

## DETECCIÓN DE *eQTLs* ASOCIADOS CON EL METABOLISMO LIPÍDICO EN EL MÚSCULO PORCINO

Puig-Oliveras, A.<sup>1,2</sup>, Revilla, M., Martínez, A.M., Folch, J.M. y Ballester, M.

<sup>1</sup>Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona. <sup>2</sup>Plant and Animal Genomics, Centre de Recerca en Agrigenòmica (CRAG), 08193 Bellaterra, Barcelona  
anna.puig@cragenomica.es

### INTRODUCCIÓN

La detección de *quantitative trait loci* asociados con los niveles de expresión génica (*eQTLs*) se ha propuesto como una buena estrategia para profundizar en el estudio de la arquitectura genética de los caracteres complejos (Schadt *et al.*, 2003). Podemos clasificar como *cis-eQTLs* aquellos que mapean cerca de la posición del gen cuya expresión ha sido analizada y *trans-eQTLs* aquellos identificados en otras regiones del genoma. A nivel biológico, la diferencia puede ser debida a que los niveles de transcripción estén afectados por el resultado de una mutación en el mismo gen (*cis-eQTLs*) o bien por otros factores que actúen lejos del gen (*trans-eQTLs*). Esta técnica permite además la identificación de zonas implicadas en la regulación de varios genes (*eQTL hotspots*).

La cantidad y composición de la grasa en el músculo están estrechamente relacionadas con la terneza y el sabor de la carne del cerdo, siendo estos dos caracteres de especial interés para la industria alimentaria (Wood *et al.*, 2008). Se trata de caracteres complejos, determinados por factores ambientales, como la dieta, y múltiples factores genéticos. La importancia de esta base genética resulta evidente en la comparación de las diferentes razas porcinas y en la identificación de *QTLs* relacionados con estos caracteres.

El cruce Ibérico × Landrace (IBMAP; Pérez-Enciso *et al.*, 2000) fue generado para el estudio de caracteres relacionados con el crecimiento, engrasamiento y composición de la grasa permitiendo la identificación de varios genes candidatos (Óvilo *et al.*, 2002; Estellé *et al.*, 2005; Mercadé *et al.*, 2005; Estellé *et al.*, 2006; Fernández *et al.* 2012; Ramayo-Caldas *et al.*, 2012a; Ramayo-Caldas *et al.*, 2012b; Corominas *et al.*, 2013; Muñoz *et al.*, 2013; Pena *et al.*, 2013; Puig-Oliveras *et al.*, 2014a; Puig-Oliveras *et al.*, 2014b; Ramayo-Caldas *et al.*, 2014; Revilla *et al.*, 2014). El presente trabajo tiene como objetivo profundizar en el estudio de genes y rutas reguladoras que juegan un papel clave en la determinación del contenido y composición de ácidos grasos en músculo mediante la detección de *eQTLs* asociados con genes del metabolismo lipídico en el músculo *Longissimus dorsi* de animales del cruce IBMAP.

### MATERIAL Y MÉTODOS

**Material animal y genotipado:** El material utilizado en este trabajo procede de un cruce entre 3 machos Ibéricos (Guadyerbos) con 31 hembras Landrace (Pérez-Enciso *et al.*, 2000). Cinco animales de la generación F1 fueron cruzados con 26 hembras Landrace obteniendo 144 animales (BC1\_LD). Se recogieron muestras del músculo *Longissimus dorsi* 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 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 músculo de 114 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, 45 genes diana y 3 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 eliminaron aquellos animales con valores de expresión atípicos (*outliers*) y se comprobó la normalidad de los datos mediante el test *Shapiro-Wilk* (Remark, 1995) de R (<http://R-project.org>). Para algunos genes los datos fueron normalizados aplicando el  $\log_2$  de los valores de RQ.

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

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

Donde:  $y_{ijklkm}$  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  $N(0, A\sigma_u)$  donde  $A$  es la matriz de parentesco y  $\sigma_u$  la varianza genética aditiva; y  $e_{ijklkm}$  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:** Los eQTLs identificados fueron clasificados como *cis* cuando se encontraban a una distancia de  $\pm 1$  Mb del inicio o final de transcripción del gen y *trans* los situados a una distancia superior a 1 Mb. También se identificaron las regiones *hotspots* de eQTLs. 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*.

## RESULTADOS Y DISCUSIÓN

Un total de 45 genes relevantes para el metabolismo lipídico fueron seleccionados teniendo en cuenta trabajos anteriores de nuestro grupo y/o identificados por otros autores mediante búsqueda bibliográfica. Con los valores de expresión de los 45 genes en 114 animales pertenecientes al retrocruce BC1\_LD (25% Ibérico y 75% Landrace) se realizó un estudio de asociación genómico utilizando un total de 40.586 SNPs que pasaron el control de calidad.

El eGWAS permitió la identificación de un total de 402 eSNPs localizados en 28 regiones cromosómicas en SSC1, SSC2, SSC3, SSC4, SSC5, SSC6, SSC8, SSC9, SSC10, SSC11, SSC13, SSC15 y SSC16, para un total de 11 genes: *ACSM5*, *CROT*, *FABP3*, *FOS*, *HIF1AN*, *PIK3R1*, *PLA2G12A*, *MGLL*, *IGF2*, *NCOA1* y *PPARA* (FDR $<0,05$ ). De los 402 eSNPs identificados, 53,75% se localizaron en regiones intergénicas, 32,5% en intrones, 6,25% en la región 5' flanqueante, 5% en la región 3' flanqueante, 0,75% en la región 3'UTR, 0,25% en la región 5'UTR, 1% en la región codificante de un gen determinando mutaciones sinónimas y 0,5% determinando mutaciones no sinónimas.

Los cromosomas SSC2, SSC6, SSC8 y SSC9 fueron los que más eQTLs presentaron (14 sobre 28). La mayoría de los eQTLs se identificaron en *trans* (25 sobre 28) observándose 4 genes asociados con más de un *trans*-eQTL. Tres de los eQTLs se encontraron en *cis* (p-valor=7,12 $\times 10^{-4}$ , p-valor=2,20 $\times 10^{-9}$ , p-valor $<1,00\times 10^{-25}$ ), sugiriendo la presencia de una mutación en el mismo gen afectando directamente a su expresión. Además, se identificaron dos *hotspots* de eQTLs en *trans*, uno alrededor de las 289 Mb en el cromosoma 1 regulando la expresión de *ACSM5* y *MGLL* y otro aproximadamente en la posición 117 Mb del cromosoma 9 afectando a la expresión de los genes *PLA2G12A* y *HIF1AN*.

La anotación de los intervalos de eQTLs permitió la identificación de genes candidatos que pueden estar asociados con los fenotipos de expresión analizados. En la región *hotspot* del cromosoma 9 se encuentra el gen *PIK3CG* que codifica para una fosfatidilinositol 3-quinasa que actúa en rutas del metabolismo lipídico (Kobayashi *et al.*, 2011). Por otra parte, el fenotipo de expresión del gen *MGLL* está *trans*-asociado con el SNP ASGA0020267 (p-valor=4,17 $\times 10^{-5}$ ) próximo al gen *CYP7A1* involucrado en la síntesis del colesterol y otros lípidos.

Siete de los 28 eQTLs detectados concuerdan con regiones GWAS para la composición de ácidos grasos en grasa intramuscular identificados en el mismo material animal (Ramayo-Caldas *et al.*, 2012a), siendo candidatos potenciales a determinar los caracteres estudiados.

Los resultados obtenidos representarán un avance importante en la identificación de genes y variantes genéticas implicadas en la determinación del contenido y composición intramuscular de ácidos grasos y, por tanto, en la calidad de la carne porcina.

## REFERENCIAS BIBLIOGRÁFICAS

Ballester, M., Cordón, R., Folch, J.M. 2013. PLoS One 8(11):e80385. • Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., *et al.*, 2013. BMC Genomics 14:842. • Estellé, J., Mercadé, A., Noguera J.L., *et al.*, 2005. J. Anim. Sci. 83(12):2723-8. • Estellé, J., Pérez-Enciso, M., Mercadé, A., *et al.*, 2006. Anim. Genet. 37(6):589-91. • Fernández, A.I., Pérez-Montarelo, D., Barragán, C., *et al.*, 2012. BMC genet. 13:41. • Kobayashi, N., Ueki, K., Okazaki, Y., *et al.*, 2001. PNAS 108(14):5753-5758. • Mercadé, A., Sánchez, A., Folch, J.M., *et al.*, 2005. J. Animal Breeding and Genetics 122, 161-164. • Óvilo, C., Oliver, A., Noguera, J.L., *et al.*, 2002. Genet. Sel. Evol. 34:465-479. • Muñoz, M., Rodríguez, M.C., Alves, E., 2013. BMC Genomics 14:845. • Pena, R.N., Noguera J.L., Casellas, J., *et al.*, 2013. Anim. Genet. 44(6):648-80. • Pérez-Enciso, M., Clop, A., Noguera, J. L., *et al.*, 2000 J. Animal Sci. 78: 2525-31. • Pérez-Enciso, M. y Misztal, I. 2011. Bioinformatics 12:202. • Puig-Oliveras, A., Ramayo-Caldas, Y., Corominas, J., *et al.*, 2014a. PLoS One 9(7):e103668. • Puig-Oliveras, A., Ballester, M., Corominas, J., *et al.*, 2014b. PLoS One 9(12):e114862. • Ramayo-Caldas, Y., Mercadé, A., Castelló, A. *et al.*, 2012a. J. Animal Sci. 90:1-11. • Ramayo-Caldas, Y., Mach, N., Esteve-Codina, A. *et al.*, 2012b. BMC Genomics 13:547. • Ramayo-Caldas, Y., Ballester, M., Fortes, M.R.S., *et al.*, 2014. BMC Genomics 15:232. • Revilla, M., Ramayo-Caldas, Y., Castelló, A., *et al.*, 2014. Genet. Sel. Evol. 46:28. • Schadt, E.E., Monks, S.A., Drake, T.A., *et al.*, 2003. Nature 422, 297-302. • Storey, J.D., y Tibshirani, R. 2003. PNAS 100(16):9440-9445 • Remark, A.S., 1995. Journal of the Royal Statistical Society 44(4):547-551. • Wood, J.D., Enser, M., Fisher, A.V., *et al.*, 2008. Meat Science 78:343-358.

**Agradecimientos:** Este trabajo ha sido financiado por el proyecto AGL2011-29821-C02 (Ministerio de Economía y Competitividad). A. Puig-Oliveras ha sido financiada con una beca de la Universidad Autónoma de Barcelona (PIF, 458-01-1/2011). M. Revilla ha sido financiado con una beca de Formació i Contractació de Personal Investigador Novell (FI-DGR) de la Generalitat de Catalunya (ECO/1639/2013). Agradecemos a J.L. Noguera (IRTA) su contribución en la obtención del material animal.

## GENOME WIDE IDENTIFICATION OF EXPRESSION QUANTITATIVE TRAIT LOCI FOR ADIPOSITY-RELATED GENES DETERMINING PIG FATNESS TRAITS

**ABSTRACT:** The aim of this work was to study the genetic basis of the expression of genes affecting the lipid metabolism in the swine muscle (*Longissimus dorsi*) in an Iberian × Landrace cross. The detection of quantitative trait loci associated with the expression level of genes (eQTLs) has been proposed as a good strategy to reduce the list of candidate genes affecting quantitative traits in segregating populations. Here, we analyzed the expression level of 45 genes in 114 animals. The eGWAS identified 402 eSNPs located in 28 chromosomal regions on SSC1, SSC2, SSC3, SSC4, SSC5, SSC6, SSC8, SSC9, SSC10, SSC11, SSC13, SSC15, and SSC16 and associated with 11 genes *ACSM5*, *CROT*, *FABP3*, *FOS*, *HIF1AN*, *PIK3R1*, *PLA2G12A*, *MGLL*, *IGF2*, *NCOA1*, and *PPARA*. Three out of 28 eQTLs were classified as cis-acting eQTLs whereas the remaining 25 eQTLs have *trans* regulatory effects on the gene expression traits. A total of 46,25% of the 402 eSNPs identified were located within a gene. Moreover, we identified two hotspots of eQTLs on SSC1 and SSC9. The obtained results will increase our knowledge in the functional mechanisms implicated in complex traits and will allow for the identification of genes and genetic variants involved in the determination of intramuscular fat content and fatty acid composition.

**Keywords:** eQTL, *Longissimus dorsi*, transcriptome, lipid metabolism