

# POTENCIALES REGULADORES DE LA EFICIENCIA ALIMENTARIA EN PORCINO IDENTIFICADOS MEDIANTE GENÉTICA DE SISTEMAS

Y. Ramayo-Caldas<sup>1</sup>, M. Ballester<sup>1</sup>, J.P. Sánchez<sup>1</sup>, R. González-Prendes<sup>2</sup>, M. Amills<sup>2</sup>, R. Quintanilla<sup>1</sup>

<sup>1</sup>Programa de Genética y Mejora Animal, IRTA, Torre Marimon, Caldes de Montbui, 08140, Spain. <sup>2</sup>Centre de Recerca en Agrigenòmica (CSIC-IRTA-UAB-UB), Universitat Autònoma de Barcelona, Bellaterra, 08193, Spain

## INTRODUCCIÓN

La eficiencia alimentaria (EA) es un objetivo de selección prioritario en porcino debido a su repercusión sobre la rentabilidad, la sostenibilidad económica y el impacto ambiental de la producción porcina. La EA está determinada tanto por factores genéticos como ambientales, habiéndose estimado heredabilidades entre 0.14 y 0.53 para diferentes indicadores de EA (Gilbert et al., 2007; Cai et al., 2008).

El objetivo de este estudio es identificar genes y rutas metabólicas asociados a eficiencia y otros caracteres de crecimiento y calidad de carne, así como los posibles reguladores de la EA en porcino, utilizando un enfoque de genética de sistemas.

## MATERIALES Y MÉTODOS

En este trabajo se analizaron conjuntamente once caracteres relacionados con EA, crecimiento y depósito de grasa. Los fenotipos medidos en 352 cerdos Duroc incluyeron: ingesta de alimento residual (RFI), índice de conversión (FCR), ganancia media diaria (ADG), ingesta diaria (DFI), espesor de grasa dorsal al sacrificio (BF), incremento diario de grasa dorsal (BFD), contenido de grasa intramuscular (IMF) o perfil de ácidos grasos (SFA, MUFA, PUFA). Inicialmente se llevó a cabo un análisis de asociación genómico (GWAS) para los once caracteres usando tras el control de calidad ~31K SNP (Porcine SNP60K BeadChip; Illumina, SanDiego, CA). Se realizó la anotación de los SNPs (genes anotados a  $\leq 10$  kb; *Sscrofa11.1*) y se seleccionaron aquellos SNPs asociados con RFI ( $p < 0.05$ ) o bien para al menos tres de los caracteres. A continuación se implementó el procedimiento de redes ponderadas de asociación (AWM - *Association Weight Matrix*) descrito por Fortes et al. (2010), construyendo la matriz que relaciona genotipos y fenotipos a partir de los efectos aditivos de los SNPs. Las interacciones entre los genes se predijeron utilizando el algoritmo de correlación parcial y teoría de la información (PCIT; Reverter y Chan, 2008). Para identificar supuestos reguladores, se exploraron combinaciones de factores de transcripción (FT) que abarcaban la mayor parte de la topología de la red con una redundancia mínima. Además, se realizó la identificación *in silico* del sitio de unión de FT (TFBS) usando *iRegulonv1.3 Cytoscape plugin* (Janky et al., 2014). La clasificación funcional y el análisis de rutas metabólicas de los genes incluidos en la red de co-asociación génica se llevó a cabo con el software Ingenuity Pathways Analysis (IPA, Ingenuity Systems, Redwood City, CA).

## RESULTADOS Y DISCUSIÓN

En este estudio se aplica un enfoque de genética de sistemas con el fin de identificar genes candidatos y rutas metabólicas asociadas con EA, crecimiento y depósito de grasa en porcino. La clusterización jerárquica (Figura 1) permite evaluar el nivel de proximidad entre los fenotipos de acuerdo a las estimas del efecto aditivo de los SNPs seleccionados. Los dos indicadores de EA (RFI y FCR) agruparon claramente. Lo mismo sucedió con el crecimiento y el consumo por un lado, y con los caracteres relacionados con el contenido de grasa intramuscular y su composición en FA por otro. Estos resultados concordaron con las relaciones fisiológicas entre los fenotipos y con las

correlaciones basadas en la AWM. De este modo, los efectos aditivos de los SNPs considerados en la AWM capturan la fuerte asociación entre los indicadores de EA (RFI y FCR;  $r = 0,89$ ), entre el crecimiento diario y la ingesta de alimento (ADG y DFI;  $r = 0,79$ ), o entre la cantidad y composición de la grasa muscular (IMF, SFA y MUFA  $r=0.51-0.75$ ; entre MUFA y PUFA  $r=-0.95$ ). Del análisis de correlación vale la pena destacar la moderada asociación entre el depósito de grasa muscular y la EA, con correlaciones positivas de IMF con RFI y FCR ( $r = 0.23$  y  $0.34$ , respectivamente). De esta forma, los efectos aditivos de estos SNPs determinan que los animales más eficientes en el uso del alimento tienden a tener menos IMF.

La red de co-asociación entre genes, construida a partir de la filas de la AWM, estuvo formada por 704 genes conectados por 35.819 interacciones. De los genes que conformaban la red, 49 fueron clasificados como FT y por tanto reguladores potenciales de la red de asociación. El estudio de FT que abarcasen la mayor parte de la topología de la red con redundancia mínima mostró que tres de estos factores, *LHX4*, *POU2AF1* y *TCF7L2*, estaban co-asociados con 397 de los 704 genes. Cabe destacar que los efectos de los SNPs incluidos en estos 397 genes explicaban conjuntamente el 51%, 48%, 46%, 35% de la varianza fenotípica para RFI, DFI, FCR y ADG, respectivamente. Asimismo, la identificación *in silico* de TFBS permitió comprobar que el 35% y el 33% de los genes potencialmente regulados por *LHX4* y *TCF7L2* estaban efectivamente conectados con sus reguladores en la red de co-asociación obtenida. El factor *TCF7L2* se ha asociado con el comportamiento y la eficiencia alimentaria en porcino (Fan et al., 2010; Do et al., 2014), mientras que *LHX4* participa en el control de la diferenciación y el desarrollo de la glándula pituitaria (Pfaeffle et al., 2008). Por su parte *POU2AF1* es esencial para la respuesta a los antígenos por parte de las células B.

La clasificación funcional de los 397 genes agrupados en la red asociada a *LHX4*, *POU2AF1* y *TCF7L2* mostró que la respuesta inmune es uno de los procesos biológicos más representados. Asimismo, el comportamiento social, el mantenimiento de los ciclos circadianos, receptores de glutamato y la regulación del crecimiento celular se identificaron como procesos significativamente sobrerrepresentados. En la línea de estos resultados, Gondret et al. (2017) identificaron la respuesta inmune como la vía biológica más representada en las diferencias de expresión entre cerdos de líneas con baja y elevada RFI. La función inmune también se señaló como un proceso biológico importante en relación a la EA en bovino (Zarek et al., 2017; Lindholm-Perry et al., 2016). Finalmente, estudios recientes de nuestro grupo, realizados en otra población de cerdos divergentes para caracteres de EA, mostraron expresión diferencial de genes relacionados con la inmunidad, el comportamiento social y el ritmo circadiano en el transcriptoma de hígado y duodeno (RNA-Seq) (Ramayo-Caldas et al., 2018; Piles et al., *en preparación*). En resumen, las predicciones de nuestra red de co-asociación están respaldadas por la bibliografía y sugieren un rol cooperativo de tres FT: *LHX4*, *POU2AF1* y *TCF7L2*, y de otros genes como por ejemplo *ARID1B*, *DNAJC6*, *MKL2*, *ESR1*, *ZNF207* y *NCOA1*. Todos estos genes y FT constituyen una red altamente interconectada, en la que quedan representados un gran número de procesos fisiológicos que parece formar parte de los mecanismos de control genético de la EA en porcino.

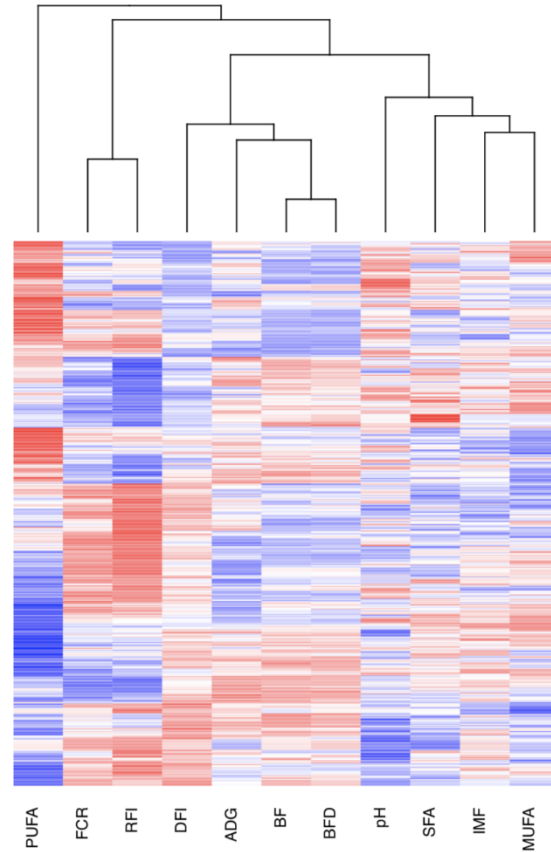
## REFERENCIAS

Fan et al. 2010. *Meat Sci.* 84:645-50 • Do et al. 2014. *Front Genet.* 5:307 • Pfaeffle et al. 2008. *J Clin Endocrinol Metab.* 93:1062-71 • Gilbert et al. 2007. *J. Anim. Sci.* 85:3182-8 • Cai et al. 2008. *J. Anim. Sci.* 86:287-8 • Fortes et al. 2010. *Proc. Natl. Acad. Sci. USA* 107:13642-7 • Reverter & Chan. 2008. *Bioinformatics* 24:2491-7 • Janky et al. 2014. *PLOS Comp. Biol.* 10:e1003731 • Zarek et al. 2017. *BMC Res Notes* 10:1 • Lindholm-Perry et al. 2016. *Anim. Genet.* 47:408-27 • Ramayo-Caldas et al. 2018. *Sci. Rep.* 8:558.

**Agradecimientos:** Este trabajo ha sido financiado con los proyectos AGL2007-66707-C02-01, AGL2010-22208-C02-01 y AGL2013-48742-C2-2-R concedidos por el Ministerio de Economía y Competitividad (MINECO). Y. Ramayo-Caldas ha sido financiado con un contrato Marie Skłodowska-Curie (P-Sphere) correspondiente a el programa H2020 de la Unión Europea. M. Ballester ha sido financiada con un contrato Ramón y Cajal (RYC-2013-12573) del MINECO.

**Figura 1.**

Clúster jerárquico de los once fenotipos analizados: ingesta de alimento residual (RFI), índice de conversión (FCR), ganancia media diaria (ADG), ingesta diaria (DFI), espesor de grasa dorsal al sacrificio (BF), incremento diario de grasa dorsal (BFD), así como pH, contenido de grasa intramuscular (IMF), ácidos grasos saturados (SFA), ácidos grasos monoinsaturados (MUFA) y ácidos grasos poliinsaturados (PUFA) en el músculo *gluteus medius*.



**A SYSTEMS GENETICS APPROACH REVEALS POTENTIAL REGULATORS OF FEED EFFICIENCY IN PIGS**

**Abstract:** We have used a systems genetics approach to identify candidate genes and pathways associated with ten traits related to feed efficiency, growth and fat deposition in 352 Duroc pigs. After a GWAS, the significant SNPs were used to build an Association Weight Matrix (AWM), relating additive effects with phenotypes. Hierarchical clustering and correlation analyses showed that AWM captures the strong association between FE traits. Gene by gene interactions were computed by PCIT; the resulting co-association network was formed by 704 genes. Several transcription factors (TF) were identified in the network, and the *in silico* identification of enriched TF binding sites was performed. Three TF were identified as the main putative regulators of the co-association network: *LHX4*, *POU2AF1* and *TCF7L2*. Our network predictions are supported by literature and suggest a cooperative role of these TF mediating a highly inter-connected regulatory cascade that seems pivotal for FE in pigs. The functional annotation of genes gathered in the gene co-expression network associated to *LHX4*, *POU2AF1* and *TCF7L2* (397 genes) showed that immune response is among the most overrepresented biological processes.

**Key words:** feed efficiency, systems genetics, association network, regulation, pig.