebas derregresadas.

## 2. Efecto poligénico

Parte de la variabilidad genética se supone que no puede ser capturada por los SNP de forma individual; por ello, incorporamos un efecto poligénico que ayuda a explicar la variabilidad mostrada por las pruebas derregresadas a partir de la genealogía y que mejora las predicciones del BLUP genómico. El porcentaje de variabilidad expresada por el efecto poligénico depende de cada carácter.

Fue necesario un estudio de validación previo para definir el porcentaje de la variabilidad genética expresada por el efecto poligénico para cada carácter, cuyo resultado fue:

Carácter	% efecto poligénico óptimo
Kilos de leche	15
Kilos de grasa	15
Kilos de proteína	30
Porcentaje de grasa	5
Porcentaje de proteína	5
Estatura	5
Anchura de pecho	30
Profundidad corporal	5
Angulosidad	40
Ángulo grupa	5
Anchura de grupa	25
Vista lateral de patas	25
Vista posterior de patas	40
Ángulo podal	40
Inserción anterior	25
Inserción posterior	25
Ligamento suspensor	40
Profundidad de ubre	5
Colocación pezones anteriores	25
Longitud pezones	5
Colocación pezones posteriores	5
Miembros y aplomos	50
Movilidad	50
Recuento de células somáticas	25
Longevidad	1
Días abiertos	5

## 3. Población de referencia

Los machos genotipados en 50K o en HD y que tienen prueba de descendencia en la última valoración Nacional o en la Internacional (MACE de Interbull) son los animales que forman la población de referencia. Dado que los animales no están valorados por descendencia para todos los caracteres al mismo tiempo, la población de referencia para cada carácter no se compone del mismo número de toros.

## 4. Fiabilidad esperada del VGD

El valor genético se suele proporcionar acompañado de la precisión de la predicción, que viene a ser la correlación entre el valor predicho y el valor verdadero o, en su caso, la correlación al cuadrado. A esta precisión se le llama Fiabilidad esperada de la prueba.

La Fiabilidad esperada se calcula invirtiendo la matriz de coeficientes de las ecuaciones

del modelo mixto para obtener la varianza del error de predicción (VEP).

$$r_u^2 = 1 - \frac{VEP_u}{\sigma_u^2}$$

donde:

$r_u^2$  : Fiabilidad esperada.

$\sigma_u^2$  : Varianza genética aditiva.

Debido a que la fiabilidad esperada está sobre-estimada en algunos caracteres cuando se compara con los estudios de validación, se multiplica la fiabilidad del GBLUP por un factor de escala S obtenido a partir de dicho estudio de validación previo. Los VGD se multiplican por  $\sqrt{S}$  con el fin de reducir dicha sobre-estimación, para obtener el VGD corregido:

$$VGD^* = \sqrt{S} * VGD$$

Y, por tanto, la fiabilidad de los VGD corregidos será:

$$Fiab^*_{VGD} = S * r_u^2$$

## 5. Combinación de pruebas (GEBV)

Tenemos dos fuentes de información a combinar, que son el VGD, y el Índice de prueba de descendencia (EBV), que en el caso de los candidatos sería un Índice de pedigrí (IP). Estas dos fuentes de información se combinan de la siguiente forma en función de la fiabilidad esperada de cada tipo de prueba.

$$GEBV = b_1 * VGD + b_2 * EBV$$

donde:

$$b_1 = \frac{\frac{fiab_{VGD}}{(1 - fiab_{VGD})}}{\frac{fiab_{VGD}}{(1 - fiab_{VGD})} + \frac{fiab_{EBV}}{(1 - fiab_{EBV})}} \quad b_2 = \frac{\frac{fiab_{EBV}}{(1 - fiab_{EBV})}}{\frac{fiab_{VGD}}{(1 - fiab_{VGD})} + \frac{fiab_{EBV}}{(1 - fiab_{EBV})}}$$

## 6. Fiabilidad de las pruebas combinadas (FGEBV)

Para calcular la fiabilidad del valor combinado, se calcula la ganancia en fiabilidad que aporta la genómica al índice genético tradicional. Dicha ganancia se calcula a través del estudio de validación.

La fiabilidad del valor combinado es:

$$Fiab_{GEBV} = Fiab_{EBV} + Fiab_{GAIN}$$

La ganancia de fiabilidad que aporta la genómica se calcula de la siguiente forma:

$$Fiab_{GAIN} = \frac{R^2_{GEBV, PDR} - R^2_{IP, PDR}}{fiab_{PDR}}$$

$$Fiab_{GEBV} = \chi + (1 - \chi) Fiab_{EBV}$$

donde:

$$\chi = Fiab_{GAIN} / (1 - Fiab_{EBV})$$

## 7. Estudio de validación cruzada

Toda valoración genómica debe ser verificada a través de una validación cruzada. Se han seguido las recomendaciones de Mäntysaari *et al.* (2010) para validar las valoraciones genómicas. Estas recomendaciones siguen la misma lógica del método 3 utilizado por Interbull para validar las valoraciones genéticas tradicionales. Este método consiste en comparar las

valoraciones genéticas de los toros obtenidas con la primera remesa de hijas con sus valoraciones finales, incluyendo toda la información disponible.

Respecto a las valoraciones genómicas, se divide la población de animales en 2 grupos: uno, usado como Población de Referencia; es decir, conociendo sus fenotipos con la información disponible cuatro años atrás, para estimar los efectos de los SNP. Otro, un grupo de animales más jóvenes (Población de Validación) cuyo fenotipo (PDR) no se incorpora a la valoración. Tras correr la valoración genómica, se comparan los valores genómicos directos (VGD) y los valores genómicos combinados (GEBV) de dicha Población de Validación con las pruebas derregresadas (PDR) disponibles de dichos animales actualmente.

Esta Población de Validación, está formada por, al menos, 150 animales con una prueba con un número de hijas efectivas superior a 20 (hijas efectivas es una medida de la cantidad de información que hay en la prueba).

Para validar las pruebas se tienen en cuenta en tres parámetros:

- La precisión de las pruebas, calculando la correlación entre los VGD estimados y las PDR conocidas.
- El sesgo que conlleva la predicción, calculando el coeficiente de regresión de los VGD sobre las PDR.
- El error cuadrático medio (MSE).

La fiabilidad realizada del modelo se estima calculando la correlación entre el VGD y la PDR tomando en cuenta la fiabilidad de las pruebas derregresadas:

$$Fiabl_{VGD} = R^2_{(VGD,PDR)} * \left( 1 + \frac{k}{EDC} \right)$$

$$k = \frac{4 - h^2}{h^2}$$

donde:

$R^2$ : Correlación al cuadrado entre los VGD estimados y las PDR conocidas.

EDC: Hijas efectivas.

$h^2$ : Heredabilidad.

La fiabilidad realizada nos permite estimar la ganancia en fiabilidad sobre la valoración tradicional que aporta la genómica, como fue detallado en los puntos anteriores.

## 8. GICO

El GICO es el índice de mérito genético total oficial de CONAFE que combina los GEBV de los distintos caracteres evaluados, según su importancia económica y sus correlaciones genéticas. Se utiliza para ordenar a los animales en los listados oficiales de los mejores toros y vacas genotipadas para mejorar la rentabilidad económica futura de las explotaciones. La fórmula del GICO es equivalente a la del ICO.

La última actualización del ICO fue en Junio 2015. La fórmula del ICO actual es la siguiente:

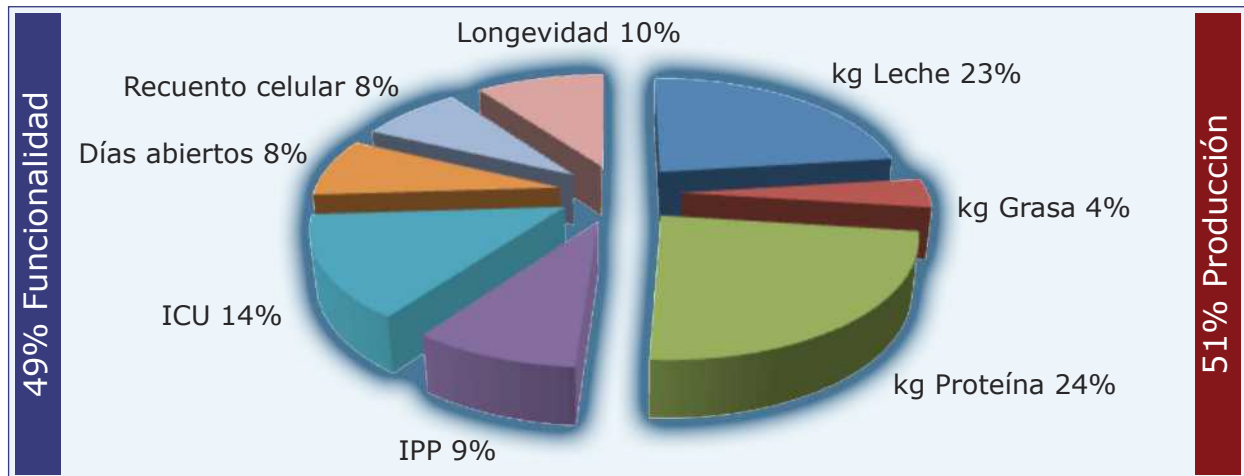
$$ICO = 1485 + 13,7 * \left( 23 * \frac{KL}{DT_{KL}} + 4 * \frac{KG}{DT_{KG}} + 24 * \frac{KP}{DT_{KP}} + 14 * \frac{ICU}{DT_{ICU}} + 9 * \frac{IPP}{DT_{IPP}} + 10 * \frac{(Longevidad-100)}{DT_{Long}} + 8 * \frac{(RCS-100)}{DT_{RCS}} + 8 * \frac{(DA-100)}{DT_{DA}} \right)$$

donde, las desviaciones típicas de cada carácter tienen los siguientes valores:

$DT_{KL} = 765$ , para kilos de leche (KL)	$DT_{ICU} = 1$
$DT_{KG} = 29$ , para kilos de grasa (KG)	$DT_{IPP} = 1$
$DT_{KP} = 22$ , para kilos de proteína (KP)	$DT_{Long} = 10$
	$DT_{RCS} = 10$
	$DT_{DA} = 10$

ICU: Índice compuesto de ubre      RCS: Recuento de células somáticas  
 IPP: Índice compuesto de patas      DA: Días abiertos

Gráfico: Pesos relativos de los distintos caracteres en el ICO



Para situar el nivel de un animal concreto respecto del resto de la población, el ICO de cada animal va acompañado de su percentil, que nos indica el porcentaje de animales evaluados con un valor inferior al que se está analizando. Se calcula un percentil para las vacas y un percentil para los toros. El percentil de los toros se calcula teniendo en cuenta sólo los toros con hijas.

## 9. Requisitos de publicación de las pruebas genómicas oficiales

Para tener una prueba genómica oficial, los toros genómicos deben:

- Ser propiedad de un programa de testaje español.
- Tener cumplidos 10 meses de edad en la fecha de la publicación de las pruebas.
- Estar registrados en CONAFE.
- Poseer un código de inseminación.
- Tener semen disponible.
- Estar genotipados en la densidad exigida por CONAFE para la prueba oficial.

Respecto a las hembras los requisitos para tener prueba oficial son

- Tener valoración genómica.
- Estar registrada en CONAFE.

También se consideran pruebas oficiales las de aquellos toros cuyo genotipo se ha recibido en los intercambios de EuroGenomics nominados como publicables por el país de origen y la de aquellos toros de centros de inseminación de terceros que han solicitado su prueba oficial a través del "Fee System" de EuroGenomics.

Los animales evaluados que no cumplen estos requisitos tendrán una valoración no oficial y no publicable sólo accesible para el propietario remitente de la muestra.

## 10. Evaluación genómica mensual

Cada mes CONAFE calcula las pruebas genómicas de los animales genotipados en el último mes. Después de cada valoración se informa tanto a ganaderos como a los programas de testaje de las pruebas de sus animales y se publican en la web de CONAFE, cuando dichas pruebas cumplen los requisitos de oficialidad.

## 11. Evaluación genómica general

Tres veces al año (Marzo, Junio y Noviembre) se recalculan las pruebas derregresadas a partir de la última evaluación MACE y se actualiza la Población de Referencia, con el objetivo de calcular la valoración genómica de todos los animales con genotipo disponible en la base de datos de CONAFE.

Se combinan los valores genómicos directos con los resultados de las pruebas tradicionales (prueba de descendencia, valor propio de las hembras o Índice de pedigrí de machos jóvenes y novillas).

Se genera un catálogo con las pruebas oficiales:

- Valores genómicos combinados (GEBV), en el caso de los animales genotipados.
- Índices de Pruebas de descendencia (EBV), en el caso de los machos no genotipados, con información suficiente de hijas para obtenerla.

Para diferenciar los animales que han sido valorados teniendo en cuenta su información genómica de los que no, las pruebas de los primeros se marcan con un tipo de prueba "G".

## **12. Información adicional a las pruebas genómicas**

El genotipado mediante SNP permite obtener información adicional a la de las evaluaciones genómicas, que son las filiaciones y la información de haplotipos, caracteres recesivos y otros genes de interés.

### Filiación por SNP:

Se puede comprobar la filiación de aquellos animales genotipados con sus progenitores, en el caso de que éstos también estén genotipados. Para ello, se contrastan todas aquellas posiciones en las cuales la muestra y la de su progenitor son homocigotas, determinando la filiación correcta si el porcentaje de incompatibilidades es inferior al 1.1%, siguiendo las directrices de ISAG para la verificación de parentesco mediante SNPs.

CONAFE comprueba la filiación mediante SNPs de todos los animales que se evalúan genómicamente y emite certificados de filiación.

### Otros caracteres:

En función de la versión y el chip usados para genotipar cada animal se puede obtener distinta información adicional como pueden ser enfermedades ligadas a caracteres recesivos como el BLAD, el CVM, haplotipos relacionados con enfermedades que afectan a la fertilidad, genes de interés como el gen "Polled" o el "Factor Rojo", el tipo de proteína de la leche o presencia de genes del cromosoma Y en hembras que provoca infertilidad.