



## Análisis de genoma completo para identificar variaciones asociadas con circunferencia escrotal en bovinos Charolais

G. Manuel Parra-Bracamonte<sup>1</sup> , Juan G. Magaña-Monforte<sup>2</sup> ,  
Xochitl F. De la Rosa-Reyna , Juan C. Martínez-González<sup>3</sup> ,  
Francisco J. Jahuey-Martínez<sup>4</sup> , Víctor R. Moreno-Medina 

Instituto Politécnico Nacional, Centro de Biotecnología Genómica. Reynosa, Tamaulipas, México

### Genome-wide association study for identification of associated variations to scrotal circumference in Charolais cattle

**Abstract.** The scrotal circumference is one of the main traits of bovine male and female fertility. Its study through genetic and genomic tools would help the discovery of favorable variation and genetic architecture to promote its selection. This study aimed to carry out a genome-wide association study using a large microarray of SNPs to identify variations and candidate genes associated with the studied trait, for which information from 141 Charolais bulls was used. Four markers found, rs110416965, rs110130953, rs43423602, and rs29003417, showed significant association at the genome level. Three of them were found within candidate genes (RBFOX2, TTF2, and CXCL2). This information contributes to the understanding of the genetic architecture of this trait.

**Key words:** Candidate gene, Fertility, Microarray, SNP.

**Resumen.** La circunferencia escrotal es una de las características relacionadas con la fertilidad del macho y la hembra bovina. Su estudio a través de herramientas genéticas y genómicas ayudaría al descubrimiento de variación favorable en su arquitectura genética para fomentar su selección. El propósito del presente trabajo fue realizar un estudio de genoma completo mediante un microarreglo de SNPs, para identificar variaciones y genes candidatos asociados con la característica, para lo que se utilizó la información de 141 toros de la raza Charolais. Cuatro marcadores encontrados, rs110416965, rs110130953, rs43423602 y rs29003417, mostraron asociación significativa a nivel de genoma. Tres de ellos se encontraron dentro de genes candidatos (RBFOX2, TTF2 y CXCL2) para ser estudiados más específicamente. Esta información contribuye al entendimiento de la arquitectura genética de esta característica.

**Palabras clave:** Fertilidad, Gen candidato, Microarreglo, SNP.

### Análise do genoma completo para identificar variações associadas à circunferência escrotal em bovinos Charolês

**Resumo.** A circunferência escrotal é uma das características relacionadas à fertilidade do macho e da fêmea bovina. Seu estudo por meio de ferramentas genéticas e genômicas auxiliaria na descoberta de variação favorável em sua arquitetura genética para promover sua seleção. O objetivo deste trabalho foi realizar um estudo do genoma completo usando um microarray de SNPs, para identificar variações e genes candidatos associados à característica, para o qual foram utilizadas informações de 141 touros Charolês. Quatro marcadores encontrados, rs110416965, rs110130953, rs43423602 e rs29003417, mostraram associação significativa no nível do genoma. Três deles foram encontrados em genes candidatos (RBFOX2, TTF2 e CXCL2) a serem estudados mais especificamente. Essas informações contribuem para o entendimento da arquitetura genética dessa característica.

**Palavras-chave:** Fertilidade, Gene candidato, Microarray, SNP.

<sup>1</sup> Autor para la correspondencia: [gparra@ipn.mx](mailto:gparra@ipn.mx)

<sup>2</sup> Universidad Autónoma de Yucatán, Facultad de Medicina Veterinaria y Zootecnia. Mérida, Yucatán.

<sup>3</sup> Universidad Autónoma de Tamaulipas, Facultad de Ingeniería y Ciencias. Victoria, Tamaulipas

<sup>4</sup> Universidad Autónoma de Chihuahua, Facultad de Zootecnia y Ecología. Chihuahua, Chihuahua.

## Introducción

La fertilidad es una de las características productivas más importantes para lograr mayor eficiencia reproductiva y rentabilidad en los sistemas de producción animal. La circunferencia escrotal es un carácter de fácil medición, que ha sido relacionado con la fertilidad en machos y en hijas de sementales. Esta característica es utilizada para predecir la calidad y cantidad de tejido productor de espermatozoides y la edad a la pubertad (Cammack et al., 2009). Con una heredabilidad de alrededor de 0.42, su respuesta a la selección se ha correlacionado con mayor tasa de preñez en vaquillas de 16 meses y un incremento de la precocidad (Terakado et al., 2015).

El mejoramiento genético contribuye a la variación del desempeño de los animales y la implementación de

herramientas genómicas para ayudar a las estrategias de selección, incrementando el cambio genético favorable de los animales reproductores en menor tiempo. El uso de miles de marcadores genéticos de tipo SNP (polimorfismos de un solo nucleótido) en estudios de genoma completo (GWAS) ayuda a la identificación de regiones genómicas, loci de características cuantitativas (QTLs) o genes candidatos que influyen caracteres de interés económico en las especies productivas domésticas (Zhang et al., 2012).

El propósito del presente estudio fue realizar un análisis de genoma completo para identificar variaciones asociadas con la circunferencia escrotal de bovinos Charolais de registro.

## Materiales y Métodos

Datos de circunferencia escrotal de 141 sementales Charolais de registro fueron utilizados en este estudio. Todos los animales tenían información de circunferencia escrotal ajustada a 365 días (CEA). Estos animales provenían de tres hatos del norte de México manejados bajo condiciones extensivas. Las muestras fueron tipificadas con el panel GeneSeek® Genomic Profiler Bovine HDTM (GGPHD) con >76 mil SNPs.

El control de calidad del panel incluyó la exclusión de los marcadores con una frecuencia de alelo menor <0.05, de la cual fueron eliminadas 4255 variantes; la evaluación de genotipos faltantes, que consideró la eliminación de 367 variantes, permitiendo el 90% de tasa de tipificación; y el análisis de locus en equilibrio de Hardy-Weinberg, considerando una  $p < 0.001$ , causó la eliminación de 741 variantes. Al final, se utilizaron 71344 SNPs para el

GWAS. Los análisis de control de calidad se realizaron en el software PLINK (Purcell et al., 2007).

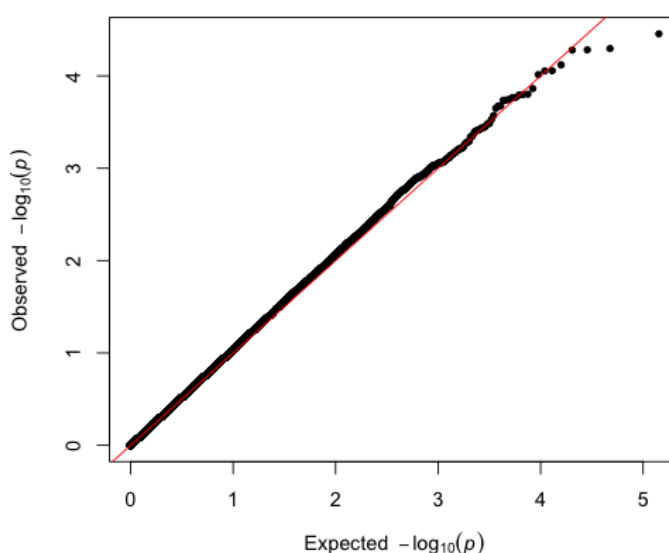
Previo al análisis de asociación, se generaron los fenotipos residuales de la variable CEA, para lo cual se ajustó un modelo lineal general, incluyendo los efectos fijos de hato, año y época de nacimiento, y edad de la madre como covariable lineal, utilizando el paquete estadístico SAS v. 9.4 (SAS, 2013).

El GWAS se realizó utilizando los valores residuales de CEA como variable y considerando el ajuste del primer componente principal de la estructura poblacional como covariable. El análisis fue realizado en el programa PLINK (Purcell et al., 2007). La visualización de los resultados y la estimación de la gráfica Q-Q se realizó en el software QQman (Turner, 2018) del entorno R (R Core Team, 2013).

## Resultados

En la Figura 1 se observa la gráfica Q-Q, en la que se muestra la distribución esperada de los resultados del GWAS, aunque con un número menor del esperado en SNPs significativos.

El Cuadro 1 muestra los 8 marcadores más importantes obtenidos por el análisis y sus características estadísticas más importantes. Los SNPs más importantes encontrados a nivel de genoma fueron rs110416965, rs110130953, rs43423602 y rs29003417. Otros marcadores cercanos a este nivel, pero asociados con nivel sugestivo, fueron rs110697381, rs42087824, rs109058292 y rs109304691.



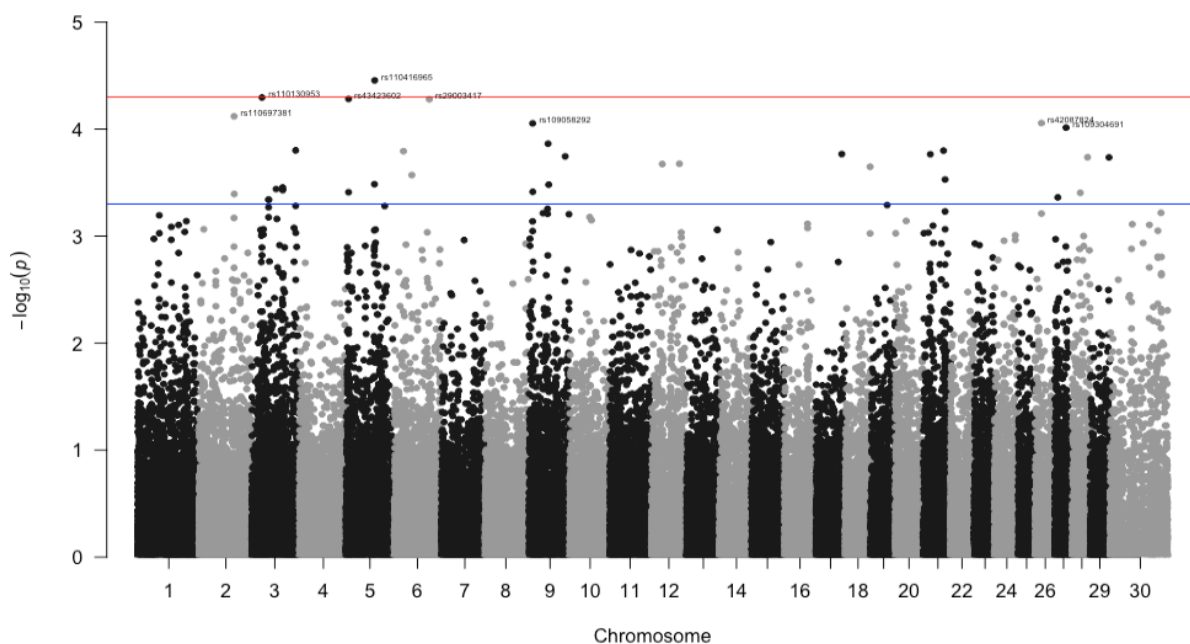
**Figura 1.** Gráfica de Q-Q del ajuste de genoma completo para circunferencia escrotal.

**Cuadro 1.** Características de los loci asociados a circunferencia escrotal en bovinos Charolais.

Locus	Crm	Posición PB	$\beta$	EE	R <sup>2</sup>	P	P-CG
rs110416965	5	74275630	-1.432	0.3322	0.1423	3.51E-05	5.68E-05
rs110130953	3	26246445	-1.227	0.2909	0.137	5.06E-05	8.04E-05
rs43423602	5	6764453	-1.242	0.2952	0.1365	5.22E-05	8.29E-05
rs29003417	6	90694839	-2.295	0.5456	0.1364	5.25E-05	8.34E-05
rs110697381	2	90714173	1.35	0.3285	0.131	7.58E-05	0.0001183
rs42087824	26	17006703	1.298	0.3188	0.1289	8.77E-05	0.0001359
rs109058292	9	9688314	1.186	0.2916	0.1288	8.83E-05	0.0001369
rs109304691	27	29037564	1.145	0.2831	0.1274	9.68E-05	0.0001492

Crm: Cromosoma, PB: Pares de bases,  $\beta$  = Coeficiente de regresión, EE: Error estándar, R<sup>2</sup>: Coeficiente de determinación, P: Valor de p de prueba de Wald, P-CG: Valor de P corregido genómicamente.

La Figura 2 muestra la gráfica de Manhattan donde se identifican los SNPs más significativos a nivel de genoma (arriba de línea azul) con un  $p\text{-log}_{10}(5e-5)$ .

**Figura 2.** Gráfica de Manhattan indicando los SNPs con efecto significativo sobre circunferencia escrotal de bovinos Charolais.

### Discusión

En el presente estudio, se realizó un GWAS para la característica de CEA en toros de la raza Charolais y se encontraron cuatro SNPs significativamente asociados a nivel de genoma. De manera importante, tres de ellos se localizaron dentro de genes y los cuatro poseen un efecto negativo sobre la CEA.

El marcador rs110416965 es un locus en el intrón del gen RBFOX2, localizado en el cromosoma 5, codificante de la proteína ligadora de RNA Homóloga 2 a Fox-1, y corresponde a una variante transcricional.

El marcador rs110130953 está localizado en un intrón del gen TTF2 (Factor de Transcripción y Terminación 2) del cromosoma 3 y también corresponde a una variante transcricional. Este gen produce dos transcritos, uno de 1163aa y otro de 1132aa. Aunque no hay información asociada con indicadores reproductivos o CEA, Mutch (2017) identificó este marcador como asociado con el consumo residual de alimento en animales Angus. El marcador rs43423602 es una variante intergénica localizada en el cromosoma 5.

El marcador rs29003417 es una variante río arriba del gen, CXCL2 (Quimocina 2 de motive C-X-C). Las quimiocinas controlan múltiples etapas de la miogénesis, y se ha demostrado que la miostatina, en particular en animales de doble musculatura, está implicada en la regulación de la expresión de CXCL2, afectando la recuperación de las células musculares (Iwasaki et al., 2013). Aunque los animales estudiados no fueron portadores de la mutación Q204X, propia de la raza Charolais, sería importante realizar otros estudios como gen candidato a CEA y su relación con la miostatina, que ha sido reportada como un síndrome relacionado a varias características correlacionadas con el doble músculo. Adicionalmente, CXCL2 ha sido sugerido como biomarcador de expresión para mastitis en ganado bovino (Sharifi et al., 2018).

De entre los marcadores asociados a nivel sugestivo, el marcador rs110697381 es una variante río arriba (5') del

gen ALS2 (Alsina Rho factor de intercambio de nucleótidos de guanina). Las mutaciones en el gen ALS2 que codifica la proteína Alsina, han sido asociadas con el inicio de esclerosis lateral amiotrófica recesiva juvenil (Devon et al., 2005). El marcador rs42087824 está localizado en un intrón del gen TCTN3 (miembro de la familia tectónica 3), en el cromosoma 26. No hay documentación relacionada en ganado bovino, pero este gen ha sido asociado a desordenes genéticos recesivos (Chen et al., 2020). Finalmente, rs109058292 y rs109304691 corresponden a variantes intergénicas cuya asociación podría estar relacionada por desequilibrio de ligamiento a otros genes causales. Estudios más detallados sobre las regiones flanqueantes de los loci asociados ayudarían a dilucidar la relación de estas variaciones encontradas con la característica asociada, adicionalmente un mayor tamaño de muestra ayudaría a incrementar la confiabilidad de los resultados.

## Conclusiones

Cuatro marcadores genéticos asociados a nivel de genoma se encuentran asociados a la circunferencia escrotal de bovinos Charolais. Tres de esos marcadores se encuentran dentro de genes y son genes candidatos a estudios más específicos sobre la característica. Esta información contribuye al entendimiento de la arquitec-

tura genética de esta característica. Un análisis de validación más amplio con un mayor tamaño de muestra ayudaría a verificar los presentes resultados que en forma interesante incluyen la identificación de genes en regiones intrónicas y regulatorias de seis genes candidatos.

## Agradecimientos

Los autores agradecen el apoyo del Instituto Politécnico Nacional, a través del proyecto SIP20230201.

## Literatura Citada

- Cammack, K.M., Thomas, M.G. and Enns, R.M. 2009. Reproductive traits and their heritabilities in beef cattle. *The Professional Animal Scientist*, 25:517-528. [https://doi.org/10.15232/S1080-7446\(15\)30753-1](https://doi.org/10.15232/S1080-7446(15)30753-1)
- Chen, H.X., Yang, Z.Y., Hou, H.T., Wang, J., Wang, X.L., Yang, Q., Liu, L. and He, G.W. 2020. Novel mutations of TCTN3/LTBP2 with cellular function changes in congenital heart disease associated with polydactyly. *Journal of Cellular and Molecular Medicine*, 24:13751-13762. <https://doi.org/10.1111/jcmm.15950>
- Devon, R.S., Schwab, C., Topp, J.D., Orban, P.C., Yang, Y.Z., Pape, T.D., Helm, J.R., Davidson, T.L., Rogers, D.A., Gros-Louis, F. and Rouleau, G., 2005. Cross-species characterization of the ALS2 gene and analysis of its pattern of expression in development and adulthood. *Neurobiology of disease*, 18:243-257. <https://doi.org/10.1016/j.nbd.2004.10.002>
- Iwasaki, S., Miyake, M., Hayashi, S., Watanabe, H., Nagasawa, Y., Terada, S., Watanabe, K., Ohwada, S., Kitazawa, H., Rose, M.T. and Aso, H. 2013. Effect of myostatin on chemokine expression in regenerating skeletal muscle cells. *Cells Tissues Organs*, 198:66-74. <https://doi.org/10.1159/000351462>
- Mutch, J.L. 2017. Identification of genes and gene sets associated with residual feed intake in Hereford, Angus, and Simmental crossbred beef cattle (Doctoral dissertation, Washington State University).
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81:559-575. <https://doi.org/10.1086/519795>
- R Core Team (2013). R: A language and environment for statistical computing.
- SAS (2013) Statistical Analysis Software. Users' Guide Statistics Version 9.4. SAS Institute Inc., Cary.
- Sharifi, S., Pakdel, A., Ebrahimi, M., Reecy, J.M., Fazeli Farsani, S. and Ebrahimi, E. 2018. Integration of machine learning and meta-analysis identifies the transcriptomic bio-signature of mastitis disease in cattle. *PLoS ONE* 13: e0191227. <https://doi.org/10.1371/journal.pone.0191227>



- Terakado, A.P.N., Boligon, A.A., Baldi, F., Silva, J.I.V. and Albuquerque, L.G. 2015. Genetic associations between scrotal circumference and female reproductive traits in Nelore cattle. *Journal of Animal Science*, 93:2706-2713. <https://doi.org/10.2527/jas.2014-8817>
- Turner, S.D. 2018. qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. *Journal of Open Source Software*, 3:731, <https://doi.org/10.21105/joss.00731>
- Zhang, H., Wang, Z., Wang, S. and Li, H. 2012. Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology*, 3:1-10. <https://doi.org/10.1186/2049-1891-3-26>