

IDENTIFICACIÓN DE GENES CANDIDATOS PARA LA REGULACIÓN DE CRECIMIENTO Y ENGRASAMIENTO MEDIANTE EL ANÁLISIS DEL TRANSCRIPTOMA DE HIPOTALAMO E HIGADO PORCINOS

Martínez-Montes, A.M.¹, Pérez-Montarelo, D., Rodríguez, M.C.¹, Barragán, C.¹, Núñez, Y.¹, Benitez, R.M.¹, Ibañez, N.², Folch, J.M.³, Silió, L.¹, Alves, E. y Fernández A.I.¹

¹Dpto. Mejora Genética Animal, INIA, Madrid,

²IRTA, Genètica i Millora Animal, Lleida

³Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB.

Dpto. de Ciència Animal i dels Aliments, Facultat de Veterinària, UAB

E-mail: martinez.angel@inia.es

INTRODUCCIÓN

La aplicación de técnicas de análisis masivo del genoma combinadas con técnicas de análisis más específicos, permite no sólo obtener datos “*omicos*” globales, sino también poder enfocar el análisis de genes concretos. Los resultados obtenidos con la técnica de RNA-Seq nos permiten no solo analizar diferencias de expresión, sino a su vez la identificación de variantes génicas (SNVs) presentes en genes candidato diferencialmente expresados (Salem et al., 2012; Cánovas et al., 2010). Mediante un correcto diseño experimental se puede utilizar esta técnica para identificar polimorfismos asociados con caracteres productivos de interés. Análisis realizados sobre un cruce experimental Ibérico x Landrace (IBMAP) han permitido analizar de forma precisa caracteres productivos, identificando varios regiones QTLs con efecto sobre crecimiento, deposición grasa y caracteres de calidad de la carne y grasa (Óvilo et al., 2002; Estellé et al., 2005; Fernández et al. 2012; Muñoz et al., 2013; Pena et al., 2013; Ramayo-Caldas et al., 2014).

El objetivo de este trabajo ha sido la identificación de genes candidatos que regulen procesos de desarrollo relacionados con caracteres productivos de interés. Para ello se han identificado polimorfismos candidatos a partir de datos de RNA-Seq en hígado e hipotálamo en individuos extremos para crecimiento y deposición grasa de un retrocruce del citado material experimental.

MATERIAL Y MÉTODOS

Animales y datos fenotípicos. 144 individuos procedentes de un retrocruce Ibérico x Landrace (F1 x Landrace) fueron utilizados para seleccionar dos grupos de individuos con fenotipos divergentes para crecimiento y deposición grasa (A y B). Para ello se llevó a cabo un análisis de componentes principales utilizando cuatro caracteres diferentes, cuyos valores medios fueron; ganancia media diaria 0,78 - 0,92 kg/d, espesor de tocino dorsal 11,6 - 16,2 mm y porcentajes de ácido linoleico (C18:2) en grasa dorsal 16,7 - 12,6 % e intramuscular 11,9 - 8,1 %.

Obtención de ARN, secuenciación. El kit *Ribopure™ of High Quality of Total RNA (Ambion)* fue utilizado para aislar el ARN de los individuos seleccionados. El proceso de secuenciación fue llevado a cabo mediante el servicio de secuenciación del CNAG (Barcelona) usando el equipo *Hi-Seq 2000 (Illumina)*, preparando pools de 3 muestras, con lecturas de 2x75 pb.

Mapeo y detección de SNVs. El software *CLC Genomics Workbench* fue utilizado para mapear las lecturas frente al genoma de referencia *Sscrofa10.2*, usando los parámetros por defecto. Una vez realizada la búsqueda de SNVs con una profundidad de lectura de 10x, se llevó a cabo una segunda búsqueda de SNPs con una cobertura de 3x para determinar el potencial genotipo de cada uno de los individuos.

Identificación de SNPs y genes candidatos. Una vez identificados potenciales SNPs, fueron filtrados en función de su relevancia y por la presencia de alelos alternativos en los individuos de los grupos A y B. La anotación funcional de los genes en los que se localizaban los SNPs se realizó utilizando la herramienta *FatiGO* de *Babelomics* que permitió seleccionar un panel de SNPs localizados en genes de interés. Una vez seleccionados los SNPs/genes candidatos fueron validados mediante secuenciación de Sanger.

Genotipado y asociación de genes candidatos. Tras la selección de los genes candidatos por función e interés biológico, se llevó a cabo el proceso de genotipado en 124 individuos del retrocruce mediante técnica de Sanger. El análisis de asociación se realizó tanto a nivel

de SNPs como de haplotipos, con los datos fenotípicos de interés; Peso de los 150 días (P150D), Grasa intramuscular (GRIM), Peso de Paletas, Jamones y Chuleteros (PAL, JAM, CHU). La identificación de los haplotipos se realizó utilizando el software PHASE. Para el análisis de asociación se utilizó el software Qxpack 5.0 (Pérez-Enciso y Misztal, 2011) usando el siguiente modelo:

$$Y = \mu + G + S + L + PC/ES + u + e$$

Donde Sexo y Lote (S y L) son los efectos fijos, Peso Canal y Edad de Sacrificio como covariables en función del carácter a analizar, G corresponde al genotipo del SNP, u es el efecto poligénico aleatorio del animal y e el residuo.

RESULTADOS

En el proceso de secuenciación se obtuvieron 839M y 877M de lecturas pareadas de 75 pb para hipotálamo e hígado. De un total 125k SNVs identificados en cada tejido aproximadamente 4.300 y 1.900 fueron seleccionados como altamente informativos en hipotálamo e hígado, de los cuales un 50% se encontraban localizados en genes anotados. En el análisis de anotación funcional se pudo ver que había un enriquecimiento funcional en procesos de interés asociados con los caracteres de deposición grasa y crecimiento, tales como regulación del desarrollo o metabolismo lipídico. En el proceso de validación mediante secuenciación clásica no solo se observaron los SNPs previamente descritos, sino que se detectaron nuevos SNPs no identificados por la técnica del RNA-Seq. Una vez identificados los bloques de SNPs que estaban cosegregando, un subconjunto de 140 SNPs informativos fue seleccionado para posteriores análisis, a partir de los cuales se eligieron los más relevantes para llevar a cabo la validación por secuenciación de 32 SNPs, 8 de ellos falsos positivos, estimándose así una tasa de error de 0,25, debido a fallos de alineamiento. A partir de estos SNPs se seleccionaron en dos de los genes más importantes, *SERPINE1* y *RETSAT*, los SNPs ENSSSCP00000026436: p.Ile150Val y ENSSSCP00000008794: p.Lys173Thr. Estos dos genes están asociados con composición y calidad de la carne y grasa (Weisz *et al.*, 2012, Moise *et al.*, 2010). El genotipado de estos polimorfismos mediante secuenciación Sanger permitieron identificar dos nuevos SNPs en el gen *SERPINE1*, ambos cosegregando con el SNP previamente descrito con un MAF=0,13 y mostrando únicamente dos genotipos, información insuficiente para el análisis de la asociación entre dichos polimorfismos y los caracteres de interés. Para el gen *RETSAT* se identificaron cuatro nuevos SNPs así como una delección en el gen *RETSAT*, con unos valores MAF entre 0,15 y 0,44. Se pudo observar que los SNPs 1 y 4 así como el 2 y 3 estaban cosegregando, por lo que los posteriores análisis de asociación se realizaron únicamente con los SNP1 y SNP2. Se pudieron identificar cuatro haplotipos diferentes para estos SNPs, pero ninguno de éstos mostró niveles de asociación significativos. Los resultados del análisis de asociación para cada SNP mostraron un efecto de 0,205 ($\pm 0,09$) del bloque compuesto por los SNPs uno y cuatro sobre el peso de los jamones, mostrando efecto contrario frente a la delección, -0,272 ($\pm 0,13$). Estos resultados, por su carácter preliminar, deben ser confirmados en una muestra más amplia de individuos del material IBMAP.

En este estudio se ha mostrado un ejemplo de la combinación del uso de una técnica de análisis masivo (RNA-Seq) junto con el estudio específico con técnicas más convencionales de genes y polimorfismos potencialmente relacionados con caracteres productivos de interés.

REFERENCIAS BIBLIOGRÁFICAS

- Salem, M., Vallejo, R.L., Leeds, T.D. et al. 2012. PlosOne 7, e36264.
- Cánovas, A., Rincon, G., Islas-Trejo, A. et al. 2010. Mamm. Genome 21,592-598.
- Óvilo, C., Clop, A., Noguera, J.L. et al. 2002. J. Anim. Sci. 80,2801-2808.
- Estellé, J., Mercadé, A., Noguera J.L., et al. 2005. J. Anim. Sci. 83(12):2723-8.
- Fernández, A.I., Pérez-Montarelo, D., Barragán, C. et al. 2012. BMC Genetics 13, 41.
- 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.
- Ramayo-Caldas, Y., Ballester, M., Fortes, M.R.S., et al. 2014. BMC Genomics 15:232.
- Pérez-Enciso, M. & Misztal, I. 2011. Bioinformatics 12:202.
- Moise,

A.R., Lobo, G.P., Erokwu, B. *et al.* 2010. FASEB J. 24(4):1261-70. • Weisz, F., Bartenschlager, H., Knoll, A. *et al.* 2012. Anim. Genet. 43(5):614-9

Agradecimientos: Este estudio se ha realizado bajo el marco del proyecto AGL2011-29821-C02. Martínez A.M. disfruta de beca FPI (BES-2012-056563).

Tabla 1. Resultados del análisis de asociación para cada uno de los cuatro bloques de SNPs (*RETSAT1*, *RETSAT2*, *RETSAT_DEL* y *RETSAT5*), mostrando el p-val, el efecto de la asociación y el error estimado.

	RETSAT1			RETSAT2			RETSAT_DEL			RETSAT5		
	P-Val	Efecto	Error	P-Val	Efecto	Error	P-Val	Efecto	Error	P-Val	Efecto	Error
W150D	0,092	1,112	1,239	0,527	-1,139	1,874	0,548	-1,230	1,929	0,217	-1,951	1,571
GRIM	0,968	0,001	0,097	0,205	0,174	0,140	0,193	-0,191	0,143	0,662	0,045	0,126
JAM	0,019	0,205	0,087	0,383	0,115	0,138	0,042	-0,272	0,132	0,180	-0,164	0,120
PAL	0,539	0,030	0,049	1,000	0,000	0,075	0,896	-0,012	0,073	1,000	0,000	0,066
CHU	0,103	0,135	0,082	0,695	-0,051	0,127	0,076	-0,220	0,123	0,220	-0,137	0,111

IDENTIFICATION OF CANDIDATE GENES FOR THE REGULATION OF GROWTH AND FATNESS IN PORCINE MUSCLE BY TRANSCRIPTOME ANALYSIS

Using RNA-Seq as a sequencing technique, not only with the purpose of differential expression analyses, but for extracting genotypic information from the obtained data. In our study we were able to identify SNPs from a RNA-Seq assay from hypothalamus and liver tissues that shown different genotypes between animals with divergent phenotypes for growth and fatness. A total of 125k of SNPs were identified in both tissues analyzed, being 6k SNPs qualified as highly informative SNPs. We focused the analysis on more specific data, using different filters to identify two relevant genes, *SERPINE1* and *RETSAT*, that contain polymorphisms associated with important production traits.

Keywords: RNA-Seq, SNP-Identification, *SERPINE1*, *RETSAT*.