

## REGIONES GENÓMICAS ASOCIADAS CON CARACTERES REPRODUCTIVOS EN CONEJO

Sosa-Madrid<sup>1</sup>, B.S., Santacreu<sup>1</sup>, M.A., Ibáñez-Escriche<sup>2</sup>, N. y Blasco<sup>1</sup>, A.

<sup>1</sup>Instituto de Ciencia y Tecnología Animal, Universitat Politècnica de València, P.O. Box 22012. Valencia, España.

<sup>2</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian EH25 9RG, Reino Unido.

msantacr@dca.upv.es

### INTRODUCCIÓN

El tamaño de camada es considerado uno de los caracteres más importante en producción cunícola. Sin embargo, la selección genética directa del tamaño de camada e indirecta a través de sus componentes (i. e. tasa ovulación y capacidad uterina) ha obtenido una baja respuesta a la selección (Khalil y Al-Saef, 2008). Desde el 2015, la disponibilidad de genotipados de alta densidad en conejo ha permitido la posibilidad de realizar estudios de asociación del genoma completo (GWAS). La identificación de las regiones del genoma que están directamente asociados con la variabilidad en los fenotipos puede ayudar a comprender mejor el control genético de los caracteres productivos. Por otra parte, los experimentos de selección divergente pueden ser un material excepcional para este tipo de estudios. El objetivo de este estudio fue identificar regiones genómicas asociadas al tamaño de camada y sus componentes, a partir de los datos de un experimento de selección divergente por capacidad uterina en conejo.

### MATERIAL Y MÉTODOS

**Material animal:** Los animales utilizados en este estudio proceden de tres líneas de conejo: línea alta de capacidad uterina (HUC), línea baja de capacidad uterina (LUC) y línea V seleccionada por tamaño de camada al destete. La población base de las líneas divergentes de capacidad uterina provienen de la generación 12 de la línea V. La capacidad uterina fue estimada como el número de nacidos totales en hembras unilateralmente ovariectomizadas. Después de diez generaciones de selección por capacidad uterina, la selección fue relajada. Para este experimento, las hembras HUC y LUC provienen de las generaciones 11 y 12. Las hembras de la línea V provienen de los descendientes de dos poblaciones de la generación 13 y 15, reconstituida a partir de embriones vitrificados (Santacreu et al., 2005). Los datos fenotípicos fueron recolectados en la segunda gestación de conejas intactas y contemporáneas. Los caracteres evaluados fueron: número de nacidos totales (NT), vivos (NV) y muertos (NM), la tasa de ovulación (TO), el número de embriones implantados (EI) y las supervivencias embrionaria (SE), fetal (SF) y prenatal (SP). A los doce días de la gestación se estimó TO por recuento de cuerpos lúteos y EI por laparoscopia en ambos cuernos uterinos. Al nacimiento se registró el tamaño de camada como número de nacidos totales (NT) y además se recolectó NV y NM. Además, se calculó SE como el cociente EI/TO, SF como NT/EI y SP como NT/TO. El número total de animales con datos de fenotipo y genotipo fue 183 hembras.

**Datos genómicos:** Los animales fueron genotipados utilizando un chip 200k Affymetrix Axiom OrcunSNP Array (Affymetrix, Inc. Santa Clara, CA, USA). El control de calidad fue realizado utilizando el programa ZANARDI (Marras et al., 2016). Las muestras con frecuencia de genotipos faltantes > 3% y errores en el test de segregación mendeliana fueron excluidas del estudio. El control de calidad de los SNPs consistió en incluir aquellos con una menor frecuencia alélica (MAF)  $\geq 3\%$ , tasa de genotipado  $\geq 95\%$ , test de Hardy-Weinberg con p-valor  $\geq 1,0E^{-7}$  y con posición conocida en el mapa genómico del conejo. Después del control de calidad, un total de 117.806 SNP (59%) fueron utilizadas en el análisis.

**Análisis estadístico:** El análisis de asociación del genoma completo fue realizado utilizando el siguiente modelo estadístico:

$$y = Xb + \sum_{j=1}^k z_j \alpha_j \delta_j + e$$

donde  $y$  es el vector de valores fenotípicos,  $X$  es la matriz de incidencia de efectos sistemáticos,  $b$  es un vector de efectos sistemáticos (año-estación, estado de lactación y

línea),  $z_j$  es un vector que indica el genotipo en el locus  $j$  ( $j = 1, \dots, k$ ),  $\alpha_j$  es el efecto de sustitución aleatorio para el locus  $j$ ,  $\delta_j$  es una variable aleatoria de 0/1 indicando la ausencia (con probabilidad  $\pi$ ) o presencia (con probabilidad  $1 - \pi$ ) del locus en el modelo. La proporción de SNPs con efecto diferente de cero en cada iteración es determinado utilizando un parámetro  $\pi$ . Los efectos  $\alpha_j$  de los SNPs incluidos en el modelo tienen una distribución normal  $N(0, \sigma_{\alpha_j}^2)$ , y  $e$  es un vector de efectos residuales con una distribución normal  $N(0, \sigma_e^2)$ . Esta modelización de los “priors” se denomina Bayes B. Los parámetros del modelo fueron estimados con el software GenSel v.4. El valor de  $\pi$  (0.9995) fue fijado en función del número de datos y de SNPs (Ros-Freixedes et al., 2016). La asociación entre las regiones genómicas y el carácter fue evaluada a través de la distribución marginal posterior del porcentaje de la varianza genómica explicada por cada región de 1 Mb (definida como una ventana) respecto a la varianza genómica total explicada por todos los marcadores. En el análisis se consideraron como ventanas asociadas al carácter aquellas que explicaban  $\geq 1\%$  de la varianza genómica total y las ventanas contiguas que explicaban al menos el 0.5% de la varianza genómica total. La identificación de los genes en estas regiones cromosómicas fue realizado con “UCSC Genome Browser” (Rosenbloom et al., 2015). Posteriormente, se realizó un análisis de las anotaciones funcionales de estos genes con las base de datos “Gen Ontology” (GO) (Ashburner et al., 2000) y DAVID v.6.8 (<http://David.abcc.ncifcrf.gov/>) (Jiao et al., 2012).

## RESULTADOS Y DISCUSIÓN

Las regiones genómicas asociadas a los caracteres reproductivos NT, NV, EI y TO se muestran en la Tabla 1. No se encontraron regiones asociadas para NM, SE, SF y SP. Las anotaciones en cuanto a la información funcional sobre los genes están incompletas en las bases de datos de conejo. Por tanto, para estudiar la función de los genes y su participación en las diferentes rutas metabólicas se utilizaron las anotaciones de los genes ortólogos en humano, rata y ratón. Los criterios para seleccionar términos GO o funciones relevantes fueron un p-valor  $< 0,01$  y estar presente en al menos dos de las especies consultadas. Los resultados de DAVID v.6.8, respecto a las funciones relevantes, fueron la respuesta celular al estímulo de prostaglandinas, la regulación de la calidad biológica, la homeostasis celular en procesos de oxidación-reducción, la regulación de la mineralización de los huesos y procesos reproductivos. Los genes más representativos involucrados en funciones reproductivas fueron *PTGDR*, *BMP4*, *STYX*, *ANKH* y *CDKN3*. El gen *PTGDR* tiene importancia para la señalización de la diferenciación de las células de Sertoli y germinales en el desarrollo embrionario de los testículos y la espermatogénesis en adultos (Blesson and Sahlin, 2014). Los genes de la familia factor-beta de crecimiento y transformación, como el gen *BMP4*, están ampliamente relacionados con el crecimiento y desarrollo folicular en mamíferos (Al-Samerria et al., 2015), el desarrollo del trofoblasto, la fijación e implantación del embrión y la formación de la placenta en humanos (Li y Parast, 2014). El gen *STYX* influyen en la degeneración de espermatogonias en los túbulos seminíferos (Matzuk y Lamb, 2002). Además en ratón, los genes *BMP4*, *ANKH* y *CDKN3* parecen estar involucrados en procesos de desarrollo embrionario (Gurley et al., 2006; Goggolidou et al., 2013).

En nuestro estudio, las ventanas encontradas en el cromosoma 17, donde se encuentran los genes *PTGDR*, *BMP4*, *STYX* y *CDKN3*, explicaron un 16% de la varianza genómica del tamaño de camada. Además, estas regiones se solaparon para NT, NV y EI lo que concuerda con las correlaciones genéticas encontradas entre estos caracteres en conejo (García y Baselga, 2002). Los resultados revelan áreas importantes para el tamaño de camada y un posible efecto pleiotrópico de algunos genes del cromosoma 17 por el solapamiento para tres caracteres. Las regiones genómicas y los genes identificados por este estudio muestran un avance en el conocimiento del tamaño de camada en conejo y ofrecen un nuevo punto de partida para futuras investigaciones sobre genómica relacionada tamaño de camada en poblaciones comerciales de conejo.

## REFERENCIAS BIBLIOGRÁFICAS

- Al-Samerria, S., I. Al-Ali, J. R. McFarlane, et al. 2015. *Reprod.* 149:403-411.
- Ashburner, M., C. A. Ball, J. A. Blake, et al. 2000. *Nat. Genet.* 25:25-29.
- Blesson, C. S. & Sahlin, L. 2014. *Recept. Clin. Investig.* 2:1-12.
- García, M. L. & Baselga, M. 2002. *Lives. Prod. Sci.*

74:45-53. • Goggolidou, P., S. Soneji, N. Powles-Glover, et al. 2013. *Mamm. Genom.* 24:459-472. • Gurley, K. A., H. Chen, C. Guenther, et al. 2006. *J. Bone and Miner. Res.* 21:1238-47. • Jiao, X., B. T. Sherman, D. W. Huang, et al. 2012. *Bioinf.* 28:1805-1806. • Khalil, M. H. & Al-Saef, A. M. 2008. *Proc. 9th World Rabbit Congr.:*1-22. • Li, Y. & Parast, M. M. 2014. *Int. J. Dev. Biol.* 58:239-246. • Marras, G., A. Rossoni, H. Schwarzenbacher, et al. 2016. *Anim. Genet.* 48:121-121 • Matzuk, M. M., & Lamb, D. J. 2002. *Nat. Cell Biol.* 4:S33-S40. • Ros-Freixedes, R., S. Gol, R. N. Pena, et al. 2016. *PLoS One* 11:e0152496. • Rosenbloom, K. R., J. Armstrong, G. P. Barber, et al. 2015. *Nucleic Acids Res.* 43:D670-D681. • Santacreu, M. A., M. L. Mocé, A. Climent & Blasco, A.. 2005. *J. Anim. Sci.* 83:2303-2307

**Agradecimientos:** El trabajo ha sido financiado por la Secretaría de Estado I+D+i del Ministerio de Economía y Competitividad de España (AGL2014-55921-C2-1-P) y B. Samuel Sosa-Madrid por la Secretaría Nacional de Ciencia, Innovación y Tecnología (SENACYT) de la República de Panamá.

**Tabla 1.** Regiones genómicas asociadas a número de nacidos totales (NT), número de nacidos vivos (NV), número de embriones implantados (EI) y tasa de ovulación (TO) en conejo.

<sup>1</sup> Cr	Caracteres ( <sup>2</sup> % Var)	Longitud ( <sup>3</sup> Mb)	Inicio/Final de posición en <sup>4</sup> bp
17	NT (14,92), NV (7,02), EI (11,68)	1,28	72.002.622-73.280.319
17	NT (1,40), NV (1,11)	1,98	70.012.491-71.994.214
11	EI (9,99)	4,00	36.000.805-39.997.550
9	TO (1,13)	0,97	47.006.518-47.977.424
9	TO (1,03)	0,96	42.008.720-42.970.607

<sup>1</sup>: cromosoma, <sup>2</sup>: porcentaje de varianza genómica; <sup>3</sup>: megabase; <sup>4</sup>: pares de bases.

## GENOMIC REGIONS ASSOCIATED WITH REPRODUCTIVE TRAITS IN RABBITS

**ABSTRACT:** The aim of this study was to identify genomic regions associated with reproduction traits: total number born (NT), number born alive (NV), number born dead (NM), ovulation rate (TO), implanted embryo (EI) and survivals: embryo (SE), fetal (SF) and prenatal (SP). GWAS analysis showed the same genomic regions associated with NT and NV. Genomic regions on chromosome 9 and 11 were found associated regions for TO and EI, respectively. The regions of chromosome 17 explained 16% of the genomic variance of NT, being themselves the most relevant genomic region. Their biological functions of genes on these regions are related to the cellular response to prostaglandin stimulus, biologic and reproductive processes. Main genes were BMP4, PTGDR, STYX, ANKH and CDKN3. Our findings provide a new starting point for further genomic studies, in order to unravel the genetic basis of litter size and its components in rabbits. These could also provide useful information for future research on other selection strategies in rabbit commercial populations.

**Keywords:** GWAS, divergent selection, rabbits, reproductive traits.