

# EVALUACIONES GENÓMICAS UTILIZANDO SECUENCIAS DEL PROYECTO DE LOS 1000 GENOMAS BOVINOS

Lakhssassi<sup>1</sup>, K., González-Recio<sup>1</sup>, O.

<sup>1</sup>Departamento de Mejora Genética Animal. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid 28040, Spain  
E-mail: lakhssassi\_kenza@hotmail.com

## INTRODUCCIÓN

En ganado lechero, la integración de la información genómica procedente de genotipados con chips de SNPs en la selección genética es uno de los mayores avances de las últimas décadas. El uso de datos genómicos permite predecir valores de mérito genético utilizados en programas de selección con mayor precisión y a edades más tempranas, reduciendo el intervalo generacional de los reproductores. Estos dos términos permiten aumentar considerablemente el progreso genético de las poblaciones (Van Der Werf, 2013). Actualmente es posible secuenciar el genoma completo de bovinos, lo cual tiene la ventaja de incluir las mutaciones causales en el ADN responsables de la variación del carácter (Meuwissen y Goddard, 2010). El efecto aditivo de los marcadores, por tanto, dejarían de depender de su desequilibrio de ligamiento con los QTL. Sin embargo el coste es aún muy alto como para que la secuenciación sea una práctica generalizada.

España, a través del INIA, se ha incorporado en el proyecto de los 1000 genomas bovinos, que recopila los esfuerzos de más de 10 países en la secuenciación de animales *Bos Taurus* y *Bos Indicus*. En este proyecto se han detectado 31,8 millones de variantes que podrían utilizarse para la detección de regiones, posiblemente causales, asociadas a rasgos de interés económico. La dificultad de tratar tal cantidad de información de una forma eficiente en las evaluaciones genéticas es un desafío en los próximos años.

El objetivo de este trabajo es el desarrollo de estrategias que incorporen información de secuencias en las evaluaciones genéticas.

## MATERIAL Y MÉTODOS

**Datos.** Se utilizaron las secuencias de 450 animales Holstein procedentes del proyecto de los 1000 genomas bovinos junto con información de su pedigrí. Los caracteres analizados fueron las pruebas MACE para proteína, Índice Global de Tipo (IGT), Recuento de células Somáticas (RCS) y Días Abiertos (DA) de 361 toros, proporcionadas por CONAFE, y ponderadas por su precisión.

**Análisis.** Se usó el programa Findhap.f90 desarrollado por VanRaden et al. (2011) para la extracción de los haplotipos en la población, descartando aquellos con frecuencia <1%. Se extrajeron los haplotipos presentes por segmentos en cada cromosoma. La longitud de estos segmentos varió en función de la magnitud del desequilibrio de ligamiento presente. Para la detección de regiones asociadas a los caracteres de interés económico estudiados en ese trabajo, se realizó un análisis de asociación para estimar los efectos de haplotipos por cromosoma a través del modelo Lasso Bayesiano corrigiendo por el efecto poligénico. Se seleccionaron los haplotipos cuya efecto fue

mayor / menor a la media más / menos 3 veces la desviación estándar de los efectos de los haplotipos para cada carácter.

## RESULTADOS Y DISCUSIÓN

El cromosoma 1 presentó el mayor número de haplotipos (194.998), mientras que el menor fue para el cromosoma 25 (55.445). El número de segmentos osciló entre 121-430 según la longitud de cada cromosoma y el número de haplotipos por segmento varió entre 449 a 470. El número de haplotipos con una frecuencia por debajo de 1% fue del 97% en todos los cromosomas. Estos haplotipos no fueron utilizados en este análisis debido a la dificultad de poder encontrar efectos estadísticos cuando el haplotipo se presenta en sólo un par de individuos de nuestra muestra. Tras este filtrado, se utilizaron 76.512 haplotipos en el análisis posterior.

Nuestro análisis estadístico detectó 515, 746, 466 y 700 haplotipos distintos que superaron el umbral para los kg de proteína, IGT, RCS y DA, respectivamente. La figura 1 muestra la distribución de las frecuencias alélicas de los haplotipos para cada carácter. Se observa que la mayoría de los haplotipos encontrados están a baja frecuencia. Por lo tanto, esperamos que nos aporten información adicional a los genotipados de SNP acerca de las variantes menos frecuentes para explorar su contribución en la variación fenotípica. La tabla 1 muestra la proporción de varianza de los valores genómicos estimados correspondiente al efecto de los haplotipos. Los haplotipos contribuyeron entre 2 - 5 % sobre la varianza total de los valores genómicos. Los haplotipos contribuyeron con un mayor porcentaje para DA (4.51%) en comparación con los otros rasgos.

Los cinco haplotipos con mayor efecto estimado sobre el carácter (en valor absoluto) para cada carácter se enumeran en la tabla 1 con sus localizaciones, frecuencias alélicas y estima del efecto. Esos haplotipos fueron localizados a través de 11 diferentes cromosomas (cromosomas 1, 2, 4, 8, 9, 10, 11, 12, 15, 17, 23). Según la tabla 2 se observa que los haplotipos 1395 y 1806 están ambos asociados con IGT y los kg de proteína. En el cromosoma 17 se detectaron 3 haplotipos de segmentos diferentes asociados con el IGT. El haplotipo que tiene mayor efecto en esta región fue el 1806 que se extiende desde el locus 321493.

## REFERENCIAS BIBLIOGRÁFICAS

- Meuwissen, T. H. E., and Goddard, M. E. 2010. *Genetics*. 185:623–31.
- Van Der Werf, J. 2013. *Methods Mol Biol*. 2013. 1019:543-61.
- VanRaden, P. M., O'Connell R. J., Wiggans, J. R., Weigel, K. A. 2011. *Gene Select Evol*. 2011, 43:10.

**Tabla1. Porcentaje de la varianza de los valores genómicos estimados que fue explicada por los haplotipos para cada carácter**

Kg Prot	IGT	DA	RCS
2,15	3,24	4,51	2,42

**Tabla 2. Haplotipos con mayor efecto obtenidos utilizando LASSO Bayesiano**

Caracter	HAP	codigo	Locus	length	Efecto hap	Freq
DA	1747	11.157.4	311845	1999	0.05544141	49.33
	3475	1.333.1	663905	2000	0.05510629	41.66
	1977	12.166.2	329846	2000	0.05100452	14.33
	1503	4.124.61	245509	1996	0.04883964	20.55
	672	15.59.1	115595	1993	0.04810013	44.00
IGT	1395	1.130.1	257905	2000	-0.00750371	34.00
	422	23.39.3	75887	1997	-0.00662909	59.11
	1806	17.162.5	321493	1997	-0.00660672	39.44
	670	17.64.5	125787	1997	-0.00582747	50.77
	1838	17.167.5	331478	1997	-0.00568824	44.11
Kg Prot	1625	10.156.3	309532	1997	-0.08133498	45.44
	1395	1.130.1	257905	2000	-0.07984824	34.00
	1806	17.162.5	321493	1997	-0.07391556	39.44
	2880	15.241.8	478324	1994	-0.06481805	20.22
	101	8.9.258	15969	1996	-0.06439777	14.77
RCS	2178	9.199.63	394658	1994	-0.03034244	15.66
	969	10.94.21	185718	1997	0.02826121	25.11
	1544	2.142.1	281437	1996	-0.02771862	17.88
	1902	15.160.8	316888	1993	-0.02762207	9.88
	19	15.3.6	3987	1993	0.02757105	7.11

**Figura 1. Distribución de las frecuencias alélicas de los haplotipos para cada carácter**

