

# IDENTIFICACIÓN DE GENES CANDIDATOS Y REDES GÉNICAS IMPLICADOS EN LA REGULACIÓN DEL METABOLISMO LIPÍDICO EN HÍGADO Y EN LA DETERMINACIÓN DE CARACTERES DE CALIDAD DE LA CARNE EN CERDO

Ballester<sup>1</sup>, M., Ramayo-Caldas<sup>1</sup>, Y., Revilla<sup>2,3</sup>, M., Corominas<sup>2,3</sup>, J., Castelló<sup>2,3</sup>, A., Estellé<sup>4</sup>, J., Fernández<sup>5</sup>, A.I. y Folch<sup>2,3</sup>, J.M.

<sup>1</sup>IRTA, Genètica i Millora Animal, Torre Marimon, 08140 Caldes de Montbui, España.

<sup>2</sup>Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Campus UAB, Bellaterra, 08193, España. <sup>3</sup>Plant and Animal Genomics, Centre de Recerca en Agrigenòmica (Consorci CSIC-IRTA-UAB-UB), Edifici CRAG, Campus UAB, Bellaterra, 08193, España. <sup>4</sup>Génétique Animale et Biologie Intégrative, INRA, AgroParisTech, Université Paris-Saclay, 78350, Jouy-en-Josas, Francia. <sup>5</sup>Departamento de Mejora Genética Animal, INIA, Ctra. de la Coruña km. 7, 28040, Madrid, España  
maria.ballester@irta.cat

## INTRODUCCIÓN

La composición de ácidos grasos de la grasa subcutánea e intramuscular son caracteres de gran importancia para la industria alimentaria debido a su efecto directo sobre la calidad de la carne (Wood et al., 2008). Además, durante los últimos años, han adquirido una mayor relevancia debido a su impacto sobre la salud humana (Lorgeril y Salen, 2012). Se trata de caracteres complejos, determinados por factores ambientales, como la dieta, y múltiples factores genéticos. En los últimos años, el uso de nuevas aproximaciones como la detección de *quantitative trait loci* asociados con los niveles de expresión génica (eQTLs) ha permitido la identificación de redes de regulación génica y ha evidenciado como la variación de estas redes explica mejor los mecanismos funcionales implicados en los caracteres complejos (Schadt et al., 2008). Estas aproximaciones unidas a los análisis de redes de co-expresión génica se han convertido en herramientas importantes para profundizar en el estudio de la arquitectura genética de los caracteres complejos (Villa-Vialaneix et al., 2013).

Uno de los principales órganos implicados en el metabolismo lipídico es el hígado. En cerdos, el hígado está implicado en la síntesis y secreción de lipoproteínas, en la síntesis *de novo* del colesterol, en la oxidación de los ácidos grasos y junto con el tejido adiposo, pero en menor grado, en la síntesis *de novo* de los ácidos grasos. Un estudio en el retrocruce BC1\_LD (25% Ibérico, 75% Landrace) ha identificado genes diferencialmente expresados en el transcriptoma de hígado de cerdos con fenotipos extremos para la composición de ácidos grasos en grasa intramuscular (Ramayo-Caldas et al., 2012a). En esta misma población se han identificado mediante GWAS regiones genómicas asociadas con la composición intramuscular de ácidos grasos (Ramayo-Caldas et al., 2012b). El presente trabajo tiene como objetivo profundizar en el estudio de genes y rutas reguladoras en hígado que puedan estar implicados en la determinación de caracteres de calidad de la carne en el BC1\_LD. Para ello se han seleccionado 44 genes candidatos relacionados con el metabolismo lipídico y se ha analizado su expresión en hígado con el fin de realizar análisis de eGWAS y de redes de co-expresión.

## MATERIAL Y MÉTODOS

**Material animal y genotipado:** El material utilizado en este trabajo procede de un cruce entre 3 machos Ibéricos (Guadyrvas) con 31 hembras Landrace (Pérez-Enciso et al., 2000). En este trabajo se han analizado 111 animales pertenecientes al retrocruce BC1\_LD obtenido del cruce de cinco machos de la generación F1 con 26 hembras Landrace. Se recogieron muestras del hígado de estos animales que fueron congeladas con nitrógeno líquido y almacenadas a -80°C. Se genotiparon los animales con el *Porcine SNP60K BeadChip* (Illumina) y se eliminaron aquellos SNPs que no pasaron los controles de calidad estándar y con frecuencia alélica mínima (MAF) < 5%.

**Extracción de ARN y análisis de la expresión génica:** Se aisló el ARN total a partir de muestras de hígado de 111 animales utilizando el kit *RiboPure™* (Ambion). El ARN total se cuantificó en un espectrofotómetro *NanoDrop ND-1000* (NanoDrop) y fue convertido a ADNc utilizando el kit *High-Capacity cDNA Reverse Transcription* (Applied Biosystems). El estudio de expresión se realizó utilizando el chip *Dynamic Array 48.48* (Fluidigm) en un sistema *BioMark* (Fluidigm). Se analizaron los niveles de expresión de 48 genes, 44 genes diana y 4 genes de referencia. Los datos fueron normalizados utilizando los dos genes endógenos

más estables, *ACTB* y *TBP*. Los datos de expresión fueron analizados con el programa *DAG Expression* (Ballester et al., 2013). Se comprobó la normalidad de los datos mediante el test *Shapiro-Wilk* de R (<http://R-project.org>) y se normalizó la expresión de aquellos genes que no presentaban una distribución normal aplicando el log2 de los valores de RQ.

**Análisis de asociación de los genotipos con los valores de expresión:** El análisis se realizó mediante el programa *Qxpak 5.0* (Pérez-Enciso y Misztal, 2011) con el modelo:

$$y_{ijkl} = \text{Sexo}_i + \text{Lote}_j + \lambda_{lk}a_k + u_i + e_{ijkl},$$

Donde:  $y_{ijkl}$  es el valor fenotípico de cada individuo;  $\text{Sexo}_i$  y  $\text{Lote}_j$  son los efectos fijos (con 2 y 5 niveles, respectivamente);  $\lambda_{lk}$  corresponde al genotipo del SNP  $k$  para el individuo  $l$ ; siendo  $\lambda = -1(\text{aa}), 0(\text{Aa}), +1(\text{AA})$ ;  $a_k$  es el efecto aditivo de sustitución alélica del SNP  $k$ ;  $u_i$  el efecto infinitesimal con distribución  $N(0, A\sigma_u)$  dónde  $A$  es la matriz de parentesco y  $\sigma_u$  la varianza genética aditiva; y  $e_{ijkl}$  es el residuo.

La corrección de los p-valores se realizó con la librería de R *q-value* (Storey y Tibshirani, 2003) considerando como significativos aquellos valores con un q-valor  $<0,05$ .

**Anotación génica, clasificación funcional y análisis de redes génicas:** Los intervalos de los eQTLs se definieron como  $\pm 1$  Mb de los SNPs más significativos y se anotaron mediante la herramienta *BioMart* de Ensembl (<http://www.ensembl.org>) utilizando la última versión del genoma de referencia porcino *Sscrofa10.2*. Los eSNPs identificados fueron clasificados como *cis* cuando se encontraban a una distancia menor o igual de 1 Mb del gen analizado y *trans* los situados a una distancia superior a 1 Mb. Para el análisis de redes génicas se utilizó el programa *IPA* (Ingenuity Systems).

**Análisis funcional y de co-expresión:** El análisis de co-expresión se realizó utilizando el algoritmo *PCIT* (Reverter y Chan, 2008). La red génica resultante fue visualizada con el programa *Cytoscape* (Shannon et al., 2003).

## RESULTADOS Y DISCUSIÓN

Para el presente trabajo se seleccionaron un total de 44 genes candidatos relacionados con el metabolismo lipídico teniendo en cuenta trabajos anteriores del BC1\_LD y/o identificados por otros autores mediante búsqueda bibliográfica. Con los valores de expresión de los genes analizados en 111 animales pertenecientes al retrocruce BC1\_LD se realizó un estudio de asociación genómico utilizando un total de 40.502 SNPs que pasaron el control de calidad.

Se detectaron un total de 92 eSNPs significativos localizados en siete regiones cromosómicas en SSC2, SSC3, SSC4, SSC8 y SSC13, para un total de ocho genes: *CROT*, *CYP2U1*, *DGAT1*, *EGF*, *FABP1*, *FABP5*, *PLA2G12A* y *PPARA*. La mayoría de los eQTLs se identificaron en *trans* (5 de 7), mientras que la expresión de los genes *FABP1* y *FABP5* se asoció con dos eQTLs en *cis*, sugiriendo la presencia de una mutación en el mismo gen afectando directamente a su expresión. Además, se identificaron dos regiones *hotspot* en el SSC8 (8:86-88 Mb y 8:116-123 Mb) asociados en *trans* con la expresión de los genes *CYP2U1* y *PPARA*; y *DGAT1*, *EGF* y *PPARA*, respectivamente. El análisis de esta región a nivel cromosómico mostró que la expresión de los genes *DGAT1* y *HADH* también estaba asociada con el *hotspot* a 86-88 Mb y la expresión de los genes *AGPAT2*, *APOB*, *CYP2U1*, *ESRRA* y *HADH* con el *hotspot* de la región 116-123 Mb. Cabe resaltar que ambas regiones han sido previamente asociadas con la composición de ácidos grasos en grasa intramuscular en *Longissimus dorsi* en un análisis GWAS realizado en el mismo BC1\_LD (Ramayo-Caldas et al., 2012b). Con el fin de estudiar el patrón de co-expresión de los genes regulados por los mismos *hotspots*, se construyó una red génica de co-expresión con el algoritmo *PCIT*, obteniéndose correlaciones positivas ( $r=0.24-0.77$ ) entre los genes regulados por el mismo eQTL (Figura 1), lo que sugiere una regulación común.

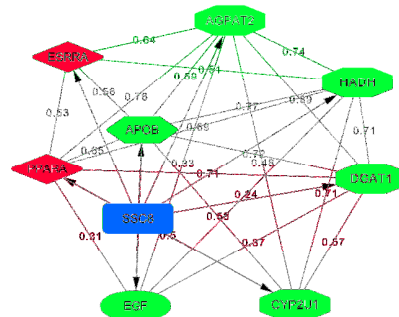
La anotación de los intervalos de eQTLs permitió la identificación de genes candidatos implicados en la regulación de los fenotipos de expresión analizados. Destacar el gen *PIK3CA*, un mediador clave en la regulación del metabolismo lipídico a través de la ruta Akt/PKB (Sopasakis et al., 2010), como gen candidato a explicar la variación de los niveles de expresión del gen *CROT*. El gen *NR3C1*, un receptor nuclear que puede activar la transcripción uniéndose directamente a los promotores de genes implicados en el metabolismo lipídico y de la glucosa (Ratman et al., 2013), como gen candidato a regular el gen *PLA2G12A*. Para la región *hotspot* a 86-88 Mb del SSC8 no se detectó ningún gen

relevante que pudiese regular la expresión de los genes asociados. No obstante, en la región *hotspot* a 116-123 Mb del SSC8, tres SNPs del gen *TBCK* estaban asociados en *trans* con la expresión de todos los genes exceptuando el *CYP2U1*, *EGF* y *HADH*. El estudio de Liu et al. (2013) muestra como la inhibición de este gen altera la ruta de señalización mTOR, la cual interviene en la regulación de la homeostasis lipídica controlando factores de transcripción clave del metabolismo lipídico como el *PPARA* y el *SREBF1*. Los resultados obtenidos incrementan nuestro conocimiento sobre los mecanismos reguladores implicados en el metabolismo lipídico y proporcionan nuevos conocimientos sobre los mecanismos moleculares implicados en los caracteres lipídicos.

### REFERENCIAS BIBLIOGRÁFICAS

• Ballester, M. 2013. PLoS One 8: e80385. • Liu, Y. 2013. PLoS One 8: e71349. • Lorgetil, M.D. 2012. BMC Med. 10: 50. • Mackay, T.F.C. 2009. Nat. Rev. Genet. 10: 565-77. • Pérez-Enciso, M. 2000. J. Anim. Sci. 78: 2525-31. • Pérez-Enciso, M. 2011. BMC Bioinformatics 12: 202. • Ramayo-Caldas, Y. 2012a. BMC Genomics 13: 547. • Ramayo-Caldas, Y. 2012b. J. Anim. Sci. 90: 2883-93. • Ratman, D. 2013. Mol. Cell. Endocrinol. 380: 41-54. • Reverter, A. 2008. Bioinformatics 24: 2491-7. • Schadt, E.E. 2008. PLoS Biol. 6: 1020-32. • Shannon, P. 2003. Genome Res. 13: 2498-504. • Sopasakis, V.R. 2010. Cell Metab. 11: 220-30. • Storey, J.D. 2003. Proc. Natl. Acad. Sci. U.S.A. 100: 9440-5. • Villa-Vialaneix, N. 2013. PLoS One 8: e60045. • Wood, J.D. 2008. Meat Sci. 78: 343-58.

**Agradecimientos:** Este trabajo ha sido financiado por los proyectos MINECO AGL2011-29821-C02 y AGL2014-56369-C2. El material animal fue generado en el contexto del proyecto INIA CPE03-010-C3 con la colaboración de investigadores del INIA, IRTA y UAB. M. Ballester está financiada con un contrato Ramón y Cajal (RYC-2013-12573) del MINECO. M. Revilla está financiado con una beca FI-DGR de la Generalitat de Catalunya (ECO/1639/2013).



**Figura 1.** Red de co-expresión de los genes asociados con los hotspots del SSC8.

### IDENTIFICATION OF DNA VARIANTS AND MOLECULAR PATHWAYS IMPLICATED IN THE FUNCTIONAL REGULATORY MECHANISMS OF MEAT QUALITY TRAITS IN PIGS

**ABSTRACT:** In the present study, liver co-expression networks and expression Genome Wide Association Study (eGWAS) were performed to identify DNA variants and molecular pathways implicated in the functional regulatory mechanisms of meat quality traits in pigs. With this purpose, the liver mRNA expression of 44 candidate genes related with lipid metabolism was analyzed in 111 Iberian x Landrace backcross animals. The eGWAS identified 92 eSNPs located in seven chromosomal regions and associated with eight genes: *CROT*, *CYP2U1*, *DGAT1*, *EGF*, *FABP1*, *FABP5*, *PLA2G12A* and *PPARA*. Furthermore, a hotspot on SSC8 associated with the gene expression of eight genes was identified and the *TBCK* gene was pointed out as candidate gene regulating it. Our results also suggested that the PI3K-Akt-mTOR pathway plays an important role in the control of the analyzed genes highlighting nuclear receptors as the *NR3C1* or *PPARA*. We provide a list of DNA variants, candidate genes and networks that increase our knowledge of the genetic architecture underlying fat composition traits.

**Keywords:** co-expression, liver, hotspot, eGWAS