

# IDENTIFICACIÓN DE REGIONES GENÓMICAS ASOCIADAS CON FERTILIDAD E INDICADORES DE LA CAPACIDAD DE FERTILIZACIÓN DE TOROS HOLSTEIN

Molina, G.<sup>1,2</sup>, Carabaño, M.J.<sup>2</sup>, Karoui, S.<sup>2</sup>, Díaz, C.<sup>2</sup>

<sup>1</sup>Universidad Politécnica de Valencia, Ciencia Animal, Camí de Vera, s/n., 46022 Valencia, Spain. <sup>2</sup>INIA, Mejora Genética Animal, Ctra. La Coruña Km 7.5, 28040 Madrid, Spain.

## INTRODUCCIÓN

Desde el siglo pasado, en los programas de mejoramiento genético de los bovinos de leche se ha realizado una intensa selección en favor del incremento en la producción de leche (DeJarnette et al., 2004). Sin embargo, dicho incremento ha llevado asociado efectos indeseables en la fertilidad de la vaca debido a la compleja y antagonista correlación entre los dos bloques de caracteres (Lucy, 2001; Dobson et al., 2008; Berry et al., 2014). Por lo tanto, los caracteres de fertilidad de la vaca han sido incluidos como objetivos de mejoramiento genético en los programas de selección en varios países (Miglior et al., 2005). La fertilidad del toro es menos considerada, aunque una evaluación fenotípica nacional de la fertilidad por servicio del toro se proporciona en Estados Unidos desde el 2008 (Kuhn y Hutchison, 2008; Abdollahi-Arpanahi et al., 2017) y la modelización para obtener evaluaciones de la fertilidad del toro ha sido considerada en otros países (Berry et al., 2011; Tiezzi et al., 2011). La fertilidad del toro tiene una pequeña pero significativa contribución al éxito de la inseminación artificial (IA) (Averill et al., 2004; Jamrozik et al., 2005) y esta ganando relevancia debido a su importancia para la industria de la IA y la posibilidad de selección directa usando características seminales con valores de heredabilidad moderados (DeJarnette et al., 2004; Druet et al., 2009; Karoui et al., 2011). La cantidad creciente de información genética disponible de esquemas de selección genética brinda valiosa información para mejorar la selección de la fertilidad en los toros mediante el uso del conocimiento de genes o regiones génicas involucradas en la fertilidad del macho.

El objetivo de este estudio fue identificar regiones génicas asociadas significativamente a caracteres de fertilidad en el toro como son el número de inseminaciones por preñez y parámetros de calidad del semen.

## MATERIALES Y MÉTODOS

### Datos

Tres tipos de fenotipos asociados con la fertilidad del toro fueron usados en el subsecuente estudio de asociación del genoma completo (GWAS), caracteres de calidad seminal (concentración (CONC), volumen (VOL), motilidad individual (IM), motilidad en masa (MM), motilidad post-descongelación (PTM) y número de espermatozoides por eyaculado (NPSZ)), nivel de fragmentación del ADN (post-descongelación o nivel basal, SDF\_0, y después de 6 horas de incubación a 37°C, SDF\_6), y los resultados de IA (EAI). Los datos rutinarios de calidad seminal fueron proporcionados por Aberekin S.A. y corresponden a semen de 502 toros obtenidos en un periodo de 17 años (Karoui et al., 2011). Los valores de fragmentación del ADN fueron obtenidos de 6 muestras de semen congelado de 201 toros del mismo centro de IA, usando el Sperm-Halomatrix kit (Halotech DNA, Madrid, Spain) en concordancia con los protocolos establecidos (Karoui et al., 2012). Los resultados de la inseminación artificial fueron obtenidos del sistema oficial del registro reproductivo monitoreado por la Federación Vasca de Asociaciones Holstein (EFRIFE). Los datos incluyeron registros de inseminaciones desde 1995 al 2008. En total, 955 toros fueron evaluados a través de 511,494 inseminaciones artificiales a 136,255 vacas. De los fenotipos reales para cada carácter, el valor de un toro fue obtenido ajustando para cada carácter un modelo lineal y el efecto permanente del toro. En todos los casos, el efecto del toro fue un efecto aleatorio con una estructura diagonal (i.e., relaciones genéticas no incluidas) para la matriz de (co)varianza (ver Karoui et al., 2011 y Karoui et al., 2012 para más detalles). Finalmente, los datos de 431, 199 y 715 toros se han utilizado para calidad seminal, fragmentación del ADN y Evaluación de la Inseminación Artificial, respectivamente. El resumen estadístico de los caracteres está en la Tabla 1.

Los genotipos del chip de Illumina para bovinos SNP50 de 754 toros que presentan uno o más fenotipos fueron utilizados en el estudio. Los SNPs con un MAF (<0.0001) y un "call rate" por SNP (<0.95) fueron eliminados. Finalmente, 51048, 45694 y 46648 marcadores para los resultados de la IA, fragmentación del ADN y calidad seminal, respectivamente, fueron usados en este estudio.

## Análisis Estadístico

Los estudios de asociación del genoma completo para los nueve caracteres fueron realizados por separado. El análisis de asociación se basa en un modelo mixto lineal que considera el efecto aditivo de cada marcador candidato (efecto fijo) para ser testado y un efecto poligénico (efecto aleatorio) capturado de la matriz de relaciones genómicas.

$$y_{ij} = a + b_i x_j + g_j + e_{ij}$$

, donde  $y_{ij}$  es el fenotipo de un carácter dado para el individuo  $j$ ,  $a$  es el término medio,  $b_i$  es el efecto aditivo del  $j^{\text{th}}$  SNP que va a ser testado para determinar asociación,  $x_j$  es el genotipo del SNP del individuo  $j$  codificado como 0, 1 o 2 y  $g_j$  es el efecto poligénico del individuo  $j$  para el carácter y  $e_{ij}$  es el término residual.  $g_j$  fue tratado como un efecto aleatorio con (co)varianzas entre los efectos poligénicos siendo la matriz de relaciones genómicas. La inclusión de la matriz de relaciones genómicas permite corregir una posible estratificación en la población. Para evitar la pérdida de poder en la detección, debido al doble ajuste del marcador candidato en el modelo y el efecto de los marcadores en desequilibrio de ligamiento (LD) con el SNP testado, el cálculo de las relaciones genómicas entre individuos fue hecho retirando todos los SNPs del cromosoma donde el SNP candidato estaba localizado.

Todos los análisis se realizaron utilizando el software *Genome-Wide Complex Trait Analysis* (GCTA) (Yang et al., 2011). Las estimas de la sustitución del efecto alélico de cada marcador, así como las pruebas estadísticas para determinar la significancia de esos efectos fueron proporcionados por el software. Para corregir por las múltiples pruebas del análisis de asociación, el procedimiento de Benjamini y Hochberg (BH) (Benjamini y Hochberg, 1995) para controlar el ratio de descubrimientos falsos (FDR) al 10% fue utilizado.

El FDR es la proporción esperada (Q) de rechazos erróneos (V) entre todos los rechazos (R) (Benjamini and Yekutieli, 2001). Desde que uno nunca conoce cuales son las señales verdaderas, el valor esperado de FDR es utilizado para fijar el p-valor que da a el nivel deseado,

$$Q_e = E(Q) = E\left(\frac{V}{R}\right) = \frac{\text{sum}(p_i \leq p)}{p * N}$$

, donde  $p$  es el umbral para los p-valores ( $p_i$ ) que producen el FDR deseado y  $N$  es el número de pruebas.

Las gráficas Manhattan del  $-\log$  (p-valores) y las correcciones BH de los p-valores nominales proporcionados por GCTA fueron obtenidos usando el lenguaje de programación R (R Development Core Team 2009). Los genes candidatos fueron buscados en ventanas de 1Mb alrededor de los SNPs que fueron declarados significantes por el procedimiento BH. La herramienta Biomart de Ensembl (cow genes UMD 3.1) fue empleada para localizar los genes.

## RESULTADOS Y DISCUSIÓN

En Fig. 1 se presentan los *Manhattan plot* para cada uno de los caracteres estudiados. En dicha gráfica se indica el umbral de significación para cada uno de ellos. Aceptamos como asociaciones significativas aquellas con un p-valor nominal entre  $4 \times 10^{-5}$  y  $5 \times 10^{-5}$  dependiendo de cada fenotipo, que se corresponde a  $\text{FDR} < 0.1$ . La fragmentación del ADN del esperma a las 0 y 6 horas después de la descongelación arrojaron señales más fuertes (menores p-valores) que los otros caracteres, indicando que esos caracteres podrían reflejar en una manera más directa la acción de los genes que determinan el fenotipo. La Fig. 2 representa el número de SNPs declarados significativos (2.A) y el número de genes encontrados en una ventana de 1Mb alrededor de dichos SNPs (2.B) para cada carácter (diagonal) y los compartidos con otros caracteres (fuera de la diagonal). Como se mencionó anteriormente, para SDF\_0 y SDF\_6 se detectaron más SNPs (24 y 19 SNPs, respectivamente) que otros caracteres. Para PTM se encontraron 9 marcadores. Para VOL, CONC, y NSPZ no se detectaron señales relevantes. Se observaron pocos SNPs significativos compartidos entre caracteres, excepto para IM y MM y los dos caracteres de fragmentación. Para los genes asociados (figura 2.B), un patrón similar fue observado, con un solo gen compartido entre los grupos de fenotipos.

Los SNPs asociados con SDF\_0 y SDF\_6 están localizados dentro o cerca de genes involucrados en funciones biológicas relacionadas con la reproducción, como *Spermatid perinuclear RNA-binding protein* gen (*SPRB*) localizado en el cromosoma 2, interleukin enhancer binding factor 2 gen (*ILF2*) localizado en el BTA 3 (238 kb desde el SNP), que participa en la generación de gametos y *WAPL cohesin release factor* gen localizado en el BTA 28 (432 kb desde el SNP), que está involucrado en el proceso de meiosis. El cromosoma 11 presenta los SNPs con mayor significación relacionados con la fragmentación del ADN del esperma. Hemos

encontrado SNPs asociados con los caracteres :EAI, PTM, SDF\_0, SDF\_6, IM and MM que están situados cerca de genes que pertenecen a la mismas familias que han sido previamente asociadas a calidad seminal de los toros y caracteres de fertilidad (Fortes et al., 2012; Peñagaricano et al., 2012). Un análisis funcional de genes podría ayudar para ir desde candidatos posicionales a funcionales en este estudio. La correlación entre SNPs con la habilidad de fertilización permitiría, a través de la selección genómica incrementar la eficiencia reproductiva en rebaños lecheros.

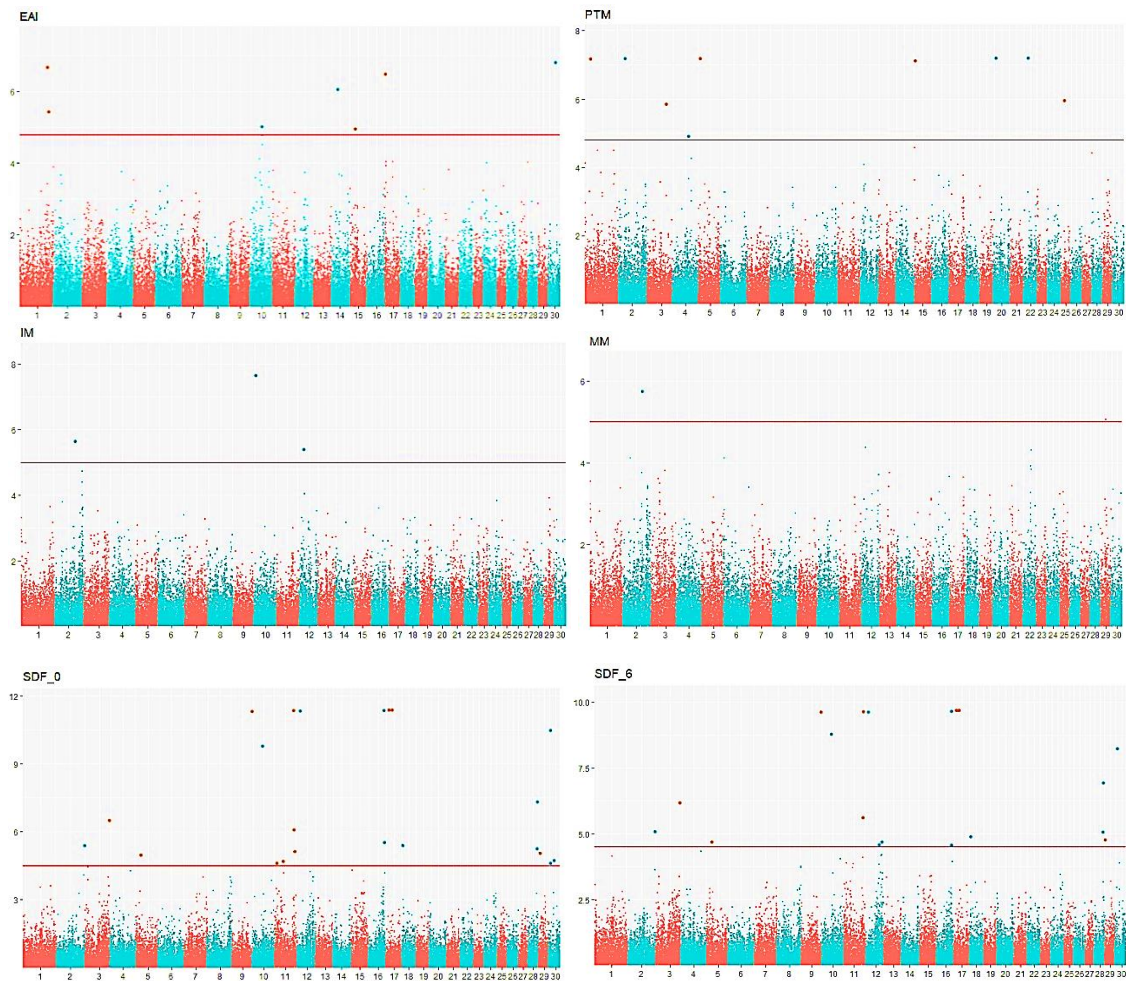
**Agradecimientos:** Los autores agradecen a Aberekin y a EFIFRE por la cesión de los datos fenotípicos, así como a CONAFE por la cesión de los genotipados para la realización de este estudio y al proyecto RTA2007-0071, que proporcionó la financiación necesaria.

#### REFERENCIAS BIBLIOGRÁFICAS

- Abdollahi-Arpanahi, Rostam, Gota Morota, y Francisco Peñagaricano. 2017. "Predicting Bull Fertility Using Genomic Data and Biological Information." *Journal of Dairy Science* 100 (12): 9656–66. doi:10.3168/jds.2017-13288.
- Averill, T a, R Rekaya, y K Weigel. 2004. "Genetic Analysis of Male and Female Fertility Using Longitudinal Binary Data." *Journal of Dairy Science* 87 (11). Elsevier: 3947–52. doi:10.3168/jds.S0022-0302(04)73534-1.
- Benjamini, Y, y Y Hochberg. 1995. "Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society B* 57 (1): 289–300. doi:10.2307/2346101.
- Benjamini, Yoav, y Daniel Yekutieli. 2001. "The Control of the False Discovery Rate in Multiple Testing under Dependency." *Annals of Statistics* 29 (4): 1165–88. doi:10.1214/aos/1013699998.
- Berry, D. P., R. D. Evans, y S. Mc Parland. 2011. "Evaluation of Bull Fertility in Dairy and Beef Cattle Using Cow Field Data." *Theriogenology* 75 (1). Elsevier Inc.: 172–81. doi:10.1016/j.theriogenology.2010.08.002.
- Berry, D. P., E. Wall, y J. E. Pryce. 2014. "Genetics and Genomics of Reproductive Performance in Dairy and Beef Cattle." *Animal* 8 (s1): 105–21. doi:10.1017/S1751731114000743.
- DeJarnette, J.M., C.E. Marshall, R.W. Lenz, D.R. Monke, W.H. Ayars, y C.G. Sattler. 2004. "Sustaining the Fertility of Artificially Inseminated Dairy Cattle: The Role of the Artificial Insemination Industry." *Journal of Dairy Science* 87 (June 2003). Elsevier: E93–104. doi:10.3168/jds.S0022-0302(04)70065-X.
- Dobson, H, S L Walker, M J Morris, J E Routly, y R F Smith. 2008. "Why Is It Getting More Difficult to Successfully Artificially Inseminate Dairy Cows?" *Animal: An International Journal of Animal Bioscience* 2 (8): 1104–11. doi:10.1017/S175173110800236X.
- Druet, T., S. Fritz, E. Sellem, B. Basso, O. Gérard, L. Salas-Cortes, P. Humblot, X. Druart, y A. Eggen. 2009. "Estimation of Genetic Parameters and Genome Scan for 15 Semen Characteristics Traits of Holstein Bulls." *Journal of Animal Breeding and Genetics* 126 (4): 269–77. doi:10.1111/j.1439-0388.2008.00788.x.
- Fortes, Marina R.S., Antonio Reverter, Rachel J. Hawken, Sunduimijid Bolormaa, y Sigrid A. Lehnert. 2012. "Candidate Genes Associated with Testicular Development, Sperm Quality, and Hormone Levels of Inhibin, Luteinizing Hormone, and Insulin-Like Growth Factor 1 in Brahman Bulls1." *Biology of Reproduction* 87 (3): 1–8. doi:10.1095/biolreprod.112.101089.
- Jamrozik, J., J. Fatehi, G.J. Kistemaker, y L.R. Schaeffer. 2005. "Estimates of Genetic Parameters for Canadian Holstein Female Reproduction Traits." *Journal of Dairy Science* 88 (6). Elsevier: 2199–2208. doi:10.3168/jds.S0022-0302(05)72895-2.
- Karoui, Sofiene, Clara Díaz, Magdalena Serrano, Roger Cue, Idoia Celorrio, y María J. Carabaño. 2011. "Time Trends, Environmental Factors and Genetic Basis of Semen Traits Collected in Holstein Bulls under Commercial Conditions." *Animal Reproduction Science* 124 (1–2): 28–38. doi:10.1016/j.anireprosci.2011.02.008.
- Karoui S, Díaz C, Amenabar M, Roy R, M Serrano, Ugarte E, y Carabaño MJ. 2012. "Is Sperm DNA Fragmentation a Good Marker for Fi Eld AI Bull Fertility ? 1," 2437–49. doi:10.2527/jas2011-4492.
- Kuhn, M.T., y J.L. Hutchison. 2008. "Prediction of Dairy Bull Fertility from Field Data: Use of Multiple Services and Identification and Utilization of Factors Affecting Bull Fertility." *Journal of Dairy Science* 91 (6). Elsevier: 2481–92. doi:10.3168/jds.2007-0743.
- Lucy, M.C. 2001. "Reproductive Loss in High-Producing Dairy Cattle: Where Will It End?" *Journal of Dairy Science* 84 (6). Elsevier: 1277–93. doi:10.3168/jds.S0022-0302(01)70158-0.
- Miglior, F., B.L. Muir, y B.J. Van Doormaal. 2005. "Selection Indices in Holstein Cattle of Various Countries." *Journal of Dairy Science* 88 (3). Elsevier: 1255–63. doi:10.3168/jds.S0022-0302(05)72792-2.
- Peñagaricano, F., K. A. Weigel, y H. Khatib. 2012. "Genome-Wide Association Study Identifies Candidate Markers for Bull Fertility in Holstein Dairy Cattle." *Animal Genetics* 43 (SUPPL. 1): 65–71. doi:10.1111/j.1365-2052.2012.02350.x.
- Tiezzi, F., M. Penasa, C. Maltecca, A. Cecchinato, y G. Bittante. 2011. "Exploring Different Model Structures for the Genetic Evaluation of Dairy Bull Fertility." *Agriculturae Conspectus Scientificus* 76 (3): 239–43.
- Yang, Jian, S. Hong Lee, Michael E. Goddard, y Peter M. Visscher. 2011. "GCTA: A Tool for Genome-Wide Complex Trait Analysis." *American Journal of Human Genetics* 88 (1). The American Society of

**Tabla 1. Número de toros (N), media, desviación típica (D.T.), mínimo (Min) y máximo (Max) valores fenotípicos de los nueve caracteres analizados.**

Carácter	N	Media	D.T.	Min	Max
CONC	431	1221	310.25	65	1980
IM	431	83.54	6.16	50	93.21
MM	431	4.11	0.45	2.81	5
PTM	431	48.68	9.44	5	67.5
NSPZ	431	5.95	1.97	0.28	14.31
VOL	431	4.97	1.43	1.5	10.19
SDF_0	199	3.34	2.32	1	20.4
SDF_6	199	3.59	2.33	1	19.5
EAI	715	0.35	0.07	0.09	0.65



**Fig 1. Gráficas Manhattan corresponden a los resultados de los estudios de asociación del genoma completo, el umbral indica el log (10) inverso del p-valor nominal de cada carácter considerando la corrección BH, para seis de los caracteres de fertilidad: EAI, PTM, SDF\_0, SDF\_6, IM, MM, PTM.**

A



B



**Fig 2.** La matriz representa en la diagonal el número de SNPs en (A) y genes en (B) asociados a cada carácter, y fuera la relación entre caracteres.

## IDENTIFICATION OF GENOMIC REGIONS ASSOCIATED WITH FERTILITY AND INDICATORS OF THE ABILITY OF FERTILIZATION OF HOLSTEIN BULLS

**ABSTRACT:** The sustainability of dairy farms is at risk due to the decline in the reproductive efficiency of cows with high rates milk production throughout the world. The cow fertility has been included in the genetic improvement programs as a selection goal. However, male fertility has been seen to a lesser extent. Therefore, the objective of this study was to identify single nucleotide polymorphisms (SNPs) significantly associated to the phenotypic evaluation of fertility of the bull and semen quality parameters. Phenotypes included 715 animals with evaluation of artificial insemination (EAI), 199 bulls with data of rate of sperm DNA fragmentation at 0 and 6 hours (SDF\_0/6) and 431 bulls with records of individual motility (IM), mass motility (MM), post thawing motility (PTM), volume (VOL), concentration (CONC) and number of spermatozoa (NSPZ). Genotypes of 754 bulls of the Illumina BovineSNP50 Bead Chip have been used in the study. SNPs were edited on their minor allele frequency ( $>0.0001$ ) and call rate ( $>0.95$ ). An analysis of association based on a linear mixed model that considers the additive effect of each SNP to be tested for association and a polygenic effect was used to estimate the effect of markers for each of the examined traits. The Benjamini and Hochberg method to control FDR at 10% was used to determine the SNPs with relevant signals for each trait. The SNPs associated with SDF\_0 and SDF\_6 are in or near genes with biological functions involved in reproduction, such as Spermatid perinuclear RNA-binding protein (SPRB) in BTA 2, ILF2 in BTA 3 (238 kb from SNP) that participates in the gamete generation and WAPL in BTA 28 (432 kb from SNP), involved in meiosis process. The association between SNPs with the bull's ability of fertilization would allow, through genomic selection, an increase of reproductive efficiency in dairy herds.

**Key Words:** *Holstein bulls, genome-wide association, male fertility, sperm DNA fragmentation, semen quality.*