

IDENTIFICACIÓN DE eQTLs IMPLICADOS EN LA REGULACIÓN DEL METABOLISMO LIPÍDICO DE LA GRASA DORSAL EN PORCINO

Revilla^{1,2*}, M., Ballester, M., Puig-Oliveras, A., Castelló, A.,
Fernández, A.I. y Folch, J.M.

¹Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), Bellaterra, España. ²Plant and Animal Genomics, Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra, España.
*manuel.revilla@cragenomica.es

INTRODUCCIÓN

Los ácidos grasos (AG) son una fuente básica de energía y son esenciales para la regulación del metabolismo lipídico. Además éstos juegan un papel crucial en las características organolépticas de la carne (Wood *et al.*, 2008). La regulación del metabolismo lipídico en porcino está influenciada por factores ambientales, como la dieta, y múltiples factores genéticos. La grasa dorsal juega un papel fundamental en la regulación de dicho metabolismo, ya que es el mayor reservorio de energía en forma de lípidos. Por lo tanto, es necesario comprender los mecanismos genéticos que intervienen en la determinación de la composición de los AG en la grasa dorsal porcina, con el objetivo de optimizar la composición de éstos.

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). Estos eQTLs pueden actuar en *cis* (mapean cerca de la posición del gen) o *trans* (mapean en otras regiones del genoma) (Michaelson *et al.*, 2009). La abundancia en el número de transcritos de un gen puede ser alterada por polimorfismos presentes en elementos reguladores localizados en el mismo gen (*cis*-eQTLs) o en otras regiones del genoma (*trans*-eQTLs).

El objetivo del presente trabajo es profundizar en el estudio de los genes y las rutas reguladoras que desempeñan un papel importante en la determinación de la composición de AG en grasa dorsal mediante la detección de eQTLs asociados con genes del metabolismo lipídico en grasa dorsal.

MATERIAL Y MÉTODOS

Material animal y genotipado: El material animal utilizado procede del cruce IBBMAP, generado por el cruce inicial de 3 verracos Ibéricos Guadyerbas y 31 cerdas Landrace (Pérez-Enciso *et al.*, 2000). Posteriormente, 5 machos de la generación F1 se cruzaron con 26 hembras Landrace obteniendo un total de 144 animales (BC1_LD; 25% Ibérico y 75% Landrace). Muestras de grasa dorsal de estos animales fueron recogidas con nitrógeno líquido y almacenadas a -80°C.

Genotipado y control de calidad: Los animales pertenecientes al BC1_LD fueron genotipados con el *Porcine SNP60K BeadChip* (Illumina). Se eliminaron los SNPs con problemas de asignación de genotipos (*cluster analysis de GenomeStudio, Illumina*) y, utilizando el programa *Plink* (Purcell *et al.*, 2007), se eliminaron aquellos SNPs con una frecuencia del alelo minoritario (MAF) inferior al 5%. Además, dos SNPs localizados en la región promotora de los genes *ACSM5* (rs331702081) y *FADS2* (rs331050552) se genotiparon para los 115 animales BC1_LD mediante la metodología *Taqman OpenArray™* en un *QuantStudio™ 12K flex Real-Time PCR System* (ThermoFisher Scientific). Así mismo, un indel localizado en el intrón 1 (*FABP4:g.2634_2635insC*, Mercadé *et al.* 2006) y un SNP localizado en la región 3'UTR (GenBank:Y16039.1:6723, Ojeda *et al.* 2006) del gen *FABP4* se genotiparon para esta misma población, siguiendo el protocolo de pirosecuenciación descrito por Mercadé *et al.* (2006) y la metodología *High Resolution Melting (HRM, ThermoFisher Scientific)*, respectivamente.

Extracción de ARN y análisis de expresión génica: El ARN total se aisló a partir de muestras de grasa dorsal de 115 animales utilizando el kit *RiboPure™* (Ambion). El ARN total se cuantificó en un espectrofotómetro *NanoDrop ND-1000* (Thermo Scientific) y se convirtió a ADNc utilizando el kit *High-Capacity cDNA Reverse Transcription* (Applied Biosystems). Los niveles de expresión génica de 45 genes, 43 genes diana y 2 genes de referencia (*ACTB* y *TBP*), fueron medidos con el chip *Dynamic Array 48.48* (Fluidigm) en un sistema *Biomark* (Fluidigm) utilizando un controlador *Nanoflex™* (Fluidigm). Los datos de expresión fueron normalizados utilizando los dos genes endógenos.

Los datos de expresión se analizaron con el programa *DAG Expression* (Ballester *et al.*, 2013). Los animales con valores de expresión atípicos se eliminaron y se comprobó la normalidad de los datos mediante el paquete *Shapiro-Wilk* de R (<http://R-project.org>). Se normalizó la expresión de aquellos genes que no presentaban una distribución normal aplicando el \log_2 de los valores de RQ.

Modelo estadístico: El análisis de asociación entre los genotipos y los valores de expresión génica se realizó mediante el programa *Qxpack 5.0* (Pérez-Enciso y Misztal, 2011) bajo el siguiente modelo:

$$y_{ijk} = \text{Sexo}_i + \text{Lote}_j + \lambda_{lk}a_k + \mu_l + e_{ijk}$$

Dónde: y_{ijk} es el valor fenotípico de cada individuo; Sexo_i y Lote_j son los efectos fijos (con 2 y 5 niveles, respectivamente); λ_{lk} corresponde al genotipo del SNP k para el individuo l ; siendo $\lambda = -1$ (aa), 0 (Aa), $+1$ (AA); a_k es el efecto aditivo de sustitución alélica del SNP k ; μ_l el efecto infinitesimal que se distribuye como $N(0, A\sigma_u)$ dónde A es la matriz de parentesco y σ_u la varianza genética aditiva; y e_{ijk} es el residuo.

Los p-valores obtenidos fueron corregidos utilizando la librería de R *q-value* (Storey *et al.*, 2003) considerando significativos aquellos eSNPs que mostraron un q-valor $< 0,05$.

Anotación de los eQTLs: Los eSNPs asociados significativamente se anotaron usando la herramienta *Variant Effect Predictor* (VEP) de Ensembl (<http://www.ensembl.org>). El número de eSNPs que pertenecen al mismo eQTL se definió como aquellos eSNPs significativos con una distancia entre ellos ≤ 10 Mb. Los eQTLs se anotaron utilizando *Genes 84 Database* mediante la herramienta *BioMart* de Ensembl. Tanto para la anotación de los eSNPs como de los eQTLs se utilizó la última versión disponible del genoma de referencia porcino *Sscrofa10.2*. Los eQTLs identificados se definieron como *cis* cuando se encontraban a una distancia de ± 1 Mb del inicio o final de transcripción del gen; y *trans* los situados a una distancia superior a 1 Mb.

RESULTADOS Y DISCUSIÓN

Teniendo en consideración estudios previos de nuestro grupo (GWAS, RNA-Seq y redes génicas) y referencias bibliográficas, se escogieron 43 genes relevantes para el metabolismo lipídico en grasa dorsal. Se realizó un estudio de asociación genómico (eGWAS) utilizando 40.460 SNPs, los cuales pasaron el control de calidad, y los valores de expresión de los 43 genes en 115 animales BC1_LD.

El eGWAS identificó un total de 224 eSNPs localizados en 21 regiones cromosómicas en SSC2-SSC4, SSC6, SSC8-SSC10 y SSC13-SSC15, para un total de 5 genes: *ACSM5*, *ELOVL6* (FDR $<0,01$), *FABP4*, *FADS2* y *FATP4* (FDR $<0,05$). Del total de 224 eSNPs identificados, 204 fueron correctamente anotados con VEP, de los cuales el 51,47% (105 eSNPs) se localizaron en regiones intergénicas. El 48,53% restante (99 eSNPs) mapearon dentro de 55 genes: 65 en intrones, 12 en la región 5' flanqueante, 16 en la región 3' flanqueante, 3 en la región 3'UTR y 3 en las regiones codificantes de genes determinando mutaciones sinónimas.

Tres de los 21 eQTLs identificados se clasificaron como *cis*-eQTLs al corresponder a los genes *ACSM5*, *FABP4* y *FADS2*, mientras que los 18 eQTLs restantes presentaron efectos regulatorios en *trans* (*ACSM5*, *ELOVL6*, *FABP4*, *FADS2* y *FATP4*). Los *cis*-

eQTLs evidenciados sugieren la presencia de mutaciones en el mismo gen que afectan directamente a su expresión. Por esta razón se procedió a identificar y genotipar un total de 4 mutaciones en el retrocruce BC1_LD para los genes que presentaban *cis*-eQTLs: 2 SNPs localizados en la región promotora del *ACSM5* (SSC3) y *FADS2* (SSC2), y 1 Indel y 1 SNP localizados en el intrón 1 y en la región 3'UTR del *FABP4*, respectivamente.

Los eGWAS realizados incorporando las mutaciones genotipadas revelaron que la asociación más fuerte para la expresión génica del *ACSM5* fue para el SNP situado en la región promotora (p -valor= $1,11 \times 10^{-17}$). Para la expresión del gen *FABP4*, el SNP de la región 3'UTR fue el más significativo (p -valor= $3,07 \times 10^{-08}$) y el indel fue también uno de los polimorfismos más asociados (p -valor= $1,27 \times 10^{-06}$). El SNP genotipado para el gen *FADS2* no mostró asociación significativa.

En consecuencia, estos SNPs son unos candidatos potenciales a explicar la variación del ARNm de los genes *ACSM5* y *FABP4*, pudiendo tener un rol en la determinación de la composición de AG en el tejido adiposo porcino. Estos resultados proporcionan recursos importantes para descifrar los mecanismos de regulación implicados en el metabolismo lipídico en grasa dorsal.

REFERENCIAS BIBLIOGRÁFICAS

- Ballester, M., *et al.*, 2013. PLoS One 8: e80385.
- Mercadé, A., *et al.*, 2006. J. Anim. Sci. 84: 2907-13.
- Michaelson, J.J., 2009. Methods 48: 265-76.
- Pérez-Enciso, M., *et al.*, 2000. J. Animal Sci. 78: 2525-31.
- Ojeda, A., *et al.*, 2006. Genetics 174: 2119-27.
- Pérez-Enciso, M., *et al.*, 2011. BMC Bioinformatics 12: 202.
- Purcell, S., *et al.*, 2007. Am. J. Hum. Genet. 81: 559-75.
- Schadt, E.E., *et al.*, 2003. Nature 422: 297-302.
- Storey, J.D., *et al.*, 2003. PNAS 100: 9440-9445.
- Wood, J.D., *et al.*, 2008. Meat Sci. 78: 343-358.

Agradecimientos: Este trabajo ha sido financiado por el proyecto AGL2014-56369-C2 y AGL2011-29821-C02 (Ministerio de Economía y Competitividad). 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.

GENE EXPRESSION ANALYSIS IN BACKFAT AND IDENTIFICATION OF eQTL REGIONS FOR FATNESS AND FATTY ACID COMPOSITION CANDIDATE GENES IN PIGS

ABSTRACT: The aim of this work was to study the genetic basis of the backfat expression of 43 lipid-related genes associated with meat quality traits in pigs. We performed an expression genome-wide association study (eGWAS) with the backfat gene expression and the PorcineSNP60 BeadChip genotype information in 115 Iberian x Landrace backcross (BC1_LD) animals. The eGWAS identified 224 eSNPs located in 21 chromosomal regions on SSC2-SSC4, SSC6, SSC8-SSC10 and SSC13-SSC15, and associated with the *ACSM5*, *ELOVL6*, *FABP4*, *FADS2*, and *FATP4* genes. Three out of 21 eQTLs corresponding to *ACSM5*, *FABP4*, and *FADS2* were classified as *cis*-acting eQTLs, whereas the remaining 18 eQTLs have *trans*-regulatory effects. In addition, four polymorphisms were identified and genotyped in the BC1_LD animals for the genes with *cis-acting* eQTLs: two SNPs in the promoter region of *ACSM5* and *FADS2*, and one Indel in intron 1 and one SNP in the 3'UTR region of *FABP4*, respectively. Noteworthy, the strongest association signal for *ACSM5* gene expression was shown for the polymorphism genotyped of this gene. For *FABP4* gene expression, the 3'UTR SNP showed the lower p -value and the indel of intron 1 was also one of the most significantly associated polymorphisms, as has been previously described. Thus, these SNPs are strong candidate polymorphisms to explain the mRNA variation of the *ACSM5* and *FABP4* genes, and may have also a role in the determination of the fatty acid composition in porcine adipose tissue.

Keywords: gene-expression, expression QTL, lipid metabolism.