Universitat de Lleida

Genes asociados con la deposición y composición de grasas en porcino: estudios de expresión génica, proteínas y genética funcional y estructural

Angela Cánovas Tienda

Departament de Ciencies Mèdiques Bàsiques – Facultat de Medicina

Departament Producció Animal – Escola Tècnica Superior d'Enginyers Agrònoms

Universitat de Lleida

Genes asociados con la deposición y composición de grasas en porcino: estudios de expresión génica, proteínas y genética funcional y estructural

Memoria presentada por

Angela Cánovas Tienda

Para obtar al **grado de Doctora**

Directora: Dra. Ramona Natacha Pena i Subirà

Turor: Dr. Martí Aldea; Dr. Francisco Ferrezuelo

Tesis Doctoral

Marzo 2011

La **Dra. Ramona Natacha Pena i Subirà**, profesora agregada del Departamento de Producción Animal de la Universidad de Lleida,

CERTIFICA:

Que **Angela Cánovas Tienda** ha realizado bajo su dirección el trabajo de investigación "*Genes asociados con la deposición y composición de grasas en porcino: estudios de expresión génica, proteínas y genética funcional y estructural*" para obtener el **grado de Doctora** por la Universidad de Lleida.

Que este trabajo se ha realizado en el Departament o de Genètica i Millora Animal del Institut de Recerca i Tecnologia Agroalimentària (IRTA).

Lleida, 1 de Febrer de 2011

Dra. Ramona Natacha Pena i Subirà

A l'Adrià;

sense el teu amor i comprensió,

aquesta Tesi no hagués estat possible.

T'estimo

A mi abuelito;

Don Joaquín Tienda Madero,

un ejemplo a seguir

Esta Tesis Doctoral se ha realizado en el Institut de Recerca i Tecnologia Agroalimentària (IRTA) de Lleida financiada con una beca predoctoral del Instituto Nacional de Investigación Agraria (INIA).

Los trabajos realizados en esta Tesis Doctoral se han llevado a cabo en el laboratori de Genètica i Millora Animal del centro IRTA-Lleida, a excepción de los trabajos expuestos en los puntos 3.4, 3.5 y 7.1 llevados a cabo durante las estancias predoctorales realizadas en Iowa State University (USA), Universidad de León, University of Bristol (UK) y University of California-Davis (USA).

Para la realización de los estudios presentados en esta Tesis Doctoral se ha recibido financiación del Ministerio de Ciencia e Innovación (MICIIN) con los proyectos AGL2002-04271-C03-02, CICYT GEN2003-20658-C05-05 y AGL2007-66707-C02-01 y del Instituto Nacional de Investigación Agraria (INIA) con la beca predoctotal.

Agradecimientos

Después de 4 años, me encuentro delante del ordenador; analizando, recordando y haciendo balance de la última etapa de mi vida.

En estos últimos 4 años dedicados básicamente al aprendizaje e inicio de mi carrera como investigadora en el ámbito de la genética y la producción animal, me vienen recuerdos de todo tipo... En resumen, ha sido una etapa llena de satisfacciones, conocimiento, búsqueda, alegrías, diversión, sorpresas, imprevistos, desilusiones y también momentos muy duros y desagradables que gracias al cariño, la amistad y el amor de las personas que me rodean se han podido sobrellevar mejor…haciéndome, sin duda, crecer tanto personal como profesionalmente.

Lo que parecía extraño y lejano 4 años atrás ahora es una realidad; llegó la hora de cerrar capítulo, no antes sin agradecer el esfuerzo, comprensión y animación de todas aquellas personas que me han ayudado y animado en algún momento durante el desarrollo de la Tesis Doctoral.

En primer lloc, m'agradaria agrair a la meva directora de Tesi, la Dra. Ramona Natacha Pena, el iniciar-me a la investigació de la genètica molecular, animant-me a realitzar el Doctorat. Gracies per transmetre'm la teva organització i perfecció a l'hora de treballar al laboratori i per ser tan exigent amb mi. Moltes gracies per formar-me com ho has fet i per esser un referent per mi en l'àmbit professional!!! Gracies per les confidències i els bons moments que hem viscut juntes. Moltes gracies per tot Romi!!!

Al Instituto Nacional de Investigación Agraria (INIA) por otorgarme la beca pre-doctoral para realizar el doctorado, además de concederme la financiación para realizar las estancia realizadas durante la Tesis Doctoral en la Universidad de León (España), University of Bristol (UK), Iowa State University (USA) y University of California-Davis (USA).

A l'Institut de Recerca i Tecnologia Agroalimentària (IRTA) de Lleida per permetre'm realitzar la Tesi Doctoral al departament de Genètica i Millora Animal, posant-me a l'abast totes les facilitats i comoditats en tot el que he pogut necessitar durant aquests anys. Gracies per facilitar-me tots els cursos de formació, màster i congressos nacionals i internacionals. Valoro molt l'esforç que aquesta institució ha fet en tota la formació rebuda durant aquests 4 anys.

A la Dra. Raquel Quintanilla, por ser más que una compañera de trabajo. Gracias por enseñarme a realizar todos análisis estadísticos y por ser la supervisora de todo ello,

además de aconsejarme, orientarme y explicarme otras dudas relacionadas con la genética en general. Por todo esto y otras muchas razones, siempre te he considerado una pieza elemental en el desarrollo de esta Tesis Doctoral. Además, agradecerte tu amistad; aconsejándome, animándome y ayudándome cuando más lo he necesitado. Muchas gracias por todo Raquelita!!!

A David Almuzara; compañero de fatigas en el laboratorio….pero sobretodo mi mejor amigo. Gracias por las largas conversaciones, por tus consejos y por exponer tu punto de vista...Gracias por tu buen humor, siempre dispuesto a sonreír. Davidorum gracias por las cenitas y "fiestas" compartidas, contigo la diversión está siempre asegurada!!! Muchas gracias hermanito (recuerda que a partir de ahora ya eres el mayor…jajaja).

A todos los integrantes del Área de Genètica i Millora Animal de l'IRTA-Lleida: Dr. Joan Tibau, Dr. Carles Rosell, Dr. Jules Hernández, Dra. Noelia Ibáñez, Dr. José Luis Noguera, Pedro López, Núria Alòs y Maite Arbonés, por su compañerismo y amistad. También mencionar a los antiguos compañeros que en su día también formaron parte de este departamento: Dr. Luis Varona, Dr. Quim Casellas y María Martínez. Gracias a todos por estar siempre dispuestos a responder cualquier duda informática, estadística o de producción!!! Gracias por los momentos tan divertidos que hemos pasado en escenarios tan diferentes como los premios Por d'Or, en los congresos y hasta en el matadero a altas horas de la madrugada!!!

A totes les companyes de Serveis Administratius per facilitar-me totes les tasques administratives i per fer tan bé la vostra feina!!! Especialment agrair a la Tonyi la seva amistat i bons consells. Moltes gracies per tot!!!!

Al Dr. Joan Estany, per facilitar-me el contacte de la Dra. Olena Dora per realitzar l'estada a University of Bristol, iniciant així, una col·laboració molt satisfactoria en tots els sentits. Ha sigut un plaer treballar amb tu Joan, i espero continuar-ho fent.
Moltes gracies per tot!!!

A la Dra. Margarita Marqués por presentarme el mundo de las células y darme la oportunidad de iniciarme en el ámbito del cultivo celular durante la estancia realizada en la Universidad de León (2007). A las compañeras de fatigas en el laboratorio Marta y Silvia (y Alejandro). Todavía recuerdo las risas a altas doras de la madrugada en el laboratorio mientras trabajábamos, vaya horas!!! Gracias por vuestra amistad y por hacerme sentir una más en León. PD: Echo de menos las pizzas de la Competencia y las sesiones de cine!!!

Thank you to Dra. Olena Doran for extends an invitation for me to spend several months internship in your laboratory at the University of Bristol (UK, 2008). During my internship I had the opportunity to work and learn about protein expression. Thank you for your help and for all. I was really well in Bristol with you. Thanks a lot Lena!!

Thank you to Dr. James M. Reecy for extends an invitation for me to spend several months internship in your laboratory at Iowa State University (USA, 2009). Thank you for offer me participation in research projects, participation in your laboratory meetings and the opportunity to meet and discuss with members of the Department of Animal Science. Thank you to Ye, Eric, Mary Su and JR for your help. This stage was my first experience in United States. I discovered a lovely country, people and culture. Thank you for all Jim!!!!.
Thank you to my "Spanish" friends Paula, Lidia and Javier por vuestra amistad en esas tierras tan lejanas, y por estar siempre dispuestos a unas cervecitas al estilo español!.

Muchas gracias al Dr. Juan F. Medrano por brindarme la oportunidad de realizar una estancia en su lab en University of California-Davis (USA, 2010). Gracias por hacerme sentir una integrante más de tu grupo de investigación y gracias por poner a mi disposición todo el equipamiento y tecnología disponible. Gracias por formarme en el ámbito del Next Generation Sequencing y el RNA-seq…ya sabes que me enamoré de esta tecnología, y que para mí ha habido un antes y un después tras los conocimientos adquiridos y las técnicas utilizadas en Davis!!! Gracias a Alma, Gonzalo, Rodrigo, Saumya y Rashida; gracias de todo corazón por tratarme como lo hicisteis y por vuestra amistad. Espero volver a veros pronto ;).
No puedo pensar en Davis sin asociarlo a los buenos momentos vividos con todos los amigos de allí, en especial con Gloria, Begoña y Rodrigo. Gracias por esas cervecitas anti estrés, y por esas fiestas tan americanas y autenticas, y como no...por las sesiones de cine. Muchas gracias chicos!!!

A toda mi familia (abuelos, tíos y primos) por preocuparos e interesaros por mi trabajo, intentando entender la "cosas raras" y "frikis" que según vosotros hago…jajaja. Gracias por enviarme e-mails durante mis estancias en el extranjero, no sabéis cuanto se valora recibir noticias cuando uno está tan lejos!!! Gracias por estar siempre ahí!!!

Gracias sobre todo a ti, abuelo; por ser como eres, por tu sabiduria, tu forma de ser, por querernos tanto, etc...Por todo esto y mucho más, te dedico esta Tesis Doctoral. Te admiro y te quiero muchísimo abuelo!!!

A la meva família política, per interessar-se pel món de la genètica i per animar-me quan ho he necessitat, en especial als meus cunyats Ricard i Jose, us estimo moltíssim!!

A todos mis amig@s!!!. Da igual la distancia o la frecuencia en que nos veamos, siempre estáis ahí. PD: Espero que me organicéis una buena fiesta para celebrar la Tesis, y que no faltéis ninguno, eh!!!jajaja

A mi hermana y Pepe, por estar siempre ahí. Tata, espero a partir de ahora tener más tiempo para hacer cosas juntas…jajaja!!! Álvaro, aunque solo tengas un añito, no te imaginas cuanto me has ayudado en este último tramo de la Tesis…Gracias por tu sonrisa, siempre contento y dispuesto a reír!!! Gracias por ayudarme a desconectar cuando lo he necesitado, simplemente riendo y jugando contigo!!! Eres el sobrinito más guapo y simpático del mundo!!!

A mis padres Antonio y Paqui. Gracias por todo. Gracias por ser unos padres modelo. Gracias por enseñarme los valores realmente importantes en la vida, y que desgraciadamente tanto escasean. Gracias por educarme como lo habéis hecho. Gracias por ser un apoyo incondicional y por estar siempre dispuestos a ayudarme. Como dice bien la canción cordobesa: "…ellos son mis dos pilares, mis dos amores…". Gracias por quererme tanto y por hacerme sentir tan orgullosa de los padres que tengo. Os quiero muchísimo.

Llegados al final de los agradecimientos, no puedo evitar emocionarme al pensar en todos vosotros, especialmente al pensar en ti Adrià.
Tu ets la persona més important i a la que més he d'agrair aquesta Tesi Doctoral. Realment aquesta Tesi es tan teva com meva i no em cansaré de dir-te que sense tu això no hagués estat possible. Moltes gracies pel teu amor incondicional, gracies pels teus ànims, gracies per la teva paciència sobretot en aquest últim any tan dur per mi. Gracies per animar-me sempre a continuar endavant i a superar-me dia a dia. Gracies per fer-me costat i animar-me a realitzar totes les estades a l'estranger, tot i el sacrifici que suposava el estar llargues temporades separats. Gracies per ser com ets i estimar-me com m'estimes. Ets la millor persona que conec i amb qui estic súper orgullosa de compartir la meva vida. Gracies per compartir amb mi una etapa important de la meva vida i sol desitjo continuar compartint amb tu totes les etapes futures de la meva vida. T'estimo una barbaritat.

Índice General

ÍNDICE

ÍNDICE FIGURAS

ÍNDICE TABLAS

SUMMARY

This PhD is part of a line of research devoted to studying the genetic basis of lipid metabolism and fat deposition in pigs with a view to producing healthy and high quality meat. The main objective is the identification of polymorphisms and regulatory mechanisms responsible for the genetic variability of these complex characters in pigs. In this sense, we have used several methods in the fields of structural (expression QTL (eQTL) maps; candidate genes studies) and functional (gene expression studies) genomics and also protein and cell studies in muscle and fat samples from pigs selected by meat quality parameters. In this context, using *microarrays* we analyzed the mRNA expression pattern in *gluteus medius* muscle samples obtained from a commercial Duroc pig population with divergent phenotypes for several parameters related to lipid deposition. As a result, we have obtained a list of genes differentially expressed between animals with divergent profiles related to lipid deposition. The ontological/functional study showed that these genes were particularly related to lipid metabolism, growth and muscle differentiation, immunity and glucose uptake in the insulin pathway. Moreover, analysis of eQTL has revealed the existence of genomic regions responsible for the variation of gene expression in porcine *gluteus medius* muscle. Some of these eQTL show positional concordance with several QTL related to meat quality and fat deposition previously identified in the same Duroc population. Additionally, we have performed a comprehensive study of four candidate genes (*ACACA*, *HMGCR*, *SCD* and *Δ6D*) directly involved in traits related to meat quality, playing an important role in fatty acids and cholesterol synthesis and desaturation. Combining the results of gene expression analysis, eQTL maps and candidate genes studied have resulted in a list of functional and positional candidate genes representing a valuable contribution to the understanding of the genetic regulation of skeletal muscle individual gene expression in swine species. This is a first step towards disentangling gene networks and molecular mechanisms involved in muscular lipid metabolism and meat quality traits in pigs.

RESUMEN

La presente Tesis Doctoral se enmarca en una línea de investigación dedicada al estudio de las bases genéticas del metabolismo de las grasas en relación a la producción de carne de porcino de alta calidad y saludable. El objetivo final es la identificación de polimorfismos y mecanismos de regulación responsables de la variabilidad genética de estos caracteres complejos en porcino. Para ello se han utilizado métodos de genómica estructural (mapas de QTL de expresión (eQTL); estudio de genes candidatos) y funcional (estudios de expresión génica) y también de análisis proteico y celular en muestras de músculo y grasa de cerdos seleccionados por sus características de calidad de carne. Así, mediante la técnica de *microarrays* se ha analizado el patrón de expresión de ARNm en muestras de músculo *gluteus medius* obtenidas a partir de cerdos de una población comercial Duroc con fenotipos divergentes para varios parámetros relacionados con la deposición de los lípidos. Como resultado, se han observado numerosos genes diferencialmente expresados entre los animales con perfiles divergentes de engorde. Un estudio ontológico/funcional mostró que estos genes estaban particularmente relacionados con el metabolismo lipídico, el crecimiento y la diferenciación muscular, la inmunidad, y la captación de glucosa en la ruta de la insulina. Por otra parte, un análisis de eQTL ha revelado la existencia de regiones genómicas responsables de la variación de la expresión génica en el músculo *gluteus medius* porcino, algunas de las cuales muestran una concordancia posicional con varios QTL para caracteres de calidad de carne y engorde detectados previamente en la misma población Duroc. Complementariamente, se ha realizado un estudio más exhaustivo de cuatro genes candidatos (*ACACA*, *HMGCR*, *SCD* y *Δ6D*) directamente implicados en caracteres relacionados con la calidad de la carne al ser los principales responsables de la síntesis y desaturación de ácidos grasos y colesterol. Combinando los resultados del análisis de expresión génica, mapas de eQTL y los genes candidatos estudiados se ha elaborado una lista de genes candidatos funcionales y posicionales que será la base de futuras investigaciones hacia el establecimiento de las redes génicas y los mecanismos moleculares implicados en el metabolismo de los lípidos musculares y los caracteres relacionados con la calidad de la carne en porcino.

RESUM

La present Tesi Doctoral s'emmarca dins d'una línia d'investigació dedicada a l'estudi de les bases genètiques del metabolisme dels greixos en relació a la producció de carn de porcí d'alta qualitat i saludable. L'objectiu final és la identificació de polimorfismes i mecanismes de regulació responsables de la variabilitat genètica d'aquests caràcters complexes en l'espècie porcina. En aquest context, s'han utilitzat mètodes de genòmica estructural (mapes de QTL d'expressió (eQTL); estudis de gens candidats) i funcional (estudis d'expressió gènica) a més a més d'anàlisis proteics i cel·lulars en mostres de múscul i greix de porcs seleccionats per les seves característiques de qualitat de carn. Així, mitjançant la tècnica de *microarrays* s'ha analitzat el patró d'expressió d'ARNm en mostres de múscul *gluteus medius* obtingudes a partir de porcs d'una població comercial Duroc amb fenotips divergents per diversos paràmetres relacionats amb la deposició dels lípids. Com a resultat, s'han observat nombrosos gens diferencialment expressats entre els animals amb perfils divergents d'engreixament. Un estudi ontològic/funcional va revelar que aquests gens estaven particularment relacionats amb el metabolisme lipídic, el creixement i la diferenciació muscular, la immunitat i la captació de glucosa en la ruta de la insulina. D'altra banda, les anàlisis d'eQTL han revelat l'existència de regions genòmiques responsables de la variació de l'expressió gènica en el múscul *gluteus medius* porcí; algunes de les quals mostren una concordança posicional amb varis QTL per caràcters de qualitat de cran i engreixament detectats prèviament a la mateixa població Duroc. Complementàriament, s'ha realitzat un estudi més exhaustiu de quatre gens candidats (*ACACA*, *HMGCR*, *SCD* i *Δ6D*), directament implicats en caràcters relacionats amb la qualitat de la carn al ser els principals responsables de la síntesis i dessaturació d'àcids grassos i colesterol. Combinant els resultats de l'anàlisi d'expressió gènica, mapes d'eQTL i els gens candidats estudiats s'ha elaborat una llista de gens candidats funcionals i posicionals que serà la base de futures investigacions cap a l'establiment de les xarxes gèniques i els mecanismes moleculars implicats en el metabolisme dels lípids musculars i els caràcters relacionats amb la qualitat de la carn en porcí.

I. Introducción

1.1.- Sector porcino

El sector porcino es el primer sector de la ganadería de nuestro país con una producción anual que supone más de 4.000 millones de euros al año, lo que le coloca como el segundo país productor de la Unión Europea detrás de Alemania con un 20% del censo porcino (MAPA, 2010). Según datos del Ministerio de Agricultura Pesca y Alimentación (MAPA), Cataluña, con un 23.3% de la producción porcina española, se mantiene como la principal área productora de España, seguida por Castilla y León (16.5%), Aragón (14.2%), Andalucía (10.4%), Murcia (8.7%), Castilla La Mancha (7.6%), Extremadura (7.2%) y la Comunidad Valenciana (5.2%). Sin embargo, cabe destacar que diversos factores como el incremento de precio del pienso industrial (la alimentación es el coste más importante de las explotaciones porcinas suponiendo el 66% de los costes totales), la aparición de enfermedades porcinas puntuales o la crisis económica actual generalizada a nivel mundial tienen repercusiones productivas muy importantes como está sucediendo actualmente en el sector porcino.

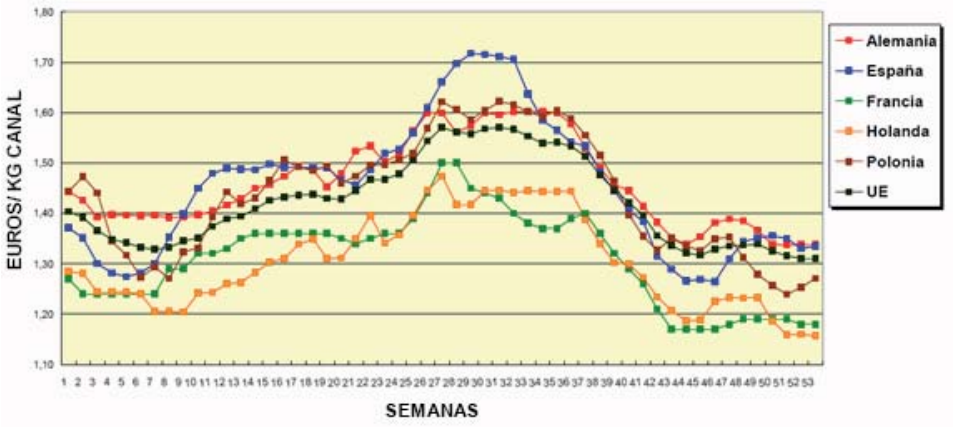


Figura 1: Comparación de precios de carne de porcino en el mercado interior comunitario durante el año 2009 (DAR; www.gencat.cat/dar/). Fuentes: Comisión de la Unión Europea, S.G. Estadística. Elaboración: S.G. Productos Ganaderos.

Según el Departament d'Agricultura, Alimentació i Acció Rural (DAR) y el Grupo de Gestión Porcina de la Universidad de Lleida, durante el segundo semestre del año 2009 el coste global para producir un cerdo engordado a 105 kg ha supuesto 109.83 €, mientras que el precio de venda medio del cerdo en matadero ha sido 110.25 €. El resultado es un margen ligeramente positivo 0.42 €/cerdo engordado. A pesar de esto, España es uno de los países con mejor precio de carne de porcino en el mercado comunitario, puntualmente superado por Alemania o Francia en las primeras y últimas semanas del año 2009 (Figura 1).

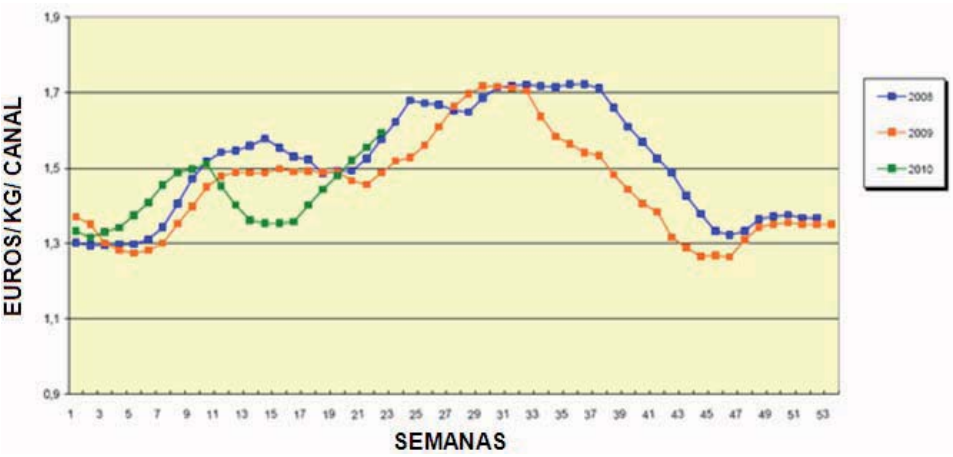


Figura 2: Comparación de precios de carne de porcino en el mercado español durante el año 2008, 2009 y primer semestre del 2010 (DAR; www.gencat.cat/dar/). Fuentes: S.G. Estadística. Elaboración: S.G. Productos Ganaderos.

Según la Comisión Europea de Agricultura y Desarrollo Rural, la producción de carne de cerdo en la Unión Europea continuará aumentando a corto plazo, aunque a un ritmo inferior al que lo hizo en los 90 (*European Commission; Directorate-General for Agriculture and Rural Development*, 2007). De esta manera, se puede observar un claro incremento tanto del precio de carne porcino como de lechones durante el año 2010

(Figura 2 y 3), observando una mejoría de los precios en el mercado español respecto al año 2009, siendo más notable en el caso del precio de lechones (Figura 3). Estos datos reflejan el liderazgo del sector porcino español a nivel europeo y refleja la gran importancia económica de este sector en nuestro país.

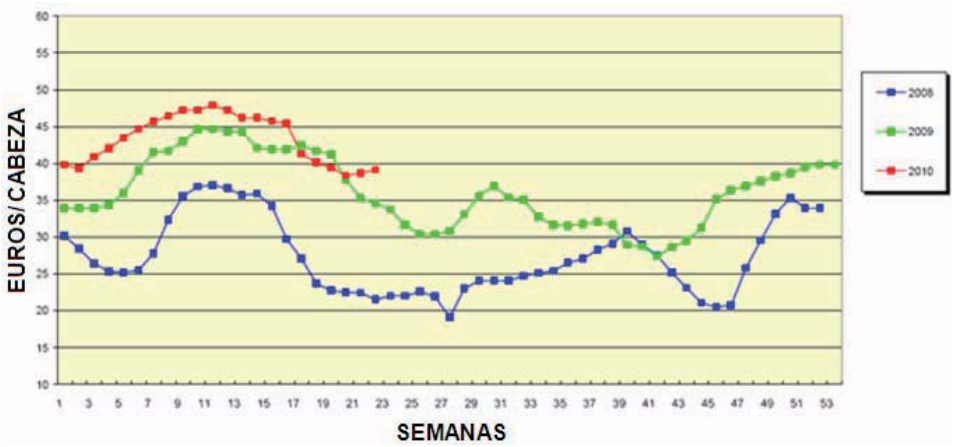


Figura 3: Comparación de precios de lechones en el mercado español durante el año 2008, 2009 y primer semestre del 2010 (DAR; www.gencat.cat/dar/). Fuentes: S.G. Estadística. Elaboración: S.G. Productos Ganaderos.

1.2.- Estructura del sector porcino y programas de mejora

En las últimas décadas, se han incorporado planes de mejora genética en la producción porcina con el objetivo de aumentar la producción y mejorar la calidad final del producto. Para ello es necesario definir estrategias de mejora genética, teniendo en cuenta la selección y los cruzamientos.

La producción porcina comprende la siguiente estructura: 1) la producción de reproductores (núcleos de selección) donde se genera el progreso genético, 2) la producción de lechones destetados (granjas de multiplicación) donde se disemina el

progreso genético y 3) el engorde de animales (granjas de producción y engorde) donde se utiliza el progreso genético (Tibau *et al*., 2004).

Para obtener buenos resultados es importante disponer de reproductores de alta calidad genética. En un esquema tradicional de selección en mejora genética porcina, el macho se selecciona teniendo en cuenta los caracteres de crecimiento (aumento del peso a edad de sacrificio y porcentaje de carne magra), mientras que la hembra se selecciona en base a los caracteres reproductivos (nacidos vivos y tamaño camada). En los últimos años, se han ido incorporando los esquemas de cruzamiento de diversas razas con aptitudes complementarias, permitiendo aprovechar la complementariedad entre las líneas y el vigor híbrido o heterosis (Varona, 2008).

Estos esquemas de selección han proporcionado una notable mejora en diferentes caracteres productivos, como son los caracteres relacionados con la reproducción y la prolificidad y los caracteres relacionados con la calidad de la carne. A pesar de la importancia económica que tienen los caracteres relacionados con la reproducción y la prolificidad, actualmente, destacan por su importancia los caracteres relacionados con la calidad de la carne, tanto en el aspecto tecnológico y sensorial de la misma como su relación con la salud humana, intentando producir carne de cerdo con unos niveles más bajos de colesterol y triglicéridos sin alterar la calidad del producto (Ventanas *et al*., 2007).

1.3.- Calidad de la carne

1.3.1. El concepto de calidad de carne

La definición de calidad de carne engloba diferentes propiedades y atributos de gran importancia para el sector productivo, la industria y el consumidor (Davoli y Braglia,

2007), definiendo la calidad de la carne como la suma de todos los aspectos y propiedades sensoriales, nutricionales y tecnológicas que influencian el valor del producto para el consumidor y la industria cárnica (Hofman, 1994 y López-Bote *et al*., 1998).

Tabla 1: Atributos relacionados con la calidad de la carne porcina según adaptado de Coma y Piquer, 1999.

Categoría	Atributos
Calidad sensorial	Color
	Terneza
	Jugosidad
	Sabor
	Olor
	Cantidad de grasa visible
	Veteado
Calidad nutritiva	Cantidad de grasa
	Composición de ácidos grasos
	Valor proteico
Calidad tecnológica	pH
	Capacidad de retención de agua
	Consistencia de la grasa
	Separación de tejidos
	Estabilidad oxidativa

Tal y como muestra la Tabla 1, son varios los atributos relacionados con la calidad de la carne, destacando entre ellos la jugosidad, la terneza, la textura, la capacidad de retención de agua y el sabor por su especial interés sobre la aceptación del consumidor.

Estos atributos dependen de la composición del músculo, es decir, del tipo de fibras musculares que está compuesto, así como del porcentaje y ratio músculo/grasa que éste posee en el momento del sacrificio y de los cambios que se derivan del proceso de maduración de músculo a carne. A su vez, la composición del músculo viene determinada por factores dietéticos, de manejo, bienestar animal y genéticos siendo la grasa intramuscular (GIM) y la composición ácidos grasos (AG) dos de los parámetros

con más efecto sobre la calidad sensorial, nutricional y tecnológica de la carne, afectando la textura, la jugosidad, la capacidad de retención de agua y el sabor de la carne (Candek-Potokar *et al*., 2002 y Olsson y Pickova, 2005). Sin embargo, la industria del porcino lleva muchos años mejorando el contenido de magro de la canal mediante selección contra el espesor de grasa dorsal (GD) (Toro y Silió, 1992), provocando una reducción de la GIM debido a la correlación genética positiva que existe entre GD y GIM (Solanes *et al*., 2009), hasta valores que mayoritariamente se sitúan por debajo de los recomendados en la producción de curados de calidad. En la última década, el mercado se ha desplazado hacia un incremento de la demanda de carne de cerdo de alta calidad debido a un interés creciente del consumidor en los aspectos relacionados con la calidad sensorial, nutricional y la salud humana. Por tanto, un objetivo del sector es, disponer de metodologías de selección que permitan manipular GD y GIM de forma independiente, así como el contenido y composición de AG muscular.

El porcentaje de GIM se asocia de forma favorable con la textura de la carne, la terneza, el sabor y la jugosidad (Huff-Lonergan *et al*., 2002). La proporción de ácidos grasos monoinsaturados (AGMI) (oleico principalmente) es un factor clave para determinar la consistencia de la grasa y el gusto de la carne, mientras que los ácidos grasos poliinsaturados (AGPI) tienden a sufrir procesos oxidativos, produciendo un olor rancio y un sabor indeseable en la carne, disminuyendo la aceptación de los consumidores.

En la especie porcina los ácidos grasos saturados (AGS) más comunes y con mayor efecto sobre la calidad tecnológica de la carne son el ácido palmítico (C16:0), el ácido esteárico (C18:0) y el ácido mirístico (C14:0) estando relacionados con la firmeza, el gusto y el aroma de la carne, mientras que el ácido oleico (C18:1), ácido palmitoleico (C16:1) y el ácido vaccénico (C18:1) son los AGMI más importantes influyendo en la correcta maduración y en el gusto y aroma de los productos curados. En el caso de los

AGPI, el ácido linoleico (C18:2), es el ácido con mayor repercusión sobre la calidad de la carne jugando un papel muy importante en la estabilidad oxidativa y el gusto de la carne (Wood *et al*., 2008).

Además del efecto que tienen los AG en el tema de la calidad de la carne, también juegan un papel muy importante en la salud humana ya que el aumento de las cantidades de AGMI y AGPI en la dieta disminuye la susceptibilidad a padecer enfermedades cardiovasculares estando asociados con una disminución de los niveles de lipoproteínas de baja densidad (LDL) y colesterol plasmático sin reducir los niveles de lipoproteínas de alta densidad (HDL) (Ventanas *et al*., 2007), mientras que los AGS tienen el efecto contrario (Ordovas, 2006). Por otro lado, la carne de cerdo es rica en grasa y particularmente en colesterol y constituye una de las fuentes principales de lípidos de las dietas occidentales siendo de especial interés ampliar el conocimiento de la base genética del metabolismo lipídico en esta especie permitiendo en un futuro seleccionar los animales para mejorar su porcentaje de colesterol y triglicéridos, haciéndolos así más saludables para el consumo humano. De hecho, la medición del contenido de GIM y de su composición es metodológicamente larga y costosa, pero, por otra parte, necesaria para seleccionar con éxito GD y GIM de forma independiente. La identificación de genes y/o biomarcadores asociados con los procesos de formación de la grasa puede ser de utilidad para el diseño de criterios de selección que permitan seleccionar contra GD a GIM constante o, alternativamente, a favor de GIM sin aumentar GD. Como resultado, se han propuesto varios genes con un papel fisiológico en el transporte de lípidos y en el metabolismo lipídico como genes candidatos para la deposición de grasa en cerdos (Davoli y Braglia, 2007) integrándose, algunos de ellos como los genes *RYR1* (*ryanodine receptor*) y *IGF2* (*insulin-like growth factor 2*), en esquemas de selección de la industria porcina (van der Steen *et al*., 2005).

1.3.2.- Relación entre depósitos de grasa y calidad de carne

Los depósitos lipídicos de los animales están formados principalmente por AG, que pueden clasificarse AGS, AGMI y AGPI en función de la presencia y número de dobles enlaces.

En porcino, los AG constituyentes de los diferentes depósitos de grasa de la canal pueden ser de procedencia exógena, es decir aquella que procede de los alimentos ingeridos, o de la biosíntesis endógena. En algunas zonas anatómicas predomina la grasa de origen exógeno (p.e área del cuello) mientras que en otras (p.e. tocino dorsal) predomina la grasa de origen endógeno. En general, los AGS proceden de la biosíntesis endógena, los AGPI son de origen exógeno y los AGMI pueden proceder de cualquiera de las dos vías (Morales, 2002).

En este contexto, los principales depósitos de grasa en esta especie son: grasa subcutánea, grasa intramuscular, grasa intermuscular, grasa abdominal y grasa perirenal. La grasa subcutánea o de cubertura se sitúa bajo la piel y representa el mayor porcentaje de grasa en el organismo depositándose formando capas separadas por tejido conjuntivo. La GIM se localiza entre y en el interior de las fibras musculares produciéndose su máxima deposición en las etapas más avanzadas del desarrollo corporal del animal (Le Dividich *et al.* 1991) mientras que la grasa intermuscular se localiza entre los músculos ocupando los espacios vacios de los puntos de unión entre músculos y huesos. Por último, mencionar que la grasa abdominal y perirenal va aumentando su porcentaje con la edad del animal.

En la composición de la grasa del cerdo se observa un gradiente de insaturación desde el centro de la canal hacia el exterior en el cual los AGS abundan en el tejido graso perirenal o GIM mientras que en la grasa subcutánea predominan los AGMI y AGPI

(Suzuki *et al*., 2006) (Tabla 2). Normalmente, el porcentaje de grasa corporal aumenta proporcionalmente al peso corporal y, asimismo, es mayor en hembras (Bucharles *et al.* 1985). De la misma manera, la castración supone cambios en el metabolismo del animal resultando en una mayor producción de grasa total, GIM y grasa intermuscular (Bonneau y Lebret, 2010). En cuanto a la edad, el contenido graso de los lechones al nacimiento es bajo (1-2% del peso vivo) y se incrementa a medida que aumentan de peso (Le Dividich *et al.* 1991).

Tabla 2: Composición de la grasa de cerdo según adaptado de Suzuki *et al*., 2006.

%	Grasa subcutánea exterior	Grasa subcutánea interior	Grasa intermuscular	Grasa intramuscular
C14:0	1.58	1.55	1.72	1.53
C16:0	25.05	26.82	27.01	26.46
C16:1	2.78	2.07	2.81	4.45
C18:0	13.81	16.79	15.35	13.43
C18:1	45.87	42.88	43.4	48.38
C18:2	10.86	9.89	9.71	5.62
Ratio insaturación AG	**59.52**	**54.86**	**55.94**	**58.44**

1.3.3.- La influencia de las razas porcinas en la calidad de carne

En la figura 4 se muestran las razas de cerdo más utilizadas en los esquemas de selección, dentro de las cuales se encuentran:

1) Large White; caracterizada por su buena aptitud maternal (carácter tranquilo, cuidado de las crías, capacidad lechera, etc.), prolificidad y fertilidad, imprescindible en cualquier cruzamiento de línea materna. Se utiliza en los programas de hibridación dando como resultado estirpes de mayor porcentaje de carnes magras en la canal. Su empleo, mayoritariamente, es en cruces como línea materna, constituyendo la principal

base genética empleada en las explotaciones españolas. Además, esta raza presenta buen rendimiento en cebo y buena calidad de carne.

2) Landrace; caracterizada por su prolificidad, fertilidad y gran aptitud materna. Presenta buena ganancia media diaria en peso y conversión alimentaría, con bajo nivel de engrasamiento, considerándose por ello una raza de tipo magro. Es una raza que se emplea en la industria cárnica por su buen rendimiento a la canal, la producción de jamones bien conformados y la calidad de su carne.

3) Pietrain; se caracteriza por ser la raza con mayor porcentaje de músculo, utilizándose en la mayoría de los cruzamientos de líneas paternas. Sin embargo, los animales de esta raza presentan malos parámetros de crecimiento, una baja prolificidad y, frecuentemente, carnes PSE (carnes pálidas, blandas y exudativas). Su producción, por tanto, está orientada hacia la obtención de productos frescos. Dando canales con unos rendimientos entre 72 - 75 %.

4) Duroc; posee unas buenas características relacionadas con la rusticidad y la buena adaptación a los climas cálidos. Destacan, a nivel productivo, por proporcionar calidad a la carne, incrementando la grasa infiltrada en los productos obtenidos en animales cruzados. A nivel reproductivo destaca su elevada prolificidad, utilizándose en los cruzamientos como línea paterna y línea materna.

5) Ibérico; destacado por su excelente calidad de carne caracterizada por poseer un mayor contenido en grasa intramuscular comparativamente con las razas de cerdo blanco al ser animales adipogénicos con unas características genéticas que les confiere una tendencia al almacenamiento de grandes depósitos de lípidos, los cuales, mediante un mecanismo biológico, se infiltran en las masas musculares, proporcionando a su

carne una incomparable untuosidad, textura y aroma. Presentan índices de conversión mayores que los de las razas blancas.

Figura 4: Razas de cerdo más utilizadas en los esquemas de selección.

Dentro de la gran variedad de razas porcinas existentes se pueden agrupar en cerdos magros y credos grasos, mostrando las siguientes características generales (Germán *et al*., 2005; Tabla 3):

Tabla 3: Características generales del las razas magras y grasas (Germán *et al*., 2005).

Característica	Raza grasa	Raza magra
Forma del cuello	Mediana y redondeada	Grande y larga
Cabeza	Pequeña	Alargada
Miembros	Cortos	Largos
Tronco	Corto, cilíndrico	Largo
Costillares	Muy arqueados	Arqueados
Línea dorsal	Recta	Arqueada
Dorso	Ancho y corto	Ancho y largo
Huesos	Finos	Menos finos

Las razas grasas juegan un papel muy importante dentro del contexto de calidad de la carne ya que el tejido graso influye directamente en la consistencia del tejido muscular,

proporcionando los componentes del aroma y sabor, previniendo contra el secado excesivo durante el procesado o cocinado y mejorando la terneza (Wood, 1989).

En los experimentos realizados en esta tesis doctoral se ha utilizado animales procedentes de la raza Duroc al presentar un alto contenido de grasa intramuscular por lo que es utilizada en la producción de jamón curado de calidad (Oliver *et al*., 1994).

Similares a otras razas de porcino, en los parámetros reproductivos, la raza Duroc, se puede equiparar a las razas Large White y Landrace, aunque es un poco inferior (Tibau *et al*., 1997). La raza Duroc se caracteriza por tener, aproximadamente, 10-10.5 lechones vivos/parto y 8-10 lechones destetados/parto. Además, la hembra de la raza Duroc, tiene una excelente aptitud reproductiva, caracterizándose por tener una muy buena aptitud maternal, por su alta prolificidad viable, por su capacidad lechera y por la calidad del lechón destetado.

La raza Duroc se emplea habitualmente como línea paterna, tanto en cruzamientos a dos como a tres vías, teniendo una ganancia media diaria (entre 20 y 95 kg de peso vivo) de 695 g/día, un índice de conversión (20-90 kg) de 3.1 kg/kg, un rendimiento de la canal a los 90 kg sin cabeza del 74%, una longitud de la canal de 93.5 cm, un porcentaje de piezas nobles del 61%, y un 52% estimado de magro en la canal (www.infocarne.com/cerdo/duroc).

Esta raza demuestra una gran capacidad de adaptación a los diferentes tipos de producción porcina y, tal y como ya se ha mencionado, su descendencia muestra una excelente calidad sensorial de la carne (Selecció Batallé). Además, el alto contenido de grasa intramuscular utilizada en la producción de jamón curado de calidad, también lo hacen un modelo idóneo para el estudio de enfermedades cardiovasculares por su

similitud con el humano en derivación coronaria vascular y lesiones ateroscleróticas (Lunney, 2007).

1.4.- Estudios genómicos en calidad de carne

El gran avance acontecido en los últimos años en el ámbito de la genómica nos permite abordar a fecha de hoy nuevas aproximaciones que contribuirán sustancialmente a conocer mejor la arquitectura genética de caracteres complejos, objeto de estudio en esta tesis doctoral. Los fenotipos complejos no son resultado directo de las secuencias genómicas lo que implica necesariamente el análisis y la combinación de las diferentes estrategias y aproximaciones con el objetivo de ampliar el conocimiento sobre la base genética de los caracteres complejos utilizando: 1) tecnologías que permitan medir la expresión génica en diferentes tejidos (por ejemplo, los *microarrays)*, 2) la detección de QTL (en inglés, *Quantitative Trait Loci*) para estos caracteres, 3) análisis de QTL de expresión (eQTL) combinando los datos de expresión génica (*microarrays*) con los estudios de ligamiento de QTL, 4) estudios de expresión proteica y 5) estudios más exhaustivos de genes candidatos (Figura 5).

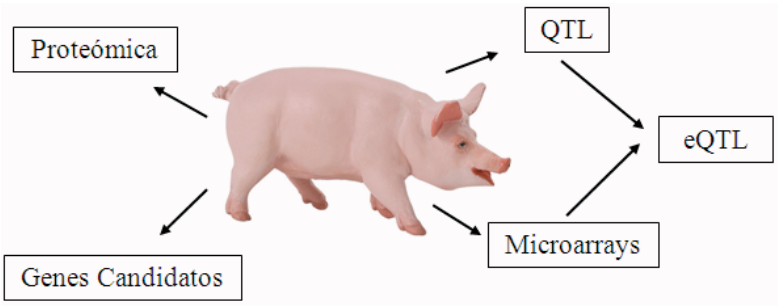


Figura 5: Esquema de la combinación de estrategias y aproximaciones utilizadas hoy en día en los estudios genómicos sobre caracteres complejos (como por ejemplo la calidad de la carne) en la especie porcina.

Sobre los años 90 se realizaron muchos estudios de QTL complementados con el estudio de genes candidatos a contener la mutación causal responsable de la variación genética. En este contexto, la metodología clásica utilizada para la detección y cuantificación de ARNm presente en una célula era el análisis por *Northern blot* y la RT-PCR (RT: reacción de transcripción reversa; PCR: reacción en cadena de la polimerasa). Estas técnicas suplieron las necesidades durante muchos años pero sufren la limitación del número de genes a reportar (Rothschild, 2003).

Hoy en día, con las nuevas tecnologías adaptadas en la especie porcina es posible realizar estudios de expresión génica mediante la tecnología de *microarrays* permitiéndonos estudiar y comparar la actividad de muchos genes simultáneamente. Tal y como se describe en los siguientes apartados, estas nuevas aproximaciones han contribuido a ampliar de manera significativa el conocimiento sobre la base genética de caracteres complejos y de especial importancia para el sector porcino como la GIM y el contenido y composición de AG. Además, recientemente se han desarrollado chips de genotipado masivo de alta densidad en la especie porcina, siendo, la aproximación de la Selección Genómica (GS) el futuro inmediato para el estudio de la variabilidad estructural a nivel genómico.

1.4.1.- Estudios de identificación de *Quantitative Trait Loci* (QTL)

Un locus cuyo polimorfismo explica una parte significativa de la variación de un carácter cuantitativo se conoce como un QTL (Lynch y Walsh, 1998). La presencia de un QTL se deduce por cartografía genética, donde la variación total está dividida en componentes ligados a varias regiones cromosómicas discretas.

Desde el punto de vista genético resulta interesante dilucidar la arquitectura de los caracteres complejos siendo un primer paso la detección de los genes o regiones cromosómicas (QTL) que controlan las diferencias fenotípicas observadas. Así, los estudios de QTL se basan en determinar si existe una asociación significativa entre la variación fenotípica de los caracteres estudiados y la variación genética analizada (Lynch y Walsh, 1998).

En las últimas dos décadas, la industria porcina está intentando definir mejor las bases genéticas de los caracteres relacionados con la calidad de la carne. En la especie porcina, la mayoría de los caracteres de interés económico son de tipo cuantitativo. Normalmente, este tipo de caracteres presentan un patrón de herencia complejo donde intervienen múltiples genes con diferentes efectos (Andersson, 2001), con lo que la detección de QTL resulta de especial importancia para comprender la arquitectura genética de éstos.

Tabla 4: Numero de QTL identificados en la especie porcina según la base de datos de QTL, QTLdb (http://www.genome.iastate.edu/cgi-bin/QTLdb/SS/browse; agosto 2010).

Cromosoma	QTLs Identificados	Cromosoma	QTLs Identificados
Y	1	9	183
X	249	10	145
1	1323	11	111
2	509	12	150
3	204	13	212
4	506	14	190
5	198	15	164
6	495	16	96
7	594	17	91
8	264	18	83

En este sentido, en la especie porcina se han identificado hasta la fecha 5.768 QTLs (Tabla 4), de los cuales más de 1.675 QTLs están influyendo en 281 caracteres diferentes relacionados con el metabolismo lipídico y la calidad de la carne (Hu y Reecy, 2007; http://www.genome.iastate.edu/cgi-bin/QTLdb/; agosto 2010).

Tal y como muestra la Tabla 5, la gran mayoría de los caracteres con más QTLs identificados en la especie porcina, están relacionados con la calidad de la carne, identificando, 89 QTL para el carácter de grasa dorsal, 31 QTL para el carácter de porcentaje de magro o 19 QTL relacionados con la profundidad de la grasa dorsal en la última costilla (Rothschild *et al*., 2007).

Tabla 5: Ranking de caracteres con más QTL identificados en la especie porcina (Rothschild *et al*., 2007).

Carácter	Nº QTL identificados
Grasa dorsal	89
Peso jamón	48
Área del lomo	47
Ganancia media diaria	46
Longitud de la canal	40
Ph	33
% magro	31
Nº pezones	27
Peso grasa dorsal	27
Grasa dorsal décima costilla	26
Grasa dorsal última costilla	25
Peso cabeza	24
Peso al nacimiento	23
Peso lomo	23
Color L	23
Diámetro fibra muscular	22
Peso canal	21
% grasa transversal	19
Profundidad grasa dorsal última costilla	19

Además, hasta la fecha se ha llevado a cabo múltiples estudios sobre QTLs con efecto sobre el contenido de GIM (de Koning et al., 1999; Paszek et al., 2001; Ovilo et al.,

2002; Rohrer et al., 2006 y Sanchez et al. 2007), sobre la composición en ácidos grasos en distintos depósitos de grasa como la GD, GIM y grasa perirenal (Pérez-Enciso *et al.* 2000; Clop *et al*. 2003; Nii *et al*. 2006; Sanchez *et al*. 2007 y Guo *et al*. 2009) y contenido de colesterol en músculo (Malek *et al*., 2001). No obstante, la identificación de la mutación causal de los QTL detectados es complicada debido al gran número de genes presentes en cada región cromosómica y al desequilibrio de ligamiento generado en las poblaciones experimentales (Varona *et al*., 2005).

En los últimos años, se han identificado mutaciones causales en algunos genes con un marcado efecto sobre caracteres productivos y económicos importantes para el sector porcino (Tabla 6).

Tabla 6: Principales genes identificados cuyas mutaciones están asociadas con la calidad de la carne y de la canal con una aplicación directa en la industria porcina según un adaptado de Davoli y Braglia, 2007.

Gen/Marcador	Carácter	Primera aplicación	Referencia
RYR1	Calidad de carne Síndrome estrés porcino Conversión pienso	1991	Fujii *et al*., 1991
MC4R	Ganancia diaria Conversión pienso % magro	1998	Kim *et al*., 2000
PRKAG3	Calidad de carne	1997	Milan *et al*., 2000
IGF2	% magro	2002	Van Laere *et al*., 2003
CAST	Calidad de carne	2003	Ciobanu *et al*., 2004

Entre las pocas mutaciones causales identificadas en la especie porcina, se encuentra la mutación no sinónima C1843T situada en el gen *RYR1* (*ryanodine receptor*) característico por estar relacionado con el síndrome de estrés porcino provocando un aumento del porcentaje de magro acompañado de una peor calidad de la carne debido al estrés (Fujii *et al*., 1991). Además, se han identificado otras mutaciones con un marcado efecto sobre la calidad de la carne en los genes *CAST* (*calpastatin*), *IGF2* (*insulin-like growth factor 2*), *MC4R* (*melanocortin receptor 4*) y *PRKAG3* (*protein kinase, AMP-activated, gamma 3 non-catalytic subunit*) relacionados todos ellos con caracteres como

el porcentaje de magro y la terneza de la carne (Tabla 6). No obstante, a pesar de los avances conseguidos, hasta la fecha no existe demasiado conocimiento sobre los genes con mutaciones responsables de QTL relevantes para caracteres de producción, manifestando así, la necesidad de continuar estudiando genes candidatos a ser responsables de la variación y regulación de caracteres de gran interés económico y productivo (Ibeagha-Awemu, 2008).

1.4.2- Estudios de la expresión génica global (*microarrays*)

La palabra *microarray* deriva del griego *mikro* (pequeño) y del inglés *array* (distribución ordenada). Los *microarrays* permiten el depósito de miles de puntos conteniendo genes o parte de genes sobre un portaobjetos para su estudio en paralelo (Harshman, 2002). De esta manera es posible tener una visión instantánea de actividad de genomas completos o de un grupo selecto de genes (Berger y Roberts, 2005) (Figura 6).

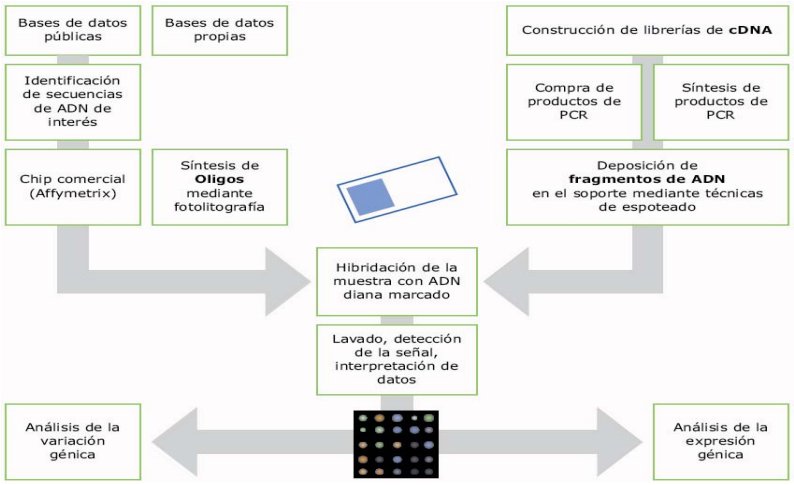


Figura 6: Proceso de fabricación y utilización de un *microarray* según el informe de vigilancia de tecnología de *microarrays* y biochips de ADN de Genoma España.

El desarrollo reciente de los *microarrays* de expresión para la especie porcina (Tuggle *et al.* 2007) aporta una herramienta de gran valor para estudiar el transcriptoma y sus mecanismos reguladores en esta especie. Actualmente, en el mercado existe más de una plataforma de *microarrays* de expresión para la especie porcina. Tuggle *et al*. (2007) realizan un estudio de comparación e integración de los datos de las dos plataformas más importantes (*QIagen* y *Affymetrix*) que existen dentro del contexto de los *microarrays* de expresión en porcino. La plataforma de *Affymetrix* contiene 23.937 sondas que se traducen en 20.201 genes, mientras que inicialmente, la plataforma de *QIagen* ofrecía *arrays* que contenían unos 8.000 genes aunque posteriormente lanzó al mercado una nueva versión que comprendía unos 20.000 genes aproximadamente. Esta última plataforma requiere una hibridación competitiva entre las clases o grupos de animales a comparar frente al sistema de *Affymetrix* donde la hibridación de cada individuo se realiza en un chip diferente, ofreciendo esta hibridación no competitiva una ventaja a la hora de analizar los datos.

Para el análisis y el estudio de expresión génica llevado a cabo en esta tesis doctoral, se utilizó *GeneChip Porcine Genome Arrays*® (*Affymetrix*) ya que ofrecía una mayor representación de sondas (20.201 genes) de 25 bases diseñadas directamente a partir de secuencias porcinas además de tener en cuenta que la tecnología *GeneChip* de *Affymetrix* hibrida una sola muestra por *chip*, de manera que para la comparación de perfiles de expresión de un control y un tratamiento se comparan los resultados para cada gen en cada *chip* (Aguado, 2007).

La creciente utilización de la tecnología de *microarrays* en especies ganaderas se ha centrado mayoritariamente en la realización de estudios orientados a detectar genes con expresión diferencial entre categorías (fenotipos, genotipos, tratamientos, etc), en la mayoría de los casos con el objetivo de detectar genes candidatos funcionales.

Tabla 7: Resumen de los principales trabajos de *microarrays* en la especie porcina en los tejidos muscular, adiposo y hepático.

Tejido	Localización	Grupos	Nº indv/ grupo	Plataforma	Nº de sondas	Test	Genes DE	Referencia
Músculo	*Longissimus Dorsi*; *Psoas*	Dos grupos de músculo divergentes (blanco y rojo respect.)	7	*Affymetrix-porcino*	23.937	*t-test*	70	Bai *et al*., 2003
	Longissimus Dorsi; *Psoas*	Dos grupos de músculo divergentes (blanco y rojo respect.)	4	*Affymetrix-porcino*	23.937	*t-test*	46 y 126 (según comparación)	da Costa *et al*., 2004
	Lomo	Dos razas (Duroc y Taoyuan) que difieren en el crecimiento del músculo	6	*Axon- humano*	9.182	ANOVA	117	Lin y Hsu, 2005
	Embrión/feto	Diferentes periodos de gestación divergentes en la formación de fibras musculares (miogénesis)	6	*Genomic Solutions- específico músculo*	1.671	*t-test;* FDR<0.05	292	Te Pas *et al*., 2005
	Embrión/feto	Dos razas (Duroc y Pietrain) que difieren en desarrollo muscular	7	*Genomic Solutions- específico músculo*	1.527	*t-test;* FDR<0.05	189	Cagnazzo *et al*., 2006
	Longissimus Dorsi	Dos grupos divergentes para valores de terneza	30	*Affymetrix-porcino*	23.937	modelo lineal*;* FDR<0.05	63	Lobjois *et al*., 2008
	Longissimus Dorsi	Dos grupos divergentes para caracteres relacionados con la calidad de la carne	74	*Affymetrix-porcino*	23.937	modelo mixto*;* FDR<0.05	226	Ponsuksili *et al*., 2009
	Longissimus Dorsi; *soleus*	Dos grupos de músculo divergentes (blanco y rojo respect.)	6	*Affymetrix-porcino*	23.937	*t-test*	550 (ratio>1.5) 159 (ratio>2)	Li *et al*., 2010
Grasa	Vascular	Dos grupos de cerdos (adultos y fetos)	4	*Telechem Inte- esp. grasa*	1.680	*t-test*	160	Hausmann *et al*., 2006
	Grasa Dorsal	Dos grupos de cerdos tratados con/sin clenbuterol	4	*China Agri. Univ.- esp. grasa*	11.521	*t-test*	336 y 507	Zhang *et al*., 2007
	Grasa Dorsal	Dos razas (Landrace y Taihu) que difieren en parámetros relacionados con el engorde	10	*National Engineering Eenter for Biochip -específico grasa*	656	ANOVA*;* FDR<0.05	140	Li *et al*., 2008
Hígado	Hígado	Dos razas (Landrace Aleman y Pietrain) que difieren en composición corporal (obesa y magra respect.)	10	*Affymetrix-porcino*	23.937	ANOVA	1298 y 2205 (según comparación)	Ponsuksili *et al*., 2007
	Hígado	Dos grupos con valores extremos de androstenona Dos razas (Duroc y Landrace Noruego)	29	*University of Aarhus*	27.774	modelo lineal*;* FDR<0.05	269	Moe *et al*., 2008

En este sentido, en la especie porcina podemos citar numerosos estudios en los que se emplean *microarrays* para determinar la expresión diferencial en los tejidos muscular (Bai *et al*. 2003; daCosta *et al*., 2004; Lin y Hsu, 2005; Te Pas *et al*., 2005; Cagnazzo *et al*., 2006; Lobjois *et al*., 2008; Ponsuksili *et al*., 2009; Li *et al*., 2010), adiposo (Hausmann *et al*., 2006; Zhang *et al*. 2007; Li *et al*., 2008), y hepático (Ponsuksili *et al*., 2007; Moe *et al*., 2008), tomando en consideración los tipos de fibras, el estado fisiológico o clínico, la raza, el fenotipo para algún carácter, o distintos tratamientos (Tabla 7). La mayoría de estos estudios se han abordado mediante el análisis de un solo tipo de tejido, en cambio, Ferraz *et al*. (2008) analizaron las diferencias de expresión entre tejidos de un mismo individuo, observando que el tipo de tejido genera más variabilidad en la expresión que el sexo o la raza.

Una de las finalidades de estos estudios de expresión global es la identificación de posibles genes candidatos funcionales para los caracteres estudiados así como la identificación de redes génicas y clústeres de expresión en el tejido estudiado. En este sentido, los estudios citados anteriormente han identificado y validado ciertos genes candidatos para varios atributos relacionados con la calidad de carne. Así, Wang *et al*. (2009) estudiaron las diferencias de expresión entre dos grupos de cerdos con alto y bajo contenido de GIM, identificando y sugiriendo que el gen *SFRS18* (*splicing factor serine-arginine rich protein*) se sobre-expresa en el tejido muscular de los cerdos con alto contenido de GIM. Con lo que respecta al tejido adiposo, Li *et al*. (2008) identificaron genes diferencialmente expresados en animales con diferencias en el contenido y composición de la GD relacionados con el metabolismo lipídico, la composición de la grasa y la calidad de la carne como es el caso de los genes *SCD* (*stearoyl-CoA desaturase*), *LPL* (*lipoprotein lipase*) y *FABP3* (*fatty acid binding protein 3*) contribuyendo a ampliar el conocimiento sobre los mecanismos de regulación

de la expresión génica en caracteres relacionados con el engrasamiento y la calidad de la carne en la especie porcina.

Dentro de este contexto, mencionar el trabajo de Bernard *et al*. (2007) como claro ejemplo de identificación de marcadores moleculares para calidad de carne en la especie bovina donde identifican 215 genes diferencialmente expresados entre los cuales se encuentra el gen *ADNJA1*, observando una correlación negativa entre la expresión de éste y la terneza de la carne explicando el 63% de la variabilidad genética. Con estos resultados, se confirma la utilidad de esta aproximación para identificar marcadores moleculares con una aplicación directa en esquemas de selección.

1.4.3.- Caracterización de *Quantitative Trait Loci* de expresión (eQTL)

Actualmente, existe un nuevo planteamiento experimental que aprovecha la información obtenida en los dos apartados explicados anteriormente con el objetivo de integrar estructural y funcionalmente los datos genómicos mejorando drásticamente la comprensión de las bases genéticas que controlan y regulan los caracteres cuantitativos en la especie porcina (Tuggle *et al*., 2007). La combinación de los datos de expresión génica obtenidos mediante *microarrays* con la información genotípica ofrece nuevas opciones para identificar los genes directamente asociados con los caracteres a estudiar. Este nuevo planteamiento experimental se conoce con el nombre *genetical genomics* y se basa en la identificación de QTL de expresión (eQTL; *Expression Quantitative Trait Loci*).

Se entiende como eQTL, una región cromosómica asociada con los niveles de expresión de uno o más genes. La aproximación metodológica es muy parecida a la descrita para los QTL en el apartado 1.4.1 utilizándose, en este caso, como fenotipos los datos de

expresión obtenidos en experimentos de *microarrays*, conjuntamente con la información genotípica (marcadores moleculares tipo microsatélites o SNPs) de los mismos individuos .

En este contexto, las regiones de eQTLs contienen varios genes asociados con la expresión de un gen siendo el mapeo de eQTL una poderosa herramienta para identificar la variación genética con efectos reguladores de los niveles de ARNm. Los eQTLs pueden representan un locus que se encuentra cerca del gen que está siendo controlado denominado *cis*-acting eQTL o *trans*-acting eQTL cuando uno o más locis están disociados del gen que está siendo controlado (Jansen y Nap, 2001).

Hasta el momento, se han publicado numerosos trabajos de detección de estas regiones genómicas con efecto sobre la expresión génica (eQTL) en organismos como el ratón (Hovatta *et al*., 2007 y van Nas *et al*., 2010), humano (Wheeler *et al*., 2009 y Franke y Jansen 2009), rata (Tesson y Jansen, 2009) y aves (de Koning *et al*., 2007) indicando claramente que la expresión génica está modulada, en parte, por el efecto de múltiples eQTL situados en *cis* y en *trans* (Cookson *et al*. 2009).

No obstante, a pesar del apogeo que vive actualmente la utilización de la técnica de *microarrays* en la especie porcina, son particularmente escasos los estudios publicados que aborden el cartografiado de eQTL a nivel genómico en esta especie, pudiendo citar tan solo algunos cartografiados parciales (Kadarmideen y Janss, 2007; Ponsuksili *et al*., 2008 y Wimmers *et al*., 2010) para un reducido número de tránscritos o cromosomas. Cabe destacar Ponsuksili *et al*. (2008) y Wimmers *et al*. (2010) por su estudio sobre caracteres de calidad de carne. En ambos trabajos los autores realizan el estudio de eQTL a partir de la selección previa de aquellas sondas que mostraron una correlación positiva con algún carácter relacionado con la calidad de la carne. Una vez realizados

los análisis de eQTL, combinaron los resultados obtenidos con los datos de QTL con el propósito de identificar genes candidatos implicados en la capacidad de retención de agua en el músculo *longissimus dorsi* porcino.

Por último, mencionar que el principal limitante en la realización de estudios de eQTL y de redes genéticas con una cierta potencia estadística es el elevado coste asociado a disponer de un tamaño muestral (número de *arrays* disponibles) suficiente. En este sentido, Schliekelman (2008) estudió la potencia ligada a la detección de eQTL y concluye que se necesitan tamaños muestrales superiores a 100 para detectar asociaciones entre los niveles de expresión de un tránscrito causal y el correspondiente fenotipo con al menos un 80% de potencia.

En el contexto de la caracterización de eQTL, las últimas tendencias muestran una mejora muy significativa en la resolución de los mapas de eQTL mediante los datos de genotipados masivos provenientes de los chips de SNP de alta densidad recientemente desarrollados en la especie porcina. En este sentido, a principios del 2009 se empezó a comercializar un chip de porcino con 60.000 SNP (PorcineSNP60; Illumina) abriendo las puertas de la selección genómica en la especie porcina esperando una revolución en el nivel de resolución y cartografiado fino y en la potencia estadística de este tipo de trabajos.

1.4.4- Caracterización de Genes Candidatos

Para desvelar y ampliar el conocimiento sobre la variación y la regulación de los diferentes caracteres que desempeñan un papel clave en la producción porcina, es necesario realizar un estudio más exhaustivo de los posibles genes responsables de estas

variaciones basándonos en diferentes criterios posicionales, funcionales y fisiológicos para la selección de los posibles genes candidatos.

Tabla 8: Principales genes candidatos cuyas mutaciones están asociadas con el engorde en la especie porcina según un adaptado de Switonski *et al*., 2010.

Gen	Polimorfismo	Carácter	Referencia
ACACA	c.4899G>A; c.5196T>C	GIM	Gallardo *et al*., 2009
APOE	c.43+170C>T	GD	Fan *et al*., 2009
CTSK	g.15G>A	GD	Fontanesi *et al*., 2010
FABP3	g.-158T>G	GD	Chmurzynska *et al*., 2007
FABP4	Intron 1	GIM; GD	Chmurzynska *et al*., 2004
GFAT1	g.101A>G	GD	Liu *et al*., 2010
GNRHR	c.-854T>G	GD	Fan *et al*., 2009
IGF1	Intron 1	GD	Estany *et al*., 2007
IGF2	g.3072G>A	GD	Oczkowicz *et al*., 2009
LEP	g.3469C>T	GD	Jiang y Gibson, 1999
LEPR	g.232T>A	GD	Mackowski *et al*., 2005
LIF	c.2604A>G	GD	Fan *et al*., 2009
MC4R	c.707A>G; c.892A>G	GD; GIM	Fan *et al*., 2009
PPARGC1A	g.1105C>A	GD	Stachowiak *et al*., 2007
RETN	g.-178G>A	GD	Cieslak *et al*., 2009
TNNI1	g.5174T>C	% grasa corporal; GD	Xu *et al*., 2010
TNNI2	g.1167C>T	% grasa corporal; GD	Xu *et al*., 2010
VDBP	c.473-164C>T	GD	Fan *et al*., 2009

En este contexto, tal y como se muestra en la tabla 8, se han identificado varios genes con polimorfismos asociados a caracteres relacionados con el engrasamiento con un efecto pequeño sobre la variación y la regulación de caracteres relacionados con la deposición de la grasa (Switonski *et al*., 2010). Entre estos genes cabe destacar los genes *IGF2* y *MC4R* responsables de una mutación causal (g.3072G>A y c.982A>G, respectivamente), siendo los primeros polimorfismos con una fuerte correlación con la deposición de la grasa en la especie porcina (Kim *et al*., 2000). No obstante, existe la necesidad de continuar estudiando genes candidatos a ser responsables de la variación y regulación de caracteres relacionados con la deposición de la grasa en el cerdo ya que

son pocos los polimorfismos identificados con una fuerte correlación con los caracteres bajo estudio.

En este sentido, una vez se seleccionan los genes candidatos, normalmente, el primer estudio que se realiza es la caracterización estructural de éstos, amplificando y secuenciando tanto la parte codificante como las regiones reguladoras y promotoras con el objetivo de identificar posibles polimorfismos responsables de la variación de los caracteres bajo estudio. Posteriormente, se suelen realizar estudios de asociación entre los polimorfismos identificados y los datos fenotípicos. Adicionalmente, es posible ampliar el estudio de este gen a la caracterización de su expresión mediante técnicas de PCR cuantitativa. Mediante los análisis de expresión génica se pueden realizar estudios de correlación y asociación entre la expresión génica del gen candidato y los diferentes datos fenotípicos.

Otra parte importante en la investigación de los genes candidatos es el estudio de la relación entre los niveles de expresión de ARNm y de proteínas así como estudios de asociación y correlación con datos fenotípicos, además de poder profundizar más en el tema estudiando la regulación de los genes candidatos antes y después de la traducción a proteína.

Además, otra de las técnicas disponibles para la caracterización de genes candidatos es el cultivo celular ya que puede utilizarse para el estudio de la funcionalidad de secuencias génicas exógenas mediante la introducción de un ADN externo por transfección. Esto se lleva a cabo para conseguir que las células expresen una proteína de interés. Más recientemente, la transfección de ARN interferentes se ha llevado a cabo como mecanismos adecuados para suprimir la expresión de un particular gen. Otra aplicación seria el estudio de la actividad transcripcional de los genes. Para ello, se

precisa preparar construcciones híbridas que contengan, por ejemplo el promotor del gen a estudiar seguido de una *cassette* de expresión con un gen *reporter* de fácil cuantificación (Figura 7). Ésta técnica permite medir la activación/inhibición de la expresión del gen como respuesta a diferentes tratamientos.

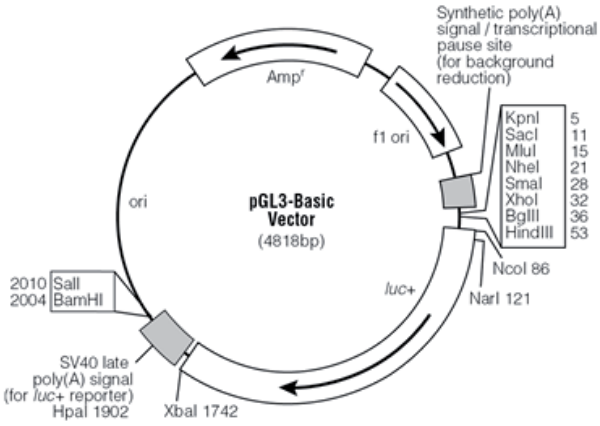


Figura 7: *Cassette* de expresión del gen *reported* Luciferasa *Renilla.L* en el vector *pGL3-Basic* frecuentemente utilizado en estudios de actividad transcripcional en cultivos celulares *in vitro*.

En esta tesis doctoral se ha incidido en el estudio de cuatro genes candidatos tanto posicionales como funcionales y fisiológicos para caracteres de calidad de carne. Siguiendo el patrón explicado anteriomente, para abordar el estudio de estos genes candidatos se ha caracterizado tanto la región codificante como la región promotora, se han identificado diferentes polimorfismos y se han realizado los correspondientes estudios de asociación y correlación. Además, según el gen, se han realizado análisis de expresión génica y/o proteica, así como estudios de funcionalidad mediante cultivo celular utilizando dos tipos diferentes de líneas celulares como las HepG2 (células de carcinoma de hígado humano) y las células C2C12 (células de músculo de ratón).

HepG2: Línea celular del tumor de hígado humano utilizada para el estudio de una variedad de funciones metabólicas específicas del hígado. La importancia de elegir una línea celular procedente del tejido hígado recae en su relevancia dentro del metabolismo de los lípidos al ser uno de los tejidos con un papel principal en el metabolismo de las lipoproteínas y del colesterol.

C2C12: Línea celular de músculo de ratón. El mioblasto es el tipo celular precursor de los miocitos (células musculares) que dará lugar a éstas por diferenciación celular. De la misma manera que las células HepG2, este tipo de células ha sido elegido debido a su especial interés en la calidad de la carne.

1.4.4.1.- *3-hidroxi-3-methilglutaril-CoA reductase* (*HMGCR*)

La enzima HMGCR actúa como enzima limitante en la síntesis del colesterol convirtiendo el HMG-CoA en mevalonato (Friesen y Rodwell, 2004) (Figura 8). La importancia de esta enzima en el metabolismo del colesterol queda reflejada en el hecho que los inhibidores de ésta inducen la expresión de los receptores LDL en el hígado, aumentan el catabolismo de las partículas LDL en plasma y reducen las concentraciones plasmáticas de colesterol (Kajinami *et al*., 2004).

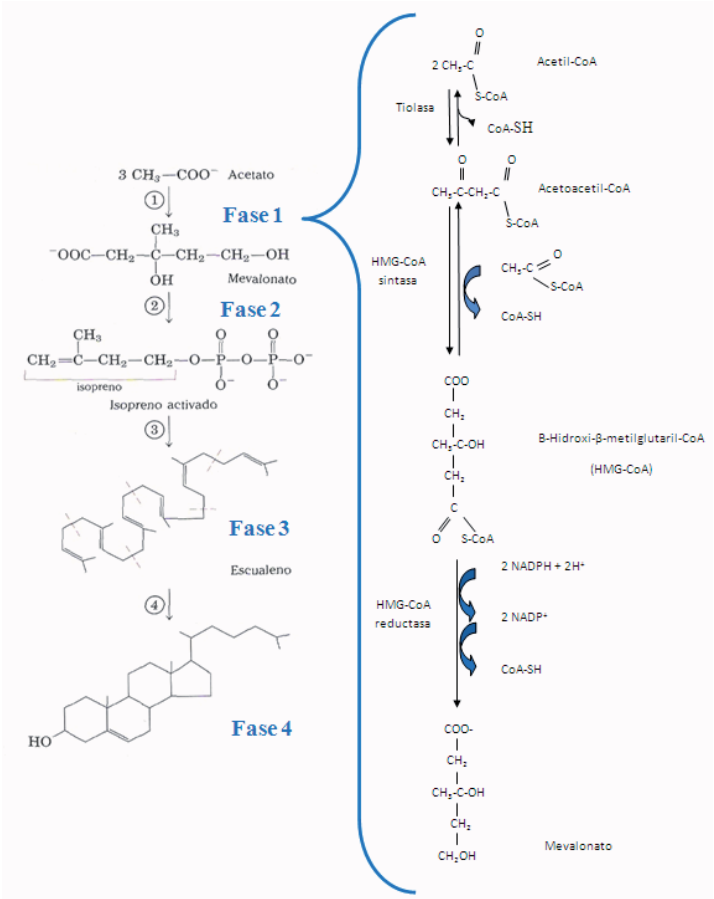


Figura 8: Formación del colesterol en cuatro fases (vía endógena). Fase 1 de la vía endógena de síntesis del colesterol donde actúa el gen *HMGCR* convirtiendo el HMGC-CoA en mevalonato.

La carne de cerdo constituye una de las fuentes principales de lípidos de las dietas occidentales, teniendo un elevado contenido de grasas saturadas aumentando el riesgo de sufrir enfermedades cardiovasculares. Debido a la importancia del problema, es necesario ampliar el conocimiento de la base genética del colesterol en esta especie con el fin de poder seleccionar los animales para mejorar su porcentaje de colesterol y triglicéridos, haciéndolos así, más saludables para el consumo humano. En este sentido, el gen *HMGCR* ha sido seleccionado como gen candidato por su efecto sobre la

composición lipídica de la carne ya que en trabajos anteriores se ha observado que elevadas concentraciones de lípidos plasmáticos pueden estar relacionados con un incremento en la deposición de la grasa en el tejido adiposo y muscular (Wilding, 2007).Además, aunque el cerdo se ha sugerido como modelo animal de enfermedades cardiovasculares, obesidad y metabolismo de las lipoproteínas y del colesterol por su similitud con el humano en derivación coronaria vascular y lesiones ateroscleróticas, hay un gran desconocimiento sobre la base genética del metabolismo del colesterol en el cerdo (Pond y Mersmann, 1996). Por este motivo, en esta tesis doctoral se ha analizado estructural y funcionalmente el gen y el promotor *HMGCR* porcino como fuente de variación de los niveles de lípidos plasmáticos con el objetivo de ampliar el conocimiento de los genes que regulan la síntesis del colesterol y la deposición de lípidos en cerdos así como su efecto sobre la calidad de la carne.

En humanos, dos polimorfismos tipo SNP (*Single Nucleotide Polymorphism*) situados en el intron 5 (SNP12) y en el intron 15 (SNP29) (cromosoma 5) de este gen se han asociado significativamente a una respuesta menor al tratamiento con estatinas (Chasman et al., 2004). Además, otras mutaciones en este gen se han asociado al aumento de niveles de triglicéridos y partículas VLDL plasmáticas (Tong *et al*., 2004). Recintemente, Burkhardt *et al*. (2008) identificaron variaciones en la cantidad de colesterol LDL plasmático producido por un procesamiento alternativo del exón 13 humano.

En cerdo, el gen *HMGCR* se localiza en el cromosoma 2 y consta de 3447 pares de bases distribuidos en 20 exones y 19 intrones según la anotación depositada en la versión Sscrofa10 del genoma porcino (ensembl id ENSSSCG00000014080) codificando una proteína de 887 aminoácidos. Además, se han identificado varios

polimorfismos en el gen *HMGCR* porcino mediante AFLPs aunque la posición de estas mutaciones no ha sido asignada (Davis *et al*., 1995).

1.4.4.2.- *Stearoyl-CoA Desaturase* (*SCD*)

La enzima lipogénica *SCD* cataliza la biosíntesis de los ácidos grasos insaturados siendo la principal responsable de convertir los AGS en AGMI (16:0 en 16:1 y 18:0 en 18:1) (Sjögren *et al*., 2008). El producto principal de *SCD* el ácido oleico (18:1) formándose por desaturación del ácido esteárico (18:0) (Figura 9).

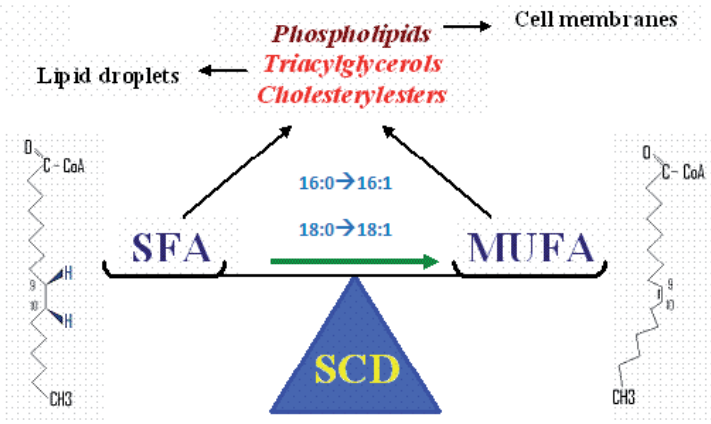


Figura 9: Esquema de la acción del gen *SCD* como principal responsable de convertir los ácidos grasos saturados en monoinsaturados.

El gen *SCD* es uno de los genes candidatos fisiológicamente relacionados con la regulación de la síntesis de los AG. Su identificación y estudio permitirán tener un mayor conocimiento sobre los mecanismos de regulación del contenido y la composición de la GIM en el cerdo, así como la deposición de la grasa subcutánea.

La actividad del gen *SCD* ha sido extensamente estudiada en modelos de ratón, indicando un considerable impacto del gen *SCD* en enfermedades como la obesidad (Ntambi y Miyazaki, 2004). El *Knockout* del gen *SCD* en ratones ha reducido la síntesis y concentración plasmática de triacilgliceridos. Por el contrario, en la literatura se pueden encontrar muy pocos datos sobre la importancia de la actividad desaturasa en el tejido adiposo humano (Sjögren *et al*., 2008).

En ratón y humano, se han identificado cuatro y dos isoformas respectivamente mientras que en cerdo se ha identificado únicamente una isoforma en el tejido adiposo, aunque el número de isoformas de *SCD* en el músculo porcino es desconocido.

En la especie porcina, el gen *SCD* está situado en el cromosoma 14, y consta de 6 exones y 5 intrones. Este gen esta descrito a nivel genómico y de cDNA, aunque la regulación tejido-específica de éste gen aun no está clara. Gondret *et al*. (2008) sugieren que la deposición de grasa en los diferentes depósitos del cerdo podría estar regulada por diferentes mecanismos y que la enzima lipogénica *SCD* juega un papel muy importante en este proceso. En este sentido, existen algunos trabajos donde estudian la respuesta específica del gen *SCD* porcino en músculo y tejido adiposo utilizando medidas de expresión proteica, observando que al reducir la cantidad de proteína en la dieta, incrementaba la expresión de la proteína de *SCD* en el músculo porcino, mientras que en la grasa subcutánea no observó ningún aumento de la expresión de la proteína, siendo desconocidas, aun, las razones de los cambios tejido-específicos de la expresión de las enzimas lipogénicas (Doran *et al*., 2006). Además, se ha observado una correlación significativa entre los niveles de expresión proteica del gen *SCD* porcino y los niveles totales de AG en la GIM (Doran *et al*., 2006), encontrando una relación similar en la especie bovina donde una alta actividad del gen *SCD* está asociada positivamente con los AGMI y negativamente con los AGS (Jiang *et al*., 2008).

Dado lo expuesto, se sugiere realizar un estudio más exhaustivo de gen *SCD* en la especie porcina debido a la directa implicación que tiene con caracteres relacionados con la calidad de la carne, siendo el principal responsable de convertir los AGS en AGMI con el consecuente efecto sobre la composición de la GIM.

1.4.4.3.- *Acetyl-CoA Carboxylase* (*ACACA*)

La Acetyl-CoA carboxilasa es una enzima que cataliza el primer paso de la biosíntesis de ácidos grasos de cadena larga, convirtiendo acetyl-CoA en malonyl-CoA, siendo éste el paso limitante en la síntesis de AGS (Abu-Elheida *et al*., 1995). Además, también juega un importante papel en el crecimiento celular y en el metabolismo de las grasas y aminoácidos.

Existen dos isoformas de acetyl-CoA carboxilasa, la alfa y la beta, codificadas por dos genes diferentes. La acetyl-CoA carboxilasa-alfa esta codificada por el gen *ACACA* y se expresa en tejidos lipogénicos como el hígado, tejido adiposo y glándula mamaria (Barber *et al*., 2001). Además, se han identificado varios promotores y procesamiento alternativo de exones, afectando en la inactivación de esta enzima (Barber *et al*., 2005).

Dada la importancia de este gen relacionado con la síntesis de AGS, ha sido seleccionado como gen candidato responsable de la variación y regulación del contenido y composición de la grasa (subcutánea e intramuscular).

Este gen está completamente secuenciado en humano y en otras especies, entre ellas, la especie porcina. En porcino, el gen *ACACA* se localiza en el cromosoma 12 y consta de 54 exones y 53 intrones codificando una proteína de 2.346 aminoácidos (Gallardo *et al*., 2009). La localización del gen *ACACA* en porcino coincide con un QTL identificado en

el cromosoma 12 que tiene un importante efecto en la composición de AG (Calvo *et al*., 2000). Se ha caracterizado y secuenciado la región codificante del gen *ACACA* (Muñoz *et al*., 2007 y Gallardo *et al*., 2009;) en cerdos de la raza Duroc y Ibérico x Landrace identificándose polimorfismos asociados a composición de AG y contenido de GIM. Por ejemplo, Gallardo *et al*. (2009) observan una asociación entre un polimorfismo situado en el gen *ACACA* y una disminución del porcentaje AGPI y omega-6, siendo una relación beneficiosa desde el punto de vista tecnológico ya que un aumento de los niveles de AGPI provoca la ocurrencia del olor y sabor rancio en la carne (Wood *et al*., 2008). Además de la importancia que tiene el gen *ACACA* sobre las características tecnológicas y de calidad de la carne, un aumento de los niveles de omega-3, omega-6, y de AGPI, disminuye la susceptibilidad de padecer enfermedades cardiovasculares (Ventanas *et al*., 2007 y Simopoulos, 2008).

Dada la importancia de este gen por su fuerte asociación con el contenido y la composición de AG en porcino, mejorando a su vez parámetros importantes que influyen sobre la salud humana, resulta interesante estudiar este gen como candidato para caracteres de calidad de la carne porcina.

1.4.4.4.- *Delta-6-Desaturase* (*Δ6D*)

La composición de los AG y el contenido de grasa no solo dependen de la lipogénesis particular de cada tejido, sino que también dependen de otros factores incluyendo el transporte de los AG desde otros sitios lipogénicos. En muchos mamíferos, la lipogénesis ocurre mayoritariamente en el hígado y/o el tejido adiposo (Girard *et al*., 1994). En cerdos, el principal sitio del metabolismo de los AG es el tejido adiposo

subcutáneo, obteniendo la mayor expresión y actividad de los principales enzimas lipogénicos como la *Δ6D* (O'Hea y Leveille, 1969) (Figura 10).

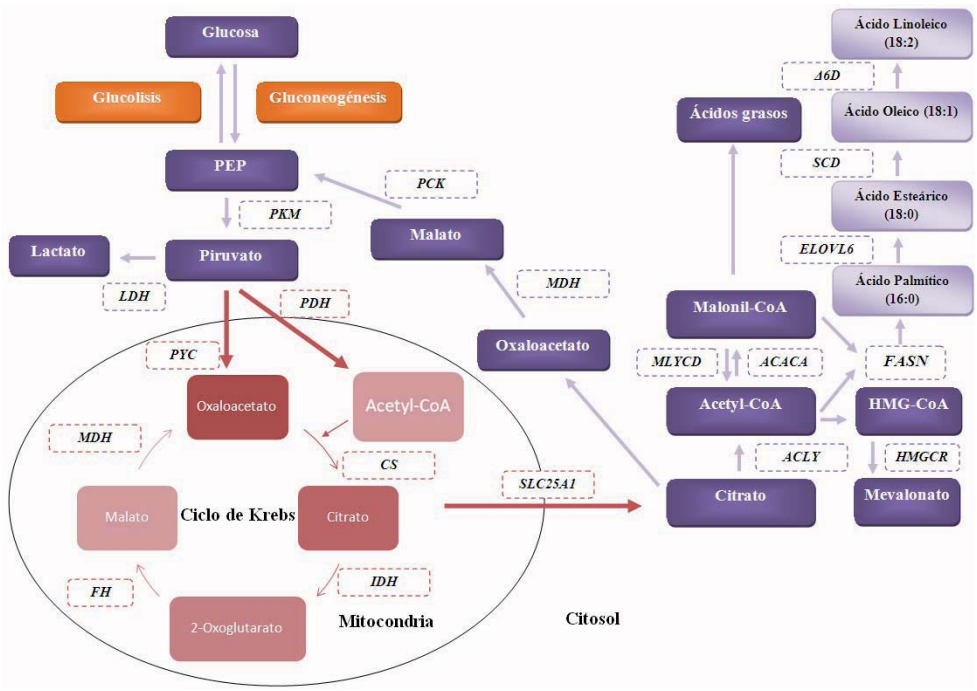


Figura 10: Esquema del proceso de la lipogénesis en un adipocito según un adaptado de Liu *et al*., 2008 y Sjögren *et al*., 2008 donde actúan los genes *Δ6D*, *ACACA*, *HMGCR* y *SCD* entre otros.

En este sentido, las enzimas desaturasa regulan la instauración de los AG mediante la introducción de dobles enlaces entre los carbonos definidos de la cadena acyl. Entre este grupo de desaturasas se encuentra la enzima *Δ6D* (también conocida como *FADS2*), enzima limitante en la síntesis de AGPI y que interviene en la conversión del ácido oleico (18:1) en linoleico (18:2) (Figura 11).

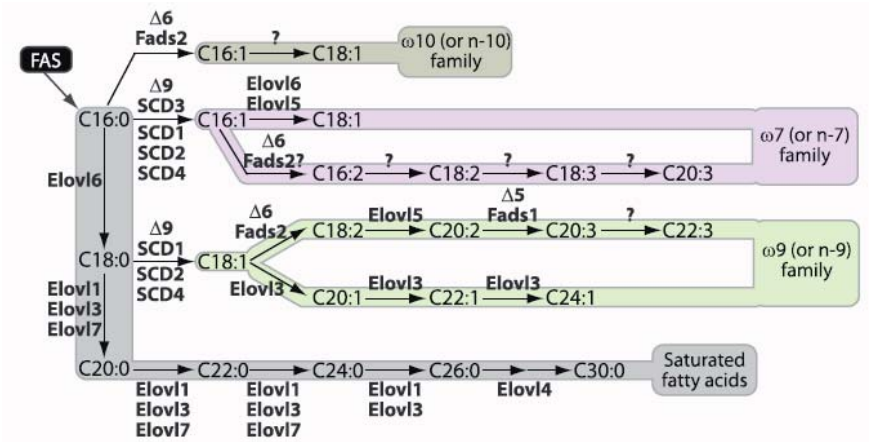


Figura 11: Esquema de la biosíntesis de AG de cadena larga (Guillou *et al*., 2010), donde la enzima *FADS2* corresponde a la enzima *Δ6D*.

Tanto en humano como en animales, las principales desaturasas son la *Δ5D* (*delta-5-desaturase*) y la *Δ6D* (*delta-6-desaturase*) caracterizadas por estar implicadas en la resistencia a la insulina y en la obesidad entre otras enfermedades aunque su papel en la actividad desaturasa del tejido adiposo es desconocida (Vessby, 2003). Ambas desaturasas coinciden en el 75% de la secuencia, mostrando homología en la organización de los exones e intrones, localizándose, además, en la misma región del cromosoma 11 humano (Guillou *et al*., 2010). En cerdo, el gen *Δ6D* se localiza en el cromosoma 2 y consta de 1252 pares de bases distribuidos en 11 exones y 10 intrones según la anotación depositada en la versión Sscrofa10 del genoma porcino (ensembl id ENSSSCG00000013072) codificando una proteína de 376 aminoácidos.

A pesar de todos los estudios realizados hasta el momento, la actividad desaturasa en la regulación de la deposición de la grasa en cerdo no está clara, proponiendo en esta tesis doctoral un estudio más exhaustivo de los enzimas involucrados en el contenido y composición de la grasa siendo un candidato el gen *delta-6-desaturase*.

II. Objetivos

El trabajo de esta tesis doctoral está enmarcado en una línea de investigación dedicada al estudio de las bases genéticas del metabolismo de las grasas en relación con la producción de carne de porcino de alta calidad y saludable, financiada a través de los proyectos AGL2002-04271-C03-02, CICYT GEN2003-20658-C05-05 y AGL2007-66707-C02-01.

El objetivo final es la identificación de polimorfismos y mecanismos de regulación responsables de la variabilidad genética de estos caracteres complejos en porcino. Con el fin de llevar a cabo el objetivo principal, el desarrollo experimental de esta tesis doctoral se ha centrado en:

- Describir los patrones de expresión génica diferencial entre animales con fenotipos divergentes en tejidos particularmente vinculados al metabolismo lipídico.
- Identificar las regiones genómicas responsables de la variación en los patrones de expresión génica (eQTL) en el músculo *gluteus medius* porcino integrando los datos de *microarrays* con los estudios de ligamiento de QTL.
- Caracterizar estructural y funcionalmente genes candidatos para caracteres de engorde y de calidad de carne en porcino.

III. Artículos

3.1. - Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits

Cánovas, A., Quintanilla, R., Amills, M., Pena, R.N

***BMC Genomics*. 11:372-392. 2010**

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

BMC Genomics

RESEARCH ARTICLE Open Access

Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits

Angela Cánovas[1], Raquel Quintanilla[1], Marcel Amills[2] and Ramona N Pena*[1]

Abstract

Background: Selection for increasing intramuscular fat content would definitively improve the palatability and juiciness of pig meat as well as the sensorial and organoleptic properties of cured products. However, evidences obtained in human and model organisms suggest that high levels of intramuscular fat might alter muscle lipid and carbohydrate metabolism. We have analysed this issue by determining the transcriptomic profiles of Duroc pigs with divergent phenotypes for 13 fatness traits. The strong aptitude of Duroc pigs to have high levels of intramuscular fat makes them a valuable model to analyse the mechanisms that regulate muscle lipid metabolism, an issue with evident implications in the elucidation of the genetic basis of human metabolic diseases such as obesity and insulin resistance.

Results: Muscle gene expression profiles of 68 Duroc pigs belonging to two groups (HIGH and LOW) with extreme phenotypes for lipid deposition and composition traits have been analysed. Microarray and quantitative PCR analysis showed that genes related to fatty acid uptake, lipogenesis and triacylglycerol synthesis were upregulated in the muscle tissue of HIGH pigs, which are fatter and have higher amounts of intramuscular fat than their LOW counterparts. Paradoxically, lipolytic genes also showed increased mRNA levels in the HIGH group suggesting the existence of a cycle where triacylglycerols are continuously synthesized and degraded. Several genes related to the insulin-signalling pathway, that is usually impaired in obese humans, were also upregulated. Finally, genes related to antigen-processing and presentation were downregulated in the HIGH group.

Conclusion: Our data suggest that selection for increasing intramuscular fat content in pigs would lead to a shift but not a disruption of the metabolic homeostasis of muscle cells. Future studies on the post-translational changes affecting protein activity or expression as well as information about protein location within the cell would be needed to to elucidate the effects of lipid deposition on muscle metabolism in pigs.

Background

Muscle lipid metabolism affects a wide diversity of meat quality traits that are of huge importance for the pig industry [1]. Intramuscular fat percentage is favourably associated with meat texture, tenderness, flavour and juiciness [2,3]. In addition, muscle fat composition has a strong effect on the sensorial, nutritional and technological properties of meat [1]. The proportion of monounsaturated fatty acids (mainly oleic) is a key factor determining meat fat consistency and taste [1], whereas polyunsaturated fatty acids have a marked tendency to be oxidized, producing a rancid odour and taste that decrease meat consumer's acceptance. From a human health perspective, increased amounts of polyunsaturated fatty acids in the diet diminish the susceptibility to suffer cardiovascular diseases, while saturated fatty acids have the opposite effect [4].

Selection for higher muscle fat content and composition might have a long-term impact on pig muscle physiology that has not been evaluated yet. For instance, in obese humans, accumulation of triacylglycerols in the myocyte is associated with the development of insulin resistance, metabolic syndrome and type II diabetes [5,6]. In this way, increases in skeletal muscle fat stores are often accompanied by a parallel reduction in the β-oxidation of fatty acids and the progressive accumulation of lipid metabolites, such as diacylglycerol and long-chain acyl-CoAs, which impair insulin-stimulated glucose transport [5,6]. Although selection for increased intramuscular fat content is not expected to impair pig health,

* Correspondence: romi.pena@irta.cat
[1] IRTA, Genètica i Millora Animal, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain
Full list of author information is available at the end of the article

it might have consequences on muscle lipid metabolism that need to be determined. Moreover, given the particularly high intramuscular fat content of Duroc pigs, they represent an interesting animal model to study the changes in gene expression that are produced in response to increased fat deposition. We have evaluated the muscle transcriptomic profiles of Duroc pigs displaying divergent fatness phenotypes to gain new insights into these fundamental questions.

Methods

Biological samples

Animals came from a commercial Duroc line used for the production of fine quality cured ham and characterised by a high intramuscular fat content. This commercial line was being selected for a compound objective including prolificacy, growth, intramuscular fat and leanness. A population of 350 half-sib castrated males was generated by mating five males with 400 females and selecting one male offspring per litter, as described in Gallardo et al.[7]. A total of 70 phenotypes on growth, fatness, feed efficiency and carcass and meat quality traits were recorded in these animals, including weight, daily food intake, fat deposition measures, and intramuscular fat content and fatty acid composition (C:12-C:22 interval) of *gluteus medius* and *longissimus thoracis et lumborum* muscles [8]. In addition, serum total cholesterol, HDL- and LDL-bound cholesterol and triglycerides were measured in two blood samples taken at 45 and 190 days of age as described in Gallardo et al. [7]. At slaughter, samples of muscle from the *gluteus medius* were collected, snap frozen in liquid nitrogen and stored at -80°C until analysed. The experimental procedures, traits recording and blood sampling were approved by the Ethical Committee of the Institution (IRTA - *Institut de Recerca i Tecnologia Agroalimentàries*).

Experimental design

With the aim of mimicking to some extent the effects of divergent selection for two extreme fatness phenotypes, a principal component analysis (PCA) that allowed synthesising global phenotypic variability recorded in our population was performed by means of the PRINCOMP procedure of SAS (SAS Inst. Inc., Cary, NC). Several preliminary analyses considering the whole set of lipid-related phenotypes (including muscle content of individual fatty acids and several carcass fat measures) allowed us to detect traits with either low variability or displaying high correlations with other ones. These traits did not contribute significantly to the phenotypic variance explained by the principal components and were discarded from further analyses. Finally, a subset of 13 measured traits (Table 1) were considered as the most relevant descriptors of total population variability related to lipid metabolism and lipid content of *gluteus medius* muscle, and were included in the PCA. The first principal component (see Table 1 for coefficients) explained 30.7% of the global phenotypic variance of these traits and was used to rank the animals. Pigs from the extremes of the ranking were selected to create the HIGH (n = 35) and LOW (n = 35) groups, as described in Figure 1. More details about the PCA results and phenotypic variation of the HIGH and LOW groups with divergent phenotypes for fatness traits is provided in the results section.

RNA isolation and microarray hybridisation

Samples of *gluteus medius* were ground with mortar and pestle in liquid nitrogen and homogenised with a mechanical rotor. RNA was isolated by the acid phenol method [9] using the RiboPure kit (Ambion, Austin, TX). RNA was quantified in a Nanodrop ND-1000 spectrophotometer and checked for purity and integrity in a Bioanalyzer-2100 (Agilent Technologies, Inc., Santa Clara, CA).

Seventy total RNA samples from the HIGH (n = 35) and LOW (n = 35) groups were individually hybridised on GeneChip Porcine Genomic arrays (Affymetrix, Inc., Santa Clara, CA). A randomised list of the samples was generated to assign hybridisation order to the 70 RNA samples. Total RNA was used to synthesize double stranded cDNA using the One Cycle cDNA Synthesis Kit (Affymetrix, Inc.) which incorporates a T7 RNA polymerase promoter. Biotin-labelled antisense cRNA was obtained using the same kit starting with 5 μg of total RNA and the oligo-(dT) primer 5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-$(dT)_{24}$.

cRNAs were purified with the GeneChip Sample Cleanup Module (Affymetrix, Inc.) and then 20 μg of cRNA were fragmented at 94°C for 30 min in 40 μl of 40 mM Tris-acetate pH 8.1, 100 mM KOAc, 30 mM $Mg(OAc)_2$, checked using the Bioanalyzer 2100 (Agilent Technologies, Inc.) and added to a hybridisation cocktail containing control oligonucleotide B2 (50 pM) and Eukaryotic Hybridization controls (BioB, BioC, BioD, Cre) at 1.5, 5, 25 and 100 pM final concentration respectively from the GeneChip Eukaryotic Hybridisation Control Kit (Affymetrix, Inc.), herring sperm DNA (0.1 mg/ml) and acetylated BSA (0.5 mg/ml). The GeneChip Porcine Genome Array was equilibrated to RT and prehybridised with 1× hybridisation buffer (100 mM MES, 1 M [Na+], 20 mM EDTA, 0.01% Tween 20) at 45°C for 10 min with rotation. The hybridisation cocktail was heated to 99°C for 5 min in a heat block, transferred to 45°C for 5 min, added to the arrays and hybridised at 45°C for 16 hours with rotation in the Affymetrix GeneChip Hyb Oven 640 (Affymetrix, Inc.). GeneChips were washed and labelled with streptavidin phycoerythrin in the Fluidics Station 450 (Affymetrix, Inc.) using the protocol EukGE-WS2-v5

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

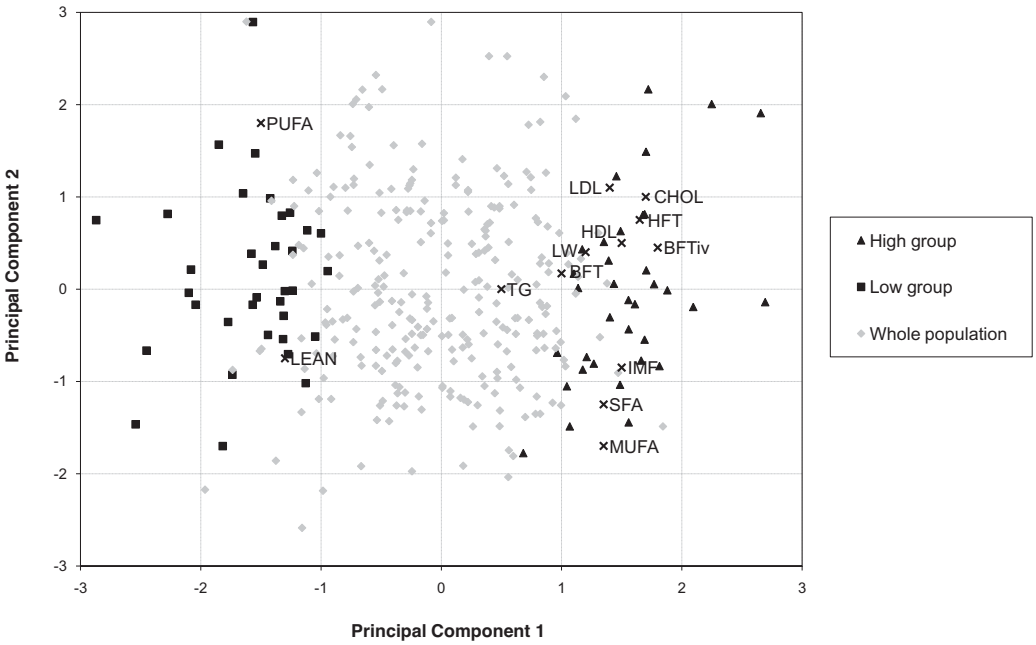


Figure 1 Graphical plot of the first and second principal components summarising phenotypic variation in the Duroc population. Summary of phenotype variation on serum lipid, growth and fatness parameters in the Duroc pig population. Relative weight of considered variables and position of animals from the HIGH and LOW groups are represented in the graph. Abbreviations are defined in Table 1.

provided by Affymetrix. GeneChips were finally scanned in an Agilent G3000 GeneArray Scanner (Agilent Technologies, Inc.).

Quality control of expression data

The "Affy" and "Sympleaffy" packages from the Bioconductor project [10] were used to implement a set of quality control metrics recommended by Affymetrix to assess the quality of RNA samples and their subsequent labelling and hybridisation steps [11]. These include comparison of the background signal, intensity scaling factor, percentage of genes called, and 3'/5' intensity ratio of control probes in all arrays. Quality control analysis performed with raw intensity data resulted in the identification of two arrays with problems in the labelling efficiency of the 5' *vs* 3' control probes. Therefore, data from these two arrays were discarded from further analysis, reducing the number of data available to 68 arrays (34 per group).

Class comparison analysis on expression data

Data pre-processing, normalisation and class comparison analysis were carried out with the BRB-ArrayTools software version 3.7.1 [12], which is available online at http://linus.nci.nih.gov/BRB-ArrayTools.html, as follows: microarray data normalisation was performed using the gcRMA algorithm, which corrects the intensity of each probe by its GC content [13]; genes showing minimal variation across the set of arrays were excluded from the analysis, that is, only genes displaying more than 20% of expression values over ± 1.5 times the median expression of all arrays were used for further analysis. From the total 23937 spots of the array, 4299 spots passed these filtering conditions. For each probe, expression fold-change was calculated as the ratio between median values in both groups.

Class comparison was performed with a dataset of 68 samples of *gluteus medius* from the HIGH (n = 34) and LOW (n = 34) groups. A global two-group t-test with a random variance model was performed. This type of t-test is an improvement over the standard one-gene *t*-test, as it permits sharing information about within-class variation among genes without assuming that all genes have the same variance [14,15]. By using these variance estimates, this method gains degrees of freedom over the standard t-test, providing greater sensitivity with no loss in specificity [14]. For each probe, the significance p-values of the group effect were calculated based on 100,000 random permutations, and the nominal significance level

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

Table 1: Total population, HIGH and LOW group mean values ± standard error for the traits used in the selection index

	Weight in PC1†	Population *n = 350*		HIGH group *n = 35*		LOW group *n = 35*	
Carcass traits							
LW - Live weight (Kg)	0.218	122.08	± 0.74	127.30	± 1.94	111.94	± 2.66
BFTiv - Backfat thickness (*in vivo*) (mm)	0.342	23.58	± 0.26	27.26	± 0.69	19.00	± 0.64
BFT - Backfat thickness 3rd-4th ribs (mm)	0.200	38.31	± 0.60	47.43	± 1.84	32.49	± 1.71
HFT - Ham fat thickness (mm)	0.328	25.71	± 0.19	28.47	± 0.41	20.94	± 0.61
LEAN - Lean %	-0.265	40.86	± 0.24	38.26	± 0.79	45.94	± 0.71
Meat quality traits (gluteus medius)							
IMF - % Intramuscular fat	0.304	5.17	± 0.10	7.46	± 0.39	3.58	± 0.15
SFA - % Saturated fatty acids	0.276	36.46	± 0.10	38.55	± 0.25	34.65	± 0.21
PUFA - % Polyunsaturated fatty acids	-0.315	20.62	± 0.32	14.37	± 0.52	27.35	± 0.81
MUFA - % Monounsaturated fatty acids	0.275	42.92	± 0.27	47.08	± 0.45	38.00	± 0.80
Serum lipid levels - 190 days							
CHOL - Total cholesterol (mg/dl)	0.335	125.15	± 1.40	157.60	± 5.02	104.03	± 2.70
HDL - HDL-cholesterol (mg/dl)	0.285	51.66	± 0.54	61.63	± 1.27	43.26	± 1.60
LDL - LDL-cholesterol (mg/dl)	0.266	63.12	± 1.09	82.65	± 4.70	50.58	± 2.53
TG - Triacylglycerides (mg/dl)	0.101	51.37	± 1.25	66.43	± 4.46	47.66	± 4.75

The weight of each trait in the first principal component (PC1) is also indicated.
†PC1 - weight of each trait in the first principal component.

of each univariate test was restricted to $p < 0.01$ [12,14]. Additionally, a multiple test correction was performed by restricting the FDR to < 0.05 [16]. The significance level of the whole experiment was also calculated by data permutation: for each permutation, the p-values were recomputed and the number of genes significant at the 0.01 level of the permutation p-value was annotated.

Gene Ontology-based analysis

We performed a functional categorization of our resulting list of differentially expressed genes using Gene Ontology (GO) information. To annotate the probes in the Affymetrix array we used the latest annotation file available [17]. First, Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov) was employed to explore functional class scoring in the resulting gene list by means of GO term enrichment analysis [18]. Significance levels were calculated following a modification of Fisher's exact test. A multiple testing-corrected p-value was also calculated using Benjamini and Hochberg algorithm [16], and only GO terms with Benjamini-corrected $p < 0.1$ were considered. DAVID was additionally used to explore the biological pathways enriched in the resulting gene list, using information from each individual gene and computing a total over-representation value for each pathway represented in the Kyoto Encyclopaedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg) and BioCarta http://www.biocarta.com.

As a complementary approach, Ingenuity Pathway Analysis, Ingenuity Systems (IPA, http://www.ingenuity.com) bioinformatic tools were employed to explore the distribution of differentially expressed genes in well reported canonical signal transduction or metabolic pathways integrating fold-change data from the microarray experiment and gene-to-gene or protein-to-protein interaction information generated from the literature databases. In this approach, statistical significance of pathway overrepresentation was established with respect to a null distribution constructed by permutations.

Systems biology studies

We have used the Pathway Express program [19] to identify pathways affected by the list of genes differentially expressed between groups. This program integrates topological information regarding the position of genes in the metabolic and gene signalling pathways defined by KEGG and the multiple roles that these genes might have depending on the pathway they belong to. This program calculates a Perturbation Factor for each gene, which is computed by normalising the expression ratio with information about its topology in the pathway (i.e., number of

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

genes interacting with it and type of interaction) [19]. In each pathway, Perturbation Factors from each individual gene are combined to calculate an Impact Factor, which is normalised according to the total number of genes in the pathway [19]. The complete list of genes in the array was used as a reference set in the analysis.

Quantitative real time PCR

Quantitative real time PCR (qPCR) was used to validate relevant genes differentially expressed between the two groups of animals with extreme values for lipid metabolism traits. With this aim, a subset of the 10 most extreme animals from each group was selected. Total RNA samples were retrotranscribed with random hexamers and SuperScript III retrotranscriptase (Invitrogen, Carlsbad, CA) following manufacturer's instructions. cDNA was diluted 1:10 in DEPC-treated H_2O prior to qPCR analysis. Primers and TaqMan probes (Additional file 1) were designed with the Primer Express software (Applied Biosystems, Foster City, CA) using pig sequences obtained by performing a BLAST search with the array probe sequence. For each gene, a standard curve was generated by amplifying serial dilutions of a control ss-cDNA to check for linearity between initial template concentration and Ct values. Quantitative real-time PCR assays were carried out in triplicate in an ABI-7500 device (Applied Biosystems) in a final volume of 5 μl containing 1× Power SYBRgreen Master mix (Applied Biosystems) and 200 nM of each primer. For the two reference genes (*HPRT* and *RPL32*), PCR reactions contained 300 nM of each primer, 200 nM TaqMan probe and 1× Universal Taqman Master Mix (Applied Biosystems). The following thermal profile was used for all reactions: 10 min at 95°C, 40 cycles of 15 s at 93°C and 1 min at 60°C, followed by a quick denaturation at 95°C for 5 min plus a slow ramp to 30°C to generate a dissociation curve to control the specificity of the amplified product. In order to quantify and normalise the expression data we used the ΔΔCt method [20] using the mean Ct value from the two reference genes and the Ct values for each test gene. For each gene, expression values between groups were compared with a t-test and differences were considered significant at $p < 0.05$. The correlation analysis between qPCR and microarray expression data was performed using the CORR procedure of SAS (SAS Inst. Inc.).

Results

Phenotypic variation in the two groups with highly divergent phenotypes for fatness traits

A principal component analysis was performed in order to summarise the global variability of traits strongly related to lipid deposition and composition. Figure 1 displays the relative weight of each variable in the two first principal components, along with the relative position of animals selected for the HIGH and LOW groups. The first principal component explained 30.7% of total variability, whereas the second and following principal components accumulated less than 16% of phenotypic variability. The first principal component (Figure 1) grouped several fatness (fat thickness) measures and serum cholesterol values, together with intramuscular fat, saturated and monounsaturated fatty acids content in *gluteus medius*, while lean percentage and *gluteus medius* polyunsaturated fatty acids content were closely located at the other extreme of the axis. It is worth mentioning that intramuscular fat, monounsaturated and saturated fatty acids content were consistently grouped together in the three first principal components, in concordance with phenotypic correlations and physiological relationships among them [21], whereas polyunsaturated fatty acids always had an opposite relative weight. We finally used the first principal component as a classification index to rank all animals in the experiment (Table 1). Pigs at the higher and lower 10% of the ranking were used to generate the HIGH (n = 35) and LOW (n = 35) groups. For all traits, phenotypic means differed significantly between groups and also with regard of the population mean (Table 1). The highest differences were observed in traits directly related to fatness and muscular fat characteristics. At the top extreme of the ranking, the classification index positioned animals with increased subcutaneous and intramuscular fat deposition. These two traits are important in the dry-cured ham manufacturing process. Monounsaturated and polyunsaturated fatty acids percentages were weighted in opposite directions, thus favouring HIGH group displaying higher oleic and lower polyunsaturated fatty acid contents in the final product. This is an advantageous feature because it reduces the rancidity potential of meat.

Differential expression between HIGH and LOW groups

A total of 1060 probes were found differentially expressed between animals from the HIGH *vs* LOW groups at a nominal $p < 0.01$ and restricting the FDR to 0.05 (Additional file 2). A very high overall significance level of the experiment was obtained ($p < 10^{-7}$), confirming differences in the expression pattern of the *gluteus medius* muscle of these two groups of animals. From the 1060 probes, most of them (n = 839) were overexpressed in the HIGH group while only 221 showed higher expression levels in the LOW group. Expression fold-change between groups ranged from 0.38 (repressed in HIGH) to 5.58 (overexpressed in HIGH) and only 388 probes displayed more than 1.5-fold change between groups. The actual range of fold-change values depends on the data pre-processing and normalisation. Indeed, the most commonly used algorithms (MAS5, RMA and gcRMA) compress data in different ways, thus affecting the proportion

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

of large fold-changes between groups simply due to the greater or smaller ranges of expression intensities across replicate values [14]. In this sense, we used here the gcRMA algorithm which moderately compresses data, less than RMA but more than MAS5. Since fold-change ranges vary depending on the algorithm employed, establishing a cutting threshold of 1.5- or 2.0-fold as biologically significant is meaningless. Consequently, we used the complete list of differentially expressed probes in the subsequent analyses. From the list of 1060 probes, 10 were not annotated, 94 were assigned to EST or ORF sequences, and 956 were assigned to genes (corresponding to 836 single genes; Additional file 2). In a microarray, some genes are represented by several probes to evaluate the consistency of gene expression measurements. In our study, 442 single genes showed differential expression for the whole set of probes (100% of consistency across probes), while for the remaining 394 single genes, only a subset of the probes in the array were found to be differentially expressed between groups. The microarray data related to all samples have been deposited in the Gene Expression Omnibus (GEO) public repository (GEO accession number: GSE19275).

Ontological analysis of differentially expressed genes

GO term annotations were used to perform a functional categorisation of the resulting list of genes generated in the class comparison analysis. The three GO categories (biological process, metabolic function and cell component) were explored with DAVID bioinformatic tools for overrepresentation of specific GO terms. Significant results (at Benjamini-corrected $p < 0.1$) are summarised in Additional file 3. The most relevant results were obtained under the Biological Process category. Within this category, the top of the list gathered several terms related to "RNA processing", while spread over the list there were a variety of general "metabolic process" terms. Interestingly, four specific GO terms related to "lipid metabolism" appeared in the middle of the list, evidencing that a substantial fraction of genes were directly associated with traits under study.

Metabolic pathways and gene networks affected by the differentially expressed genes

The analysis of metabolic pathways and gene transactivation networks where the differentially expressed genes play a prominent role is essential to extract biological information and conclusions from a microarray experiment. In order to extract most information from our between-groups gene expression data, results from three exploration tools were compared.

First, the IPA software was used to survey the pathways overrepresented in our gene list. A total of 12 pathways were above the significance threshold (Figure 2A). Strikingly, the *Antigen presentation pathway* was the most significant route. In addition, several gene transactivation cascades involving the RXR factor were also included (Figure 2A). Besides, the *IGF-1 signalling pathway*, which determines the fat *vs* muscle growth, resulted very significant ($p < 0.001$) in this analysis. As regards the regulatory networks (Figure 2B), several lipid metabolism-, energy uptake- and muscle growth and function-related pathways were amongst the most relevant networks represented by the list of differentially expressed genes.

A second exploration tool, DAVID Functional Annotation tool, identified none of the BioCarta pathways as affected by the list of differentially expressed genes. In contrast, 12 KEGG pathways ($p < 0.1$; Table 2) were identified and, remarkably, eight of them played a relevant role in lipid metabolism or muscle and/or adipose tissue differentiation (arrows in Table 2).

A third method, Pathway Express [19], was used to explore the existence of signalling networks connecting differentially expressed genes. A total of 19 pathways were inferred from the list of genes with different mRNA levels in both groups ($p < 0.05$; Table 3). Among these, we found the 12 KEGG pathways previously described by DAVID's Functional Annotation tool but with important differences in the ranking of relevance. A KEGG diagram showing genes in each pathway is presented in the Additional file 4A to H. Many of the significant pathways were very relevant to the muscular and/or adipose physiology (arrows in Table 3). Like in IPA, the pathway with the highest impact factor was *Antigen presenting and processing*, followed then by *Phosphatidylinositol signalling* and *Biosynthesis of unsaturated fatty acids*. Regarding the latter, seven out of the 23 genes in this route, which encode enzymes directly involved in the elongation and desaturation of fatty acids, were differentially expressed between the HIGH and LOW groups (Table 3, Additional file 4C). Only one of them (*GPSN2* - a reductase that participates in the desaturation of long and very long chain fatty acids) is overexpressed in the LOW group whereas the other six have higher levels in the HIGH group. The *Insulin signalling* and *Type II diabetes* pathways (Additional file 4D and 4), which are deeply associated with glucose metabolism, were differentially expressed in both groups. Among genes that take part in (or are connected with) the insulin signalling cascade, one should emphasise the following groups, all of them expressed at higher levels in the HIGH group: (1) *mitogen-activated protein kinase* (*AMPK*), *acetyl-coA carboxylase α* (*ACACA*) and *fatty acid synthase* (*FASN*) genes that have a lipogenic role; (2) *PKA* and *hormone-sensitive lipase (LIPE*) genes which display a lipolytic action; (3) *TBC1 domain family, member 1* gene (*TBC1D1*), which plays a role in the translocation of the glucose transporter 4 (GLUT4) from the endoplasmic reticulum to the cell membrane in response

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

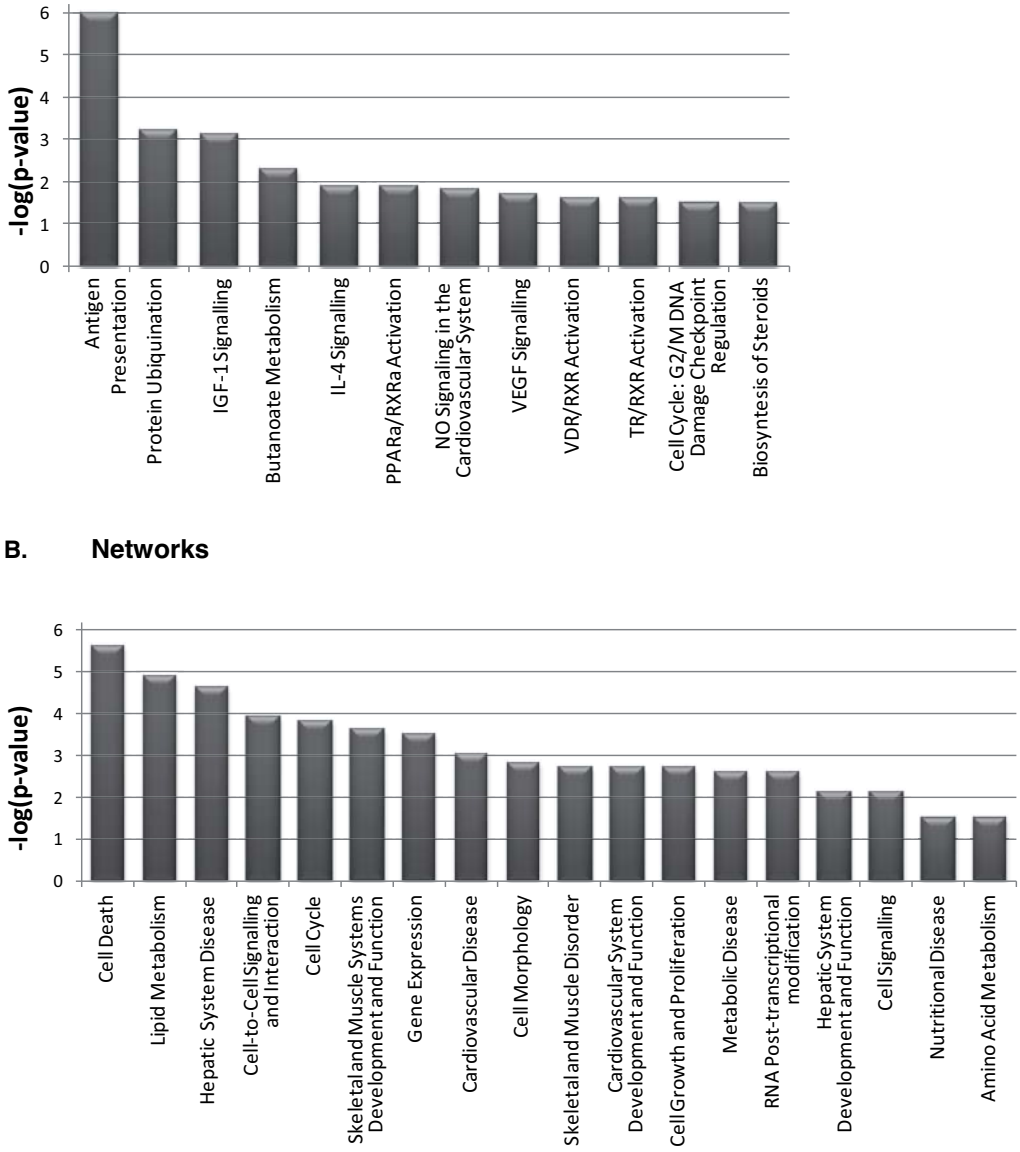


Figure 2 Functional categorisation analysis of the list of genes differentially expressed between the HIGH and LOW groups obtained with IPA. A: List of most significant pathways represented in the list of genes ($p < 0.01$). B: List of most significant networks affected by the joint effect of represented pathways ($p < 0.01$).

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

Table 2: DAVID analysis of pathways significantly enriched in the list of differentially expressed genes

KEGG id	KEGG pathway name	Genes (n)	Genes	Fold†	p-value*
?01040	Polyunsaturated fatty acids biosynthesis	7	*SCD, TECR, ELOVL5, ELOVL6, FASN, PTPLB, ACAA1*	8.3046	0.00001
05222	Small cell lung cancer	15	*LAMB1, PIK3CA, LAMA5, LAMA2, CYCS, BIRC3, MAX, PIK3R1, ITGB1, COL4A1, ITGAV, CKS1B, LAMC2, BIRC4, RXRG*	3.2728	0.0001
04612	Antigen processing and presentation	12	*HSPA2, HSPA4, HLA-A, PSME1, HSP90AA1, TAPBP, HLA-DRA, B2M, HLA-G, CD74, NFYA, HLA-DMB*	2.8473	0.0027
04510	Focal adhesion	21	*PIK3CA, LAMB1, LAMA5, VEGFA, LAMA2, RAP1A, COL5A3, BIRC3, PIK3R1, ITGB1, IGF1R, COL4A1, ITGB3, THBS4, ITGAV, CAV1, LAMC2, EGFR, BIRC4, ROCK1, ROCK2*	2.0031	0.0034
?04910	Insulin signalling pathway	15	*PIK3CA, ACACA, PPARGC1A, RHOQ, EIF4E, LIPE, PIK3R1, PRKAR2B, IRS2, PPP1R3B, FASN, SORBS1, PRKAB2, IRS1, PRKAG2*	2.1248	0.0097
?03320	PPAR signalling pathway	9	*SCD, OLR1, FABP4, PPARG, ADIPOQ, ACAA1, RXRG, PPARD, SORBS1*	2.3727	0.0334
?04512	ECM-receptor interaction	10	*LAMB1, LAMA5, THBS4, ITGAV, LAMA2, COL5A3, LAMC2, ITGB1, COL4A1, ITGB3*	2.1570	0.0396
?04070	Phosphatidylinositol signalling system	9	*PIK3CA, ITPR1, DGKZ, SYNJ1, PLCD4, FN3K, PIK3R1, PIK3C2A, DGKG*	2.2479	0.0444
04150	mTOR signalling pathway	7	*ULK2, PIK3CA, VEGFA, EIF4E, PIK3R1, ULK3, RICTOR*	2.6054	0.0488
?04930	Type II diabetes mellitus	6	*PIK3CA, IRS2, CACNA1C, ADIPOQ, PIK3R1, IRS1*	2.5309	0.0844
?04920	Adipocytokine signalling pathway	8	*IRS2, PPARGC1A, ADIPOQ, ADIPOR2, RXRG, IRS1, PRKAB2, PRKAG2*	2.0802	0.0862
?04350	TGF-beta signalling pathway	9	*BMP5, FST, THBS4, CUL1, ZFYVE16, TGFB2, ROCK1, SMAD5, ROCK2*	1.8982	0.0978

†Fold enrichment of the KEGG pathways in the list of differentially expressed genes with regard to the background dataset; *p-value from Fisher's exact test. Arrows indicate pathways relevant to muscle/adipose tissue function.

to the insulin signal; and (4) *peroxisome proliferator-activated receptor-γ coactivator 1 gene* (*PPARGC1A*), an adapter of the PPAR family of nuclear receptors, which takes part in the transactivation of effector genes. In addition, it is worth to mention the differential expression of the *insulin receptor substrate 2* (*IRS2*) gene, first effector downstream the insulin receptor. The effect on this substrate is, therefore, directly transmitted to the rest of the cascade.

The PPAR signalling pathway is one of the most relevant routes during the process of adipocyte tissue development, differentiation and activation of lipogenesis (Additional file 4F). In the HIGH group we observed increased mRNA levels of peroxisome proliferator-activated receptor δ (PPARD), γ (PPARG) and PPARGC1 while retinoid × receptor γ (RXRG) mRNA was downregulated. Moreover, the fatty acid binding protein 4 (FABP4) gene, a key player in the uptake of fatty acids, exhibited higher mRNA levels in pigs of the HIGH group. We also observed changes in the Adipocytokine signalling pathway (Additional file 4G), particularly the mRNA levels for ADIPOQ and adiponectin receptor 2 (ADIPOR2) were overexpressed in the HIGH group (Additional file 4G) alongside with AMPK, RXRG, PPARGC1A which indirectly affect the oxidation and degradation of lipids in mitochondria through the carnitine palmitoyltransferase 1 gene (CPT1).

Two other pathways related to myocyte growth and differentiation happened to be affected by the genes differentially expressed between the HIGH and LOW groups. First, the *Extracellular matrix interaction* pathway (Additional file 4H), which consists of a complex mixture of structural and functional macromolecules (integrins, proteoglycans, CD36, and other cell-surface-associated components) with an important role in tissue and organ morphogenesis and function [22]. From the genes that take part in this pathway, we have identified expression changes in *α* and *β integrins*, *collagen* and *transforming growth factor β* (*TGFβ*) *receptor* genes. The second pathway of interest was the *TGFβ signalling* route (Additional file 4I) that participates in many cellular processes like cell differentiation, proliferation and apoptosis. The

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

TGFβ factor differentially expressed in the HIGH group was *TFGB2*, whose role in the muscular or adipose tissue development is not currently known.

Quantitative PCR validation of microarray results

The information relative to the gene ontology and pathways analysis was used to select 31 genes with a relevant function in muscle and adipose physiology and lipid metabolism (Additional file 5). The selection was also based on literature and database searches for each individual gene from the resulting a 879-genes list. In this way, genes functionally related to muscle and fat cell metabolism were selected to be validated by qPCR in a subset of 10 animals from each group. In order to confirm the probe annotation information, we first conducted a BLAST search with the sequences used by Affymetrix to generate the microarray probes. Three of the 31 probes analysed were wrongly annotated and were, therefore, excluded from further analysis. Efficient qPCR assays were established for 25 out of the 28 remaining genes. For all 25 genes (Table 4), the fold-change ratios between HIGH and LOW groups were consistent in both assays except for *TBC1D1* and *RXRG* genes. In addition, in most cases the expression ratios were higher in the qPCR experiment than in microarray based analyses (Figure 3) [23]. A high level of correlation (ranging from 0.48 to 0.96; $p < 0.05$) was observed between qPCR and Affymetrix microarray data for all genes except for *TBC1D1* and *RXRG* (Table 4).

All together, validation of microarray data by qPCR revealed a high correspondence between both analyses, confirming differential expression for 19 out of 25 genes tested ($p < 0.05$). The significance level of the differences between HIGH and LOW groups tended to be greater in the microarray experiment, which might be due to the higher background noise of the qPCR assay.

Discussion

The ultimate goal of selection is to fix a set of genotypes that are advantageous from a production, economic or nutritional point of view [24]. This might involve significant changes in the physiology and metabolism of individuals under selection [25]. In pigs, traditional selection has been focused to obtain leaner individuals but this practice has had detrimental consequences on meat quality due to a reduction in intramuscular fat and a parallel increase in polyunsaturated fatty acids [26], which are highly susceptible to lipid oxidation [1]. Genetic and phenotypic correlations suggest that selection for higher intramuscular fat (as a way to improve the palatability and juiciness of meat) might increase fatness, serum lipids and meat monounsaturated fatty acids and saturated fatty acids contents [3]. In summary, selection for increasing either leanness or intramuscular fat might lead to two extreme phenotypes with divergent muscular metabolic profiles. We have been able to capture these two phenotypic classes (HIGH and LOW groups) in a Duroc commercial population that had been selected aiming to increase intramuscular fat while maintaining carcass fatness. The utilization of a PCA approach seems to have been particularly efficient in detecting individuals that are different for a global multitrait phenotype rather than for a particular trait, as supported by the high overall significance achieved in the differential expression experiment. In the next sections we will discuss changes of muscle gene expression that are associated with each one of these composite phenotypes and their physiological and metabolic implications.

Lipid metabolism

We have observed significant changes in the expression of genes and pathways related to lipid metabolism (Tables 4 and Additional file 1). Fatty acid uptake genes such as *FABP4* and *Apolipoprotein E* (*APOE*) were upregulated in the HIGH group. Fatty acid-binding proteins are small cytosolic molecules that bind saturated and unsaturated fatty acids. Mice where *FABP4* expression has been abolished are less susceptible to develop insulin resistance and dislipidemias [27]. APOE remove chylomicron remnants from circulation by interacting with low density lipoprotein receptor (LDLR), LDLR-related protein 1 (LRP1) and other receptors thus favouring lipid deposition [28]. In fact, *APOE* knockout mice have less body fat stores, smaller adipocytes and they are more resistant to diet-induced obesity [28]. This increased expression of fatty acid uptake molecules might represent a compensatory response to the higher levels of serum lipids in the HIGH group (Table 1). *De novo* lipogenesis genes (*i.e. ACACA*, *FASN*) were also upregulated (Table 4 and Additional file 1). Malonyl-CoA, a powerful inhibitor of the carnitine/palmitoyl shuttle system for fatty acid β-oxidation and the substrate of FASN [29], is synthesised by ACACA, so increased expression of this enzyme might lead to an accumulation of muscle fat stores [29]. Two other genes promoting lipid deposition, *stearoyl CoA desaturase* (*SCD*) and *diacylglycerol O-acyltransferase 2* (*DGAT2*), displayed an increased expression in the HIGH group (Tables 4 and Additional file 1). *SCD* plays a key role in the synthesis of monounsaturated fatty acids which, in turn, regulate other fundamental components of lipid metabolism such as sterol regulatory element-binding protein (SREBP) and MAPK. In human, a high SCD activity is associated with a metabolic state favoring hepatic triglyceride accumulation and expansion of adipose triglyceride stores [30]. In consequence, a higher *SCD* expression in the muscle might partition fatty acids towards storage rather than to oxidation [30]. DGAT2 is the main enzyme involved in the synthesis of triglycerides

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

Table 3: Pathway Express analysis of metabolic pathways and gene networks significantly affected by differentially expressed genes

KEGG pathway-id	KEGG pathway name	IF†	input/pathway genes	pathway genes on chip	p-value*
04612	Antigen processing and presentation	61.066	15/88	45	1.87E-25
?04070	Phosphatidylinositol signalling system	18.350	10/77	60	2.08E-07
?01040	Biosynthesis of unsaturated fatty acids	10.791	7/23	18	2.43E-04
05222	Small cell lung cancer	8.686	14/87	71	1.64E-03
04510	Focal adhesion	7.117	20/199	169	6.58E-03
05332	Graft-versus-host disease	6.966	6/42	22	7.52E-03
?04910	Insulin signalling pathway	6.430	16/138	114	1.20E-02
05310	Asthma	6.323	5/30	17	1.31E-02
04520	Adherents junction	5.957	6/75	69	1.80E-02
04514	Cell adhesion molecules (CAMs)	5.952	13/133	95	1.81E-02
05330	Allograft rejection	5.813	6/38	26	2.04E-02
?04930	Type II diabetes mellitus	5.556	6/44	33	2.53E-02
?04512	ECM-receptor interaction	5.211	10/87	72	3.39E-02
05320	Autoimmune thyroid disease	4.955	6/53	31	4.20E-02
04940	Type I diabetes mellitus	4.955	6/44	31	4.20E-02
04150	mTOR signalling pathway	4.932	7/51	42	4.28E-02
?04920	Adipocytokine signalling pathway	4.779	8/72	64	4.66E-02
?03320	PPAR signalling pathway	4.684	9/69	61	4.75E-02
?04350	TGF-beta signalling pathway	4.662	9/89	74	4.95E-02

†Impact Factor; *p-value from Fisher's exact test. Arrows indicate pathways relevant to muscle/adipose tissue function.

and its increased expression might also favour lipid storage [31]. Similarly, CCAAT/enhancer-binding protein δ (CEBPD) and PPARG, whose expression is upregulated in HIGH pigs, are transcription factors which promote fat deposition [32,33]. This pattern of gene expression suggests that the increased lipid deposition of HIGH pigs is the result of several metabolic processes rather than the consequence of a single biochemical event.

There is substantial evidence that the aforementioned genes are differentially expressed in adipose tissue cells and myocytes (Novartis Gene Expression Atlas, http://biogps.gnf.org/#goto=welcome), while their status in intramuscular adipocytes remains largely unknown [34]. An obvious and unavoidable limitation of our experiment was that we were unable to separate myocytes and intramuscular adipocytes before isolating RNA. The influence of this feature on our expression data is difficult to evaluate. Intramuscular fat (LOW: 3.5%, HIGH: 7.5%) comes from two different sources: small adipocytes localized between muscular fibres and lipid droplets contained within myocytes [34]. This means that the percentage of intramuscular adipocytes cannot be simply calculated based in intramuscular fat content. Moreover, gene expression profile of intramuscular adipocytes is mostly unknown, although there is compelling evidence that in pigs is clearly different from that of subcutaneous and perirenal adipocytes [34,35]. For instance, *FASN*, *PPARG* and *LIPE* mRNAs are 600, 200 and 800 times less expressed in intramuscular than in subcutaneous adipocytes of 160-days-old pigs [34]. Proteomic studies have also evidenced that lipogenic, lipolytic, glycolytic and fatty acid oxidation pathways are clearly downregulated in intramuscular adipocytes when compared with fat cells from other depots [35]. These results suggest that differences in the proportion of intramuscular adipocytes between HIGH and LOW groups might have a minimum impact on the patterns of muscle gene expression. This interpretation is strongly supported by the fact that gene expression changes detected in our experiment are very consistent with a plethora of data obtained in human indicating that fatty acid uptake, lipogenesis and triglyceride synthesis genes are upregulated in the muscle cells of obese individuals [5].

Fat accumulation in the skeletal muscle might be quite harmful for this tissue as well as for the whole organism [6,36]. Obese sedentary individuals have an increased rate of lipid peroxidation and higher levels of certain lipid metabolites, such as diacylglycerol, ceramides and long-

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

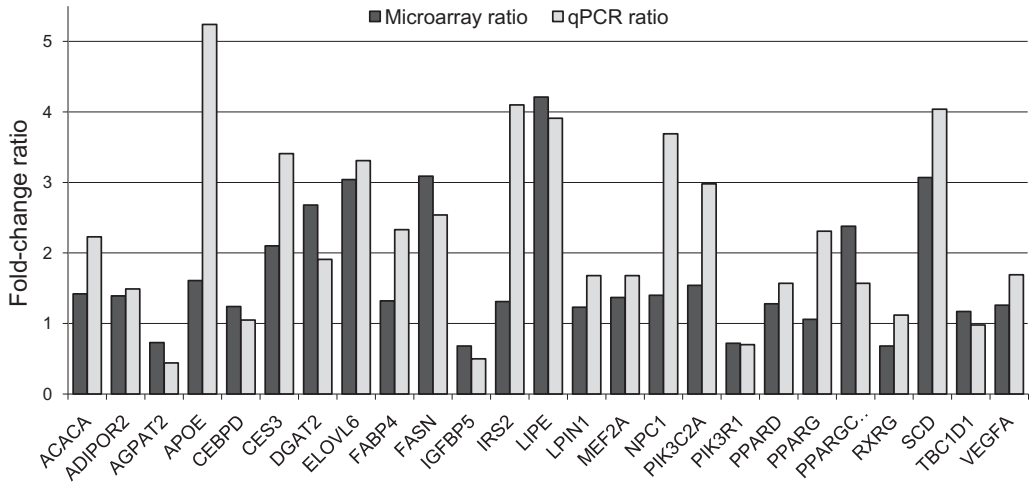


Figure 3 Comparison of microarray and qPCR expression ratios (HIGH vs LOW groups) for the list of selected genes.

chain fatty acids in the myocyte cytoplasm [6]. Both features are well known to promote the development of insulin resistance [6]. In this sense, animals in the HIGH group had upregulated levels of *lipin 1* (*LPIN1*) mRNA, which induces the synthesis of diacylglycerol from phosphatidic acid [37,38], and of *elongation of very long chain fatty acids-like 5* and *6* (*ELOV5, ELOVL6*) mRNA which promote the elongation of fatty acids [39]. These genes might increase the production of metabolites that are well known to have negative effects on insulin metabolism.

In this context, it might be argued that selection for increasing intramuscular fat content might have detrimental consequences on glucose and lipid homeostasis in the skeletal muscle. However, we have obtained substantial evidence of the activation of pathways which might counteract the accumulation of intracellular lipids (Table 4 and Additional file 1). In this way, increased *LIPE* and *CES3* mRNA expression were detected in the HIGH group. These two lipases are deeply involved in the hydrolysis of triacylglycerols [37]. The remarkable expression differences observed in the HIGH and LOW groups (about 3-4 fold in the qPCR analysis) cannot be explained in terms of adipocyte abundance, but as the consequence of a physiological change in muscle cells. In this regard, the simultaneous upregulation of genes with lipogenic (*ACACA*, *FASN*, *SCD*, *DGAT2*) and lipolytic (*LIPE*, *CES3*) actions would involve the existence of a futile cycle where triacylglycerols are continuously degraded and resynthesised. It has been suggested that, ultimately, intramuscular fat accumulation results from a balance between uptake, synthesis and degradation of triglycerides, rather than the upregulation of a single pathway [40]. Our results seem to point out also in this direction. Moreover, the upregulated *LIPE* and *CES3* mRNA expression might have a protective effect in the muscle cell by degrading intracellular lipids (such as diacylglycerol) which might otherwise inhibit insulin signalling. Another line of defence against lipid overload would be featured by *ADIPOQ*, *PPARD* and *PPARGC1* mRNAs, which are upregulated in the HIGH group and promote the catabolism of fatty acids. *PPARGC1A* is an adapter of nuclear receptors which plays a central role in the modulation of gene transactivation in several signalling routes, such as insulin, adipokines and bioactive lipids [41]. In the skeletal muscle, *PPARGC1A* avoids the development of insulin-resistance by increasing the proportion of oxidative fibres and stimulating mitochondrial biogenesis, two features that promote fatty acid oxidation [42]. PPARD, a transcription factor primarily expressed in muscle cells, also enhances fatty acid oxidation and energy expenditure [33]. Finally, adiponectin binding to its receptor also stimulates fatty acid oxidation and decreases triglyceride storage in muscle, which may explain, in part, the insulin-sensitising effect of this hormone [43]. This process relies on the activation of AMPK, a master regulator of cellular energy balance [43]. Activation of AMPK, in turn, depends partly on a phospholipase C/$Ca2^+$ $Ca2^+$/Calmodulin-dependent protein-kinase kinase-dependent pathway [44], a feature that agrees well with our observations that *phospholipase C* and *calmodulin 2* mRNA are upregulated in the HIGH group. Taken together, these results suggest that genes activating the catabolism of fatty acids are upregulated in

Cánovas *et al. BMC Genomics* 2010, **11**:372
http://www.biomedcentral.com/1471-2164/11/372

Table 4: Validation of microarray expression data by quantitative real-time PCR for 25 selected genes

GENE	MICROARRAY		qPCR		Correlation	
	n = 68(34H *vs* 34L)		n = 20(10H *vs* 10L)			
	Ratio†	p-value	Ratio†	p-value	r*	p-value
ACACA	**1.42**	**0.003**	**2.23**	**0.047**	**0.77**	**0.0002**
ADIPOR2	1.39	0.001	1.49	0.354	0.59	0.0108
AGPAT2	**0.73**	**0.01**	**0.44**	**0.043**	**0.60**	**0.0099**
APOE	**1.61**	**0.001**	**5.24**	**0.017**	**0.73**	**0.0006**
CEBPD	1.24	0.004	1.05	0.078	0.69	0.0016
CES3	**2.1**	**0.000003**	**3.41**	**0.008**	**0.83**	**< 0.0001**
DGAT2	**2.68**	**0.001**	**1.91**	**0.048**	**0.83**	**< 0.0001**
ELOVL6	**3.04**	**0.005**	**3.31**	**0.042**	**0.88**	**< 0.0001**
FABP4	**1.32**	**0.0002**	**2.33**	**0.037**	**0.92**	**< 0.0001**
FASN	**3.09**	**0.0008**	**2.54**	**0.046**	**0.88**	**< 0.0001**
IGFBP5	**0.68**	**0.0007**	**0.50**	**0.025**	**0.83**	**< 0.0001**
IRS2	**1.31**	**0.001**	**4.1**	**0.012**	**0.53**	**0.0231**
LIPE	**4.21**	**0.000002**	**3.91**	**0.001**	**0.76**	**0.0002**
LPIN1	**1.23**	**0.008**	**1.68**	**0.005**	**0.68**	**0.0018**
MEF2A	**1.37**	**0.0002**	**1.68**	**0.037**	**0.63**	**0.0051**
NPC1	**1.4**	**0.00006**	**3.69**	**0.018**	**0.48**	**0.0437**
PIK3C2A	**1.54**	**0.009**	**2.98**	**0.021**	**0.68**	**0.0020**
PIK3R1	**0.72**	**0.01**	**0.70**	**0.020**	**0.48**	**0.0424**
PPARD	**1.28**	**0.008**	**1.57**	**0.021**	**0.53**	**0.0248**
PPARG	**1.06**	**0.005**	**2.31**	**0.041**	**0.69**	**0.0016**
PPARGC1A	2.38	0.004	1.57	0.152	0.96	< 0.0001
RXRG	0.68	0.002	1.12	0.621	-0.26	0.2911
SCD	3.07	0.004	4.04	0.156	0.96	< 0.0001
TBC1D1	1.17	0.0002	0.98	0.552	-0.16	0.5287
VEGFA	**1.26**	**0.009**	**1.69**	**0.032**	**0.48**	**0.0444**

Expression ratios and p-values are indicated for both assays and correlation coefficients between qPCR and microarray results are also shown. H = HIGH group; L = LOW group; †Expression ratios between H and L groups; *correlation coefficients between qPCR and microarray results; in **bold**, genes whose difference in expression between groups has been validated by qPCR. Arrows up and down indicate gene overexpression and downregulation in the HIGH group, respectively.

the HIGH pìgs, although we have not found direct evidence that genes integrated in the *β-oxidation* pathway are differentially expressed in the HIGH and LOW groups. Probably, ADIPOQ and PPARD exert their action by introducing post-translational modifications in the β-oxidation enzymes rather than modifying their levels of expression. Finally, it is worth mentioning that our findings do not agree with the observation that, in human, fatty acid oxidation is reduced in obese individuals [5].

Insulin signalling and glucose metabolism

We have detected differential expression of genes related to the insulin signalling pathway between the HIGH and LOW groups. *MADS-box transcription enhancer factor 2 polypeptide A* (*MEF2A*, overexpressed in the HIGH group) is abundantly expressed in muscle cells and it binds to the promoter of the *GLUT4* gene upregulating its expression and glucose uptake [45]. Interestingly, *insulin receptor substrate 2* (*IRS2*), one of the two major insulin receptor substrates, was also upregulated in the HIGH group. Impaired function of IRS2 has been related to a defective activation of phosphoinositide 3-kinase (PIK3), a reduction in glucose uptake and insulin resistance in certain studies [46] but not in others [47]. It is also worth mentioning that we have found that *PIK3 regulatory subunit 1* (*PIK3R1*) and *PIK3 class 2 alpha polypeptide*

(*PIK3C2A*) mRNAs are down- and upregulated, respectively. Free, monomeric PIK3R1 can act as a negative regulator of PIK3 signalling downstream the insulin and IGF1 receptors by competing with the PIK3R1/PIK3C heterodimer for binding to tyrosine-phosphorylated IRS1 [48], while PIK3C2A is a subunit of PIK3. As a whole, pigs of the HIGH group, which are fatter than their LOW counterparts, show an mRNA profile compatible with an increased glucose uptake. This does not necessarily mean that the functionality of the molecules involved in this pathway is identical in both groups, because we have not evaluated post-translational changes, which are known to have a critical impact on their activity [49]. At most, we can affirm that glucose uptake and insulin signalling pathways are not downregulated in the *gluteus medius* muscle of HIGH pigs. This observation agrees very well with the fact that *de novo* lipogenesis, which can use glucose as a substrate, is upregulated in this group.

Immunity

The most striking difference amongst the HIGH and LOW groups was that expression of antigen-presenting molecules was downregulated in the former (Additional file 4A). Interestingly, a significant reduction in the expression levels of genes related to the immunity in muscles from chickens selected for fat content has also been described in other microarray studies [50]. A possible mechanism explaining our results would be related to the increased expression of the *adiponectin receptor 2* in the muscle cells of HIGH pigs, which increases the synthesis of the antiinflammatory cytokine interleukin (IL) 10, inhibits the synthesis of proinflammatory cytokines, such as IL1β, IL6 and tumor necrosis factor (TNF), and activates AMPK, a kinase that can disrupt *NFκB* transcriptional activation, resulting in downregulation of *MHC* expression [51]. *Stearoyl CoA desaturase* overexpression has also been reported to have antiinflammatory activity through the conversion of palmitate to oleate [52].

Muscle growth and differentiation

We have observed an upregulation of the *TGFB2* and *TGFBR3* mRNAs in the HIGH group (Additional file 2 and 4I). The TGFβ signalling pathway participates in many cellular processes like cell differentiation, proliferation and apoptosis. In the skeletal muscle, TGFβ is a powerful inhibitor of the muscular proliferation and regeneration from satellite cells [53]. For instance, myostatin, which is upregulated by TGFB1, inhibits muscular proliferation [54]. Inactivation of this gene produces an excessive growth of the muscular tissue that results in the "double muscle" phenotype observed in Blue Belgian or Asturiana cattle [55]. Moreover, transforming growth factor β molecules are positive regulators of extracellular matrix proteins expression [56]. In this sense, we have detected increased mRNA levels of *integrins α5* (*ITGAV*) and *β1* (*ITGB1*) as well as of several collagen molecules (*COL4A5*, *COL12A1*, *COL4A1*, *COL5A3*). The ability of satellite cells to successfully participate in growth and repair of the skeletal muscle is influenced heavily by surrounding extracellular matrix factors [56]. In this way, the formation of multinucleated myotubes requires the migration of myoblasts, a process regulated by integrins that ensure the adhesion of these precursor cells to the extracellular matrix [56]. Finally, it is worth mentioning that we have detected increased mRNA levels of myogenic factors such as *insulin growth factor-binding protein 5* (*IGFBP5*) and *ADIPOR2*. These findings might be biologically significant because there are evidences that some level of postnatal muscle growth, relying on the activation, differentiation and proliferation of precursor satellite cells, takes place in pigs [57].

Conclusions

We have examined the mRNA expression profile of muscle samples obtained from pigs with extreme phenotypes for several fatness parameters. The HIGH phenotype corresponded to pigs with high levels of intramuscular fat, serum lipids and muscle saturated and monounsaturated fatty acid content, while LOW pigs were leaner and showed increased levels of muscle polyunsaturated fatty acids. This approach allowed us to detect striking differences in the expression profile of both groups, with the HIGH group showing upregulated mRNA levels of genes related to lipogenesis, lipolysis and glucose uptake and decreased levels of mRNAs encoding antigen-presenting molecules. As a whole, our data suggest that selection for increasing intramuscular fat content in pigs would not lead to a disruption of the metabolic homeostasis of muscle cells. An evident limitation of our experiment is that mRNA concentrations are not necessarily a faithful reflection of protein levels or activities. In this regard, Maier et al. [58] have reported that correlations between mRNA and protein levels are moderate (r_P = 0.36) to high (r_P = 0.76), being influenced by gene-specific (protein half-life, translation efficiency etc.) and technical factors. In consequence data about post-translational changes affecting protein activity or expression as well as information about protein location within the cell would be needed to fully understand how lipid deposition affects muscle physiology in pigs [57].

Additional material

Additional file 1 Sequence of primer sets used to validate microarray results of selected genes by quantitative real-time PCR. Gene symbol and GenBank accession number for each pig gene are indicated. Putative exon location is based on human gene structure information.

Additional file 2 List and annotation of differentially expressed probes between HIGH and LOW groups. List and annotation of the 1060 probes differentially expressed between *gluteus medius* muscle of HIGH and LOW pigs. Parametric, permutation p-values and 5% FDR are indicated. Means and medians of $\log_2$ intensities are used to calculate ratios of fold-change expression.

Additional file 3 Gene ontology terms. List of significantly overrepresented gene ontology terms (distributed over the three main family categories) in the list of 1060 probes differentially expressed between *gluteus medius* muscle samples of HIGH and LOW pigs. Parametric p-values are indicated along with Bonferroni, Benjamini and 5% FDR multiple test corrections.

Additional file 4 Graphical representation of KEGG diagrams. Graphical representation of microarray results over the nine KEGG pathways significantly affected by the genes differentially expressed between the HIGH and LOW groups ($p < 0.05$). A: Antigen Processing and Presentation; B: *Phosphatidylinositol Signalling*; C: *Biosynthesis of Unsaturated Fatty Acids*; *D: Insulin Signalling*; E: *Type II Diabetes Mellitus*; F: *PPAR Signalling*; G: *Adipocytokine Signalling*; H: *ECM-receptor Interaction*; I: *TGF-beta Signalling*. In red, genes overexpressed in the HIGH group. In blue, genes overexpressed in the LOW group.

Additional file 5 List of genes selected for real time qPCR validation. Genes were selected for their involvement in lipid metabolism or muscle/fat development. A brief summary describing gene function and involved in KEGG pathways is provided.

Authors' contributions

AC participated in the RNA isolation, initial microarray analysis and validation of the results by qPCR. RQ developed the experimental design and performed the statistical analysis. MA participated in the analysis and interpretation of the data and helped drafting the manuscript. RNP participated in RNA isolation, microarray analysis and ontology analysis. All authors have been involved in drafting and revising the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We are thankful to *Selección Batallé S.A.* for providing the animal material, to J Reixach, JL Noguera and D Almuzara for their cooperation in the experimental protocol, and to I Díaz (IRTA, Tecnologia dels Aliments), JM Prat-Cuffi (Hospital de Palamós) and the Facilities in the Vall d'Hebron Research Centre for their technical support. A Cánovas received a predoctoral scholarship from INIA. RN Pena received a contractual grant from INIA. This project was financed by the Spanish Ministry of Education and Science (projects GEN2003-20658-C05-05 and AGL2007-66707-C02-01).

Author Details

[1]IRTA, Genètica i Millora Animal, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain and [2]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Received: 9 December 2009 Accepted: 11 June 2010
Published: 11 June 2010

References

1. Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, Hughes SI, Whittington FM: **Fat deposition, fatty acid composition and meat quality: A review.** *Meat Science* 2008, **78:**343-358.
2. Candek-Potokar M, Monin G, Zlender B: **Pork quality, processing, and sensory characteristics of dry-cured hams as influenced by Duroc crossing and sex.** *J Anim Sci* 2002, **80:**988-996.
3. Huff-Lonergan E, Baas TJ, Malek M, Dekkers JC, Prusa K, Rothschild MF: **Correlations among selected pork quality traits.** *J Anim Sci* 2002, **80:**617-627.
4. Ordovas JM: **Genetic interactions with diet influence the risk of cardiovascular disease.** *Am J Clin Nutr* 2006, **83:**443S-446S.
5. Kiens B: **Skeletal muscle lipid metabolism in exercise and insulin resistance.** *Physiol Rev* 2006, **86:**205-243.
6. Moro C, Pillard F, de Glisezinski I, Klimcakova E, Crampes F, Thalamas C, Harant I, Marques M, Lafontan M, Berlan M: **Exercise-induced lipid mobilization in subcutaneous adipose tissue is mainly related to natriuretic peptides in overweight men.** *Am J Physiol Endocrinol Metab* 2008, **295:**E505-513.
7. Gallardo D, Pena RN, Amills M, Varona L, Ramirez O, Reixach J, Diaz I, Tibau J, Soler J, Prat-Cuffi JM, *et al.*: **Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs.** *Physiological Genomics* 2008, **35:**199-209.
8. Gallardo D, Quintanilla R, Varona L, Díaz I, Ramirez O, Pena R, Amills M: **Polymorphism of the pig acetyl coenzyme A carboxylase α (ACACA) gene is associated with fatty acid composition in a Duroc commercial line.** *Anim Genet* 2009, **40:**410-417.
9. Chomczynski P, Sacchi N: **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** *Analytical biochemistry* 1987, **162:**156-159.
10. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al.*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5:**R80.
11. Wilson CL, Miller CJ: **Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis.** *Bioinformatics* 2005, **21:**3683-3685.
12. Xu X, Zhao Y, Simon R: **Gene Set Expression Comparison kit for BRB-ArrayTools.** *Bioinformatics* 2008, **24:**137-139.
13. Wu Z, Airizarry R, Gentleman R, Martinez-Murillo F, Spencer F: **A Model-Based Background Adjustment for Oligonucleotide Expression Arrays.** *Journal of the American Statistical Association* 2004, **99:**909-917.
14. Murie C, Woody O, Lee AY, Nadon R: **Comparison of small n statistical tests of differential expression applied to microarrays.** *BMC Bioinformatics* 2009, **10(45):**41-45. 18
15. Pavlidis P, Qin J, Arango V, Mann JJ, Sibille E: **Using the gene ontology for microarray data mining: a comparison of methods and application to age effects in human prefrontal cortex.** *Neurochem Res* 2004, **29:**1213-1222.
16. Benjamini Y, Yekutieli D: **The control of the false discovery rate in multiple testing under dependency.** *Annals of Statistics* 2001, **29:**1165-1188.
17. Wang Y, Couture OP, Qu L, Uthe JJ, Bearson SM, Kuhar D, Lunney JK, Nettleton D, Dekkers JC, Tuggle CK: **Analysis of porcine transcriptional response to Salmonella enterica serovar Choleraesuis suggests novel targets of NFkappaB are activated in the mesenteric lymph node.** *BMC Genomics* 2008, **9:**437.
18. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4:**44-57.
19. Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R: **A systems biology approach for pathway level analysis.** *Genome Res* 2007, **17:**1537-1545.
20. Yuan JS, Reed A, Chen F, Stewart CN Jr: **Statistical analysis of real-time PCR data.** *BMC Bioinformatics* 2006, **7:**85.
21. Solanes FX, Reixach J, Tor M, Tibau J, Estany J: **Genetic correlations and expected response for intramuscular fat content in a Duroc pig line.** *Livestock Science* 2009, **123:**63-69.
22. van der Flier A, Sonnenberg A: **Function and interactions of integrins.** *Cell Tissue Res* 2001, **305:**285-298.
23. Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C, Hunkapiller K, Jensen RV, Knight CR, Lee KY, *et al.*: **Evaluation of DNA microarray results with quantitative gene expression platforms.** *Nat Biotechnol* 2006, **24:**1115-1122.
24. Falconer D, Mackay T: **Introduction to Quantitative Genetics.** 3rd edition. Harlow, Essex, UK: Longmans Green; 1996.
25. Jones P, Lichtenstein A, Schaefer E, Namchuk G: **Effect of dietary fat selection on plasma cholesterol synthesis in older, moderately hypercholesterolemic humans.** *Arterioscler Thromb* 1994, **14:**542-548.
26. Tibau J: **Genetic improvement of production traits in pigs.** *Porci* 1992, **7:**11-21.
27. Makowski L, Hotamisligil G: **Fatty acid binding proteins--the evolutionary crossroads of inflammatory and metabolic responses.** *J Nutr* 2004, **134:**2464S-2468S.
28. Kypreos K: **Mechanisms of obesity and related pathologies.** *FEBS J* 2009, **276:**5719.
29. Wakil S, Abu-Elheiga L: **Fatty acid metabolism: target for metabolic syndrome.** *J Lipid Res* 2009, **50(Suppl):**S138-143.
30. Flowers MT, Ntambi JM: **Role of stearoyl-coenzyme A desaturase in regulating lipid metabolism.** *Curr Opin Lipidol* 2008, **19:**248-256.

31. Yen C, Stone S, Koliwad S, Harris C, Farese RJ: **Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis.** *J Lipid Res* 2008, **49:**2283-2301.
32. Morrison RF, Farmer SR: **Hormonal signaling and transcriptional control of adipocyte differentiation.** *J Nutr* 2000, **130:**3116S-3121S.
33. Tontonoz P, Spiegelman BM: **Fat and beyond: the diverse biology of PPARgamma.** *Annu Rev Biochem* 2008, **77:**289-312.
34. Gardan D, Gondret F, Louveau I: **Lipid metabolism and secretory function of porcine intramuscular adipocytes compared with subcutaneous and perirenal adipocytes.** *Am J Physiol Endocrinol Metab* 2006, **291:**E372-380.
35. Gardan D, Gondret F, Van den Maagdenberg K, Buys N, De Smet S, Louveau I: **Lipid metabolism and cellular features of skeletal muscle and subcutaneous adipose tissue in pigs differing in IGF-II genotype.** *Domest Anim Endocrinol* 2008, **34:**45-53.
36. Blaak E: **Metabolic fluxes in skeletal muscle in relation to obesity and insulin resistance.** *Best Pract Res Clin Endocrinol Metab* 2005, **19:**391-403.
37. Brasaemle DL: **Thematic review series: adipocyte biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis.** *J Lipid Res* 2007, **48:**2547-2559.
38. He XP, Xu XW, Zhao SH, Fan B, Yu M, Zhu MJ, Li CC, Peng ZZ, Liu B: **Investigation of Lpin1 as a candidate gene for fat deposition in pigs.** *Mol Biol Rep* 2009, **36:**1175-1180.
39. Moon Y, Shah N, Mohapatra S, Warrington J, Horton J: **Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins.** *J Biol Chem* 2001, **276:**45358-45366.
40. Hocquette JF, Gondret F, Baeza E, Médale F, Jurie C, Pethick DW: **Biological Markers of Intramuscular Fat (IMF) Content.** *Proceedings of the 62nd American Meat Science Association: 21-24 June 2009; Rogers, Arkansas (USA); American Meat Science Association* 2009:1-5.
41. Desvergne B, Michalik L, Wahli W: **Transcriptional regulation of metabolism.** *Physiol Rev* 2006, **86:**465-514.
42. Bonen A: **PGC-1alpha-induced improvements in skeletal muscle metabolism and insulin sensitivity.** *Appl Physiol Nutr Metab* 2009, **34:**307-314.
43. Dyck DJ, Heigenhauser GJ, Bruce CR: **The role of adipokines as regulators of skeletal muscle fatty acid metabolism and insulin sensitivity.** *Acta Physiol (Oxf)* 2006, **186:**5-16.
44. Zhou L, Deepa S, Etzler J, Ryu J, Mao X, Fang Q, Liu D, Torres J, Jia W, Lechleiter J, *et al.*: **Adiponectin activates AMP-activated protein kinase in muscle cells via APPL1/LKB1-dependent and phospholipase C/ Ca2+/Ca2+/calmodulin-dependent protein kinase kinase-dependent pathways.** *J Biol Chem* 2009, **284:**22426-22435.
45. Im S, Kwon S, Kim T, Kim H, Ahn Y: **Regulation of glucose transporter type 4 isoform gene expression in muscle and adipocytes.** *IUBMB Life* 2007, **59:**134-145.
46. Vollenweider P, Ménard B, Nicod P: **Insulin resistance, defective insulin receptor substrate 2-associated phosphatidylinositol-3' kinase activation, and impaired atypical protein kinase C (zeta/lambda) activation in myotubes from obese patients with impaired glucose tolerance.** *Diabetes* 2002, **51:**1052-1059.
47. Huang C, Thirone A, Huang X, Klip A: **Differential contribution of insulin receptor substrates 1 versus 2 to insulin signaling and glucose uptake in l6 myotubes.** *J Biol Chem* 2005, **280:**19426-19435.
48. Engelman J, Luo J, Cantley L: **The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism.** *Nat Rev Genet* 2006, **7:**606-619.
49. Bevan P: **Insulin signalling.** *J Cell Sci* 2001, **114(Pt 8):**1429-1430.
50. Wang HB, Li H, Wang QG, Zhang XY, Wang SZ, Wang YX, Wang XP: **Profiling of chicken adipose tissue gene expression by genome array.** *BMC Genomics* 2007, **8:**193.
51. Gabler N, Spurlock M: **Integrating the immune system with the regulation of growth and efficiency.** *J Anim Sci* 2008, **86(14 Suppl):**E64-74.
52. Peter A, Weigert C, Staiger H, Machicao F, Schick F, Machann J, Stefan N, Thamer C, Häring H, Schleicher E: **Individual stearoyl-coa desaturase 1 expression modulates endoplasmic reticulum stress and inflammation in human myotubes and is associated with skeletal muscle lipid storage and insulin sensitivity in vivo.** *Diabetes* 2009, **58:**1757-1765.
53. Schabort E, van der Merwe M, Loos B, Moore F, Niesler C: **TGF-beta's delay skeletal muscle progenitor cell differentiation in an isoform-independent manner.** *Exp Cell Res* 2009, **315:**373-384.
54. Budasz-Rwiderska M, Jank M, Motyl T: **Transforming growth factor-beta1 upregulates myostatin expression in mouse C2C12 myoblasts.** *J Physiol Pharmacol* 2005, **56(Suppl 3):**195-214.
55. Kollias HD, McDermott JC: **Transforming growth factor-beta and myostatin signaling in skeletal muscle.** *J Appl Physiol* 2008, **104:**579-587.
56. Velleman S: **The role of the extracellular matrix in skeletal muscle development.** *Poult Sci* 1999, **78:**778-784.
57. Patruno M, Caliaro F, Martinello T, Mascarello F: **Expression of the paired box domain Pax7 protein in myogenic cells isolated from the porcine semitendinosus muscle after birth.** *Tissue & Cell* 2008, **40:**1-6.
58. Maier T, Guell M, Serrano L: **Correlation of mRNA and protein in complex biological samples.** *Febs Letters* 2009, **583:**3966-3973.

doi: 10.1186/1471-2164-11-372
Cite this article as: Cánovas *et al.*, Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits *BMC Genomics* 2010, **11**:372

3.2. - Expression Quantitative Trait Loci and their relationship with lipid deposition in *gluteus medius* muscle of Duroc pigs

Cánovas, A., Pena, R.N, Gallardo, D., Amills, M., Quintanilla, R

Plos One **(en preparación)**

Expression Quantitative Trait Loci and their relationship with lipid deposition in *gluteus medius* muscle of Duroc pigs

*Cánovas, A., *[1,2]Pena, R.N., [¥]Gallardo, D., [¥]Amills, M., [1]*Quintanilla, R.

[*]IRTA, Genètica i Millora Animal, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain.

[¥]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[1]Corresponding authors:

Raquel Quintanilla

IRTA - Genètica i Millora Animal

Centre UdL-IRTA, 191 Av Alcalde Rovira Roure, 25198 Lleida (Spain).

E-mail: Raquel.Quintanilla@irta.cat

Ramona N. Pena

Departament de Producció Animal, Campus ETSEA

Universitat de Lleida, 191 Av Alcalde Rovira Roure, 25198 Lleida (Spain).

E-mail: romi.pena@prodan.udl.cat

[2]Current address: Departament de Producció Animal, Campus ETSEA, Universitat de Lleida, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain.

ABSTRACT

In the last 20 years, there has been significant research towards defining the genetic basis of lipid metabolism and meat quality related traits in pigs. Nowadays, the tools of global transcriptome analysis allow going far beyond in the genetic dissection of these complex traits. In the present study, the landscape of genome transcriptional regulation in the pig *gluteus medius* muscle has been described by means of a genome-wide eQTL scan. *GeneChip Porcine Genome* arrays *(Affymetrix)* were used to determine the gene expression levels of *gluteus medius* samples from 105 commercial Duroc pigs belonging to two groups with divergent phenotypes for muscle, plasma and carcass lipid traits. This experimental design aimed to favour detection of eQTL affecting genes related to lipid metabolism and meat quality traits. The whole genome scan with a panel of 110 microsatellites allowed us detecting 613 genome-wide significant eQTL unequally distributed across the pig genome. Chromosome SSC3 and SSC5 harboured the highest number of eQTL. Moreover, 11 eQTL hotspots affecting the expression levels of a high number of genes have been described. After mapping target probes and discarding low quality probes, a total of 59 *cis*- and 396 *trans*-acting eQTL were retained for further analyses. A Gene Ontology (GO) term study showed that lipid-related GO terms were the most enriched in the list of eQTL-regulated genes. On the basis of these results, a comprehensive list of 29 positional and functional candidate genes was elaborated through 1) the positional concordance between eQTL and QTL linkage maps and 2) a phenotypes/expression correlation study. These results represent a valuable contribution to the understanding of genetic regulation of skeletal muscle individual gene expression in swine species, and a first step towards disentangling gene networks and molecular mechanisms involved in lipid metabolism and meat quality traits.

INTRODUCTION

Genetical genomics studies aiming at mapping expression quantitative trait loci (eQTL) represent a valuable approach towards identifying regulatory regions and DNA sequence variants affecting the expression levels of genes, gaining new insights into gene regulation of complex traits. Recently, the chromatin occupancy of several transcription factors have been

characterised at global genomic level in order to better understanding the key role of gene expression regulation in biological processes (Cheung and Spielman 2009).

In the last couple of decades, thorough investigations have been conducted to better define the genetic basis of meat quality traits in pigs. In this sense, more than 1675 quantitative trait loci (QTL) that influence 281 different traits dealing with lipid metabolism and meat quality have been identified in the pig species (Hu and Reecy, 2007). Combining gene expression data obtained through microarray technique with QTL linkage studies offers new options to better understanding the genetic basis and the putative genes directly associated with these traits.

So far, eQTL studies in swine species are particularly scarce and have been limited to some meat quality traits. Thus, Wimmerset al. (2010) applied functional and genetical genomics approaches in order to identify functional networks of genes related to meat performance in *longissimus dorsi* muscle, whereas Ponsuksili et al. (2008) combined eQTL analysis with QTL data in order to identify candidate genes involved in water holding capacity. These eQTL studies in swine species were performed on a previous selection of probes whose expression levels had showed a positive correlation with some meat quality traits.

On the basis of their physical distance from the targeted gene, eQTL are usually classified in two categories (Cheung and Spielman 2009): *cis*-acting eQTL and *trans*-acting eQTL. Unfortunately, published studies focus mainly in*cis*-acting eQTLwhereas the construction of tissue-specific genetic networks should also take *trans*-acting eQTL into account (de Koningand Haley, 2005).

In the present study, we have carried out an eQTL study in order to characterise the landscape of mRNA regulatory regions in the pig skeletal muscle and investigate their association with several lipid metabolism and meat quality traits. To tackle this issue, we have used microarray data from *gluteus medius* muscle of Duroc pigs belonging to two groups with divergent phenotypes for lipid deposition and composition (Cánovas et al., 2010a) combining the following approaches: 1) genome-wide eQTL scans for 6139 *Affymetrix* probes displaying gene expression variability across a set of 105 arrays; 2) functional classification of the eQTL-regulated genes and their encoded protein

products; and 3) combining the eQTL results with QTL linkage maps and with correlation analyses between gene expression levels and phenotypes related to lipid metabolism and fat deposition.

MATERIAL AND METHODS

Animal material

An experimental pig population (350 castrated males) distributed in five half-sib families and four contemporary batches were generated from a commercial Duroc line. This commercial Duroc line was used in the production of fine quality cured ham characterized for high intramuscular fat content. These animals were castrated and controlled at IRTA-CCP experimental station. A number of phenotypes dealing with fatness, serum lipid levels, and intramuscular fat content and composition were recorded and measured as previously described by Gallardo et al. (2008, 2009). To date, several investigations including QTL mapping for lipid metabolism and fat deposition traits (Gallardo et al. 2008; Quintanilla et al. 2010) and a differential expression study (Cánovas et al. 2010a) had been conducted on this experimental population. The experimental procedures, traits recording and blood sampling were approved by the Ethical Committee of the Institution (IRTA- *Institut de Recerca i Tecnologia Agroalimentàries*).

Microsatellite genotyping

The commercial Duroc pig population described previously was genotyped for 110 informative microsatellites covering the whole genome with an average interval space between markers of 20.5 cM. All the information related to genotyped microsatellites and the description of the linkage map can be found in Gallardo et al. (2008).

Expression data and normalization

We have analyzed the global mRNA expression profile of *gluteus medius* muscle samples obtained from 105 individuals belonging to two groups with divergent lipid metabolism profile (53 and 52 individuals in High and Low group respectively). Groups were established on the basis of multivariant analyses for several lipid deposition traits as described in Cánovas et al. (2010a). RNA isolation and

hybridization in *GeneChip Porcine Genome*® arrays (Affymetrix, Inc., Santa Clara, CA) procedures were also reported in Cánovas et al. (2010a).

Microarray data normalisation was assessed through the gcRMA algorithm, which corrects the intensity of each probe by its GC content (Wu et al., 2004), using the BRB-ArrayTools software version 3.7.1 (Xu et al., 2008), which is available online at http://linus.nci.nih.gov/BRB-ArrayTools.html.

Microarray data analysed in the present study have been deposited in the Gene Expression Omnibus (GEO) public repository with GSE19275 and GSE26091 GEO accession numbers.

Expression quantitative trait loci analyses

A genome-wide eQTL scan was carried out for 6139 *Affymetrix* probes witch displayed more than 20% of expression values over ± 1.5 times the median expression of all arrays. These eQTL analyses were performed for each probe, using the aforementioned panel of 110 microsatellites, by means of the *GridQTL* software (Seaton et al. 2006), available at http://www.gridqtl.org.uk/.

The common model used for eQTL searching was:

$$y_{ijk} = \mu + b_i + l_j + \sum_{sire=1}^{5} \alpha_{sire} p_{k(sire)} + e_{ijk}$$

where: y_{ijk} is the expression data of individual k; b_i is the effect of i^{th} contemporary batch of fattening (4 levels); l_j is the laboratory j effect (2 levels); α is the regression coefficient (mean allele substitution effect); p_k is the probability of individual k inheriting a given allele from its common parent.

Genome-wide significance thresholds for the F-values (eQTL model *vs* no eQTL model) were approximated with Bonferroni correction as described in Gallardo et al. (2008). The resulting *P*-values corresponding to 95% and 99% genome-wide significant levels were p=0.0009 and p=0.0002 respectively.

Reference assembly and mapping analyses

For those target probes with significant eQTL, we have analyzed the *Affymetrix* porcine probe-sets (11 probes per probe-set). Significant eQTL *Affymetrix* porcine probes were assembled to the pig genome sequence with Ensembl annotation (http://www.ensembl.org/info/data/ftp/index.html) using CLCBio workbench

software (CLC Bio, Aarhus, Denmark), and quality of all significant probes was studied. Mapping analysis was performed by taking into account the fraction of mismatches with the reference sequence; the probes were classified in uniquely mapped probes and non-specifically mapped probes.

In order to locate all probes showing significant eQTL and to align these eQTL on pig genome, we used the GBrowse tool available in pig QTLdb (http://www.animalgenome.org/cgi-bin/gbrowse/pig/). The alignment of eQTL genome locations against the target transcripts or genes was accomplished by converting the linkage eQTL map (cM) to genome sequence assembly locations (bp) as described in Hu et al. (2007) This was done with references of anchoring markers that are mapped on both (linkage and genome) maps. In such cases where eQTL boundaries are not exactly at the anchoring marker locations, relative genome locations were calculated with an algorithm that takes into account the chromosomal length, the cM *vs* bp ratio, and the offset of the eQTL location to that of anchoring marker (Hu et al., 2007). The comparison between eQTL and their corresponding target genes locations allowed us to classify eQTL in *cis*-acting eQTL (close to the target gene) and *trans*-acting eQTL (in different chromosome or far from the target gene). As in Ponsuksili et al. (2008), *cis*-acting eQTL were considered when the target gene position fall within the interval delimited by the two flanking markers of the eQTL peak.

Gene Ontology analyses

Gene Ontology (GO) describes the basic term hierarchies and relationships between terms within the context of biology. The three GO categories (biological process, cellular component and molecular function) were explored with several bioinformatic tools.

GO term enrichment analyses for overrepresentation of specific GO terms were performed with DAVID (Database for Annotation, Visualization and Integrated Discovery) bioinformatic package, available at http://david.abcc.ncifcrf.gov. Significance levels were calculated following a modification of Fisher's exact test. Multiple testing-corrected P-values were also calculated by using the Benjamini and Hochberg algorithm, and only GO terms with Benjamini-corrected P-values <0.1 were considered.

DAVID was additionally used to explore the biological pathways enriched in the resulting eQTL-regulated gene list, computing a total overrepresentation value for each pathway represented in the Kyoto Encyclopaedia of Genes and Genomes (KEGG, www.genome.jp/kegg).

Correlation analyses between gene expression levels and muscle fat deposition traits

Correlation analyses between expression levels and several phenotypes related to lipid metabolism and fat deposition were additionally carried out for a number of eQTL-regulated genes selected on the basis of the former analyses. The correlation study was performed using the CORR procedure of SAS (SAS Institute Inc.) after adjusting phenotypes and expression levels for the environmental significant effects considered in previous analyses (residual correlations). Two correlations analyses were carried out: 1) correlation analyses considering all data set (n=105 for each correlation); and 2) within-group correlation analyses, considering separately the two groups of individuals with divergent lipid metabolism profile (n(HIGH)=53 and n(LOW)=52 for each group).

RESULTS AND DISCUSSION

Identification of expression quantitative trait loci in *gluteus medius* muscle

In the whole-genome scan carried out for 6139 transcripts, a total of 613 eQTL were identified at genome-wide significance level. There was an uneven distribution of eQTL across the 18 autosomes, SSC5 and SSC3 harbouring the highest number of eQTL (107 and 74 respectively), whereas only 7 and 3 eQTL were observed on SSC14 and SSC11, respectively (Figure 1). Out of the 613 total eQTL -targeted probes, 490 were successfully mapped to the pig genome (Table 1). Most of them were uniquely mapped probes (478), while a small fraction corresponded to non-specifically mapped probes (12). The high number of unmapped probes (123 probes; 20%) is due to the incomplete annotation of porcine genome available at the moment. The relative position of the eQTL and the target transcripts/genes was studied in order to classify them into *cis*-acting and *trans*-acting eQTL. From the 478 mapped eQTL, only 63 could be catalogued as a *cis*-acting eQTL on the basis of their close physical distance

from the target gene, whereas the remaining 415 eQTL were assigned as *trans*-acting eQTL (Table 1). A high proportion of both *cis*- and *trans*-acting eQTL target probes showed high quality binding of the probe-sets (i.e., in most cases, six or more out of 11 probes specifically assembled to a unique region; Table 1). Based on this probe quality analysis, only 4 and 19 *cis*- and *trans*-acting eQTL target probes, respectively, were discarded thus leaving 59 *cis*-acting and 396 *trans*-acting eQTL for further analyses (Table 1).

The proportion of *cis*-acting and *trans*-acting eQTL varies from study to study. Whereas some studies report more *cis*-eQTL (Goring et al., 2007; Schadt et al., 2008), other authors have found a higher proportion of *trans*-eQTL (Morley et al., 2004; Myers et al., 2007; Ponsuksili et al., 2008). Differences in sample size and significance thresholds are probably among causes for these disagreements (Cheung and Spielman, 2009). Verdugo et al. (2010) and Schadt et al. (2008) suggest that the smaller proportion of *trans*-acting eQTLis related to a lack of power to detect small and moderate eQTL effects. In the present study, a higher fraction of *trans*-acting eQTL than *cis*-acting eQTL (396 *vs* 59) was obtained. In the line of our results, also Ponsuksili et al. (2008) observed a much higher proportion of *trans*-eQTL (96 *trans*- *vs* 8 *cis*-acting) among eQTL correlated with water holding capacity.

A detailed analysis of eQTL affecting the expression level of a high number of genes lead us to define 11eQTL hotspot (Figure 2) on SSC1, SSC2, SSC3, SSC5, SSC6, SSC7, SSC12 and SSC18. The number of genes regulated by these eQTL hotspots regions ranged from 4 to 16 (Table 2). The comparison of these regions with published pig QTL maps (http://www.animalgenome.org/cgi-bin/QTLdb) showed that a number of QTL linked to carcass, growth, feed intake, fatness and meat quality traits have been mapped on all these eQTL hotspot regions, with the only exception of SSC6 (124 cM). The positional concordance with QTL for lipid metabolism and muscle fat deposition traits is detailed more thoroughly below.

Functional study of genes regulated by eQTL

GO term annotations were used in order to perform a functional categorization of the target genes showing genome-wide significant eQTL. The three main GO categories

(biological process, molecular function and cellular component) were analysed in order to constitute a full coverage of the GO spectrum (Figure 3).

An important presence of GO lipid-related terms was particularly palpable under the biological process GO category (Figure 3A), gathering GO terms such as *regulation of fatty acid oxidation, fatty acid biosynthetic process, regulation of fatty acid metabolic process*, *regulation of lipid metabolic process*, and *lipid biosynthetic process*. This evidences that a substantial fraction of the eQTL-regulated genes could be associated with traits under study. Additionally, a variety of developmental and morphological processes related to muscle development and functionality (e.g. *actin filament-based process*, *cytoskeleton organization*, and *muscle organ development*) were also spread over the list of most significantly enriched GO terms. Other metabolic processes such as*regulation of glucose metabolic process* and *regulation of insulin-like growth factor receptor signalling* appeared in the middle of the list.

Concerning the molecular function GO category (Figure 3B), *carboxylic acid binding*, *acetyl-CoA carboxylase activity* and *lipid metabolism* were among GO terms most significantly enriched by the list of eQTL-regulated genes. The cellular component GO classification (data not show) only indicated that most genes encoded intracellular components.

On the basis of these results, eQTL regulating genes potentially associated with muscle fat deposition and meat quality traits were further investigated. First of all, we performed a more detailed study of the proteins encoded by those eQTL-regulated genes (Tables 4 and 5). These included the genes specifically enriching the five lipid metabolism-related GO terms previously mentioned: 8 *cis*- (*ABAT*, *ACER3, ANXA8L1*, *APOBEC1*, *EHD1*, *FIG4*, *PAFAH1B3* and *SNX14*) and 16 *trans*-regulated (*ACLY, ACACA*, ACOT6, *CAST*, *CS, CYP24A1*, *EHD1*, *FABP5*, *HRC, IGFBP5, IGF1, LRP6*, *MOP-3*, *PRKAA1*, *PRKAA* and *RRMB2*) genes. Taken as a whole, the *cis*-eQTL either (i) regulated genes modulate the intracellular level of bioactive lipids that function as second messengers to cell and differentiation events or (ii) affect the trafficking of lipids by lipoprotein. In the list of 16 *trans*-regulated genes, there are well-studied genes related to fat deposition (*ACACA,*

FABP5, IGF1, *ACLY* and *CS*) or meat tenderisation (*CAST*); other genes function through an indirect action over lipid transport or cholesterol synthesis.

Complementary to this, we also performed a pathway analysis with the same list of genes. Pathways involved in lipid metabolism and/or adipose/muscle tissue differentiation were amongst the most significant (in bold in Table 3; P<0.1). These included the *hypertrophic cardiomyopathy* and the *adipocytokine signalling* pathway on the top of the list. The *insulin signalling pathway*, associated to glucose metabolism, was also enriched by several genes directly involved in lipogenesis and cholesterogenesis process like *ACACA*, *PRKAA1* and *PRKAA2* genes. Also *PPAR signalling pathway* appeared in the middle of the list and contained genes associated to traits under study such as *ACSL1*, *ACSL5* and *FABP5* (Table 3). Finally, the *TGF-beta signaling* pathway regulates the switch of myocyte/adipocyte differentiation.

In addition to the main block of genes belonging to adipose related functions, the list of eQTL-regulated genes also contained an important fraction of genes functionally related to cell growth and differentiation, in both embryonic and adult tissues. These included GO terms specifically related to muscle development and functional processes, such as the formation of actin filaments and microtubules and the cytoskeleton organization (Figure 3A). Among genes enriching these categories it is worth mentioning *DNAJB6* (DnaJ (Hsp40) homolog, subfamily B, member 6), *HOXA10* (homeobox A10), *MYO6* (myosin VI) and *PPP1CB* (protein phosphatase 1 gene) genes, which regulate a variety of cellular processes such as cell division, glycogen metabolism and muscle contractility.

The sub-population analysed for eQTL scan consisted on two groups of animals with divergent phenotypes for fatness, serum lipids concentrations and FA profile (Cánovas et al. 2010a). This experimental design was *a priori* expected to be particularly suited to find eQTL regulating genes functionally related to lipid metabolism and fat deposition.As a matter of conclusion of this section, results provided by the functional study agree well with that was expected, *i.e.*a predominance of lipid-related functions among genes found to be eQTL-regulated in our experiment. In addition and under the hypothesis that expression differences

are modulated by DNA polymorphisms, results are also consistent the transcriptomic study of these two groups of animals (Cánovas et al., 2010a), where differential expression for genes related to fat metabolism, but also and more predominantly, for cell differentiation and energy balance genes was detected between these groups.

Positional concordance between eQTL and QTL for lipid metabolism and fat deposition traits

Combining eQTL and QTL linkage maps represents an interesting approach to find polymorphisms regulating gene expression and consequently phenotypic variation. In order to gain new insights into gene regulation of muscle fat deposition traits, a positional concordance study between the QTL detected in the present study and QTL for IMF content and FA profile performed in the same population (Quintanilla et al, 2010) was carried out (Tables 4 and 5). From these, 14 corresponded to *cis*-acting eQTL and 66 to *trans*-acting eQTL. Of notable interesting are those eQTL/QTL pairs which involve genes functionally associated to lipid metabolism and meat quality (Table 4 and 5).

In this regard, the IMF content is one of the most studied meat quality traits due to its strong correlation with sensory and technological meat parameters. We have found positional concordance for five QTL for IMF and eQTL involving lipid metabolism and adipose function genes. Amongst them, it is of notable interest the co-localisation of the chromosome-wide IMF QTL at SSC7 (132 cM) and *trans*-acting eQTL regulating the levels of three genes well-known for their direct role in fat metabolism: *ACACA*, *ACLY* and *FABP5*. Polymorphisms on these three genes have been associated to IMF and fat deposition traits in pig. Moreover, in our Duroc population two *ACACA* polymorphisms are associated with IMF and FA profile (Gallardo et al. 2009), and animals with divergent phenotypes for fatness traits show differential *ACACA* and *FABP5* expression (Cánovas et al. 2010a). This SSC7 eQTL hotspot (Figure 2) coincides with a genome region particularly relevant for fat deposition, as multiple QTL had been found for BFT and lean percentage by us (Quintanilla et al., 2010) and others (Nagamine et al., 2003). Another interesting region is the genome-wide QTL for IMF which co-localises with an eQTL hotspot (Figure 2) at SSC3

(24cM) involving at least seven genes (Tables 4 and 5) including *RRM2B*, which encodes a protein that provides the energy for beta-oxidation and lipid synthesis (Huang et al., 2009). The relevance of this IMF QTL at SSC3 is highlighted by concordant findings in unrelated pig populations (de Koning et al., 2003; Ponsuksili et al., 2008; Quintanilla et al., 2010).

Muscle FA content and composition are fat deposition traits with a strong impact on sensorial, technological and particularly nutritional quality of meat (Wood et al., 2008). QTL affecting different FA (such as vaccenic, myristic and palmitic FA) at SSC2, SSC6, SSC8, SSC14 and SSC18 matched with several *cis*-and *trans*-acting eQTL. Amongst the genes regulated by these eQTL are *LRP6*, *CAST* and *CYP24A1* regulated genes. *LRP6* is involved in receptor-mediated endocytosis of lipoprotein and protein ligands (Gustafson and Smith, 2010); *CAST* encodes a calpain inhibitorhighly correlated with muscle protein accumulation in neonatal pigs (Li et al., 2009); *CYP24A1* belong to the family of mitochondrial monooxigenases regulating the synthesis of cholesterol, steroids and other lipids. Despite the repercussions to human nutrition and the marketability of muscle cholesterol content, this trait has rarely been included in production trait studies in pigs. Interestingly, a QTL for muscle cholesterol content at SSC11 co-localised with an eQTL affecting the expression of *MOP-3*, involved in lipid synthesis activity (Table 5), in our Duroc resource population.

In the light of these results, eQTL hotspots described here constitute interesting candidate genome regions to be responsible for lipid metabolism and IMF fat content and composition variation. Further investigations aiming to detect causal polymorphisms associated to these variations in gene expression and phenotypes will be worthy.

Correlation between target gene expression and muscle fat deposition traits

The correlation coefficients between probe expression levels and muscle fat deposition traits were computed for the 80 target genes located in the confidence interval of QTL previously identified. Correlation coefficients that resulted significant in at least one of the three calculations (whole dataset, within HIGH and LOW groups of lipid traits) are shown in

Table 6. On the whole, the correlation coefficients between gene expression levels and muscle fat content and composition phenotypes were moderate (positive and negative), their absolute values ranging from 0.25 to 0.44 (Table 6). There was not a general consistency between the correlation coefficients obtained within the two groups of animals with divergent profiles for fat deposition and lipid composition. Fewer correlations were consistent across the three groups of animals: these include the positive relationship between *NEDD4L* mRNA and myristic FA content and the positive correlations of *CYP24A1* expression levels with total palmitic FA.

The RNA levels of *ACACA* and *ACLY* genes showed positive correlation coefficients with IMF content that reached the statistical significance only in the group of animals with lower levels of fat deposition. To this regard it is also important to take into account the positional concordance between the SSC7 *trans*-eQTL regulating these genes and a highly significant QTL for IMF (Quintanilla et al. 2010). In the same line, muscle gene expression of *IL12RB2* was positively correlated with IMF content in the group of animals with high fatness and lipid profile, and its position matched with a QTL for IMF at SSC6.

Interestingly two genes (*CYP24A1* and *FABP5*) showed expression levels that negatively correlated with the muscle cholesterol content, a trait scarcely studied despite its high impact on meat nutritional value. On the other hand, several genes (*e.g. FABP5*, *HOXA10*, *DNAJB6*, *NUDT6*, *LRP6*, *PPP1CB*, *RRM2B*, *NEDD4L* and *CAST*) showed a functional correlation with the muscle content of several FA, but only in one of the two groups of animals. In other cases, even an opposite relationship depending upon the group of animals considered was observed (*e.g.* correlation of *CS*, *CAST*, *LRP6* and *RRM2B* mRNA levels with linoleic FA muscular content). Finally it is worth mentioning that *CAST,* involved in meat tenderness, was positively (in HIGH group) and negatively (in LOW group) correlated with myristic FA acid, its *trans*-eQTL at SSC8 co-localising with a QTL for myristic FA content.

Taken as a whole, results from correlation study partially corroborate the functional relationship between some of these genes and traits related to fat deposition and FA metabolism. Jointly with the QTL-eQTL positional concordance, this represents an

additional argument supporting that polymorphisms responsible for eQTL could be also associated to phenotypes. Nevertheless, differences found between the two groups of animals suggest that these functional relationships could be regulated by the metabolic profile of individuals, which in turn is probably associated to transcriptomic activity of other genes. In a previous study of muscle transcriptomic profiles of these two groups with divergent phenotypes for fatness traits (Cánovas et al. 2010a), important differences in the expression of cell differentiation, energy balance and fat metabolism genes were described between these groups of pigs. However, the complete understanding of molecular mechanisms modulating gene expression and complex phenotypes requires the study of how genes are connected. To this extreme, our eQTL study represents a first step towards elucidate the structure of gene networks involved in lipid metabolism and meat quality traits.

CONCLUSION

The present study aimed to gain insight into genetic regulation of skeletal muscle gene expression, focusing on genes related to lipid metabolism and meat quality related traits. The genome-wide eQTL scan for gene expression in *gluteus medius* samples of Duroc pigs has allowed detecting 59 *cis*- and 396 *trans*-acting eQTL unequally distributed across the pig genome. Moreover, 11 genome regions with eQTL affecting the expression levels of a high number of genes (eQTL hotspots) have been described. The relationship of this transcriptomic regulation with lipid metabolism and meat quality traits has been investigated combining eQTL results with functional classifications, QTL linkage maps and correlation analyses between phenotypes and expression levels. As a result, several candidate genes positional and functionally associated to lipid metabolism and pork quality have been identified.

These results represent an additional step towards understanding the genetic regulation of skeletal muscle gene expression in swine species and its relationship with fat deposition traits. Future studies integrating genetic mapping results with co-expression studies will be needed to disentangle interactions amongst genes as well as detecting gene networks and master

genes modulating their topology and the phenotypes.

ACKNOWLEDGMENTS

We acknowledge Dr. James M. Reecy (Iowa State University, Ames, USA), Dr. Juan F. Medrano and Dr. Gonzalo Rincón (University of California-Davis, Davis, USA) for their help in statistical analysis, David Almuzara (IRTA, Lleida, Spain) for his cooperation in the experimental protocol. Angela Cánovas received a predoctoral scholarship from INIA. This project has been funded by the Spanish Ministry of Science (projects GEN2003-20658-C05-05 and AGL2007-66707-C02-01).

REFERENCES

Botelho RJ, Efe, J.A., Teis, D.,Emr, S.D. (2008) Assembly of a Fab1 Phosphoinositide Kinase Signaling Complex Requires the Fig4 Phosphoinositide Phosphatase. Molecular Biology of the Cell 19: 4273-4286.

Cánovas A, Estany, J., Tor, M., Pena, R.N., Doran, O. (2009) Selection for decreased backfat thickness at restrained intramuscular fat content is related to reduction in acetyl-CoA carboxylase and stearoyl-CoA desaturase protein expression in pig subcutaneous adipose tissue. J Anim Sci 87: 3905-3914.

Cánovas A, Quintanilla, R., Amills, M., Pena, R.N. (2010a) Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. BMC Genomics 11: 372-315.

Cánovas A, Quintanilla, R., Gallardo, D., Díaz, I., Noguera, J.L., Ramírez, O., Pena, R.N. (2010b) Functional and association studies on the pig HMGCR gene, a cholesterol-synthesis limiting enzyme. Animal 4: 224-233.

Cheung VG, Spielman, R.S. (2009) Genetics of human gene expression:mapping DNA variants that influence gene expression. Nature Reviews Genetics.

de Koning DJ, Pong-Wong, R., Varona, L., Evans, G.J., Giuffra, E., Sanchez, A., Plastow, G., Noguera, J.L., Andersson, L., Haley, C.S. (2003) Full pedigree quantitative trait locus analysis in commercial pigs using variance components. Journal of Animal Science 81: 2155-2163.

de Koning DJ, Haley, C.S. (2005) Genetical genomics in humas and model organism. . Trends Genet 21: 377-381.

de Koning DJ, Cabrera CP, Haley CS (2007) Genetical genomics: combining gene expression with marker genotypes in poultry. Poult Sci 86: 1501-1509.

Drake TA, Schadt, E.E., Lusis, A.J. (2006) Integrating genetic and gene expression data: application to cardiovascular and metabolic traits in mice. Mamm Genome 17: 466-479.

Franke L, Jansen RC (2009) eQTL analysis in humans. Methods Mol Biol 573: 311-328.

Gallardo D, Canovas E, Lopez-Bejar M, Ramirez O, Pena R, et al. (2008) Alternative splicing at exon 28

of the acetyl-coenzyme A carboxylase alpha gene in adult pigs and embryos. Anim Genet 39: 205-206.

Gallardo D, Quintanilla R, Varona L, Diaz I, Ramirez O, et al. (2009) Polymorphism of the pig acetyl-coenzyme A carboxylase alpha gene is associated with fatty acid composition in a Duroc commercial line. Anim Genet 40: 410-417.

Gilad Y, Pritchard, J.K., Thornton, K. (2009) Characterizing natural variation using next-generation sequencing technologies. Trends Genet 25: 463-471.

Göring HH, Curran CJ, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JB, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, MacCluer JW, Kissebah AH, Collier GR, Moses EK, Blangero J. (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 39: 1208-1216.

Hu ZL, Bao J, Reecy JM. (2008) CateGOrizer: a web-based program to batch analyze gene ontology classification categories. Online Journal of Bioinformatics 9: 108-112.

Hu ZL, Reecy JM (2007) Animal QTLdb: beyond a repository. A public platform for QTL comparisons and integration with diverse types of structural genomic information. Mamm Genome 18: 1-4.

Hu ZL, Fritz ER, Reecy, JM. (2007). Animal QTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. Nuc Acid Res 35: 604-609.

Huang CC, Hsu, C.H. (2009) Mitochondrial disease and mitochondrial DNA depletion syndromes. Acta Neurol Taiwan 18: 287-295.

Lambert CB, Spire, C., Claude, N., Guillouzo, A. (2009) Dose- and time-dependent effects of phenobarbital on gene expression profiling in human hepatoma HepaRG cells. Toxicol Appl Pharmacol 234: 345-360.

Liu G, Kim JJ, Jonas E, Wimmers K, Ponsuksili S, et al. (2008) Combined line-cross and half-sib QTL analysis in Duroc-Pietrain population. Mamm Genome 19: 429-438.

Mao C, Obeid, L.M. (2008) Ceramidases: regulators of cellular responses mediated by ceramide, sphingosine, and sphingosine-1-phosphate. Biochim Biophys Acta 1781: 424-434.

20. Meyers, S.N., Beever, J. E. (2008) Investigating the genetic basis of pork tenderness: genomic analysis of porcine CAST. Animal Genetics 39: 531-543.

Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG. (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430: 743-747.

Myers AJ, Gibbs GJ, Webster JA, Rohrer K, Zhao A, Marlowe L, Kaleem M, Leung D, Bryden L, Nath P, Zismann VL, Joshipura K, Huentelman MJ, Hu-Lince D, Coon KD, Craig DW, Pearson JV, Holmans P, Heward CB, Reiman EM, Stephan D, Hardy J. (2007) A survey of genetic human cortical gene expression. Nature Genet 39: 1494-1499.

Nagamine Y, Haley, C.S., Sewalem, A., Visscher, P.M. (2003) Quantitative trait loci variation

for growth and obesity between and within lines of pigs (Sus scrofa). Genetics 164: 629-635.

Nagao K, Yanagita, T. (2008) Bioactive lipids in metabolic syndrome. Prog Lipid Res 47: 127-146.

Nicot, AS, Laporte, J. (2008) Endosomal Phosphoinositides and Human Diseases. Traffic 9: 1240-1249.

Pena RN, Cánovas, A., Varona, L., Díaz, I., Gallardo, D., RamÍrez, O., Noguera, J. L., Quintanilla, R. (2009) Nucleotide sequence and association analysis of pig apolipoprotein-b and ldl-receptor genes. Anim Biotechnol 20: 110-123.

Ponsuksili S, Jonas E, Murani E, Phatsara C, Srikanchai T, et al. (2008) Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. BMC Genomics 9: 367.

Quintanilla R, Pena, R.N., Gallardo, D., Cánovas, A., Ramírez, O., Díaz, I., Noguera, J.L., Amills, M. (2010) Porcine intramuscular fat content and composition are regulated by quantitative trait loci with muscle-specific effects. Physiological Genomics (submitted).

Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich R. (2008) Mapping the genetic architecture of gene expression in human liver. PLoS Biol 6: e107.

Schug M, Heise, T., Bauer, A., Storm, D., Blaszkewicz, M., Bedawy, E., Brulport, M., Geppert, B., Hermes, M., Föllmann, W., Rapp, K., Maccoux, L., Schormann, W., Appel, K.E., Oberemm, A., Gundert-Remy, U., Hengstler, J.G. (2008) Primary rat hepatocytes as in vitro system for gene expression studies: comparison of sandwich, Matrigel and 2D cultures. Arch Toxicol 82: 923-931.

Seaton G, Hernandez, J., Grunchec, J.A., White, I., Allen, J., de Koning, D.J., Wei, W., Berry, D, Haley, C., Knott, S. (2006) GridQTL: A grid portal for QTL mapping of compute intensive datasets. . In *Proc 8th WCGALP* 27: 633-634.

Tesson BM, Jansen RC (2009) eQTL analysis in mice and rats. Methods Mol Biol 573: 285-309.

Tomaszewski M, Charchar, F.J., Barnes, T., Gawron-Kiszka, M., Sedkowska, A., Podolecka, E., Kowalczyk, J., Rathbone, W., Kalarus, Z., Grzeszczak, W., Goodall, A.H., Samani, N.J., Zukowska-Szczechowska, E. (2009) A common variant in low-density lipoprotein receptor-related protein 6 gene (LRP6) is associated with LDL-cholesterol. Arterioscler Thromb Vasc Biol 29: 1316-1321.

Uemoto Y, Sato, S., Ohnishi, C., Terai, S., Komatsuda, A., Kobayashi, E. (2009) The effects of single and epistatic quantitative trait loci for fatty acid composition in a Meishan x Duroc crossbred population. Journal of Animal Science 87: 3470-3476.

van Wijk HJ, Dibbits, B., Baron, E.E., Brings, A.D., Harlizius, B., Groenen, M.A., Knol, E.F., Bovenhuis, H. (2006) Identification of quantitative trait loci for carcass composition and pork quality traits in a commercial finishing cross. Journal of Animal Science 84: 789-799.

Verdugo RA, Farber, C.R., Warden, C.H., Medrano, J.F. (2010) Serious limitations of the QTL/microarray approach for QTL gene discovery. BMC Biol 8: 96.

Wimmers K, Murani E, Ponsuksili S (2010) Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. Brief Funct Genomics 9: 251-258.

Wood JD, Enser, M, Fisher A.V, Nute, G.R, Sheard, P.R, Richardson, R.I, Hughes, S.I, Whittington, F.M (2008) Fat deposition, fatty acid composition and meat quality: A review. Meat Science 78: 343-358.

Wu Z, Airizarry, R., Gentleman, R., Martinez-Murillo, F., Spencer, F. (2004) A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association* 99: 909-917.

Xu X, Zhao, Y., Simon, R. (2008) Gene Set Expression Comparison kit for BRB-ArrayTools. *Bioinformatics* 24: 137-139.

TABLES & FIGURES

Table 1: Number of high and low quality porcine affymetrix probes showing *cis*- and *trans*-acting eQTL regulating their expression levels.

	High quality probes	Low quality probes	Number of probes
***cis*-acting eQTL**	59	4	63
***trans*-acting eQTL**	396	19	415
Total	455	23	478

Table 2: Number and list of genes regulated in each eQTL hotspots regions.

Chromosome	Position	Nº regulated genes	List of regulated genes
SSC1	10 cM	11	ARPC1A, CBPG, CKS2, CLDND1, CXXC5, G3BP1, IGFBP5, MYO6, RBM25, THOC1, WDR70
SSC2	122 cM	7	ACOT6, EHD1, UFD1L, SEC22B, FUCA1, ZNF236, C10orf28
SSC3	24 cM	16	PPP1CB, LCMT1, MEPCE, WASL, CPEB4, TANK, EDEM3, RRM2B, RDX, KCNMB3, KDELR2, THUMPD1, CASD1, PLOD1, TRPM7, EMSP1
	30 cM	14	EHD1, IGF1, TRIB2, NNAT, VCPIP1, BDKRB1, GUSB, RBM39, SAP30, GATAD2A, DUSP7, EBAG9, EMD, FAM20C
SSC5	114 cM	16	PRKAG2, TRPM7, LYRM2, ALDH1L2, EIF3A, MCFD2, HNRNPH1, PRKAA2, HIGD1A, AGRP, ISOC2, PSMC6, PRPF39, CYCS, THUMPD1
	124 cM	13	SLC5A1, TMEM44, CUL1, NGAP, APOBEC1, NRXN1, HSD11B1, SF3B1, HMGB2, HRAS, MRPS18C, PFMK, TAF1D
SSC6	66 cM	10	LRP6, PAFAH1B3, GARS, KIAA1217, DR1, MTSS1, ARF3, LOC654323, SIGLEC10, TM2D2
	124 cM	9	CEBPG, HRC, FST, MED10, TRIM44, RBM12B, HDGF2, SS18, CAST
SSC7	134 cM	14	ACACA, FABP5, CKAP4, PDE7A, NADK, BAG3, PUS1, ACLY, SLC5A6, ZAN, CNKSR3, CLDN9, GPR160, PTPLB
SSC12	88 cM	13	TAOK1, COL1A1, TMEM98, TRIP11, IGFBP6, PAFAH1B3, C20orf79, OBFC1, CCDC39, TRIM27, FAM98C, CUX1, IL18BP
SSC18	40 cM	4	CYP24A1, DNAJB6, HOXA10, TPPP3

Table 3: DAVID analysis of pathways significantly enriched in the list of *cis*- and *trans*-regulated genes significant at genome-wide level.

KEGG id	KEGG pathway name	Genes	Genes (n)	p-value	Fold Enrichment
05410	**Hypertrophic cardiomyopathy (HCM)**	LAMA2, ITGA6, ITGAV, PRKAG2, IGF1, PRKAA1, PRKAA2, TTN, TGFB2	9	0.002	3.739
04920	**Adipocytokine signalling pathway**	ACSL1, PRKAG2, PRKAA1, ACACB, PRKAA2, AGRP, IRS1, ACSL5	8	0.003	4.216
04960	Aldosterone-regulated sodium reabsorption	HSD11B1, IGF1, NEDD4L, ATP1A2, IRS1	5	0.027	4.306
05219	Bladder cancer	HRAS, MMP9, MDM2, CDK4, MMP1	5	0.029	4.204
04144	Endocytosis	STAMBP, HRAS, DNM1L, TGFBR1, RAB22A, SH3KBP1, ADRBK2, MDM2, NEDD4L, ZFYVE20, EHD1	11	0.034	2.111
04910	**Insulin signalling pathway**	HRAS, PPP1R3B, PRKAG2, ACACA, PRKAA1, ACACB, PRKAA2, PPP1CB, IRS1	9	0.036	2.354
03320	**PPAR signalling pathway**	ACSL1, SCD, MMP1, FABP5, ACOX3, ACSL5	6	0.044	3.071
04510	Focal adhesion	LAMA2, CAV2, HRAS, ITGA6, PDGFA, TNC, ITGAV, IGF1, COL1A1, PPP1CB, PARVB	11	0.056	1.933
04110	Cell cycle	YWHAZ, MAD2L1, CDC23, MDM2, CDK7, CDK4, CUL1, TGFB2	8	0.061	2.260
05200	Pathways in cancer	HRAS, PDGFA, MMP9, TGFBR1, IGF1, CDK4, MMP1, TGFB2, LAMA2, CCDC6, ITGA6, ITGAV, MDM2, TPR, GSTP1	15	0.076	1.615
04350	TGF-beta signalling pathway	TGFBR1, FST, ID4, BMPR1B, CUL1, TGFB2	6	0.097	2.435

Blood indicate pathways related to lipid metabolism and fat deposition.

Table 4: *Cis*-acting eQTL identified at genome-wide significant level in *gluteus medius* muscle of a commercial Duroc pig population. QTL for muscle fat deposition traits fitting in the confidence interval of eQTL are indicated.

	eQTL location and signification				QTL positional concordance in the same population
Cis-regulated gene	Chromosome	Position	F-value	Nominal p-value	Linked Trait
CNKSR3	SSC01	000cM	5.28	0.000229	-
TBP	SSC01	004cM	5.99	0.000065	-
MYO6	SSC01	010cM	6.74	0.000017	-
*SNX14	SSC01	028cM	4.81	0.000533	-
*FIG4	SSC01	040cM	5.10	0.000316	-
*NEDD4L	SSC01	064cM	5.03	0.000359	IMF
CALR	SSC02	028cM	12.85	0.000000	-
*EHD1	SSC02	122cM	6.16	0.000048	-
GPRC5B	SSC03	006cM	4.59	0.000794	-
*ABAT	SSC03	010cM	5.62	0.000125	IMF; VACCENIC
GSPT1	SSC03	010cM	5.60	0.000129	IMF; VACCENIC
LCMT1	SSC03	024cM	7.78	0.000003	IMF
MEPCE	SSC03	024cM	5.15	0.000289	IMF
PPP1CB	SSC03	024cM	4.64	0.000725	IMF
TRIB2	SSC03	028cM	4.68	0.000675	-
KCMF1	SSC03	052cM	4.76	0.000584	-
SELENBP1	SSC04	054cM	4.53	0.000885	-
MLLT11	SSC04	072cM	4.64	0.000725	-
TMEM5	SSC05	072cM	5.11	0.000311	-
CRY1	SSC05	102cM	5.91	0.000075	-
ALDH1L2	SSC05	114cM	5.92	0.000073	-
*APOBEC1	SSC05	124cM	4.95	0.000414	-
YEATS4	SSC05	130cM	5.12	0.000305	-
PDCD2L	SSC06	044cM	4.83	0.000515	-
*PAFAH1B3	SSC06	070cM	5.00	0.000379	VACCENIC
*IL12RB2	SSC06	092cM	5.13	0.000300	IMF
*HRC	SSC06	124cM	5.21	0.000260	-
CEBPG	SSC06	126cM	4.88	0.000470	-
SS18	SSC06	126cM	4.85	0.000496	-
HIST1H2BK	SSC07	056cM	4.66	0.000699	-
CLIC5	SSC07	058cM	10.06	0.000000	-
SBAB-591C4.1	SSC07	060cM	5.44	0.000172	-
SLA-1	SSC07	072cM	5.35	0.000202	PALMITOLEIC
LRFN2	SSC07	108cM	4.53	0.000885	-
LCORL	SSC08	052cM	6.18	0.000046	-
SFRP2	SSC08	058cM	5.19	0.000269	-

SLC30A9	SSC08	058cM	4.52	0.000901	-
*NUDT6	SSC08	076cM	6.94	0.000012	MYRISTIC
C11orf17	SSC09	002cM	7.60	0.000004	-
CTSC	SSC09	034cM	6.30	0.000038	-
APLP2	SSC09	100cM	4.58	0.000808	-
*ACER3	SSC09	138cM	4.68	0.000675	-
RAB3GAP2	SSC10	062cM	5.53	0.000147	-
DCLK1	SSC11	002cM	7.26	0.000007	-
DCLK1	SSC11	008cM	4.90	0.000454	-
TAOK1	SSC12	088cM	6.95	0.000012	-
COL1A1	SSC12	088cM	4.70	0.000651	-
TMEM98	SSC12	088cM	4.59	0.000794	-
CCDC50	SSC13	078cM	5.21	0.000260	-
C21orf7	SSC13	084cM	4.58	0.000808	-
C2CD2	SSC13	124cM	6.31	0.000037	-
*ANXA8L1	SSC14	044cM	4.78	0.000563	PALMITOLEIC
*ITGAV	SSC15	068cM	4.56	0.000838	MUFA; PUFA
ATIC	SSC15	090cM	4.59	0.000794	-
CXCR7	SSC15	108cM	4.56	0.000838	-
RAD1	SSC16	12cM	6.26	0.000040	-
ALP	SSC17	034cM	10.76	0.000000	-
DNAJB6	SSC18	040cM	4.98	0.000393	PALMITIC; SFA
HOXA10	SSC18	040cM	4.68	0.000675	PALMITIC; SFA

*Genes related to lipid metabolism

Table 5: *Tans*-eQTL identified at genome-wide significant level in *gluteus medius* muscle whose target gene is functionally related to lipid metabolism and fat deposition. QTL and positional genes fitting in the confidence interval of eQTL are indicated.

	eQTL location and signification					QTL positional concordance in the same population
***Trans*-regulated gene**	**SSC**	**Position**	**F-value**	**Nominal p-value**	**Positional genes**	**Linked Trait**
IGFBP5	SSC01	10cM	4.69	0.000663	AKAP12	-
YWHAZ	SSC01	68cM	4.58	0.000808	FBXL4	IMF
NFE2L1	SSC02	092cM	5.03	0.000359	CAST	VACCENIC
ACOT6	SSC02	122cM	4.53	0.000885	ACSL6	-
RRM2B	SSC03	024cM	4.58	0.000808	XYLT1	IMF
NNAT	SSC03	028cM	5.44	0.000172	LITAF	IMF
EHD1	SSC03	034cM	5.49	0.000157	CREBBP	-
IGF1	SSC03	058cM	4.96	0.000407	CTNNA2	-
PRKAA2	SSC05	114cM	5.26	0.000237	ALG12	-
LRP6	SSC06	066cM	5.53	0.000147	PPT1	VACCENIC
MTSS1	SSC06	070cM	5.11	0.000311	LPIN2	VACCENIC
CS	SSC06	094cM	4.77	0.000573	AK5	IMF
PRKAA1	SSC06	144cM	5.88	0.000079	RNU6-1	-
FABP5	SSC07	132cM	4.64	0.000725	PPP1R13B	IMF
ACACA	SSC07	136cM	4.85	0.000496	ALPK3	IMF
ACLY	SSC07	136cM	4.53	0.000885	ALPK3	IMF
CAST	SSC08	074cM	4.75	0.000594	MGST2	MYRISTIC
MOP-3	SSC11	034cM	5.22	0.000255	KLHL1	MUSCLE CHOLESTEROL
CYP24A1	SSC18	038cM	8.18	0.000001	NPSR1	PALMITIC; SFA

Table 6: Significant correlation coefficients (r) between muscle gene expressions of several *cis*- and *trans*- regulated genes and intramuscular fat content (IMF), fatty acid (FA) profile and cholesterol content of *gluteus medius* muscle.

			High group		Low group		Whole dataset	
	Regulated gene	QTL trait	r	p-value	r	p-value	r	p-value
***cis*-acting eQTL**	**DNAJB6**	Linoleic FA	-0.08	0.5729	**-0.30**	**0.0339**	-0.18	0.0670
	HOXA10	Palmitoleic FA	**0.39**	**0.0037**	0.01	0.9643	0.19	0.0526
		Linoleic FA	**0.35**	**0.0100**	-0.22	0.1250	0.08	0.4191
		Stearic FA	**-0.32**	**0.0227**	-0.09	0.5429	**-0.21**	**0.0388**
	IL12RB2	IMF GM	**0.37**	**0.0062**	0.16	0.2565	**0.32**	**0.0008**
		Total Cholesterol	0.16	0.2377	**-0.31**	**0.0261**	0.13	0.1713
	NEDD4L	Myristic FA	**0.28**	**0.0367**	**0.30**	**0.0310**	**0.33**	**0.0006**
		Palmitoleic FA	**0.28**	**0.0425**	-0.04	0.7929	-0.03	0.7984
	NUDT6	Palmitic FA	0.04	0.7829	**-0.33**	**0.0293**	-0.14	0.1602
	PPP1CB	Palmitic FA	**0.29**	**0.0380**	-0.13	0.4082	0.14	0.1712
		Linoleic FA	-0.22	0.1035	**0.36**	**0.0099**	-0.02	0.8783
***trans*-acting eQTL**	**ACACA**	IMF GM	0.15	0.2669	**0.43**	**0.0016**	**0.42**	**0.0001**
	ACLY	IMF GM	-0.02	0.8688	**0.31**	**0.0248**	0.18	0.0623
		Palmitic FA	0.04	0.7942	**0.37**	**0.0123**	**0.23**	**0.0264**
	CAST	Palmitoleic FA	**-0.32**	**0.0189**	-0.25	0.0713	**-0.32**	**0.0010**
		Myristic FA	0.25	0.0708	**-0.29**	**0.0381**	0.09	0.3623
		Linoleic FA	-0.23	0.0832	**0.35**	**0.0107**	-0.06	0.5432
		Palmitic FA	**0.31**	**0.0297**	-0.04	0.7828	**0.22**	**0.0310**
	CS	Palmitoleic FA	**0.29**	**0.0318**	-0.03	0.8497	0.15	0.1237
		Linoleic FA	**0.39**	**0.0039**	**-0.33**	**0.0188**	0.07	0.4947
	CYP24A1	Palmitic FA	**0.25**	**0.0468**	**0.38**	**0.0108**	**0.26**	**0.0100**
		Total Cholesterol	**-0.32**	**0.0172**	**-0.34**	**0.0134**	-0.17	0.0890
	FABP5	Palmitoleic FA	**-0.41**	**0.0023**	0.18	0.2045	0.05	0.6454
		Linoleic FA	**-0.35**	**0.0086**	0.12	0.4153	0.07	0.4594
		Stearic FA	**0.44**	**0.0014**	0.01	0.9986	0.15	0.1449
		Cholesterol GM	**-0.27**	**0.0453**	0.05	0.7523	-0.12	0.2088
	LRP6	Palmitoleic FA	**0.27**	**0.0444**	0.08	0.5585	0.08	0.4374
		Myristic FA	0.01	0.9834	**0.31**	**0.0286**	0.16	0.1023
		Linoleic FA	0.24	0.0859	**-0.41**	**0.0027**	-0.15	0.1392
	RRM2B	Linoleic FA	**-0.29**	**0.0312**	**0.40**	**0.0032**	0.04	0.6706
		Stearic FA	**0.29**	**0.0385**	0.05	0.7690	0.18	0.0755
	YWHAZ	Palmitoleic FA	**0.33**	**0.0148**	0.21	0.1494	**0.28**	**0.0040**
		Linoleic FA	**0.33**	**0.0151**	-0.17	0.2205	0.11	0.2434

Figure 1

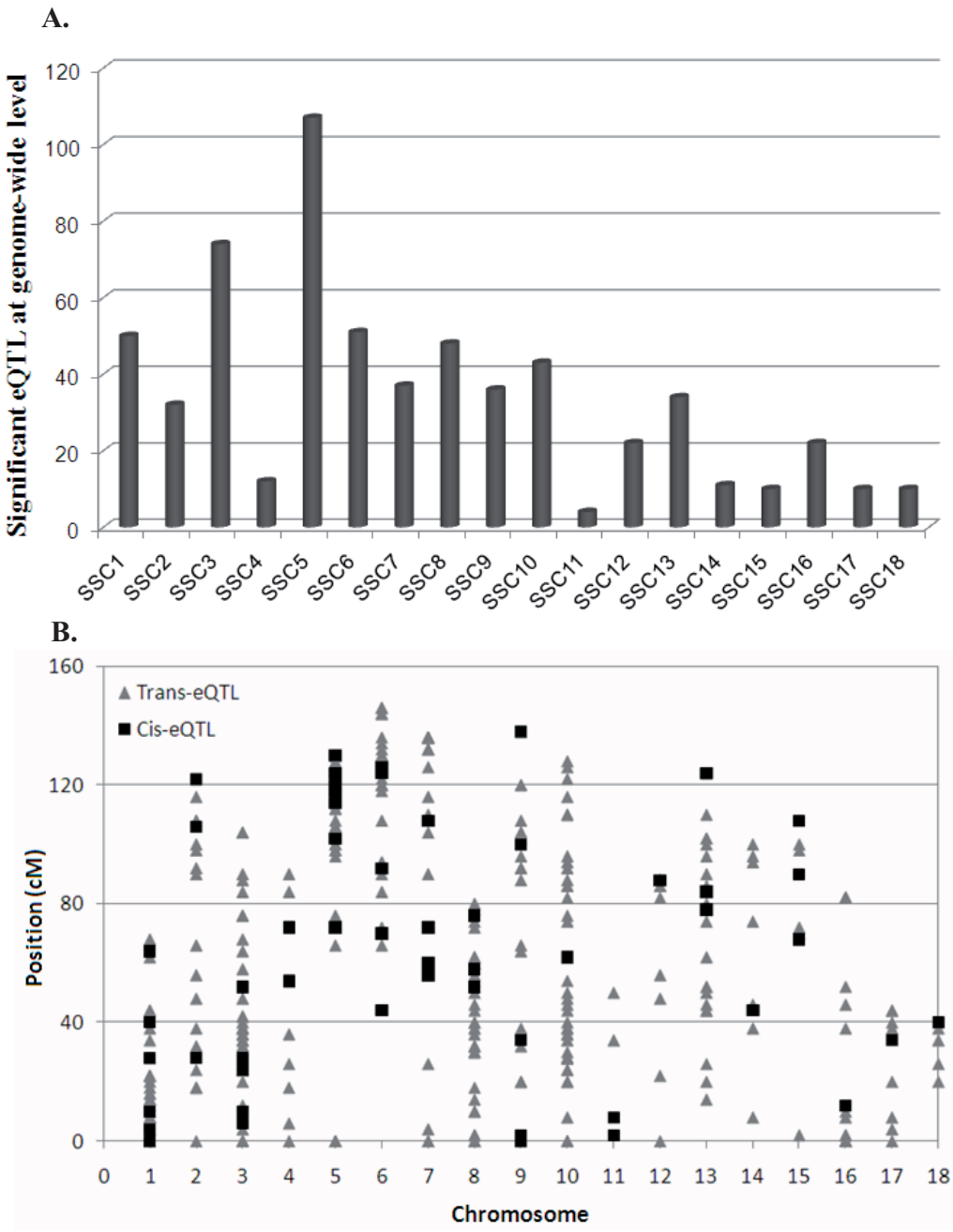
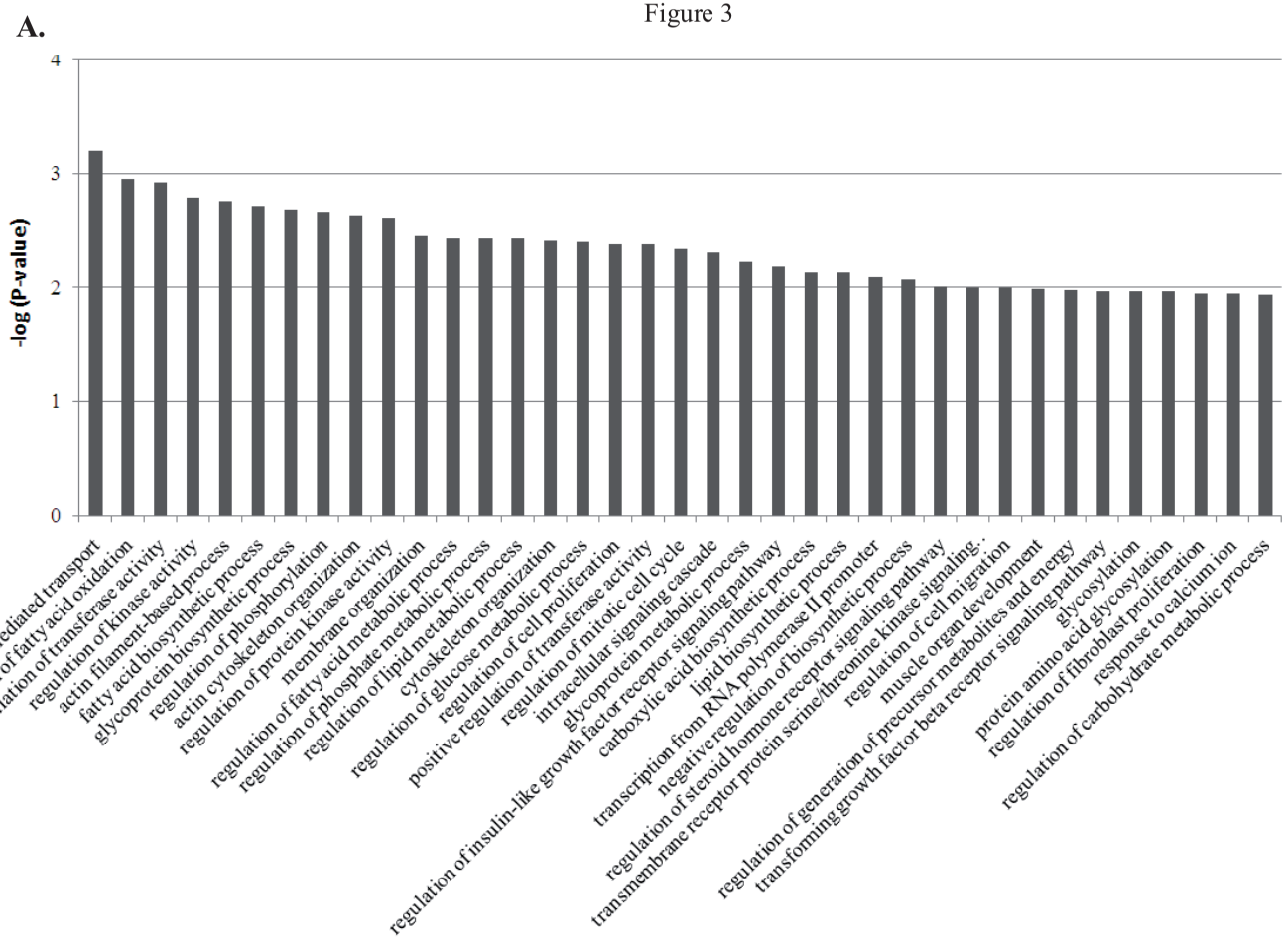


Figure 1: **(A)** Number of eQTL detected at genome-wide level in each chromosome. **(B)** Distribution of all genome-wide significant eQTL across the pig genome; ■ corresponded to *cis*-acting eQTL and ▲ corresponded to *trans*-acting eQTL.

Figure 2

Figure 2: Pig genome regions with high concentration of eQTL (eQTL hotspot).

Figure 3


A.
-log (P-value)
0
1
2
3
4
vesicle-mediated transport
regulation of fatty acid oxidation
regulation of transferase activity
regulation of kinase activity
actin filament-based process
fatty acid biosynthetic process
glycoprotein biosynthetic process
regulation of phosphorylation
actin cytoskeleton organization
regulation of protein kinase activity
membrane organization
regulation of fatty acid metabolic process
regulation of phosphate metabolic process
regulation of lipid metabolic process
cytoskeleton organization
regulation of glucose metabolic process
regulation of cell proliferation
positive regulation of transferase activity
regulation of mitotic cell cycle
intracellular signaling cascade
glycoprotein metabolic process
regulation of insulin-like growth factor receptor signaling pathway
carboxylic acid biosynthetic process
transcription from RNA polymerase II promoter
lipid biosynthetic process
negative regulation of biosynthetic process
regulation of steroid hormone receptor signaling pathway
transmembrane receptor protein serine/threonine kinase signaling...
regulation of cell migration
muscle organ development
regulation of generation of precursor metabolites and energy
transforming growth factor beta receptor signaling pathway
glycosylation
protein amino acid glycosylation
regulation of fibroblast proliferation
response to calcium ion
regulation of carbohydrate metabolic process

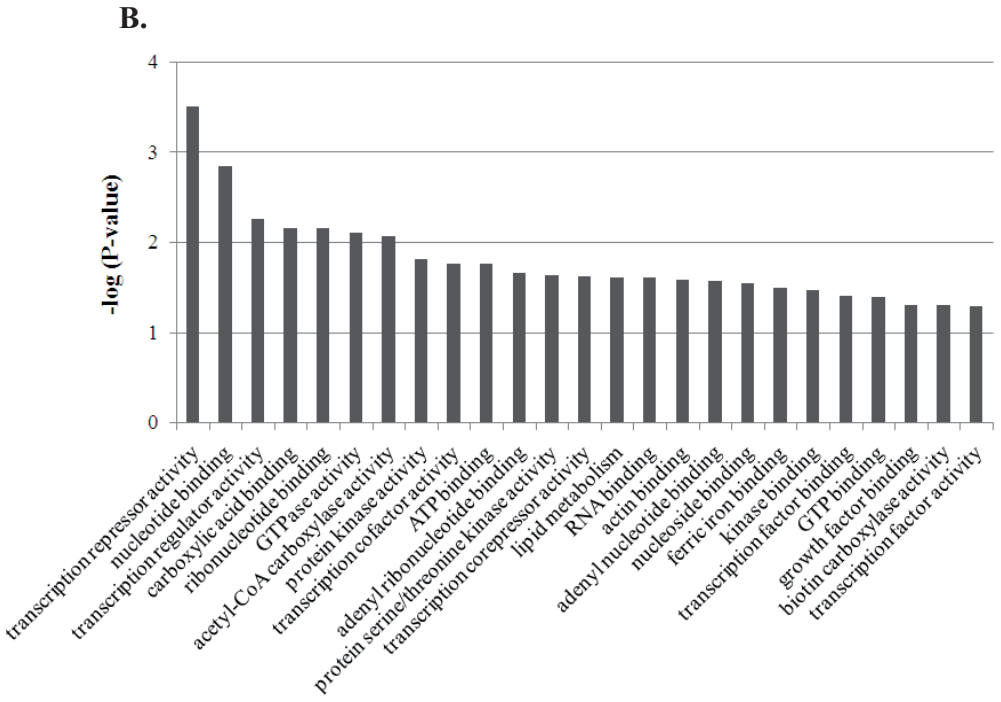


Figure 3: Global GO enrichment analysis performed with DAVID software on the list of eQTL-regulated genes. **(A)** List of most significant GO terms represented in biological process category ($P<0.01$). **(B)** List of most significant GO terms represented in molecular function category ($P<0.05$).

3.3. - Functional and association studies on the pig *HMGCR* gene, a cholesterol-synthesis limiting enzyme

Cánovas, A., Quintanilla, R., Gallardo, D., Díaz, I., Noguera, J.L., Ramírez, O., Pena, R.N

***Animal*. 4:224-233. 2010**

Animal (2010), **4:2**, pp 224–233 © The Animal Consortium 2009
doi:10.1017/S1751731109991145

animal

Functional and association studies on the pig *HMGCR* gene, a cholesterol-synthesis limiting enzyme

A. Cánovas[1], R. Quintanilla[1], D. Gallardo[2], I. Díaz[3], J. L. Noguera[1], O. Ramírez[2] and R. N. Pena[1†]

[1]*IRTA, Genètica i Millora Animal, 191 Alcalde Rovira Roure, 25198 Lleida, Spain;* [2]*Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain;* [3]*IRTA, Tecnologia dels Aliments, Finca Camps i Armet, 17121 Monells, Spain*

(Received 21 January 2009; Accepted 12 August 2009; First published online 23 October 2009)

The 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) is the rate-limiting enzyme in the biosynthesis of cholesterol. We have studied the role of the HMGCR *gene in pig lipid metabolism by means of expression and structural analysis. We describe here the complete coding region of this gene in pigs and report two synonymous single nucleotide polymorphisms in the coding region. We have, additionally, studied the association of one of these polymorphisms (*HMGCR:c.807A>C*) with several lipid deposition- and cholesterol-related traits in a half-sib population generated from a commercial Duroc line, showing in some families a positive relationship of* HMGCR:c.807A *allele with serum low-density lipoprotein (LDL)-bound cholesterol and triglyceride levels, and also with intramuscular fat (IMF) content of* gluteus medius *muscle. We have also assessed the expression levels in muscle and in liver from 68 Duroc individuals corresponding to the most extreme animals for the analysed traits. Liver* HMGCR *expression correlated negatively with the serum high-density lipoprotein (HDL) levels, carcass lean percentage and stearic acid content, while muscle expression correlated also negatively with the carcass lean percentage, stearic and linoleic acids content, but showed a positive correlation with the serum lipid cholesterol (HDL, LDL and total cholesterol), IMF and muscle oleic and palmitic fatty acid content. With this information, we have performed an association analysis of expression data with lipid metabolism phenotypic levels and the* HMGCR *genotype. The results indicate that* HMGCR *expression levels in muscle are different in the two groups of pigs with extreme values for fat deposition and total cholesterol levels, and also between animals with the different* HMGCR *genotypes.*

Keywords: meat quality, fatty acid, intramuscular fat, gene expression, porcine

Implications

This is the first study linking pig 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR)* expression levels with carcass and meat quality traits. This gene, which codes for the limiting enzyme in the *de novo* synthesis of cholesterol, regulates plasma lipid levels and cholesterol uptake in humans. Our study suggests that it also affects local lipids uptake and deposition in pig peripheral tissues, as muscle *HMGCR* expression correlates with carcass lean percentage, muscle intramuscular fat (IMF) content and individual fatty acid composition. Moreover, we describe a polymorphism in the coding region of this gene significantly associated with plasma lipid levels and IMF content.

Introduction

The rate limiting step of the *de novo* biosynthesis of cholesterol occurs at the 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) catalysed step, which converts HMG-CoA into mevalonate (Friesen and Rodwell, 2004). Expression of this gene represses the expression of low-density lipoprotein (LDL) receptors in liver, altering the serum cholesterol levels (reviewed in Kajinami *et al*., 2004). In agreement with this, polymorphisms on the human *HMGCR* gene have been associated with changes in the plasma cholesterol and triglyceride (TG) levels (Tong *et al*., 2004). Also, mutations resulting in alternative splicing of human exon 13 are a source of serum LDL-cholesterol variation (Burkhardt *et al*., 2008). Moreover, differences in liver HMGCR activity were reported between pigs selected for thin or thick backfat thickness at 80 Kg body weight for more than 14 generations (Pond and Mersmann, 1996).

This gene has been extensively studied in humans as HMGCR inhibitors, known as statins, are the standard treatment for hypercholesterolemic patients. Its sequence and promoter transactivation has also been characterised in model animals such as rat (Osborne *et al*., 2004), mouse

(Datta *et al*., 2006) and opossum (Chan *et al*., 2008) with a view to studying the genetics of serum cholesterol and TG variation and their consequences over atherosclerosis. The pig has also been used as a model for atherogenesis, obesity and lipoprotein and cholesterol metabolism due to anatomic and pathogenic similarities with humans (reviewed in Pond and Mersmann, 1996). HMGCR activity is regulated at both post-transcriptional and post-translational levels. Regarding the pig, hepatic HMGCR enzyme activity seems to be regulated throughout development but inhibition of hepatic reductase activity by dietary cholesterol is particularly evident in young pigs (McWhinney *et al*., 1996).

In addition to the pathological consequences, high-serum lipid concentrations can also contribute to increased fat deposition in adipose tissue and muscle. A classical example of this is type-2 diabetes mellitus-related obesity (Wilding, 2007). However, to our knowledge the relationship between *HMGCR* polymorphism, expression levels and fat deposition has not yet been attempted in humans or in any other species.

Following years of intensive selection for leaner pigs, in the last decade the market has shifted to an increasing demand for better quality pork. Two parameters, intramuscular fat (IMF) and fatty acid composition, have a marked effect over several meat quality traits such as tenderness, water holding capacity, juiciness and flavour (Olsson and Pickova, 2005). A number of genes with a physiological role on lipid transport and metabolism have been proposed as candidate genes for fat deposition in pigs (Davoli and Braglia, 2007) and some of these have been integrated in pig industry's selection schemes (van der Steen *et al*., 2005). With all the above in mind and in order to contribute to the knowledge of genes that control cholesterol and lipid deposition in pigs, we have focussed the present work on the *HMGCR* gene as a source of genetic variation for traits related to serum lipid levels and fat deposition in pigs. With this aim, we have analysed the level of *HMGCR* expression in muscle and liver in two groups of commercial Duroc pigs displaying extreme values for carcass fat deposition and serum and muscle lipid traits. We have also characterised the pig *HMGCR* coding region, where we have described two polymorphisms. Subsequently, we have studied the relationship between liver and muscle expression data, *HMGCR* genotype and several serum and muscle lipid parameters.

Material and methods

Animal material and phenotypic data

Animals came from a high-IMF commercial Duroc line used in the production of fine quality cured ham. An experimental population was generated on the basis of a half-sib design, by mating five parental boars with 385 females and taking only one male offspring per litter to a total of 385 animals as described before (Gallardo *et al*., 2008). These animals were castrated and kept under normal intensive conditions, being all subjected to the same management procedures at IRTA-CCP control station.

Two blood samples were taken at approximately 45 and 190 days of age to measure total serum cholesterol (TC), LDL, high-density lipoprotein (HDL) and TG concentrations (Gallardo *et al*., 2008). Animals were slaughtered according to a commercial protocol around 190 days of age (approximately 122 kg of live weight). Samples of 200 g *gluteus medius* muscle were collected immediately after slaughter for meat analyses carried out at the Centre of Meat Technology (*CTC-IRTA*). Analyses of *gluteus medius* lipid components included the determination of percentage of IMF, cholesterol content and fatty acid composition (in the C12 to C22 interval) were performed as follows: IMF content was determined by Near Infrared Transmittance (NIT; Infratec® 1625, Tecator AB, Hoganas, Sweden) (Plastow *et al*., 2005, Pinhu *et al*., 2008) fatty acid composition was analysed by gas chromatography of methyl esters as described in (Mach *et al*., 2006), and cholesterol content was measured following (Cayuela *et al*., 2003). The most relevant phenotypic traits analysed in this study, along with their abbreviations and statistics in the studied Duroc population, are shown in Table 1 (complete data set was finally obtained from 338 animals). Differences between group means were computed by a two-tailed *t*-test. All experimental procedures were approved by the Ethics Committee for Animal Experimentation of IRTA.

Isolation of DNA and RNA samples

Liver and *gluteus medius* muscle samples were collected at slaughter, snap frozen in liquid nitrogen and stored at −80°C for subsequent RNA analysis. To isolate total RNA, samples were ground with mortar and pestle in liquid nitrogen and homogenised with a mechanical rotor. RNA was isolated by the acid phenol method (Chomczynski and Sacchi, 1987) using the RiboPure kit (Ambion, Austin, TX, USA). DNA was isolated from blood samples using standard procedures (Sambrook and Russell, 2001).

Characterisation of coding region and polymorphisms

Primer design was carried out with the Primer Express software (Applied Biosystems, Foster City, CA, USA) using pig sequences as a template when available or human cDNA sequences that displayed high homology with other available mammalian sequences. For the amplification of the coding region, a total of five primer sets were designed (Table 2). Total RNA (0.5 to 1 μg) was retrotranscribed to cDNA using an Oligo(dT) primer and MMuLV RT enzyme (Fermentas, Glen Burnie, MD, USA) at 37°C for 1 h (Sambrook and Russell, 2001). PCR reactions were carried out in a PTC-100 thermocycler (MJ Research, Waltham, MA, USA) in a volume of 25 μl with 1 × buffer, 200 μM dNTP mix, 2.0 mM $MgCl_2$, 400 nM of each primer, 1U of Taq polymerase (ECOGEN SL, Barcelona, Spain) and 0.5 μl of cDNA. Thermal profile was as follows: denaturation 5 min at 95°C; 35 cycles of 15 s at 95°C, annealing 30 s at temperature indicated in Table 2, 40 s at 72°C; final extension 5 min at 72°C. PCR products were directly sequenced using the BigDye Terminator Sequencing kit v3.1 (Applied Biosystems) in an ABI-3100 capillary electrophoresis system (Applied Biosystems). Sequences obtained were edited

Table 1 *Mean (s.e.) of phenotypic records in blood, carcass and muscle samples for the whole Duroc pig population and for the two extreme groups of animals selected for HIGH and LOW values for cholesterol and lipid metabolism traits*

	Total population $n = 338$	HIGH group $n = 34$	LOW group $n = 34$
Serum lipids – 45 days			
TC	77.09^{a,b} *(0.65)*	82.06^{a} *(0.74)*	72.22^{b} *(1.16)*
HDL	30.40 *(0.34)*	31.88 *(0.38)*	28.61 *(0.53)*
LDL	37.86 *(0.43)*	40.46 *(0.48)*	35.60 *(0.68)*
TG	43.99 *(0.90)*	48.63 *(1.06)*	39.94 *(1.64)*
Serum lipids – 190 days			
TC†	124.60^{A} *(1.03)*	158.24^{B} *(1.40)*	105.29^{C} *(1.63)*
HDL†	51.77^{A} *(0.53)*	61.62^{B} *(0.55)*	43.40^{C} *(0.42)*
LDL†	62.46^{A} *(0.90)*	83.14^{B} *(1.07)*	51.44^{C} *(1.53)*
TG†	51.37^{A} *(1.52)*	67.18^{B} *(1.27)*	48.94^{A} *(1.44)*
Carcass traits			
Carcass weight†	94.58^{A} *(0.59)*	101.13^{A} *(0.58)*	86.19^{B} *(0.48)*
Lean %†	40.84^{A} *(0.23)*	38.30^{A} *(0.25)*	45.81^{B} *(0.26)*
BFTr†	25.12^{A} *(0.18)*	27.57^{B} *(0.17)*	21.00^{C} *(0.12)*
HFT†	25.72^{A} *(0.20)*	28.50^{B} *(0.19)*	21.03^{C} *(0.13)*
Meat fat traits of GM muscle			
Cholesterol	65.20^{A} *(0.51)*	63.65^{A} *(0.60)*	66.11^{A} *(0.61)*
IMF†	5.18^{A} *(0.05)*	7.24^{B} *(0.11)*	3.58^{C} *(0.10)*
% Myristic acid	1.38^{A} *(0.01)*	1.55^{B} *(0.01)*	1.20^{C} *(0.01)*
% Palmitic acid†	23.35^{A} *(0.07)*	24.62^{B} *(0.07)*	22.20^{C} *(0.05)*
% Palmitoleic acid	2.83^{A} *(0.02)*	3.18^{B} *(0.03)*	2.48^{C} *(0.03)*
% Stearic acid†	11.29^{A} *(0.06)*	11.89^{B} *(0.06)*	11.02^{C} *(0.06)*
% Oleic acid†	34.96^{A} *(0.23)*	38.60^{B} *(0.24)*	30.57^{C} *(0.15)*
% Vaccenic acid	4.06^{A} *(0.02)*	3.99^{A} *(0.02)*	4.03^{A} *(0.02)*
% Linoleic acid†	14.92^{A} *(0.20)*	10.88^{B} *(0.21)*	19.30^{C} *(0.12)*
% Arachidonic acid	3.25^{A} *(0.07)*	1.76^{B} *(0.09)*	5.04^{C} *(0.04)*

TC = total serum cholesterol; LDL = low-density lipoprotein; HDL = high-density lipoprotein; TG = serum triglyceride levels; BFTr = backfat thickness measured between the 3rd and 4th ribs; HFT = ham fat thickness; GM = *muscle gluteus medius*; IMF = intramuscular fat.
†Traits used in the principal component analysis to select animals for the HIGH and LOW groups. Different superscripts denote significant differences between group means (uppercase: $P < 0.001$; lowercase: $P < 0.05$).

using the Sequencing software (Applied Biosystems) and aligned with the ClustalW program (Chenna *et al.*, 2003).

Polymorphism genotyping

To genotype the *HMGCR:c.807A > G* polymorphism, primers were designed on exon eight and intron eight of the *HMGCR* gene (SNP1_A > G_F and SNP1_A > G_R; Table 2). Genotyping was performed by restriction enzyme cleavage of the 650 bp-long PCR product with *Hha*I (Fermentas). The G allele (two fragments of 450 and 200 bp) but not the A allele was cleaved and the fragments resolved in a 1.5% agarose gel.

Microarray hybridisation design

A differential expression study comparing animals with extreme characteristics for the lipid metabolism traits was carried out in our commercial Duroc population by means of hybridisation in oligo-based microarrays. With this purpose, from the complete set of phenotypic data available for the 385 Duroc pigs, the most important phenotypes (13 traits; see Table 1) were selected to perform a principal component analysis by means of PRINCOMP procedure of SAS (SAS Institute Inc., Cary, NC, USA). The first principal component, which accumulated 30.7% of total phenotypic variance, was defined by several fatness and serum cholesterol measures, together with IMF, monounsaturated and saturated fatty acid content in *gluteus medius*, and, in opposite direction, by *gluteus medius* polyunsaturated fatty acid content and carcass lean percentage. We used the values from the first principal component as a selection index to rank all studied individuals. Animals at both extremes of the ranking were selected to form the HIGH and LOW groups (34 animals per group).

Thus, RNA from 68 samples of *gluteus medius* muscle was individually hybridised on GeneChip Porcine Genome arrays (Affymetrix, Boston, MA, USA). Expression data were normalised using the Robust Multichip Average (RMA) algorithm. Expression values corresponding to the Ssc.16088.1.S1 probe (representing the porcine *HMGCR* gene) were extracted and analysed as explained below.

Quantitative PCR analysis (qPCR)

In order to study the expression of *HMGCR* gene in liver, and also to validate the muscle *HMGCR* microarray expression data in a subset of muscle samples, we carried out real-time qPCR essays in an ABI-7500 device (Applied Biosystems) in a final volume of 5 μl. Primers and TaqMan probes were designed using the Primer Express v2.0 software (Applied Biosystems; Table 2). The *HMGCR* cDNA was amplified in a

Table 2 *Primers used for the characterisation and genotyping of the pig* HMGCR *gene*

Primer name	Sequence 5′ → 3′	Position†	T°ann‡	Size§
F1	CTAGTGAAACCAGGAGGATC	E1/E2	58	590
R1	GTGCCAACTCCAATCACAAG	E6		
F2	GAATATTGCTCGTGGAATGG	E6	58	600
R2	GATTCCGTCTCTGCTTGTTC	E10		
F3	GTTATTACCCTAACTTTGGCTC	E10	58	600
R3	CATACCAAGGAGTAATTATAGTC	E12/13		
F4	TCGCCGACAGTTACTTTCC	E12	58	620
R4	CCCTCTATCCAGTTTACAGC	E16		
F5	GGAGATTCTGGCAGTCAGTG	E16	58	650
R5	AAGCCCGTGTTTCTGTTCAG	E20		
SNP1_A > G_F	CAAATCCTGTTACTCAGAGAG	E8	56	650
SNP1_A > G_R	CAGGAGCATAGCGTGTTATG	I9		
HPRT_E6_F	AAGATGGTCAAGGTTGCAAGCT	E6		
HPRT_E6_R	ATTTCAAATCCAACAAAGTCTGGTCTA	E7	60	83
HPRT_E6 probe	FAM-TGGTGAAAAGGACCCCTCGAAGTGTTG-TAMRA	E6/E7		
RPL32_F	CACCAGTCAGACCGATATGTCAA	E1		
RPL32_R	CGCACCCTGTTGTCAATGC	E2	60	70
RPL32 probe	FAM-TAAGCGGAACTGGCGGAAACCCA-TAMRA	E1/E2		

†Position according to exon/intron distribution in human gene.
‡Annealing T° used in the PCR reactions.
§Size in base pair of the amplified product.

PCR reaction containing 200 nM primers (F3 and R2; Table 2) and 1× SYBR Green Master mix (Applied Biosystems). For the two endogenous controls (*HPRT* and *RPL32*), PCR reaction contained 300 nM primers, 200 nM TaqMan probe and 1× Universal Taqman Master Mix (Applied Biosystems). The following thermal profile was used for all reactions: 10 min at 95°C, 40 cycles of 15 s at 93°C, 1 min at 60°C and 30 s at 72°C, followed by a quick denaturation at 95°C × 5 min plus a slow ramp to 30°C to generate a dissociation curve to control the specificity of the amplified product. In order to quantify and normalise the expression of the samples we used the ΔΔCt method (Yuan *et al*., 2006).

Statistical analyses

Association analysis of polymorphisms detected in the HMGCR gene. Association studies of polymorphism detected in the *HMGCR* gene with the complete set of available phenotypic data from the experimental Duroc population ($n = 338$; traits detailed in Table 1) were carried out by means of the GLM procedure of SAS (SAS Institute Inc.). The model used for these analyses was:

In the whole-population analysis:

$$y_{ijklm} = \mu + f_i + p_j + q_k + g_l + e_{ijklm},$$

In the within-family analysis:

$$y_{ijlm} = \mu + f_i + p_j + g_l + e_{ijlm},$$

where y_{ijklm} and y_{ijlm} represents phenotypic observation for each individual; μ is the overall mean; f_i is the systematic effect farm of origin, with three levels (only for serum lipid concentrations at 45 days); p_j is the effect batch of fattening, with four levels; q_k is the effect of belonging to the *j-th* half-sib family, with five levels; g_l is the effect of individual genotype (three levels, the two homozygous and the heterozygous); e_{ijklm} and e_{ijlm} are the residual effects.

In all analysed traits, least-squares (LS) means for the three genotypes were computed and compared between them.

Finally, an adjustment for multiple testing was carried out by Bonferroni correction i.e. as a solution to $P_{exp} = 1-(1-P)^N$, where P_{exp} is the experiment-wide P value and N is the number of independent analyses. Nevertheless, an important correlation between performed association tests is expected due to correlation between analysed traits. The fact that Bonferroni correction is overly conservative under these circumstances must be taken into account when interpreting data.

HMGCR differential expression and association with HMGCR *genotypes*

HMGCR expression data ($n = 68$) from both muscle (microarray data) and liver (qPCR results) were used to investigate the relationship between the *HMGCR* expression level and genotype.

An initial differential expression analysis was performed with the following model:

$$y_{jklmn} = \mu + p_j + q_k + g_l + t_m + e_{ijklmn},$$

where y_{jklm} represents the individual *HMGCR* expression in muscle or liver; t_m represents the effect of belonging to one

of two groups (HIGH- and LOW-lipid metabolism) *and* p_j, q_k, g_l and e_{jklmn} are the same effects as in the association analysis for the *HMGCR* polymorphisms. LSmeans for the group and genotype effects were computed.

Correlation study between HMGCR *gene expression and phenotypic data*

In order to further investigate the relationship between lipid metabolism and *HMGCR* expression, we performed a correlation study between the phenotypes analysed (Table 1) and *HMGCR* expression levels both in liver and muscle. The correlation study was performed using the CORR procedure of SAS (SAS Institute Inc.) after adjusting both, phenotypes and expression levels, for the environmental significant effects considered in previous analyses. In order to test possible differences between groups, two correlation analyses were carried out: 1) a conjoint correlation analysis considering all data set ($n = 68$ for each correlation) after correcting for all environmental factors including group, and 2) a within-group (HIGH and LOW lipid metabolism) correlation analyses ($n = 34$ for each correlation) after correcting for environmental factors.

Results

Identification and genotyping of polymorphisms in the porcine HMGCR *coding region*

We successfully amplified 2691 bp of the pig *HMGCR* cDNA, which included the entire coding region, in five partially overlapping PCR fragments (submitted to GenBank with accession number EU726797). This sequence is identical to the reference sequence for porcine *HMGCR* cDNA (NM_001122988.1) but for two synonymous polymorphisms we identified in our commercial Duroc animals: *HMGCR:c.807A > G* and *HMGCR:c.2247C > T*, situated in putative exons 9 and 17, respectively (Figure 1). For the first polymorphism, the studied Duroc population exhibited an intermediate allele frequency for allele G ($q = 0.33$). Moreover, three of the five founder sires were heterozygous for this polymorphism. We also genotyped this single nucleotide polymorphism (SNP) in a number of animals from different breeds (Table 3). Allele frequencies varied widely between populations: white pig breeds displayed low (Landrace) and intermediate (Large White, Pietrain) allelic frequencies for allele G, while it did not segregate in three other populations where either the A allele (Iberian and a Duroc line of another origin) or the G allele (Meishan pigs) was fixed. The second polymorphism could not be genotyped in a larger set of animals as the lack of information on the pig genomic sequence around the mutation prevented setting up a genotyping protocol on genomic DNA.

Association between HMGCR *genotype and lipid metabolism traits*

Table 4 shows results corresponding to the association study between the *HMGCR:c.807A > G* genotype and several serum and meat lipid composition traits, at both population and within-family levels. Concerning the serum lipid concentrations, the A allele was related to higher TC and LDL levels at 45 days in family 3, being these differences significant at experiment-wide level. A similar tendency, although not statistically significant ($0.5 < P < 0.1$; not indicated in tables), for LDL at 190 days was observed in family 3. Also an experiment-wide significant association of the A allele with higher LDL at 190 days was found in family 5, affecting also the TC levels and the ratio of HDL *v.* LDL cholesterol (in these two cases at nominal level). At nominal level of significance, the AA genotype showed higher serum TG concentrations at 45 days in the whole-population analysis. Except for family 4, this effect was not significant when a within-family analysis was conducted, although a similar tendency was observed in most cases.

As regards the lipid content of *gluteus medius* muscle, no experiment-wise significant associations were obtained, but at population level, we observed (nominal) significantly

Table 3 *Allelic frequency for the* HMGCR:c.807A > G *polymorphism in several porcine breeds*

		HMGCR:c.807A > G				
Breed	*n*	AA	AG	GG	p(A)	q(G)
Landrace	50	38	12	–	0.88	0.12
Pietrain	44	12	20	15	0.43	0.57
Duroc†	47	47	–	–	1	0
Iberian	7	7	–	–	1	0
Large White	20	9	9	2	0.67	0.32
Meishan	17	–	–	17	0	1
Total	185	110	41	34	0.79	0.21

†Population of different origin from the analysed Duroc population.

Table 4 *Least-squares means for association study of* HMGCR:c.807A >G *genotype with serum and carcass and meat lipid traits*

	Duroc population $n = 338$			Family 1 $n = 43$			Family 2 $n = 86$			Family 3 $n = 73$			Family 4 $n = 75$			Family 5 $n = 61$		
	AA $n = 158$	AG $n = 151$	GG $n = 29$	AA $n = 3$	AG $n = 30$	GG $n = 10$	AA $n = 39$	AG $n = 38$	GG $n = 9$	AA $n = 55$	AG $n = 18$	GG $n = 0$	AA $n = 30$	AG $n = 37$	GG $n = 8$	AA $n = 31$	AG $n = 28$	GG $n = 2$
[§]*Serum lipids – 45 days*																		
TC	4.34	4.31	4.28	4.34	4.35	4.26	4.33	4.31	4.29	4.37^{A}	4.24^{B}	–	4.31	4.29	4.24	4.33	4.34	4.46
HDL	3.39	3.37	3.35	3.34	3.36	3.35	3.34	3.35	3.31	3.45	3.38	–	3.37	3.39	3.32	3.40	3.34	3.50
LDL	3.62	3.59	3.55	3.88	3.67	3.51	3.58	3.61	3.61	3.65^{A}	3.48^{B}	–	3.57	3.54	3.49	3.59	3.36	3.72
TG	3.75^{a}	3.69^{ab}	3.59^{b}	3.48	3.82	3.60	3.86	3.76	3.65	3.66	3.56	–	3.75^{a}	3.58^{b}	3.65^{ab}	3.67	3.74	3.61
HDL/LDL	0.25	0.24	0.20	0.58^{a}	0.32^{ab}	0.18^{b}	0.26	0.28	0.31	0.21	0.10	–	0.22	0.17	0.17	0.19	0.31	0.22
[§]*Serum lipid – 190 days*																		
TC	4.83	4.79	4.84	4.99	4.81	4.89	4.85	4.85	4.91	4.82	4.72	–	4.71	4.77	4.70	4.91^{a}	4.79^{b}	4.85^{ab}
HDL	3.94	3.91	3.96	3.93	3.89	4.05	3.98	3.93	4.04	3.94	3.90	–	3.87	3.90	3.73	3.96	3.92	3.98
LDL	4.13	4.08	4.15	4.37	4.09	4.20	4.09	4.15	4.22	4.12	3.97	–	3.95	4.04	4.01	4.28^{a}	4.07^{b}	4.11^{AB}
TG	3.87	3.83	3.81	4.19	3.93	3.71	3.96	3.94	3.86	3.71	3.60	–	3.81	3.91	4.06	3.86	3.71	4.07
HDL/LDL	0.19	0.17	0.2	0.44	0.20	0.15	0.11	0.23	0.17	0.18	0.07	–	0.08	0.14	0.29	0.32^{a}	0.15^{b}	0.13^{ab}
Carcass traits																		
Carcass weight	94.82	94.53	94.74	97.71	90.24	95.05	94.47	95.84	91.17	95.22	96.06	–	96.77	94.82	96.61	94.00	93.79	96.08
Lean %	40.65	40.94	39.33	43.67	40.17	39.06	41.25	41.30	38.64	40.82	42.01	–	39.97	41.08	40.04	41.04	40.90	41.24
Fat meat composition of GM muscle																		
Cholesterol	65.10	63.73	63.24	72.7	68.62	65.87	64.60	65.42	62.25	65.61	62.81	–	62.01	62.02	65.59	70.29	66.35	79.48
IMF	5.66^{a}	5.17^{b}	4.72^{b}	5.70	6.22	5.79	6.40	5.83	5.82	4.89	4.54	–	4.83	4.19	3.59	5.09	4.63	4.31
% Oleic acid	35.76	35.35	34.44	35.75	36.55	36.32	35.49	35.92	33.87	35.16	35.09	–	35.23^{a}	33.03^{b}	32.81^{ab}	34.30	34.81	34.74
% Linoleic acid	14.34	14.88	15.49	12.61	13.60	13.77	14.03	13.91	15.53	14.88	15.41	–	15.38	17.15	17.08	15.68	15.57	15.40
% Stearic acid	11.24	11.20	11.11	11.67	11.16	10.70	11.62	11.92	11.58	11.40	11.24	–	10.81	10.58	10.80	10.95	10.73	11.17
% Palmitic acid	23.27	23.19	23.04	24.57	23.41	23.59	23.76	23.79	23.46	23.24	23.04	–	23.00	22.67	22.69	22.76	22.97	22.84

TC = total serum cholesterol; TG = serum triglyceride levels; LDL = low-density lipoprotein; HDL = high-density lipoprotein; IMF = intramuscular fat; GM = *muscle gluteus medius*.
[§]Means of serum lipid traits with log transformed data. Different superscripts denote significant ($P < 0.05$) differences at nominal level. Superscripts in uppercase letter denote significant differences at experiment-wide level ($P_{exp} < 0.05$), corresponding to a nominal $P < 0.006$.

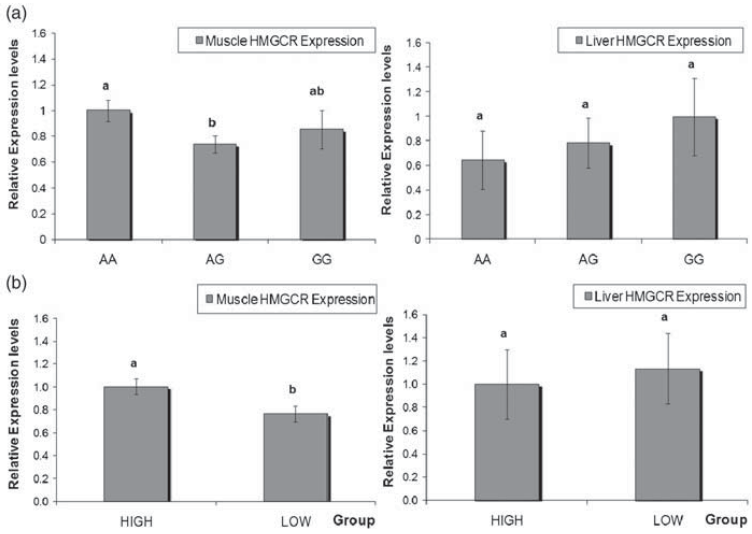


Figure 2 Graphical display of relationship of pig *HMGCR* muscle and liver expression levels with (a) *HMGCR:c.807A > G* genotype ($P < 0.05$; n(AA) = 29, n(AG) = 33, n(GG) = 6) and (b) lipid metabolism phenotypic group ($P < 0.001$; n(HIGH) = 34; n(LOW) = 34). Mean values are represented. Error bars indicate s.e.

higher IMF in individuals carrying the A allele. The same trend was observed in all within-family analysis except in family 1, but these results did not reach the significance level. Concerning the fatty acid composition of *gluteus medius* muscle, the A allele was also significantly ($P < 0.05$) associated with higher oleic content and suggestively ($P = 0.06$; not indicated in tables) with lower linoleic percentage in family 4.

Association between HMGCR *genotype and expression levels*

Results from the association analyses between *HMGCR: c.807A > G* genotype and *HMGCR* mRNA levels in muscle and liver are shown in Figure 2a. These results revealed significant differences in muscle *HMGCR* expression levels between animals with genotypes AA and AG ($P < 0.05$), whereas individuals with genotypes GG (only six animals) showed intermediate expression levels and no differences from the other genotypes. In contrast, liver expression levels were uniform amongst genotypes (Figure 2a). The different behaviour of this gene in these two tissues was also reflected by a lack of correlation between liver and muscle *HMGCR* mRNA levels (data not shown).

HMGCR *gene expression in animals with extreme lipid metabolism and its relationship with phenotypic measures*

Figure 2b presents the comparison of *HMGCR* muscle and liver expression levels in two groups of Duroc pigs corresponding to the most extreme animals for 13 traits related to growth, fattening and cholesterol metabolism traits (HIGH and LOW groups; Table 1). Results indicate that *HMGCR* expression varied according both to the tissue type and to the group of analysis. In this sense, muscle *HMGCR* expression was 1.31-fold higher in animals from the HIGH than from the LOW group ($P < 0.001$). These differences were, however, specific of muscle, as liver expression levels did not differ between groups. We validated the microarray results of muscle *HMGCR* gene by qPCR of a subset of ten muscle samples from each group. The expression ratios between HIGH and LOW groups remained in the same sense in both assays (microarray and qPCR) obtaining identical significance level ($P < 0.001$) with both datasets.

In addition, we studied the within-group correlation coefficients between *gluteus medius* and liver *HMGCR* expression levels and serum, carcass and muscle phenotypes (Table 5). With some punctual exceptions, we obtained no significant differences between both sets of within-group (HIGH and LOW) correlation coefficients, and neither with the global correlation coefficients obtained with pooled data set (results not shown). In all cases, muscle *HMGCR* expression was positively correlated with serum TC and HDL levels, and also with IMF, oleic and palmitic fatty acid content in *gluteus medius*, and negatively correlated with linoleic and stearic fatty acid content, and with lean percentage. In contrast, a positive correlation between serum LDL levels and muscle expression levels was detected only in the HIGH group.

As far as *HMGCR* expression in liver is concerned, it shared with muscle expression the negative correlation with stearic acid content and lean carcass percentage, but, in contrast, a negative interaction with serum HDL levels was also detected. Liver *HMGCR* mRNA levels displayed significant correlation with no other traits analysed.

Table 5 *Correlations between muscle and liver HMGCR expression values and serum lipid concentrations, carcass traits and muscle fat content and composition*

	Lipid serum traits				Carcass traits			Meat composition traits					
	TC (190 days)	HDL (190 days)	LDL (190 days)	TG (190 days)	Carcass weight	Lean %	Chol GM	IMF	% Oleic acid	% Linoleic acid	% Stearic acid	% Palmitic acid
Ms expression *HMGCR*	**H: 0.349*** **L: 0.374***	**H: 0.305*** **L: 0.325***	**H: 0.305*** L: −0.248	H: −0.125 L: 0.132	H: 0.112 L: −0.330	**H: −0.301*** **L: −0.335***	H: 0.106 L: 0.170	**H: 0.313*** **L: 0.326***	**H: 0.245*** **L: 0.267***	**H: −0.280*** **L: −0.292***	**H: −0.347*** **L: −0.399***	**H: 0.297*** **L: 0.314***
Lv expression *HMGCR*	H: −0.125 L: −0.311	**H: −0.386*** **L: −0.432***	H: −0.057 L: 0.014	H: −0.166 L: −0.101	H: 0.150 L: 0.275	**H: −0.271*** **L: −0.393***	H: −0.173 L: 0.301	H: 0.165 L: 0.137	H: −0.062 L: 0.045	H: −0.049 L: −0.069	**H: −0.437*** **L: −0.428***	H: 0.268 L: 0.031

Ms = muscle; Lv = liver; TC = total serum cholesterol; TG = serum triglyceride levels; Chol GM = cholesterol content in *muscle gluteus medius*; IMF = intramuscular fat; H = HIGH group; L = LOW group.
*Significant results ($P<0.05$) are reported in **bold**.

Discussion

Despite the crucial role of the microsomal membrane-bound glycoprotein HMGCR enzyme in determining serum cholesterol levels in humans (reviewed in Schmitz and Langmann, 2006), the information regarding this gene in pigs is sparse. Previous studies have focused on the regulation of HMGCR hepatic activity by dietary fat indicating that this enzyme is particularly inhibited by dietary cholesterol reductase activity in young pigs (McWhinney *et al*., 1996; Pond and Mersmann, 1996) and that the modulation of hepatic reductase activity is related, at least in guinea pigs, to the ratio of saturation to insaturation of the dietary fat (Fernandez and McNamara, 1991). Regarding the porcine *HMGCR* gene, restriction fragment length polymorphism (RFLP) polymorphisms in the porcine gene have been previously described by Southern blotting (Davis *et al*., 1995). Therefore, the position of these mutations was unassigned. Moreover, no study on the relationship between this gene and livestock production traits has yet been reported.

We report here the complete coding region for the pig *HMGCR* cDNA and describe two synonymous polymorphisms located at position +807 and +2247 of the coding region. The former of these polymorphisms (*HMGCR:c.807A > C*) was segregating in commercial pig breeds (Table 3) and showed significant whole-population associations, although at nominal level, with *gluteus medius* IMF content and serum TG at 45 days of age (Table 4). Similar effects have been described in humans, where analyses of an 830A/C polymorphism in human *HMGCR* gene resulted in higher levels of circulating low-density lipoprotein and TG in patients with the AA genotype (Tong *et al*., 2004). Considering the correction for multiple testing, the association tests with LDL-bound cholesterol reached the significance at experiment-wide level ($P < 0.006$) in family 3 (45 days) and family 5 (190 days), and also with TC in family 3 (45 days). The fact that Bonferroni correction is overly conservative when correlated association tests are performed must, however, be taken into account. In fact, other associations at nominal level were also observed but they were not consistent across all within-family analyses. These inconsistencies are probably caused by the limited sample size of some families, and the corresponding loss of power, but could also be due to family differences in the linkage disequilibrium phase of *HMGCR:c.807A > C* with the mutation responsible for these differences. More powerful experiments with a larger number of animals would be worthy to disentangle these inconsistencies, but the considerable management and economic difficulties involved in performing this kind of studies in commercial pig populations constitute an obstacle not easy to overcome.

We have also carried out a functional study on the pig *HMGCR* gene, investigating the relationship of *HMGCR* expression in liver and *gluteus medius* muscle with serum lipid levels and fat muscle content. Correlations obtained in this study suggested associations between *HMGCR* expression levels and phenotypic variability of several fatness traits and serum and meat lipid content, but differences were

not always consistent across tissues. Thus, carcass lean and stearic fatty acid percentages in *gluteus medius* were negatively correlated with both liver and muscle *HMGCR* expression, whereas HDL levels were negatively correlated with liver expression levels, possibly reflecting the role of the liver in the excretion of HDL-cholesterol particles, but exhibited a positive correlation with muscle *HMGCR* mRNA levels. In contrast, only muscle *HMGCR* expression showed a significant relationship with serum TC and LDL levels, IMF and other fatty acids (oleic, linoleic and palmitic) content in muscle. The positive correlation of cholesterol serum concentrations (TC, HDL and LDL) with muscle expression levels allow to hypothesise a relationship between changes in lipoprotein transport levels and muscle gene transactivation, as a results of the lipids uptaken from circulating lipoproteins (Nagao and Yanagita, 2008). The liver plays a critical role in orchestrating cholesterol biosynthesis and transport to and from peripheral tissues. In addition to this established role, we anticipated that local activity of this enzyme in peripheral tissues would potentially regulate local levels of lipid uptake and storing. In this context, results shown in this study suggest that high-serum lipid concentrations can also contribute to enhance fat deposition in adipose and muscle tissue.

The association analysis between structural and functional variations in *HMGCR* gene (Figure 2a) revealed a relevant association between *HMGCR* expression levels in *gluteus medius* muscle and *HMGCR:c.807A>G* genotype. In contrast, no significant differences for liver expression were found between genotypes. In humans, although a number of intronic SNP have been reported to lower the response to statins (Chasman *et al.*, 2004), structural mutations of this gene have not been associated to changes at liver expression level. Moreover, *HMGCR* mRNA levels are unaffected by changes in intronic splicing events (Medina *et al.*, 2008). Consistent to this, the slight decrease in liver *HMGCR* mRNA levels of hemizygous *HMGCR* mice resulted in no differences in cholesterol or TG content in liver or blood (Ohashi *et al.*, 2003). In this same line, divergent selection on pigs for serum total cholesterol (seven generations) was not associated with differences in hepatic or jejunum HMGCR activity or with *HMGCR* RFLP polymorphisms, at least in young animals, but rather with *CYP7* polymorphisms (Schoknecht *et al.*, 1994). In contrast with these results, in our Duroc pig population a 1.3-fold change in the muscular expression of this gene is associated with plasma and muscle lipid traits (Figure 2b), such as those defining the HIGH and LOW groups (Table 1). With respect to the LOW group, animals from the HIGH group had overall higher serum lipid concentrations, backfat thickness, IMF and monounsaturated fatty acid content (oleic acid), together with lower lean content and polyunsaturated fatty acid (linoleic and arachidonic acid) content. This is in agreement with a higher frequency of AA animals in the HIGH ($p(AA) = 0.53$) *v.* LOW ($p(AA) = 0.32$) group, and conversely for the lower frequency of AG animals ($p(AG) = 0.38$ in HIGH *v.* $p(AG) = 0.59$ in LOW). The most probable reason is that serum plasma lipids determine the amount of intracellular lipids in muscle and, therefore, the activation of steroids and other signalling molecules affecting the transactivation of the *HMGCR* promoter, which in turn affects the number of active LDLR (low-density lipoprotein receptor) and, subsequently, the uptake and accumulation of lipids in muscle (Telford *et al.*, 2003). Indeed, steroids (including cholesterol and mevalonate) promote a negative feed-back over the transcription of *HMGCR* and a positive enhancement of enzyme ubiquination and degradation (Ness and Chambers, 2000).

Conclusion

In summary, in this study, we have characterised the coding region and two synonymous polymorphisms for the cholesterol metabolism-related *HMGCR* porcine gene, one of them showing significant associations with serum LDL concentrations. Moreover, we have for the first time described relevant correlations between the *HMGCR* expression and fat deposition and meat composition traits, such as carcass lean percentage, IMF or meat fatty acid composition. Also, we have established a relationship between structural variations in this gene and expression levels in the *gluteus medius* muscle. As a whole, results obtained in this study allow us to conclude that the *HMGCR* gene displays a relationship with not only lipid serum traits but also with commercially important pig meat quality and production traits, and can be shown as an interesting candidate gene for gene assisted selection in pigs.

Acknowledgements

The authors thank Selección Batallé S.A. for providing the animal material and for their cooperation in the experimental protocol and D. Almuzara for technical support. A. Cánovas received a predoctoral scholarship from INIA. RN Pena received a contractual grant from INIA. This project has been financed by the Spanish Ministry of Education and Science (projects GEN2003-20658-C05-05/CICYT, AGL2002-04271-C03-02 and AGL2007-66707-C02-01).

References

Burkhardt R, Kenny EE, Lowe JK, Birkeland A, Josowitz R, Noel M, Salit J, Maller JB, Pe'er I, Daly MJ, Altshuler D, Stoffel M, Friedman JM and Breslow JL 2008. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. Arteriosclerosis, Thrombosis and Vascular Biology 28, 2078–2084.

Cayuela JM, Garrido MD, Banon SJ and Ros JM 2003. Simultaneous HPLC analysis of alpha-tocopherol and cholesterol in fresh pig meat. Journal of Agriculture and Food Chemistry 51, 1120–1124.

Chan J, Donalson LM, Kushwaha RS, Ferdinandusse S, VandeBerg JF and VandeBerg JL 2008. Differential expression of hepatic genes involved in cholesterol homeostasis in high- and low-responding strains of laboratory opossums. Metabolism 57, 718–724.

Chasman DI, Posada D, Subrahmanyan L, Cook NR, Stanton VP Jr and Ridker PM 2004. Pharmacogenetic study of statin therapy and cholesterol reduction. The Journal of the American Medical Association 291, 2821–2827.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG and Thompson JD 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research 31, 3497–3500.

Chomczynski P and Sacchi N 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Annals on Biochemistry 162, 156–159.

Datta S, Wang L, Moore DD and Osborne TF 2006. Regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase promoter by nuclear receptors liver receptor homologue-1 and small heterodimer partner: a mechanism for differential regulation of cholesterol synthesis and uptake. Journal of Biology and Chemistry 281, 807–812.

Davis AM, White BA and Wheeler MB 1995. Rapid communication: a *Dra*I restriction fragment length polymorphism at the porcine 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCoAR) locus. Journal of Animal Science 73, 310.

Davoli R and Braglia S 2007. Molecular approaches in pig breeding to improve meat quality. Brief Functional Genomic Proteomic 6, 313–321.

Fernandez ML and McNamara DJ 1991. Regulation of cholesterol and lipoprotein metabolism in guinea pigs mediated by dietary fat quality and quantity. Journal of Nutrition 121, 934–943.

Friesen JA and Rodwell VW 2004. The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. Genome Biology 5, 248.1–248.7.

Gallardo D, Pena RN, Amills M, Varona L, Ramirez O, Reixach J, Diaz I, Tibau J, Soler J, Prat-Cuffi JM, Noguera JL and Quintanilla R 2008. Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs. Physiological Genomics 35, 199–209.

Kajinami K, Takekoshi N, Brousseau ME and Schaefer EJ 2004. Pharmacogenetics of HMG-CoA reductase inhibitors: exploring the potential for genotype-based individualization of coronary heart disease management. Atherosclerosis 177, 219–234.

Mach N, Devant M, Diaz I, Font-Furnols M, Oliver MA, Garcia JA and Bach A 2006. Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. Journal of Animal Science 84, 3039–3048.

McWhinney VJ, Pond WG and Mersmann HJ 1996. Ontogeny and dietary modulation of 3-hydroxy-3-methylglutaryl-CoA reductase activities in neonatal pigs. Journal of Animal Science 74, 2203–2210.

Medina MW, Gao F, Ruan W, Rotter JI and Krauss RM 2008. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. Circulation 118, 355–362.

Nagao K and Yanagita T 2008. Bioactive lipids in metabolic syndrome. Progress in Lipid Research 47, 127–146.

Ness GC and Chambers CM 2000. Feedback and hormonal regulation of hepatic 3-hydroxy-3-methylglutaryl coenzyme A reductase: the concept of cholesterol buffering capacity. Procedings of the Society of Experimental Biology in Medicine 224, 8–19.

Ohashi K, Osuga J, Tozawa R, Kitamine T, Yagyu H, Sekiya M, Tomita S, Okazaki H, Tamura Y, Yahagi N, Iizuka Y, Harada K, Gotoda T, Shimano H, Yamada N and Ishibashi S 2003. Early embryonic lethality caused by targeted disruption of the 3-hydroxy-3-methylglutaryl-CoA reductase gene. Journal of Biology and Chemistry 278, 42936–42941.

Olsson V and Pickova J 2005. The influence of production systems on meat quality, with emphasis on pork. Ambio 34, 338–343.

Osborne AR, Pollock VV, Lagor WR and Ness GC 2004. Identification of insulin-responsive regions in the HMG-CoA reductase promoter. Biochemistry and Biophysics Research Communications 318, 814–818.

Pinhu L, Park JE, Yao W and Griffiths MJ 2008. Reference gene selection for real-time polymerase chain reaction in human lung cells subjected to cyclic mechanical strain. Respirology 13, 990–999.

Plastow GS, Carrión D, Gil M, Garcla-Regueiro JA, i Furnols MF, Gispert M, Oliver MA, Velarde A, Guàrdia MD, Hortós M, Rius MA, Sárraga C, Diaz I, Valero A, Sosnicki A, Klont R, Dornan S, Wilkinson JM, Evans G, Sargent C, Davey G, Connolly D, Houeix B, Maltin CM, Hayes HE, Anandavijayan V, Foury A, Geverink N, Cairns M, Tilley RE, Mormède P and Blott SC 2005. Quality pork genes and meat production. Meat Science 70, 409–421.

Pond WG and Mersmann HJ 1996. Genetically diverse pig models in nutrition research related to lipoprotein and cholesterol metabolism. In Advances in Swine in Biomedical Research (ed. ME Tumbleson and LB Schook), pp. 843–861. Plenum Press, New York.

Sambrook J and Russell D 2001. Molecular Cloning: A Laboratory Manual Cold Spring Harbor. Laboratory Press, Cold Spring Harbor, New York.

Schmitz G and Langmann T 2006. Pharmacogenomics of cholesterol-lowering therapy. Vascular Pharmacology 44, 75–89.

Schoknecht PA, Ebner S, Pond WG, Zhang S, McWhinney V, Wong WW, Klein PD, Dudley M, Goddard-Finegold J and Mersmann HJ 1994. Dietary cholesterol supplementation improves growth and behavioral response of pigs selected for genetically high and low serum cholesterol. Journal of Nutrition 124, 305–314.

Telford DE, Edwards JY, Lipson SM, Sutherland B, Barrett PH, Burnett JR, Krul ES, Keller BT and Huff MW 2003. Inhibition of both the apical sodium-dependent bile acid transporter and HMG-CoA reductase markedly enhances the clearance of LDL apoB. Journal of Lipid Research 44, 943–952.

Tong Y, Zhang S, Li H, Su Z, Kong X, Liu H, Xiao C, Sun Y and Shi JJ 2004. 8302A/C and (TTA)n polymorphisms in the HMG-CoA reductase gene may be associated with some plasma lipid metabolic phenotypes in patients with coronary heart disease. Lipids 39, 239–241.

van der Steen HAM, Prall GFW and Plastow GS 2005. Application of genomics to the pork industry. Journal of Animal Science 83, E1–E8.

Wilding JP 2007. The importance of free fatty acids in the development of Type 2 diabetes. Diabetes and Medecine 24, 934–945.

Yuan JS, Reed A, Chen F and Stewart CN Jr 2006. Statistical analysis of real-time PCR data. BMC Bioinformatics 7(85), 1–12.

3.4.- A polymorphism in the pig *HMGCR* promoter affects the transcriptional activation differently in liver and muscle cells

Cánovas, A., Quintanilla, R., Reecy, J.M., Marqués, M., Amills, M., Pena, R.N

Biochimie **(en preparación)**

A polymorphism in the pig *HMGCR* promoter affects the transcriptional activation differently in liver and muscle cells.

A. Cánovas[a], R. Quintanilla[a], J.M. Reecy[b], M.M. Marqués[c], M. Amills[d] and R.N. Pena[a,e]

[a]IRTA, Genètica i Millora Animal, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain,

[b]Animal Science, Iowa State University, 2255 Kildee Hall, Ames, 50011 IA, USA

[c]INDEGA, Universidad de León, Campus de Vegazana, 24071 León, Spain

[d]Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[e]Departament de Producció Animal, Universitat de Lleida, Campus ETSEA, 191 Av Alcalde Rovira Roure, 25198 Lleida, Spain.

Running head: *Transcriptional characterisation of the pig HMGCR promoter*

[1]To whom correspondence should be addressed.

E-mail address: romi.pena@prodan.udl.cat

Abstract

We have performed a functional and structural characterisation of the pig *HMGCR* proximal promoter. This gene encodes the rate-limiting enzyme in the biosynthesis of cholesterol and its muscular, but not hepatic, expression levels have been associated with meat quality-related traits in pigs. Expression analysis in six tissues related to lipid uptake, transport and storing showed the highest levels in the duodenum and subcutaneous fat, the lowest levels in heart and liver while skeletal muscle samples displayed intermediate levels. Sequencing of 520 bp of the proximal promoter revealed that in pigs this is a TATA-less promoter containing several transcription factor binding sites well conserved in other mammals. In addition, two SNP identified at positions -225 and -1 bp could potentially affect the binding site of bHLH and Ets-2 transcription factors, respectively. In the presence of several activators and inhibitors (insulin, 25-hydroxycholesterol and forskolin) the transcriptional activity of the pig *HMGCR* promoter differed significantly between liver (HepG2) and muscle (C2C12) cells and also between promoter genotype at the *HMGCR:g.-225G>T* SNP. This could be related to the differential CpG methylation pattern of the pig *HMGCR* promoter between muscle and liver, suggesting the existence of tissue-specific differences in the union of transcription factors to this promoter.

Keywords: Cholesterol, Gene expression, Methylation, Porcine, Regulatory region, Transcriptional activity

Abbreviations: CRE, cAMP response element; HMGCR, 3-hydroxy-3-methylglutaryl-CoA reductase; LRH-1, liver receptor homolog-1; SHP, small heterodimer partner; SRE, sterol response element; USF-2, upstream stimulatory factor-2

1. Introduction

The 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) enzyme catalyses the rate limiting step of cholesterol biosynthesis and is the target of statin therapy used with hypercholesterolemic patients [1]. Studying the regulation of this gene is of particular interest as it has been shown that polymorphisms in the regulatory regions [2-3] and splicing signals [4] that alter *HMGCR* mRNA levels are associated with changes in statin responsiveness and serum lipid concentration, therefore affecting the development of cholesterol-related pathologies in these patients. Moreover, in a previous work we have shown that *HMGCR* expression levels in muscle (but not in liver) differ in pigs with extreme values for fat deposition and serum cholesterol levels, and also between animals with different *HMGCR* genotypes at the *HMGCR:c.807A>G*, showing significant associations with lipid serum content and commercially important pig meat quality and production traits such as intramuscular fat and oleic acid content in muscle [5].

In other mammals, most regulatory elements concentrate in the first 350 bp of the proximal *HMGCR* promoter [4, 6]. In particular, expression of this gene in the rat promoter is under sterol and insulin control through the action of the upstream stimulatory factor-2 (USF-2) and the liver receptor homolog-1 (LRH-1) binding sites, both at position -327 [4, 6], a sterol response element (SRE) at -160, a Sp1-like site at -140, a cAMP response element (CRE) at -104 and NF-Y sites at -245 and -70. These sites are well conserved in the human promoter and are critical for maximal response to enhancing (LRH-1) and inhibitory (small heterodimer partner, SHP) signals *in vivo* [6] and *in vitro* [7]. In addition, other transcription factors of the basal transcription machinery have been shown by ChIP-PCR to bind this promoter [8].

In order to further the study of cholesterol metabolism in pigs we performed a structural and functional study of the proximal promoter of the pig *HMGCR* gene. With this aim, we have characterised the promoter sequence, where we have identified two polymorphisms, studying the transcriptional activity of *HMGCR* in promoter genotype at the *HMGCR:g.-225G>T* SNP in hepatic and muscular cell lines in response to activators and inhibitors. Also, we have identified two CpG-rich stretches in the promoter, both of which are differentially methylated in

muscle and liver samples. In addition, we have also studied the transcriptional activity of *HMGCR* in several tissues related to lipid metabolism.

2. Materials and methods

2.1. Animal material

Animal material came from a high intramuscular fat commercial Duroc line used in the production of fine quality cured ham [9]. Tissue samples were collected at slaughter, snap frozen in liquid nitrogen and stored at -80ºC until analysed. All experimental procedures were approved by the Ethics Committee for Animal Experimentation of IRTA.

2.2. Isolation of DNA and RNA

To isolate total RNA, liver, fat, heart, duodenum, *gluteus medius* and *longisimus dorsi* muscle samples were ground with mortar and pestle in liquid nitrogen and homogenised with a mechanical rotor. RNA was isolated by the acid phenol method [10] using the RiboPure kit (Ambion, Austin, TX). DNA was isolated from blood samples using standard procedures [11].

2.3. Promoter characterisation and polymorphism discovery

A 613 bp-long fragment of the *HMGCR* promoter and 5'UTR region was amplified from 20 ng of genomic DNA with primers described in Table 1 (HMGCRprom_F and HMGCRprom_R). Primer design was carried out using the Primer Express software (Applied Biosystems, Foster City, CA) using trace pig sequences as templates. Primers were positioned in sequences that displayed high homology with other mammalian sequences available. PCR reactions were carried out in a PTC-100 thermocycler (MJ Research, Waltham, MA) in a volume of 25 µl with 1x buffer, 200 µM dNTP mix, 2.0 mM $MgCl_2$, 400 nM of each primer, 1U of Taq polymerase (Ecogen SL, Spain) and 20 ng of genomic DNA. Thermal profile was as follows: 5 min at 95ºC; 35 cycles of 15 s at 95ºC, 30 s at 66ºC, 40 s at 72ºC; 5 min at 72ºC. PCR products were directly sequenced using the BigDye Terminator Sequencing kit v3.1 (Applied Biosystems) in an ABI-3100 capillary electrophoresis

system. Sequences were edited using the Sequencing software (Applied Biosystems) and aligned with the ClustalW program [12]. Transcription factor analysis of promoter regions was performed with the TESS web tool (http://www.cbil.upenn.edu/cgi-bin/tess/tess; [13]).

2.4. Estimates of evolutionary divergence between sequences

For across-species comparisons, alignment files built with ClustalW were used on MEGA4 [14]. Positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons. Evolutionary distances were computed using the Maximum Composite Likelihood method [15]. Nucleotide identities were calculated as (1 – distance).

2.5. Quantitative PCR analysis

Real-time PCR essays were performed with conditions and primers as described in [5]. We studied the mRNA expression profile of the porcine *HMGCR* in six tissues involved in lipid transport and metabolism: liver, fat, *gluteus medius* and *longissimus thoracis et lumborum* muscles, heart and duodenum. Relative expression levels were quantified and normalised using the ΔΔCt method with *hypoxanthine phosphoribosyl-transferase* (*HPRT*) and *ribosomal protein L32* (*RPL32*) genes as endogenous controls [16]. These two endogenous controls were previously checked with three more candidate reference genes, resulting the most stable genes across tissues [17].

2.6. Plasmid construction

A PCR product from pig genomic DNA containing 520 bp of *HMGCR* promoter and 93 bp of 5' untranslated region was amplified from animals with TT or GG genotype at the *HMGCR:g.-225G>T* polymorphic site using primers (HMGCRclon_F and HMGCRclon_R) indicated in Table 1. The resulting *HMGCR* PCR product was digested with *Hin*dIII and *Xho*I enzymes and cloned into the pGL3 basic vector (Promega, Madison, WI) using standard techniques.

2.7. In vitro promoter analysis

The response of this promoter to insulin, 25-hydroxycholesterol and forskolin was studied in two cell lines of hepatic (HepG2) and muscular (C2C12) origins. Both cell lines were grown in DMEM (Dulbecco's modified Eagle's) medium supplemented with 10% foetal bovine serum, penicillin/streptomycin and glutamine. To transfect the DNA, HepG2 ($2.6x10^5$) and C2C12 ($6x10^4$) cells were plated in 24-well dishes. After 24 h, cells were transfected with 0.5 µg of plasmid DNA/well using Lipofectamine LF2000 (Invitrogen, Carlsbad, CA) with the luciferase reporter vector plus a CMV-β-galactosidase expression construct [18] as a normalisation control. Five hours post-transfection, cells were washed with PBS and new medium was added with different treatments: insulin (0.1 µM, 1 µM and 10 µM), 25-hydroxycholesterol (5 µM, 12.5 µM and 25 µM) and forskolin (1 µM, 10 µM and 100 µM). After 24 h with treatments, cellular extracts were prepared using the Reporter Lysis Buffer and the luciferase activity was measured following manufactures procedure (Promega) in a TD20/20 luminometer (Turner Designs,Sunnyvale, CA). In the C2C12 differentiation experiment, single (myoblastic) cells were incubated 48 h in DMEM plus 2% horse serum before adding the treatments, in order to induce multinuclear myotubes through cell fusion. Values of luciferase activity were normalised by CMV-β-galactosidase activity. β-galactosidase assay was performed by adding a diluted cell extract sample to an equal volume of 2x Assay buffer that containing ONPG (o-nitrophenyl-galactopyranoside) (Promega). Samples were incubated at 37ºC x 30 min. The reaction was terminated by adding sodium carbonate and the absorbance was read at 420 nm with a spectrophotometer.

2.8. Statistical Analyses

Significance of differences between genotypes and treatments was determined by a pairwise t-test comparison of the means. Results were considered statistically significant at $P < 0.05$. Data were analysed with SAS (SAS Inst. Inc.,Cary, NC).

2.9. DNA methylation analyses

A total of 10 Duroc pigs with TT or GG genotype at the *HMGCR:g.-225G>T* polymorphic site were selected to perform DNA methylation analyses in two tissues (liver and muscle) using the bisulfite method [19]. Around 200 to 500 ng of purified genomic DNA from each sample was processed using EZ DNA Methylation kit (Zymo Research, Orange, CA). The presence of CpG islands in the pig *HMGCR* promoter was examined using MethPrimer software available at www.urogene.org/cgi-bin/methprimer [20]. The region of interest was amplified using primers design over the resulting methylated sequence using MethPrimer (HMGCRmeth_F and HMGCRmeth_R, Table 1). PCR amplification was performed with the QIAgen Multiplex PCR kit (QIAgen, Hilden, Germany). The following thermal profile was used for all reactions: 15 min at 95ºC; 35 cycles of 15 s at 95ºC, 90 s at 60ºC and 90 s at 72ºC; 60 s at 72ºC. Methylated PCR products were directly sequenced in duplicate and along both strands using the BigDye Terminator sequencing kit v3.1 (Applied Biosystems) in an ABI-3100 capillary electrophoresis system. Sequences were edited using the Sequencing software (Applied Biosystems) and the BiQ Analyzer software was used to perform quality control and to analyze the selective conversion of unmethylated cytosine to uracil by bisulfite treatment [21].

3. Results

3.1. Characterisation of the porcine HMGCR promoter region

Using sequence information from human, mouse and rat *HMGCR* promoters, we identified a clone from the trace archive of the NCBI database which corresponded to the pig *HMGCR* promoter. Primers were designed based on this sequence to amplify a fragment of 613 bp in animals from a commercial Duroc pig population, which included 520 bp of proximal promoter and 93 bp of the 5' untranslated region (GenBank accession number EU729728). Comparison of the pig promoter region with rat, mouse, cow and human sequences revealed an overall high level of conservation (72 to 86% sequence identity; Fig. 1) with conserved motives for several transcription factors. In particular,

the binding sites for the transcription factors USF-2 and LRH-1 (both at position -338), SRE (-176), CRE (-116), and NF-Y (-80 and -205), well-characterised in the human and rat promoters [4, 6-7], were conserved in sequence and relative position (Fig. 1).

As in humans and hamster [22-23], the pig *HMGCR* promoter is a TATA-less DNA polymerase II promoter. Unlike TATA-containing promoters, TATA-less promoters are characterised by having multiple start sites which generate a variety of transcripts that differ in the length of the 5'UTR [24]. Taking as reference the transcription initiation site recorded in Ensembl for the pig *HMGCR* gene, two SNP polymorphisms were identified in the pig promoter at positions -225 bp and -1 bp (Fig. 1) characterised by G>T and T>C substitutions, respectively. *In silico* analyses revealed that the distal polymorphism could potentially affect the binding of transcription factors of the basic helix-loop-helix (bHLH) family of myogenesis regulators. The more proximal mutation at position -1 alters a TTCCTT motif, which is an Ets-2 potential binding-site.

3.2. Expression of the pig HMGCR gene in six lipogenic tissues

In order to study the expression profile of the pig *HMGCR* gene, relative expression levels were measured in six lipogenic tissues such as liver, fat, *gluteus medius* and *longissimus thoracis et lumborum* muscles, heart and duodenum (Fig. 2). *HMGCR* expression was detected in all samples, with relative higher levels in duodenum and subcutaneous fat, while heart and liver exhibited the lowest *HMGCR* mRNA expression levels. The three samples of striate muscle displayed intermediate expression levels.

3.3. Transcriptional activity of the pig HMGCR promoter

A functional study of the pig *HMGCR* proximal promoter was performed in order to study its responsiveness to different treatments (insulin, 25-hydroxycholesterol and forskolin) *in vitro* in two cell lines of hepatic (HepG2) and muscular (C2C12) origins. The two alleles of the *HMGCR:g.-225G>T* polymorphism were analysed, as it alters the

potential binding site for the myogenic bHLH family of transcription factors.

Liver and muscle cells responded differently to treatments. Thus, when insulin was added to HepG2 cells, it inhibited the basal expression of the G allele at low concentrations ($P<0.01$, Fig. 3A). In contrast, in this cell line the T allele of the promoter was very responsive to higher concentrations of insulin (10 μM; $P<0.01$). In the myoblastic C2C12 cell line, the G allele of the promoter exhibited a progressive transactivation in response to increasing levels of insulin concentration (0.1 μM, 1 μM and 10 μM; $P<0.01$; Fig. 3B). This effect was less pronounced in differentiated C2C12 cells, where insulin treatment increase the response of this promoter in G the allele of the pig *HMGCR* promoter only at 10 μM ($P<0.01$; Fig. 3C).

An inhibitory effect of 25-hydroxycholesterol over the transcriptional activity of pig *HMGCR* promoter was obvious both in HepG2 and C2C12 cells ($P<0.01$; Fig. 4). When the effect of the SNP allele promoter was analyzed, we observed that in the HepG2 cell line, 25-hydroxycholesterol inhibited more the expression of pig *HMGCR* in the G allele while, in C2C12 cells (single and differentiated cells), the T allele obtained the lowest level of *HMGCR* promoter activity (Fig. 4). C2C12 cells also showed a more gradual response to increasing 25-hydroxycholesterol concentrations.

The clearest difference in treatment responsiveness between cell lines was observed when forskolin was added to the medium (Fig. 5). In HepG2 cells, forskolin gradually inhibited the transcriptional activity of *HMGCR* promoter equally in both alleles ($P<0.01$; Fig. 5). In contrast, in C2C12 myoblast, overall neither of the two alleles of the *HMGCR* promoter responded to forskolin treatment (Fig. 5B), although, when differentiated to myotubes, the transcriptional activity of both alleles was inhibited by low concentrations of forskolin and activated by the highest levels ($P<0.01$; Fig. 5C).

3.4. DNA methylation pattern of the pig HMGCR promoter

DNA bisulfite conversion and methylation-specific PCR methods were used to compare the

methylation patterns in the pig *HMGCR* promoter between liver and muscle tissue and between the two alleles of the *HMGCR:g.-225G>T* SNP. Two CpG islands of 113 and 237 bp were identified in the 613 bp sequenced fragment (Fig. 6A). Bisulfite-converted specific primers were selected to amplify a 424 bp of this promoter, which included the two CpG islands (36 CpG motifs), both in liver and muscle samples from five pigs with GG or TT genotype.

Liver and muscle genomic DNA samples clearly differed in the percentage and distribution of methylated CpG sites (average of 3.33% versus 45.55% methylated CpG motifs in liver and muscle, respectively; Fig.6B). Sixteen CpG sites were consistently methylated in all liver DNA samples, one of which was located in the sterol-response element (SRE) at position -176. Four of these methylated CpG sites were adjacent to the CRE element. However, and quite remarkable, the CpG motif within the CRE element required for insulin stimulation of *HMGCR* gene was unmethylated in both tissues (Fig. 6B).

The genotype at *HMGCR:g.-225G>T* did not affect the level or distribution of promoter methylation as, within each tissue, the distribution of methylated CpG sites did not differ between genotypes.

4. Discussion

We have described here the proximal promoter of the porcine *HMGCR* gene which encodes for a limiting enzyme in the *de novo* cholesterol biosynthesis pathway. As in humans and mouse, this is a TATA-less promoter which probably supports multiple transcription start sites. In this sense, treatment with statins changed the rate of utilisation of the different transcription initiation sites in the hamster *HMGCR* gene [22]. Moreover, the length of the *HMGCR* 5'UTR had a significant role in regulating its translational efficiency.

Studies in human, rat and mouse indicate that the first 350 bp of this promoter are critical for maximal transcriptional activity *in vitro* and *in vivo* [6] and it contains binding sites for several transcription factors including USF-2, LRH-1, SRE, CRE and NF-Y. *HMGCR* gene expression has been reported to be under hormonal (insulin, thyroid hormone), dietary (cholesterol,

bile acids) and environmental (fasting, circadian clock) control [1, 4, 25-27]. In humans, this gene is expressed ubiquously although differences in expression levels between tissues have been reported [28]. We have detected pig *HMGCR* expression in six tissues involved in cholesterol biosynthesis, transport and storing (Fig.2). In this sense, significant differences were found between tissues; maximal expression was detected in duodenum and fat while liver exhibited the lowest expression levels. These results agree well with the fact that, in most mammals, the most important fraction of cholesterol biosynthesis occurs in liver, duodenum and adipose tissue [29]. However, in pigs the most important lipogenic organ is the adipose tissue rather than the liver [30], which might also explain the relatively lower *HMGCR* expression is pigs' liver compare to humans (herein and [31]). The difference in the expression pattern between tissues is also represented by the transcription factor binding sites in the promoter of this gene. Most studies of transcriptional activation of this gene have been performed in liver (live animals) or in hepatic cell lines. These studies have resulted in the identification of sequences important for the effect of several compounds in liver *HMGCR* expression. For instance, the LRH-1 binding site is needed for maximal liver expression in rats [6] and to mediate bile acid effects on this promoter; the CRE response element required for insulin stimulation of *HMGCR* expression in hepatic cell lines [32]; and mutation of the SRE, NF-Y or USF-2 sites essentially abolished the response to thyroid hormone [4].

The liver plays a very relevant role in synthesis and transport of cholesterol and fatty acids to peripheral tissues including the muscle. It is known that *HMGCR* expression affects not only directly in *de novo* cholesterol biosynthesis but also the *LDLR* gene expression through an indirect mechanism. Thus, *HMGCR* can regulate the level of cholesterol and triglycerides in liver as well as in peripheral target tissues such as skeletal muscle by regulating cholesterol synthesis and uptake. In addition, in livestock animals, muscle fat content and composition largely defines the quality of meat from a sensory, technological and health point of view [33]. In this study, the three striate pig muscles analysed (*longissimus thoracis et lumborum*, *gluteus medius* and heart) exhibited intermediate *HMGCR* expression levels. Few studies have

examined variation in the pig *HMGCR* gene in relation to muscle biology and meat quality. Despite the little knowledge related to cholesterol biosynthesis activity in muscle, in a previous work we have shown that the levels of muscle (but not liver) *HMGCR* expression differ between pigs with extreme serum, carcass and muscle lipid levels [5]. Given the relative level of *HMGCR* expression observed between these two tissue types (Fig. 2) we decided to characterise the functionality of the promoter is both tissues. Moreover, during the characterisation of the pig promoter, two SNP polymorphisms were identified at positions -225 and -1 which affected two potential transcription factor binding sites. Although SNPs affecting gene expression can occur in all regions of the genome, transcription factor binding sites are attractive regions to search for functional regulatory SNPs [34] since they can increase, decrease or abolish the binding of transcription factors or even create a novel binding site. As a consequence, polymorphisms within regulatory regions can regulate the level and/or timing of transcription activity, sometimes in a tissue-specific manner [34]. Regarding the pig *HMGCR* promoter, the C allele of the more proximal polymorphism affects an Ets-2 motif. Ets-2 is a protooncogene factor related to cell growth and differentiation, with an established role in the regulation of genes related to cholesterol excretion in liver *via* bile salts [35] but whose role related to muscle physiology has not been established beyond the neuromuscular synapse. In contrast, the presence of the G allele in the more distal polymorphism led to the creation of a motif for the family of myogenenic bHLH regulators (CTGGTGCT), which includes myogenin, myoD, Myf5 and MRF4 [34]. Therefore, this SNP was selected for further analysis on the effect on gene regulation in liver and muscle cell lines and its response to different treatments such as insulin, 25-hydroxycholesterol and forskolin. Insulin was selected as previous experiments suggested that it altered the transcriptional activity of *HMGCR* promoter *in vivo* [6] and *in vitro* [1, 32] through its involvement in determining the energetic balance of cells. Also, 25-hydroxycholesterol, final product of the cholesterol biosynthesis pathway, has been reported to be a direct inhibitor of *HMGCR* expression through a negative feed-back loop [37]. Finally, forskolin treatment is related to intracellular cAMP levels, a marker of the energetic balance of cells [38].

On the whole, the transcriptional activation of the pig *HMGCR* promoter differed between hepatic and muscular cell lines and also between promoter genotypes. This was especially evident for insulin and forskolin treatments, both of which had more relevant effects in liver than in muscular cells. Insulin preferentially activated the T allele of the *HMGCR:c.225G>T* SNP in liver cells and the G allele in muscle myoblast and myotubes (Fig.3). Previous studies indicate that the CRE element is necessary but no sufficient to transmit the response to insulin in rat hepatome cells [32]. In addition, alterations in the occupancy of the SRE element and the more distal NF-Y site of the *HMGCR* promoter have been reported in live diabetes rats which might be due to the binding of a specific repressor [1]. The accessibility to these binding sites might vary between muscle and liver tissues, explaining their different response to insulin. However, the *HMGCR:c.225G>T* polymorphism does not affect any of these sites above (Fig.1) and it is not adjacent to differentially methylated CpG islands. It is premature to postulate at this point the mechanism by which these differences take place, but the results advocate for furthering the study of the potential myogenic bHLH binding site.

A gradual and consistent inhibitory effect of the end-product 25-hydroxycholesterol was observed over the transcriptional activity of the pig *HMGCR* promoter in all experimental conditions. Differences between genotypes were only observed in HepG2 cells where the promoter carrying the T allele was more resistant to the inhibitory action of the 25-hydroxycholesterol than the G allele. These results matched with previous experiments describing inhibition of the *HMGCR* expression by 25-hydroxycholesterol in HepG2 cells and in mouse fibroblast L cells [39].

Forskolin has been reported as an activator of the *HMGCR* transcriptional activity in cell lines such as the thyroid follicular FRTL-5 cells [40]. When HepG2 cell line was used to check the action on pig *HMGCR* promoter, opposite results were obtained showing that the pig *HMGCR* transcriptional activity was inhibited after forskolin addition independently of the promoter genotype. Activation of the promoter was only seen in differentiated myotubes after exposing the cells to high concentrations of forskolin. These results indicate a cell line-specific effect

of forskolin, acting as an activator of pig *HMGCR* transcriptional activity in thyroid follicular FRTL-5 cells and in differenciated C2C12 cells and as an inhibitor of transcriptional activity in liver HepG2 cells.

The different behaviour of this promoter in hepatic and muscular cells *in vitro* (results herein) and *in vivo* [5] suggest the existence of tissue-specific differences in the accessibility of transcription factors to this promoter and/or the presence of tissue-specific factors. The modification of cytosine residues within CpG dinucleotides represents an important epigenetic mark that plays important roles in gene regulation, imprinting and is generally associated with gene repression [41]. A CpG island is a stretch of DNA in which the frequency of CpGs is higher than expected. CpG islands have been shown to colocalise with the promoters of all constitutively expressed genes and approximately 40% of those displaying a tissue restricted expression profile [41]. A proportion of all CpG islands are differentially methylated between tissues, representing differences in gene expression and cellular functions [42]. However many genes display a relatively poor correlation between CpG island hypermethylation and the transcriptional activity level. This seems to be the case of the pig *HMGCR* gene, which clearly showed tissue-specific differences in the methylation status and distribution of the promoter region. This might explain the differences in promoter response to insulin, 25-hydroxycholesterol and forskolin found between the three cell type analysed (hepatic, myoblasts and myotubes) but since the methylation status did not seem to be influenced by the polymorphism at position -225 it is unlikely that this epigenetic marker was the cause of the differences observed between promoter genotypes. The role by which this specific polymorphism affects gene expression differently in these cell lines remains, therefore, a challenge to be solved.

Acknowledgments

We acknowledge M Fernández of Instituto de Desarrollo Ganadero, León, Spain, MS Meyers of Iowa State University, Ames, USA and D Almuzara of IRTA, Lleida, Spain. A Cánovas received a predoctoral scholarship from INIA. This project has been funded by the Spanish Ministry of Science (projects GEN2003-20658-C05-05, AGL2002-04271-C03-02 and AGL2007-66707-C02-01).

References

[1] W.R. Lagor, E.D. de Groh, G.C. Ness, Diabetes alters the occupancy of the hepatic 3-hydroxy-3-methylglutaryl-CoA reductase promoter, J Biol Chem 280 (2005) 36601-36608.

[2] Y.C. Chen, Y.D. Chen, X. Li, W. Post, D. Herrington, J.F. Polak, J.I. Rotter, K.D. Taylor, The HMG-CoA reductase gene and lipid and lipoprotein levels: the multi-ethnic study of atherosclerosis, Lipids 44 (2009) 733-743.

[3] L. Keller, C. Murphy, H.X. Wang, L. Fratiglioni, M. Olin, M. Gafvels, I. Bjorkhem, C. Graff, S. Meaney, A functional polymorphism in the HMGCR promoter affects transcriptional activity but not the risk for Alzheimer disease in Swedish populations, Brain Res 1344 (2010) 185-191.

[4] L.R. Boone, M.I. Niesen, M. Jaroszeski, G.C. Ness, In vivo identification of promoter elements and transcription factors mediating activation of hepatic HMG-CoA reductase by T3, Biochem Biophys Res Commun 385 (2009) 466-471.

[5] A. Cánovas, R. Quintanilla, D. Gallardo, I. Diaz, J.L. Noguera, O. Ramirez, R. Pena, Functional and association studies on the pig HMGCR gene, a cholesterol-synthesis limiting enzyme, Animal 4 (2010) 224-233.

[6] W.R. Lagor, R. Heller, E.D. de Groh, G.C. Ness, Functional analysis of the hepatic HMG-CoA reductase promoter by in vivo electroporation, Exp Biol Med (Maywood) 232 (2007) 353-361.

[7] S. Datta, L. Wang, D.D. Moore, T.F. Osborne, Regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase promoter by nuclear receptors liver receptor homologue-1 and small heterodimer partner: a mechanism for differential regulation of cholesterol synthesis and uptake, Journal of Biology and Chemistry 281 (2006) 807-812.

[8] M. Naito, K. Bomsztyk, R.A. Zager, Renal ischemia-induced cholesterol loading: transcription factor recruitment and chromatin remodeling along the HMG CoA reductase gene, Am J Pathol 174 (2009) 54-62.

[9] D. Gallardo, R.N. Pena, M. Amills, L. Varona, O. Ramirez, J. Reixach, I. Diaz, J. Tibau, J. Soler, J.M. Prat-Cuffi, J.L. Noguera, R. Quintanilla, Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs, Physiological Genomics 35 (2008) 199-209.

[10] P. Chomczynski, N. Sacchi, Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, Analytical biochemistry 162 (1987) 156-159.

[11] J. Sambrook, D. Russell, Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2001.

[12] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, J.D. Thompson, Multiple sequence alignment with the Clustal

series of programs, Nucleic Acids Research 31 (2003) 3497-3500.

[13] J. Schug, Using TESS to Predict Transcription Factor Binding Sites in DNA Sequence, in: A.D. Baxevanis (Ed.), Current Protocols in Bioinformatics, J. Wiley and Sons, 2003.

[14] K. Tamura, J. Dudley, M. Nei, S. Kumar, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, Mol Biol Evol 24 (2007) 1596-1599.

[15] K. Tamura, M. Nei, S. Kumar, Prospects for inferring very large phylogenies by using the neighbor-joining method, Proc Natl Acad Sci U S A 101 (2004) 11030-11035.

[16] J.S. Yuan, A. Reed, F. Chen, C.N. Stewart, Jr., Statistical analysis of real-time PCR data, BMC Bioinformatics 7 (2006) 85.

[17] M. Martinez, J.L. Noguera, A. Tomàs,A.Fernández,R.N. Pena, Time-course analysis of ovary, uterus and pituitary gland expression profiles in pregnant sows, Quantitative Proteomics; Applications in Biology and Medicine, Sheffield, UK, (2006), 10^{th}-11^{th} October.

[18] N. Rosenthal, Identification of regulatory elements of cloned genes with functional assays, Methods Enzymol 152 (1987) 704-720.

[19] M. Frommer, L.E. McDonald, D.S. Millar, C.M. Collis, F. Watt, G.W. Grigg, P.L. Molloy, C.L. Paul, A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands, Proc Natl Acad Sci U S A 89 (1992) 1827-1831.

[20] L.C. Li, R. Dahiya, MethPrimer: designing primers for methylation PCRs, Bioinformatics 18 (2002) 1427-1431.

[21] C. Bock, S. Reither, T. Mikeska, M. Paulsen, J. Walter, T. Lengauer, BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing, Bioinformatics 21 (2005) 4067-4068.

[22] A.K. Gayen, D.M. Peffley, The length of 5'-untranslated leader sequences influences distribution of 3-hydroxy-3-methylglutaryl-coenzyme A reductase mRNA in polysomes: effects of lovastatin, oxysterols, and mevalonate, Arch Biochem Biophys 322 (1995) 475-485.

[23] T.A. Ince, K.W. Scotto, A conserved downstream element defines a new class of RNA polymerase II promoters, J Biol Chem 270 (1995) 30249-30252.

[24] Y. Lin, T.A. Ince, K.W. Scotto, Optimization of a versatile in vitro transcription assay for the expression of multiple start site TATA-less promoters, Biochemistry 40 (2001) 12959-12966.

[25] H.M. de Vogel-van den Bosch, N.J. de Wit, G.J. Hooiveld, H. Vermeulen, J.N. van der Veen, S.M. Houten, F. Kuipers, M. Muller, R. van der Meer, A cholesterol-free, high-fat diet suppresses gene expression of cholesterol transporters in murine small

intestine, Am J Physiol Gastrointest Liver Physiol 294 (2008) G1171-1180.

[26] J. Acimovic, M. Fink, D. Pompon, I. Bjorkhem, J. Hirayama, P. Sassone-Corsi, M. Golicnik, D. Rozman, CREM modulates the circadian expression of CYP51, HMGCR and cholesterogenesis in the liver, Biochem Biophys Res Commun 376 (2008) 206-210.

[27] G.C. Ness, C.M. Chambers, Feedback and hormonal regulation of hepatic 3-hydroxy-3-methylglutaryl coenzyme A reductase: the concept of cholesterol buffering capacity, Procedings of the Society of Experimental Biology in Medicine 224 (2000) 8-19.

[28] R. Burkhardt, E.E. Kenny, J.K. Lowe, A. Birkeland, R. Josowitz, M. Noel, J. Salit, J.B. Maller, I. Pe'er, M.J. Daly, D. Altshuler, M. Stoffel, J.M. Friedman, J.L. Breslow, Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13, Arteriosclerosis, Thrombosis and Vascular Biology 28 (2008) 2078-2084.

[29] P.C. Champe, R.A. Harvey, D.R. Ferrier, Cholesterol and Steroid Metabolism, Biochemistry, Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia 2008.

[30] G.L. Allee, D.H. Baker, G.A. Leveille, Fat utilization and lipogenesis in the young pig, J Nutrition 101 (1971) 1415-1422.

[31] R. Shyamsundar, Y.H. Kim, J.P. Higgins, K. Montgomery, M. Jorden, A. Sethuraman, M. van de Rijn, D. Botstein, P.O. Brown, J.R. Pollack, A DNA microarray survey of gene expression in normal human tissues, Genome Biol 6 (2005) R22.

[32] A.R. Osborne, V.V. Pollock, W.R. Lagor, G.C. Ness, Identification of insulin-responsive regions in the HMG-CoA reductase promoter, Biochemistry and Biophysics Research Communications 318 (2004) 814-818.

[33] J.D. Wood, M. Enser, A.V. Fisher, G.R. Nute, P.R. Sheard, R.I. Richardson, S.I. Hughes, F.M. Whittington, Fat deposition, fatty acid composition and meat quality: A review, Meat Science 78 (2008) 343-358.

[34] B.N. Chorley, X. Wang, M.R. Campbell, G.S. Pittman, M.A. Noureddine, D.A. Bell, Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies, Mutat Res 659 (2008) 147-157.

[35] T. Kanayama, M. Arito, K. So, S. Hachimura, J. Inoue, R. Sato, Interaction between sterol regulatory element-binding proteins and liver receptor homolog-1 reciprocally suppresses their transcriptional activities, J Biol Chem 282 (2007) 10290-10298.

[36] V. Sartorelli, G. Caretti, Mechanisms underlying the transcriptional regulation of skeletal myogenesis, Curr Opin Genet Dev 15 (2005) 528-535.

[37] J.D. Horton, J.L. Goldstein, M.S. Brown, SREBPs: activators of the

complete program of cholesterol and fatty acid synthesis in the liver, J Clin Invest 109 (2002) 1125-1131.

[38] A. Romani, A. Scarpa, Hormonal control of Mg2+ transport in the heart, Nature 346 (1990) 841-844.

[39] R. Gerber, J.D. Ryan, D.S. Clark, Cell-based screen of HMG-CoA reductase inhibitors and expression regulators using LC-MS, Anal Biochem 329 (2004) 28-34.

[40] B. Perillo, I. Tedesco, C. Laezza, M. Santillo, A. Romano, S.M. Aloj, M. Bifulco, Regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase gene expression in FRTL-5 cells. II. Down-regulation by v-K-ras oncogene, J Biol Chem 270 (1995) 15237-15241.

[41] R.S. Illingworth, A.P. Bird, CpG islands--'a rough guide', FEBS Lett 583 (2009) 1713-1720.

[42] R. Illingworth, A. Kerr, D. Desousa, H. Jorgensen, P. Ellis, J. Stalker, D. Jackson, C. Clee, R. Plumb, J. Rogers, S. Humphray, T. Cox, C. Langford, A. Bird, A novel CpG island set identifies tissue-specific methylation at developmental gene loci, PLoS Biol 6 (2008) e22.

Table 1

Primers used for the amplification and methylation analysis of the pig *HMGCR* proximal promoter.

Primer Name	Sequence 5' → 3'	Position	annT[a]	Size[b]
HMGCRprom_F	AACTTCTCTGCCACCTTCCC	promoter (-525)	66	613
HMGCRprom_R	CACCGCCCGCCTCACCTC	E1		
HMGCR_qF	GTTATTACCCTAACTTTGGCTC	E10	58	65
HMGCR_qR	GATTCCGTCTCTGCTTGTTC	E10		
HMGCRclon_F[c]	GCA*CTCGAG*AACTTCTCTGCCACCTTCCC	promoter (-525)	64	580
HMGCRclon_R[c]	TCG*AAGCTT*CACCGCCCGCCTCACCTC	E1		
HMGCRmeth_F	GTAGGTTAGGTTTTGGGTTGTAGG	promoter (-332)	60	424
HMGCRmeth_R	TCCTCACCTCCTAATTTCACTAAAA	E1		

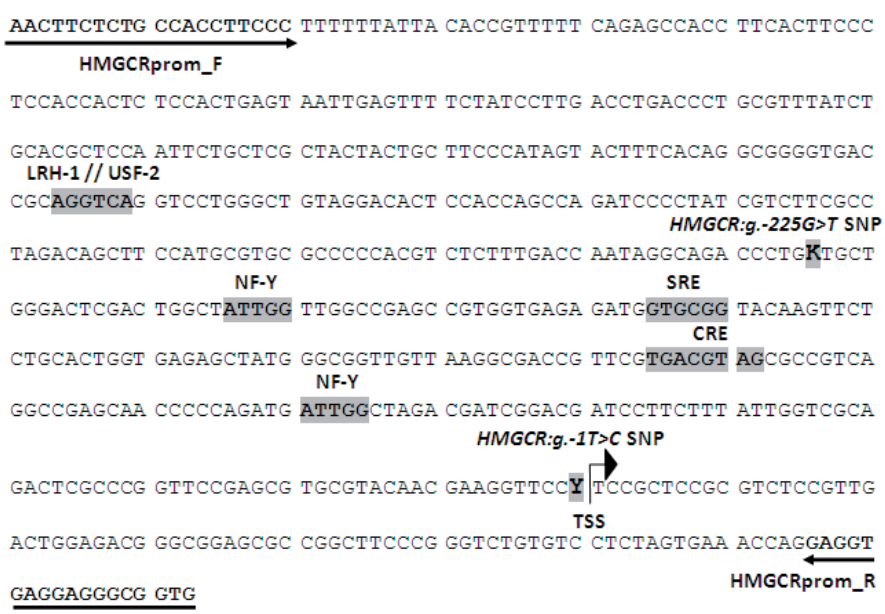
[a] Annealing T° used in the PCR reactions
[b] Size in bp of the amplified product
[c] Inserted *Hin*dIII and *Xho*I target sites are underlined.

Figure legends

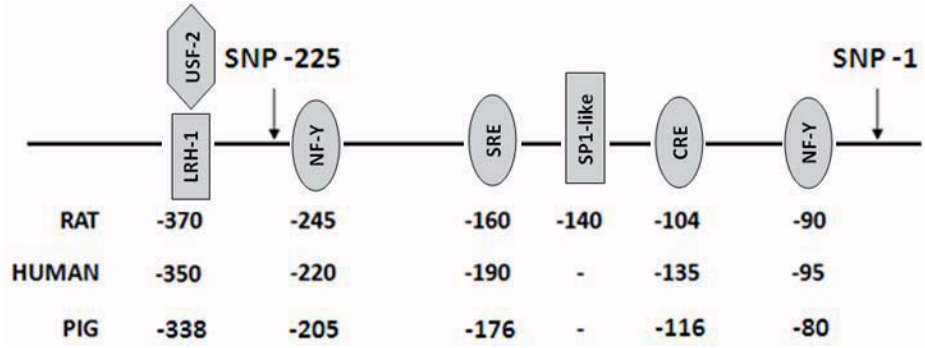
Fig. 1. (A) Graphical representation of sequence of the *HMGCR* pig promoter. Primers used for the amplification of *HMGCR* promoter (HMGCRprom_F and HMGCRprom_R), SNPs identified (*HMGCR:g.-225G>T* SNP and *HMGCR:g.-1T>C* SNP) and regulatory elements are represented *(USF-2*, upstream stimulatory factor-2; *LRH-1*, liver receptor homologue-1; *NF-Y*, nuclear factor-Y; *SRE*, sterol-response element; *CRE*, cyclic AMP-response element; *TSS*, transcription start site). **(B)** Comparison of transcription factor-binding sites in the rat, pig and human *HMGCR* promoter (*SP1-like*, simian-virus-40-protein-1) and localisation of the two new SNPs. **(C)** Sequences similarities with other mammal species, shown as sequence identities. All results are based on the pairwise analysis of sequences.

Figure 1

A.



B.



C.

	Sequence similarity in the promoter region				
	[1]	[2]	[3]	[4]	[5]
[1] S.scrofa					
[2] B.taurus	0.8608				
[3] H.sapiens	0.8282	0.7902			
[4] M.musculus	0.7197	0.7125	0.6980		
[5] R.norvergicus	0.7342	0.7251	0.7251	0.9096	

Figure 2

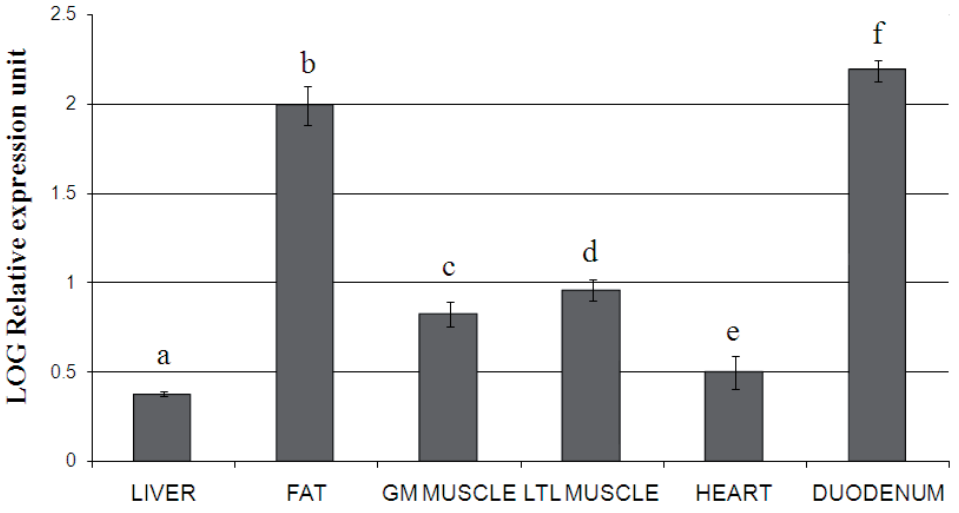


Fig. 2. Relative expression of porcine *HMGCR* gene in samples from six lipogenic tissues in Duroc pigs. Mean of four animals per tissue is shown. Error bars indicate s.e. Different superscripts designate significant differences at $P < 0.05$.

Figure 3

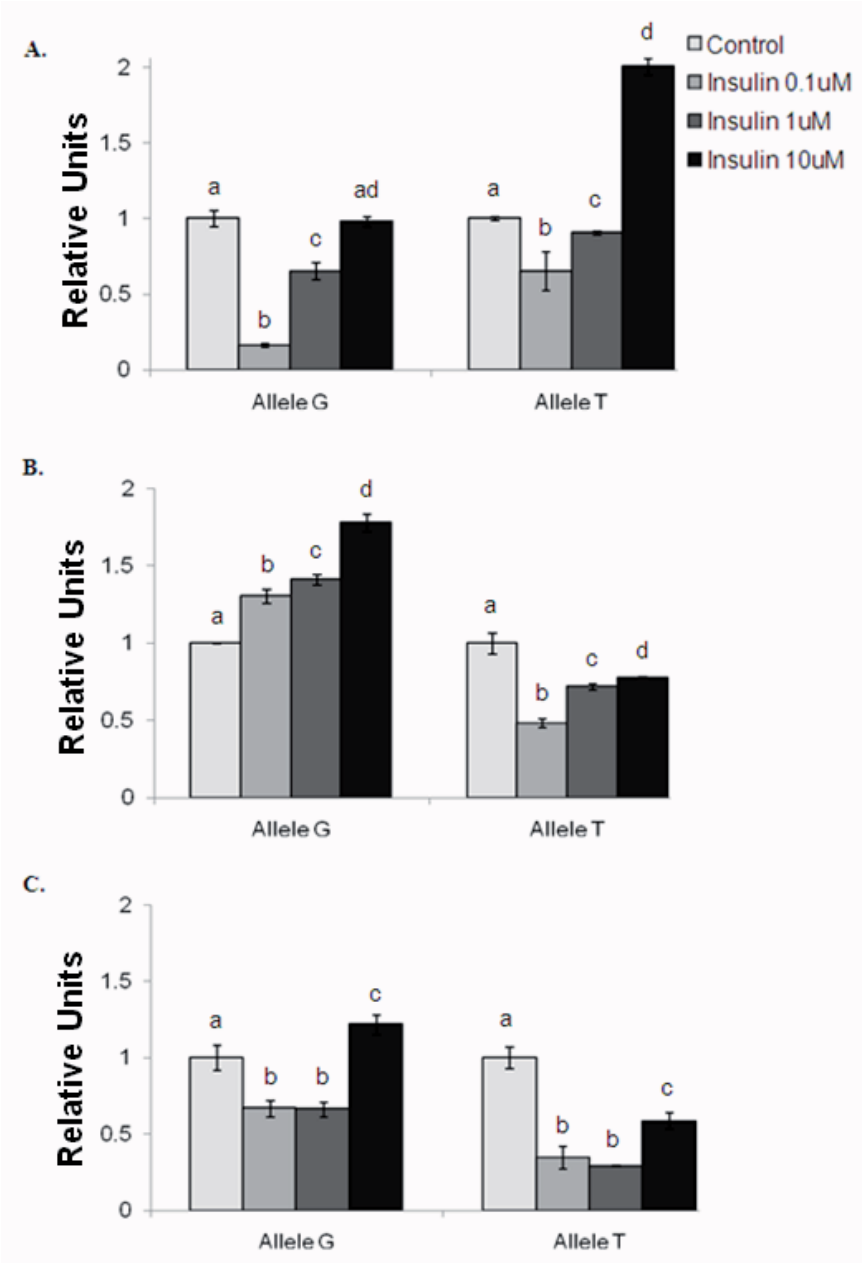


Fig. 3. Response of the transcriptional activity of the two alleles of the *HMGCR:g.-225G>T* polymorphism of the pig *HMGCR* promoter to insulin treatment in transfected cells: (A) HepG2 cells, (B) C2C12 myoblasts and (C) C2C12 differentiated myotubes. Mean on triplicate measures for two experiments is shown. Error bars indicate s.e. Different superscripts designate significant differences at $P < 0.05$.

Figure 4

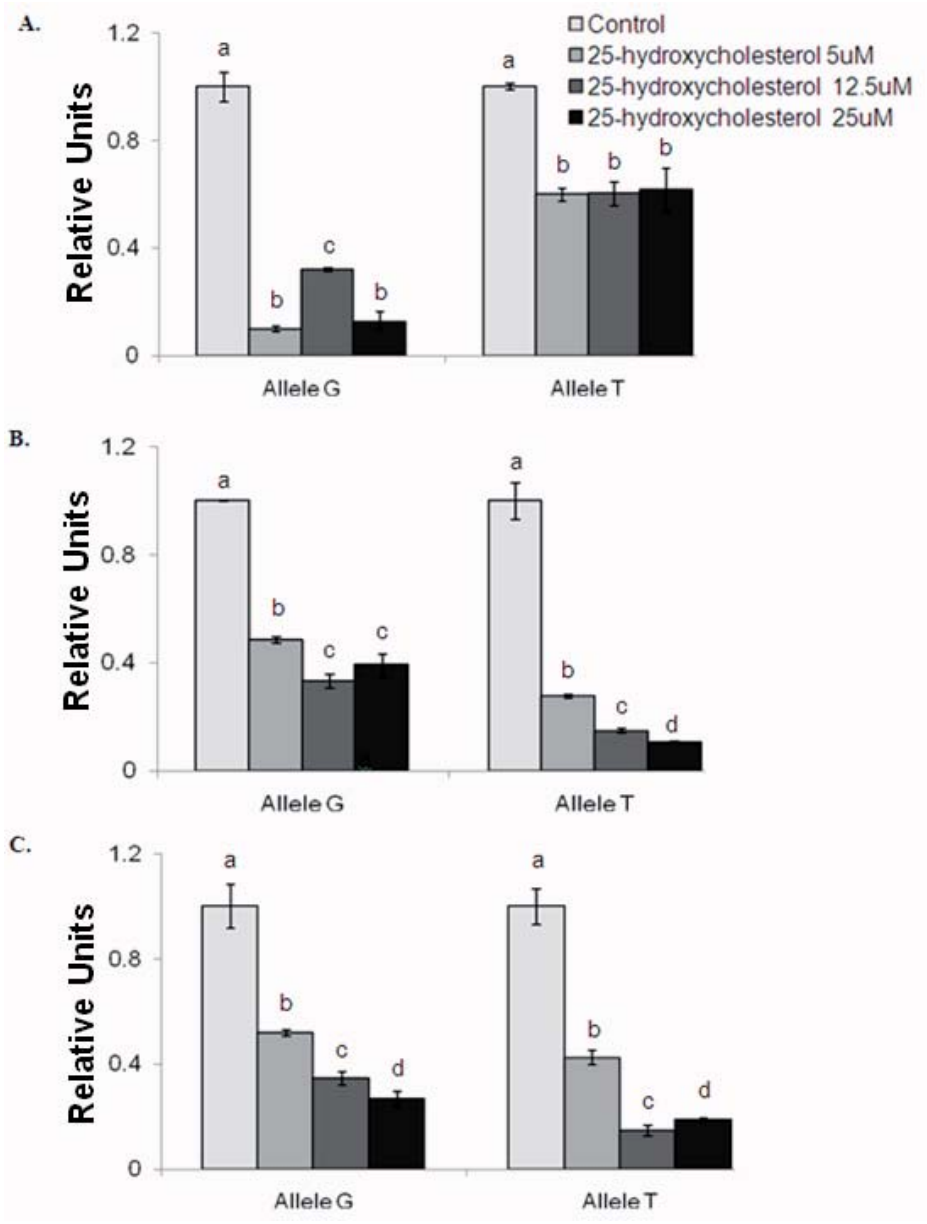


Fig. 4. Response of the transcriptional activity of the two alleles of the *HMGCR:g.-225G>T* polymorphism of the pig *HMGCR* promoter to 25-hydroxycholesterol treatment in transfected cells: **(A)** HepG2 cells, **(B)** C2C12 myoblasts and **(C)** C2C12 differentiated myotubes. Mean on triplicate measures for two experiments is shown. Error bars indicate s.e. Different superscripts designate significant differences at $P < 0.05$.

Figure 5

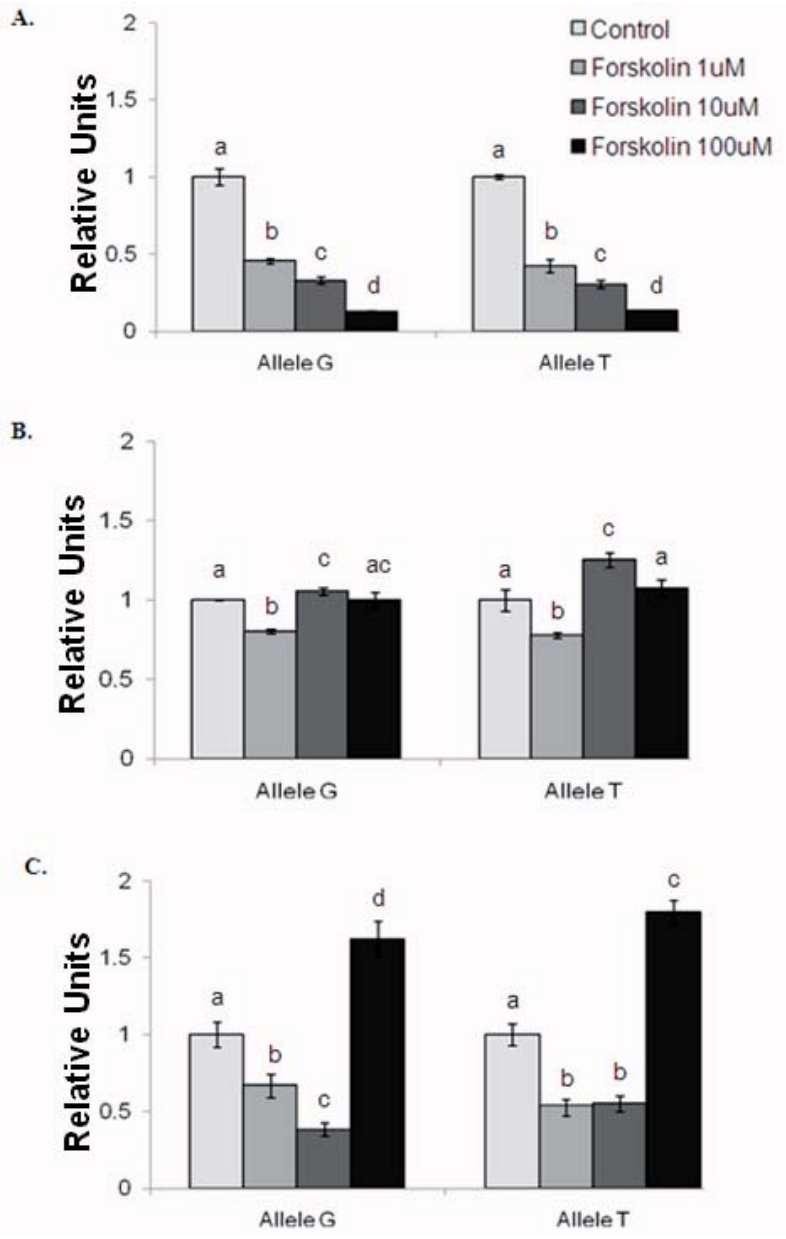


Fig. 5. Response of the transcriptional activity of the two alleles of the *HMGCR:g.-225G>T* polymorphism of the pig *HMGCR* promoter to forskolin treatment in transfected cells: **(A)** HepG2 cells, **(B)** C2C12 myoblasts and **(C)** C2C12 differentiated myotubes. Mean on triplicate measures for two experiments is shown. Error bars indicate s.e. Different superscripts designate significant differences at $P < 0.05$.

Figure 6

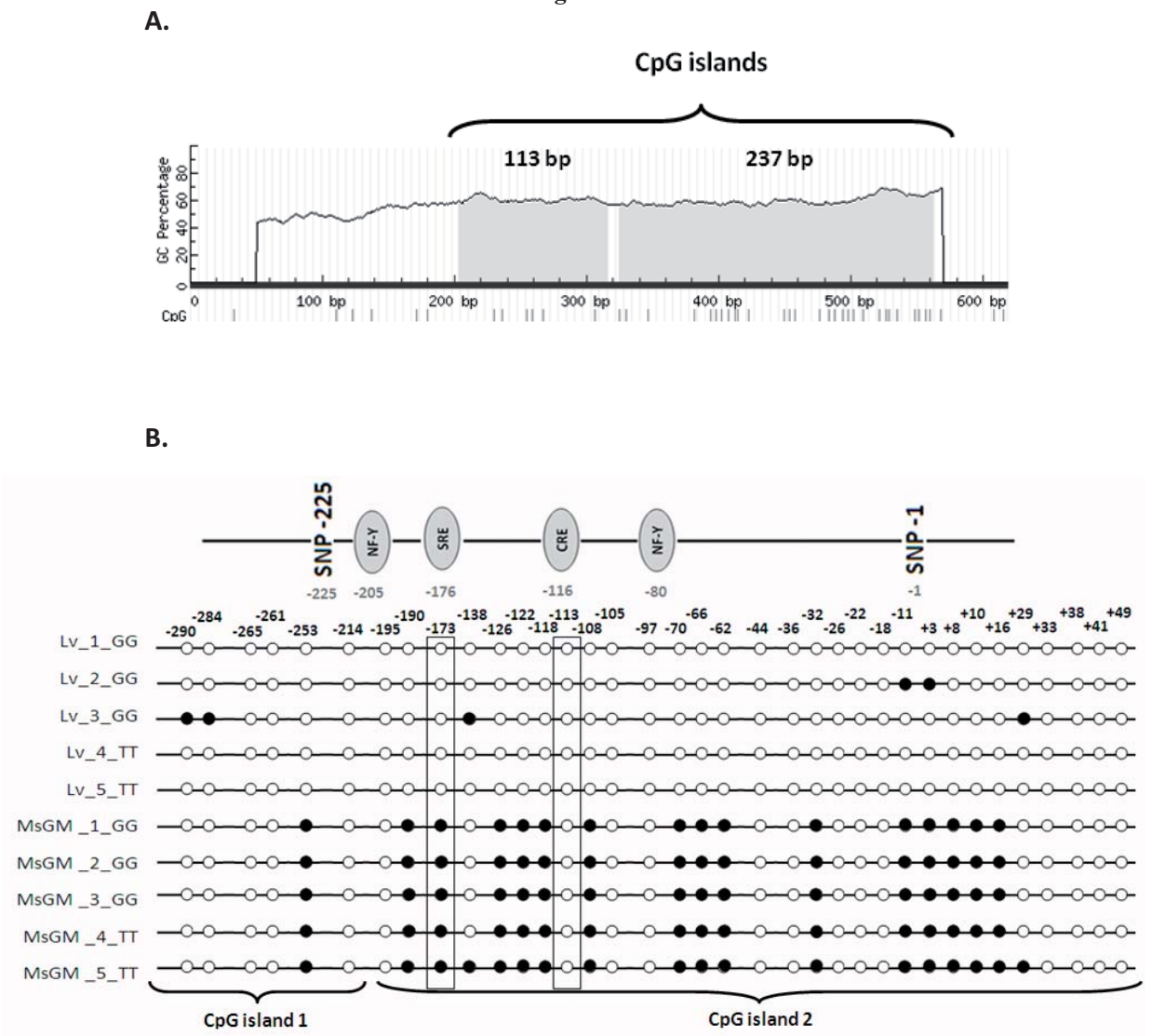


Fig. 6. (A) Graphical representation of two CpG islands of 113 bp and 237 bp-long identified in the pig *HMGCR* promoter. **(B)** DNA methylation pattern of the pig *HMGCR* promoter represented by Lollipop-style chart in muscle and liver tissues. Filled (black) circles correspond to methylated CpG motifs and unfilled (white) circles correspond to unmethylated CpG motifs (*NF-Y*, nuclear factor-Y; *SRE*, sterol-response element *CRE*, cyclic AMP-response element).

3.5.- Selection for decreased backfat thickness at restrained intramuscular fat content is related to reduction in acetyl-CoA carboxylase and stearoyl-CoA desaturase protein expression in pig subcutaneous adipose tissue

Cánovas, A., Estany, J., Tor, M., Pena, R.N., Doran, O

Journal of Animal Science. **87: 3905-3914. 2009**

Acetyl-CoA carboxylase and stearoyl-CoA desaturase protein expression in subcutaneous adipose tissue is reduced in pigs selected for decreased backfat thickness at constant intramuscular fat content[1]

A. Cánovas,* J. Estany,† M. Tor,† R. N. Pena,* and O. Doran‡[2]

*IRTA, Genètica i Millora Animal, 191 Rovira Roure, 25198 Lleida, Spain; †University of Lleida, Department of Animal Production, 191 Rovira Roure, 25198 Lleida, Spain; and ‡School of Life Sciences, University of the West of England, Frenchay Campus, Coldharbour Lane, BS16 1QY Bristol, United Kingdom

ABSTRACT: The objectives of this study were 1) to determine whether selection toward less subcutaneous fat thickness at constant intramuscular fat content in pigs is related to tissue-specific changes in the expression of lipogenic enzymes acetyl-CoA carboxylase (ACC), stearoyl-CoA desaturase (SCD), and Δ^6-desaturase (Δ6d); and 2) to investigate tissue specific distribution of the porcine ACC, SCD, and Δ6d. The study was conducted on 20 purebred Duroc barrows. Ten animals were from a group selected for decreased subcutaneous fat thickness at constant intramuscular fat content (experimental group). The other 10 animals were from the unselected (control) group. Distribution of ACC, SCD, and Δ6d was investigated in semimembranosus muscle (SM), subcutaneous adipose tissue (SA), liver (L), kidney (K), heart (H), diaphragm (D), rectus capitis muscle (RCM), and abdominal fat (AF). The enzyme expression was studied in 10 animals in the case of SM and SA and in 4 animals in the case of other tissues. The following expression pattern was established for ACC: SM $\leq$ H = K $\leq$ D $<$ RCM $<$ L $<$ AF = SA, whereas the expression patterns for SCD and Δ6d proteins were SM $<$ H $<$ RCM $<$ D $<$ L $<$ K $<$ AF = SA and RCM = SM = D $<$ L $\leq$ H $<$ SA $<$ K $<$ AF, respectively. Expression of ACC and SCD proteins was less in subcutaneous adipose tissue of the experimental animals when compared with the control group ($P < 0.001$). However, no difference ($P > 0.1$) in ACC and SCD protein expression between the control and experimental groups was observed in SM. Expression of Δ6d protein did not differ between the control and experimental groups for SA ($P = 0.47$) or SM ($P = 0.31$). There was a positive relationship between muscle SCD protein expression and intramuscular fat content ($r = 0.48$, $P < 0.05$). Intramuscular fat content did not correlate with ACC or Δ6d protein expression ($P = 0.23$ and $P = 0.80$, respectively). We conclude that SCD might be an effective potential biomarker for intramuscular fat deposition.

Key words: intramuscular fat, lipogenic enzyme expression, meat quality, pig, subcutaneous fat

J. Anim. Sci. 2009. 87:3905–3914
doi:10.2527/jas.2009-2091

INTRODUCTION

An increase in carcass lean content has been considered a major objective of the pig industry. This has been achieved via selective breeding against subcutaneous fat thickness (McPhee and Trout, 1995; Hermesch et al., 2000). However, selection toward leaner carcasses has also resulted in a reduction in intramuscular fat (**IMF**), which is known to contribute to eating quality of pork (Ellis et al., 1996; Verbeke et al., 1999; Ruiz et al., 2000). Therefore, a challenge is to produce pigs with less subcutaneous fat thickness without reduction of IMF below the level required for optimum eating quality. It has been suggested that fat deposition in different depots might be regulated by different mechanisms (Gardan et al., 2006; Gondret et al., 2008) and that the lipogenic enzyme stearoyl-CoA desaturase (**SCD**) plays the key role in this process (Da Costa et al., 2004; Doran et al., 2006). Contribution of other lipogenic enzymes to the tissue-specific regulation of fat deposition in pigs remains unclear. Moreover, the mechanisms regulating IMF and subcutaneous fat

[1]We acknowledge Teresa Giró, Anna Ñaco, and Laura Frutos of the Department of Animal Production, University of Lleida, Spain, for assistance with sampling and laboratory analyses. A. Cánovas is the recipient of a PhD scholarship from INIA. This research was partially funded by the Spanish Ministry of Education and Science, Spain (CICYT grant AGL2006-01243). We thank the staff of Selección Batallé for cooperation in this experiment. Lipogenic enzyme expression was analyzed at the facilities of the University of Bristol, UK.

[2]Corresponding author: Olena.Doran@uwe.ac.uk

Received April 30, 2009.
Accepted August 12, 2009.

deposition are largely unknown. Knowledge of these mechanisms would allow identification of physiological candidate genes that could be used for developing of tests for evaluation of effectiveness of genetic selection or dietary manipulations.

The objectives of the present study were 1) to investigate whether selection toward less backfat thickness at constant IMF content in pigs is related to tissue-specific changes in the expression of the lipogenic enzymes acetyl-CoA carboxylase (**ACC**), SCD, and Δ^6-desaturase (**Δ6d**), and 2) to characterize tissue distribution of ACC, SCD, and Δ6d. This would contribute to our knowledge about possible involvement of different tissues in fat deposition in pigs.

MATERIALS AND METHODS

All the experimental procedures were approved by the Ethics Committee for Animal Experimentation of the University of Lleida.

Animals and Sample Collection

The experiments were conducted on 20 purebred Duroc barrows, which were selected from previously established experimental and control groups (Reixach et al., 2008, 2009). Ten of these animals were randomly chosen from a genetic group selected for decreased backfat thickness at constant BW and IMF content (experimental group). The other 10 animals were randomly chosen from the unselected (control) group. The experimental group differed from the control group in backfat thickness but not in IMF content and composition (Reixach et al., 2008). Mid-parent BLUP breeding values for backfat depth (based on approximately 37,000 records) and IMF content in gluteus medius (based on 3,000 records) were used to separate newborn litters into 2 groups differing in backfat depth but not in IMF content. This 1-generation selection process was repeated in 4 consecutive batches. Pigs in the present manuscript were obtained from the third batch. The groups were constituted according to the mid-parent (litter) breeding values for backfat thickness and BW at 180 d and for IMF content in the gluteus medius muscle adjusted for carcass weight, which were predicted using the model described in Solanes et al. (2009). Litters in the experimental group were selected against backfat thickness while maintaining IMF content and BW to the values most similar to those in the control group. Linear programming was used to select the litters satisfying the constraints above. At the age of 11 wk, 2 barrows per litter were taken at random and moved to a finishing facility until the age of 200 d. During the test period, pigs had ad libitum access to a commercial diet (Esporc, Ruidarenes, Girona, Spain; Table 1). Feed analyses were performed in triplicate. Dry matter was determined by oven-drying at 100 to 102°C for 24 h; ash was determined by muffle-heating at 550°C until constant weight. Crude protein was analyzed by the Kjeldahl method (AOAC, 2000), crude lipid was determined by Soxhlet fat analysis (AOAC, 2000), and crude fiber was analyzed by acid and alkaline digestion with a Dosi-Fiber extractor (Selecta, Barcelona, Spain; AOAC, 2000). Analyses of fatty acid composition were performed as described below for tissue samples after extracting the total lipids by the method of Hanson and Olley (1963).

Table 1. Composition of the diet fed during the finishing period

Item	Amount
DM, g/kg	893.2
CP, g/kg of DM	164.5
Crude lipid, g/kg of DM	66.1
Crude fiber, g/kg of DM	61.6
Ash, g/kg of DM	60.0
Fatty acid composition, mg/g of fatty acids	
C12:0	1.2
C14:0	12.2
C16:0	208.8
C16:1	25.2
C18:0	60.1
C18:1	353.2
C18:2	293.5
C18:3	1.8
C20:0	28.3
C20:1	8.0
C20:2	5.4
C20:4	2.3
SFA	310.6
MUFA	386.4
PUFA	303.0

Pigs were performance-tested at around 180 and 200 d of age. This included the measurement of BW and backfat thickness. The backfat thickness was determined ultrasonically at 5 cm off the midline at the position of the last rib (Piglog 105, SFK-Technology, Herlev, Denmark). Pigs were slaughtered at 203 d in a commercial abattoir. Carcass backfat thickness at 6 cm off the midline between the third and fourth last ribs and carcass lean percentage were measured by an on-line ultrasound automatic scanner (AutoFOM, SFK-Technology, Herlev, Denmark). Carcass lean content was estimated on the basis of 35 measurements of AutoFOM points by using the official approved equation (Decision 2001/775/CE, 2001). The differences that have been observed for backfat thickness in live animals (at 200 d) and for carcasses (at 203 d) are related to the use of different measurement techniques. Samples from 10 pigs from the same slaughtering batch (each originated from a different litter) were collected. Intramuscular fat content was measured as described previously (Bosch et al., 2009).

Immediately after slaughter, samples of semimembranosus muscle (**SM**) and subcutaneous adipose tissue were collected at the level of the third and fourth ribs. In addition, samples of liver, kidney, heart, diaphragm, rectus capitis muscle, and abdominal fat were collected from 4 randomly chosen animals and were used to in-

vestigate tissue distribution of lipogenic enzymes. The animals used in these experiments were the first 4 pigs on the slaughter line. Three of these animals were from the experimental group, and 1 animal was from the control group.

All the tissue samples were snap-frozen and stored at −80°C until analyzed. It has been previously demonstrated that these storage conditions do not affect expression of lipogenic enzymes (Doran et al., 2006).

Fatty Acid Analysis

Once defrosted, SM and subcutaneous adipose tissue samples were freeze-dried and thoroughly homogenized by mixing with sand using a glass stirring rod. Due to the small sample size, DM was calculated as the weight difference before and after freeze-drying. Fatty acid composition analysis was performed in duplicate by quantitative determination of the individual fatty acids by gas chromatography. Fatty acid methyl esters were directly obtained by transesterification using a solution of 20% boron trifluoride in methanol (Rule et al., 1997). Methyl esters were determined by gas chromatography using a capillary column SP2330 (30 m × 0.25 mm, Supelco, Bellefonte, PA) and a flame ionization detector with helium as carrier gas. The analytical column was coated with a 0.20-μm film.

The oven temperature program was increased from 150 to 225°C (by 7°C per min). The injector and detector temperatures were 250°C (Tor et al., 2005). Fatty acid quantification was carried out via normalization of the area under appropriate picks after adding 1,2,3-tripentadecanoylglycerol into each sample as an internal standard. Intramuscular fat content in SM and gluteus muscles was calculated as a sum of individual fatty acids expressed as triglyceride equivalents (AOAC, 2000) on a dry tissue basis (Bosch et al., 2009).

Isolation of Microsomal and Cytosolic Fractions

Expression of SCD and Δ6d proteins was analyzed in microsomal fraction; ACC protein expression was analyzed in cytosolic fraction. Microsomes and cytosol were isolated by differential centrifugation with Ca^{2+} precipitation (Schenkman and Cinti, 1978) with minor modifications. In brief, approximately 10 g of frozen tissue (2 g for liver) was homogenized in 20 mL of cold sucrose buffer (10 m*M* Tris-HCl, 250 m*M* sucrose, pH 7.4) and centrifuged at 12,000 × *g* for 10 min at 4°C. The supernatant (or infranatant in the case of adipose tissue) was collected and mixed with 8 m*M* $CaCl_2$ to facilitate sedimentation of the microsomal fraction. Microsomes were obtained by centrifugation at 25,000 × *g* for 35 min at 4°C. The supernatant (cytosolic fraction) was collected and the remaining microsomal pellet was resuspended in a buffer containing 10 m*M* Tris-HCl, 250 m*M* KCl (pH 7.4), and inhibitors of proteolytic enzymes (1.5 μ*M* antipain, 1.5 μ*M* pepstatin, and 2 μ*M* leupeptin; Sigma, Dorset, UK). Total microsomal and cytosolic protein content was determined by the Bradford method using BSA as the standard (Bradford, 1976).

Protein Expression

Expression of the lipogenic enzymes was analyzed by western blotting. Microsomal and cytosolic proteins (6 μg) were separated by SDS-PAGE and electroblotted onto a nitrocellulose membrane as described previously (Nicolau-Solano et al., 2006). The membrane was incubated with 1 of the following antibodies: goat polyclonal anti-human ACC (Santa Cruz Biotechnology Inc., Santa Cruz, CA), rabbit polyclonal anti-bovine adipose tissue SCD (custom-made at the University of Bristol, Bristol, UK), or rabbit polyclonal anti-Δ6d IgG (Sigma Genosys Ltd., Cambridge, UK). Incubation with primary antibody was followed by incubation with secondary antibody that was horseradish peroxide-linked donkey anti-rabbit IgG for SCD and Δ6d (GE Healthcare, Amersham, Bucks, UK) or donkey anti-goat IgG for ACC (Santa Cruz Biotechonology). Blots were developed using an Enhanced Chemiluminescent Reagent (GE Healthcare) and quantified using the ImageQuant program (Molecular Dynamics, Sunnyvale, CA). A microsomal or cytosolic preparation from 1 particular animal was present on all the blots (the reference sample). Intensity of the reference sample signal was taken as 100 arbitrary units, and the intensities of other samples on the blot were expressed as fractions of the reference sample. In the case of the genetic selection effects study, we have used subcutaneous adipose tissue and SM reference samples for adipose tissue and muscle Western blots, respectively. In the case of between-tissue enzyme expression comparison, the reference sample was prepared from liver and has been continuously used on all the relevant blots. All gels and blots were done in duplicate. The duplicate samples were run on different blots. The average intensity of the duplicates was calculated.

Statistical Analyses

The effect of selection for reduced backfat thickness at restrained BW and IMF content on lipogenic enzyme expression was analyzed by comparing mean values between the control and experimental groups for each tissue using a *t*-test. Regression analysis was used to test the association between fat content and expression of the lipogenic enzymes. A regression × group interaction was included to test if associations differed across the groups. Results were considered statistically significant at $P < 0.05$. Data were analyzed with SAS (SAS Inst. Inc., Cary, NC) using PROC MIXED procedure. No significant diversion from normality according to the Shapiro-Wilk test was found ($P < 0.05$). Between-tissue differences in enzyme expression were

Table 2. Backfat thickness, BW, and carcass characteristics in the control (n = 10) and experimental (n = 10) groups

	Group		
Trait	Control	Experimental[1]	SEM
Live measurement at 180 d			
Age, d	177.5	179.4	
BW, kg	110.6	107.0	3.2
Backfat thickness,[2] mm	19.0[a]	16.1[b]	1.0
Live measurement at 200 d			
Age, d	199.5	201.4	
BW, kg	125.7	121.3	3.3
Backfat thickness,[2] mm	21.9[a]	18.3[b]	1.1
Carcass measurement			
Carcass weight, kg	96.9	94.1	2.9
Backfat thickness,[3] mm	28.8[a]	25.0[b]	1.3
Lean percentage, % of fresh tissue	41.8	44.8	1.3
Intramuscular fat in GM,[4] % of fresh tissue	5.0	5.0	0.4
Intramuscular fat in SM,[4] % of fresh tissue	2.2	1.7	0.3

[a,b]Means within a row with different superscript differ ($P < 0.05$).

[1]Pigs in the experimental group were selected for decreased backfat depth at restrained intramuscular fat content.

[2]Backfat thickness on live animal was measured ultrasonically at 5 cm off the midline at the position of the last rib using the Piglog technology (SFK-Technology, Herlev, Denmark).

[3]Backfat thickness of carcasses was determined at 6 cm off the midline between the third and fourth last ribs using the AutoFOM automatic scanner (SFK-Technology).

[4]GM = gluteus muscle; SM = semimembranosus muscle.

analyzed by 1-way ANOVA with 8 levels and post-hoc comparison of means by the Tukey test.

RESULTS

The results of meat quality traits analysis in pigs from experimental and control groups are shown in Table 2. Backfat thickness in the experimental group was less ($P < 0.05$) at all time points (i.e., in live animals at 180 d, live animals at 200 d, and in carcasses at 203 d by 15.3, 14.4, and 13.2%, respectively), whereas IMF content in the gluteus muscles did not change. Fatty acid composition of the SM and subcutaneous adipose tissue for the control and experimental groups did not differ (Table 3).

Effect of Selection on Expression of Lipogenic Enzymes

To determine whether the reduction in subcutaneous fat thickness in experimental animals is related to inhibition of lipogenic enzyme expression, the expression of ACC, SCD, and Δ6d proteins was analyzed. Figure 1 shows that there was a decrease in the expression of

Table 3. Fatty acid composition of semimembranosus muscle and subcutaneous adipose tissue in the control (n = 10) and experimental (n = 10) groups

	Semimembranosus muscle			Subcutaneous adipose tissue		
Fatty acid, mg/g of DM[1]	Control	Experimental[2]	SEM	Control	Experimental[2]	SEM
Total SFA	28.95	22.20	3.27	240.32	247.09	6.20
C14:0	1.17	0.89	0.14	10.83	10.37	0.25
C16:0	18.75	14.21	2.11	157.81	159.01	3.89
C18:0	8.91	7.02	1.03	70.47	76.44	2.41
C20:0	0.11	0.08	0.02	1.22	1.26	0.05
Total MUFA	37.72	26.17	4.34	340.43	324.67	8.03
C16:1	2.52	1.47	0.30	17.87	16.88	0.81
C18:1	34.65	24.32	4.12	315.91	301.41	7.21
C20:1	0.55	0.38	0.07	6.65	6.38	0.18
Total PUFA	13.26	12.69	0.85	143.68	139.25	4.61
C18:2	10.14	9.56	0.72	125.49	122.13	4.19
C18:3	0.46	0.42	0.06	9.80	9.13	0.32
C20:2	0.41	0.35	0.12	6.62	6.33	0.12
C20:4	2.24	2.35	0.08	1.77	1.67	0.10

[1]DM determined by freeze-drying to constant weight.

[2]Pigs in the experimental group were selected for decreased backfat depth at restrained intramuscular fat content.

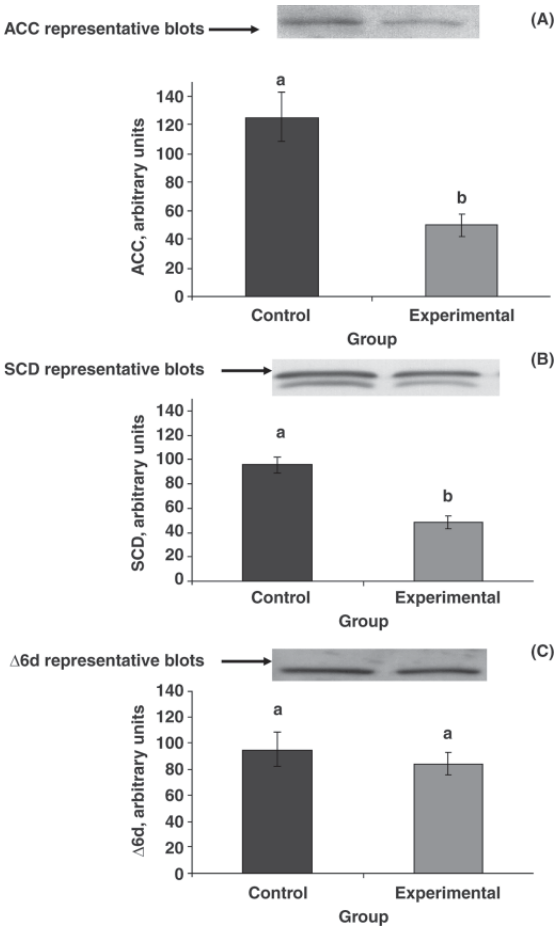


Figure 1. Representative blots and expression of acetyl-CoA carboxylase (ACC), stearoyl-CoA desaturase (SCD), and Δ^6-desaturase (Δ6d) proteins in subcutaneous adipose tissue of pigs in the control and experimental groups. Pigs in the experimental group were selected for decreased backfat depth at restrained intramuscular fat content. Bars represent average of measurements for 10 animals. All measurements were done in duplicate. Error bars represent SEM. [a,b]Means without a common letter differ ($P < 0.001$).

ACC and SCD proteins in the subcutaneous adipose tissue from the experimental group. The ACC protein level was about 60% less ($P < 0.001$; Figure 1A) and SCD expression was about 50% less ($P < 0.0001$, Figure 1B) than the control group. No differences between control and experimental groups were observed in the case of Δ6d protein expression ($P = 0.47$; Figure 1C). There was no relationship between the ACC, SCD, and Δ6d protein expression and the content of SFA, MUFA, and PUFA, respectively, in adipose tissue for the whole set of animals (control plus experimental groups, data not shown). However, a regression × group interaction analysis showed that selection against backfat thickness triggered a change in the association pattern between SCD protein expression and subcutaneous adipose tissue MUFA content. Thus, there was a negative relationship between SCD protein expression and MUFA content in the control ($r = -0.68$, $P < 0.05$), but not in the experimental group ($r = 0.48$, $P = 0.20$; Figure 2).

Results of ACC, SCD, and Δ6d proteins expression analysis in SM are presented in Figure 3. There were no differences ($P > 0.1$) between the control and experimental groups for the enzymes investigated. Moreover, no regression × group interaction of IMF, SFA, MUFA, and PUFA content on ACC, SCD, and Δ6d protein expression was found ($P > 0.05$, data not shown). When the results were analyzed as 1 data set, a positive relationship was found between SCD protein expression and IMF content ($r = 0.48$, $P < 0.05$; Figure 4A). Similar relationship was also found between SCD expression and the product of SCD-catalyzed reaction, MUFA ($r = 0.53$, $P < 0.05$; Figure 4B); and SCD expression and C18:1/C18:0 ratio ($r = 0.61$, $P < 0.01$; Figure 4C). No relationship was observed between IMF content and expression of ACC ($P = 0.23$) or Δ6d ($P = 0.80$) in muscle.

Tissue Distribution of ACC, SCD, and Δ6d Proteins

Although subcutaneous and intramuscular adipose tissues are the most important fat depots in terms of meat quality traits, understanding the mechanism regulating the whole body fat distribution is important for designing strategies for manipulation of fat partitioning. In this study, we have investigated tissue-specific distribution of the key lipogenic enzymes, ACC SCD, and Δ6d, catalyzing the biosynthesis of SFA, MUFA, and PUFA, respectively. Immunoreactive bands for all 3 proteins were detected in the liver, subcutaneous adipose tissue, abdominal fat, rectus capitis muscle, SM, diaphragm, heart, and kidney samples. Expression profiles of ACC, SCD, and Δ6d proteins are presented in Figure 5. Expression of ACC protein was the greatest in subcutaneous adipose tissue and abdominal fat, with a decrease in liver, followed by rectus capitis muscle; the smallest level occurred in SM, heart, and kidney ($P < 0.05$; Figure 5A). Similar to ACC, the greatest SCD expression level was observed in subcutaneous adipose tissue and abdominal fat. In contrast to ACC, the next organ with the greatest SCD expression was kidney, followed by liver, diaphragm, heart, rectus capitis muscle, and SM ($P < 0.05$; Figure 5B). The greatest Δ6d protein expression was observed in abdominal fat, kidney, subcutaneous adipose tissue, liver, and heart. The smallest Δ6d protein level was observed in diaphragm, rectus capitis muscle, and SM ($P < 0.05$; Figure 5C).

DISCUSSION

It has been demonstrated that that reduction in IMF negatively affects meat juiciness, tenderness, and flavor (Fernandez et al., 1999), and therefore increasing IMF content without affecting backfat thickness is one of the challenges of the pig industry. Positive correlation be-

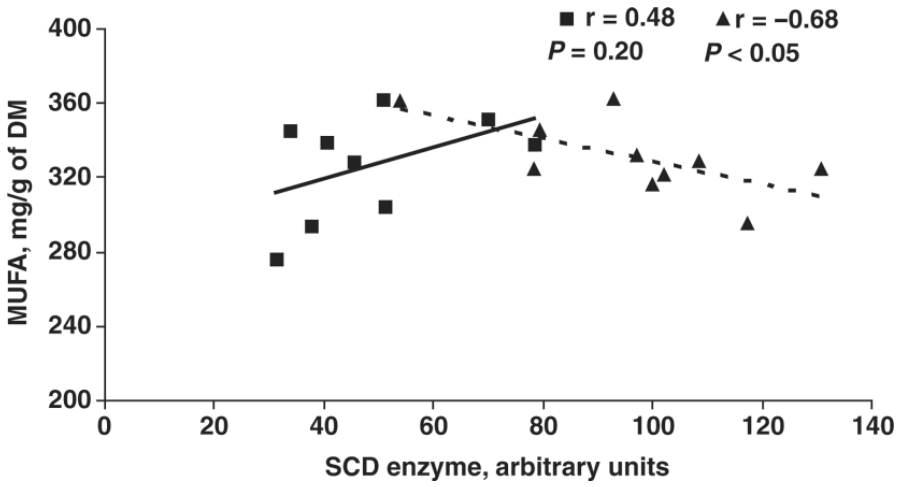


Figure 2. Relationship between stearoyl-CoA desaturase (SCD) protein expression and MUFA content (mg/g of DM) in subcutaneous adipose tissue of pigs in the control (▲, $P < 0.05$) and experimental groups (■, $P = 0.20$). Pigs in the experimental group were selected for decreased backfat depth at restrained intramuscular fat content.

tween IMF and pork eating quality has been found in a range of IMF from 0.7 to 4.5% (Goransson et al., 1992; Eikelenboom et al., 1996). In spite of a large number of reports supporting the positive impact of IMF on eating quality of pork, some reported a lack of correlation between IMF and perceived juiciness, which might be related to the age and eating habits of the particular group of consumers involved in this study (Ventanas et al., 2007).

In the present study, maximum differences in the backfat thickness between the control and experimental groups were observed at 180 d. This is consistent with the results of our previous larger study (188 and 172 pigs for control and experimental groups, respectively), which demonstrated that the difference in backfat depth between the 2 groups was already significant at 120 d, reached the maximum at around 180 d, and remained relatively stable, or even decreased, by 210 d (Reixach et al., 2009).

The present study demonstrated that selection for reduced subcutaneous fat in pigs with constant IMF is accompanied by significant decreases of ACC and SCD protein expression in subcutaneous adipose tissue but not in muscle. Stearoyl-CoA desaturase is the enzyme involved in the biosynthesis of MUFA from SFA (Enoch et al., 1976), whereas ACC catalyzes the first step in the SFA biosynthesis. Acetyl-CoA carboxylase is considered to be a rate-limiting enzyme of lipogenesis in animal tissues, and in pig tissues in particular (Scott et al., 1981). Tissue-specific responses of porcine ACC and SCD have been previously reported in dietary trials. Thus, Doran et al. (2006) established that a reduced protein diet increases the expression of SCD protein (and to a less extent ACC protein) in pig LM, but not in subcutaneous adipose tissue. The reasons for tissue-specific changes in lipogenic enzyme expression are not clear. This could be related, at least in part, to variations in the level of transcription factors regulating the enzyme expression. It is known that IMF fat is a later-maturing tissue when compared with subcutaneous fat (Gardan et al., 2006), and there are significant morphological and metabolic differences between these 2 depots, including differences in transcription factors level (Gardan et al., 2006; Gondret et al., 2008).

The other possible reason for tissue-specific responses of the porcine ACC and SCD could be tissue-specific expression of ACC and SCD isoforms. It is known that more than 1 ACC and SCD isoforms exist in mice, rats, and some other species (Thiede et al., 1986; Miyazaki and Ntambi, 2003; Miyazaki et al., 2003). It has also been demonstrated that SCD isoforms are tissue-specific, distributed with SCD1 being the predominant isoforms in liver and adipose tissue (Ntambi et al., 1988; Kim et al., 2002), whereas SCD2, SCD3, and SCD4 have been found in brain, skin, and heart, respectively (Kaestner et al., 1989; Zheng et al., 2001; Miyazaki et al., 2003). So far, 2 SCD isoforms have been reported in pigs: SCD1, which is preferentially expressed in subcutaneous adipose tissue (Ren et al., 2004); and SCD5, which has been recently reported to be expressed at very high levels in pig brain (Lengi and Corl, 2008). The SCD isoform spectrum in pig muscles and other tissues remains unknown.

In the present study, a decrease in the expression of adipose tissue SCD in the experimental group was not accompanied by changes in the amount of the product of SCD catalyzed reaction, namely C16:1 and C18:1, although there was a trend toward decrease in the amount of these fatty acids. A lack of differences might be related to large between-individual variations (especially in the case of C18:1) and to a relatively small number of animals.

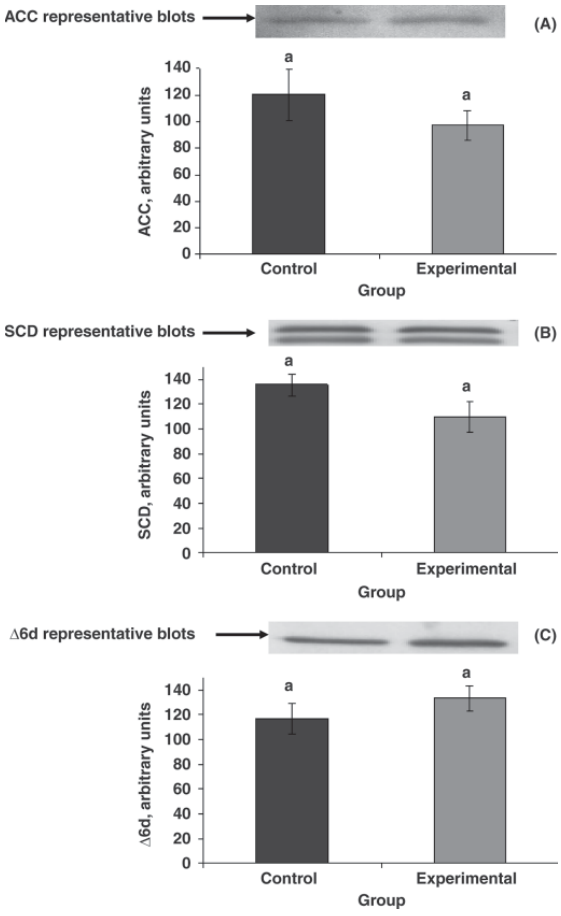


Figure 3. Representative blots and expression of acetyl-CoA carboxylase (ACC), stearoyl-CoA desaturase (SCD), and Δ^6-desaturase (Δ6d) proteins in semimembranosus muscle of pigs from control and experimental groups. Pigs in the experimental group were selected for decreased backfat depth at restrained intramuscular fat content. Bars represent average of measurements for 10 animals. All measurements were done in duplicate. Error bars represent SEM. [a]Means without a common letter differ ($P < 0.05$).

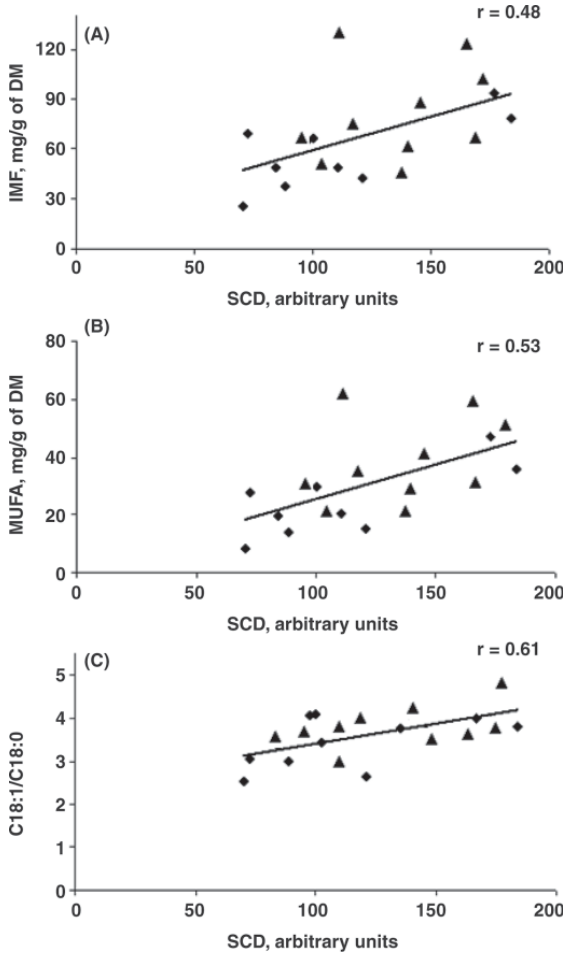


Figure 4. Relationship between stearoyl-CoA desaturase (SCD) protein expression and intramuscular fat (IMF) content (A), MUFA content (B), and C18:1/C18:0 ratio (C) in semimembranosus muscle of pigs of the control (▲) and experiment (◆) groups.

In the present study, the selection for reduced backfat thickness with constant IMF was not accompanied by any change in the expression of muscle or subcutaneous fat Δ6d protein. Delta-6 desaturase is one of the enzymes that catalyzes the conversion of the essential fatty acids (linoleate and α-linolenate) into long-chain PUFA in animal tissues (Cho et al., 1999). The fact that we did not observe any changes in Δ6d in this study indicates that biosynthesis of PUFA might have less input in the regulation of fat partitioning in pig when compared with the biosynthesis of MUFA and SFA.

Fatty acid composition and fat content depend on not only the rate of de novo lipogenesis in a particular tissue, but also on several other factors, including the rate of fatty acid transport from other lipogenic sites. In most mammals, lipogenesis occurs predominantly in the liver and adipose tissue (Girard et al., 1994). In pigs, the major site of fatty acid metabolism is subcutaneous adipose tissue, which has the greatest expression and activity of the key lipogenic enzymes (O'Hea and Leveille, 1969). The input of other tissues in regulation of fat deposition in pigs remains unclear, and tissue-specific distribution of lipogenic enzymes is unknown. In this study, we have characterized the distribution of ACC, SCD, and Δ6d proteins in 8 tissues from organs with diverse physiological functions. Immunoreactive bands corresponding to all 3 enzymes investigated were detected in liver, subcutaneous adipose tissue, abdominal fat, rectus capitis muscle, SM, diaphragm, heart, and kidney. The greatest expression of the lipogenic enzymes was found in subcutaneous adipose tissue and abdominal fat, which is consistent with the key role of these enzymes in lipid biosynthesis and other processes

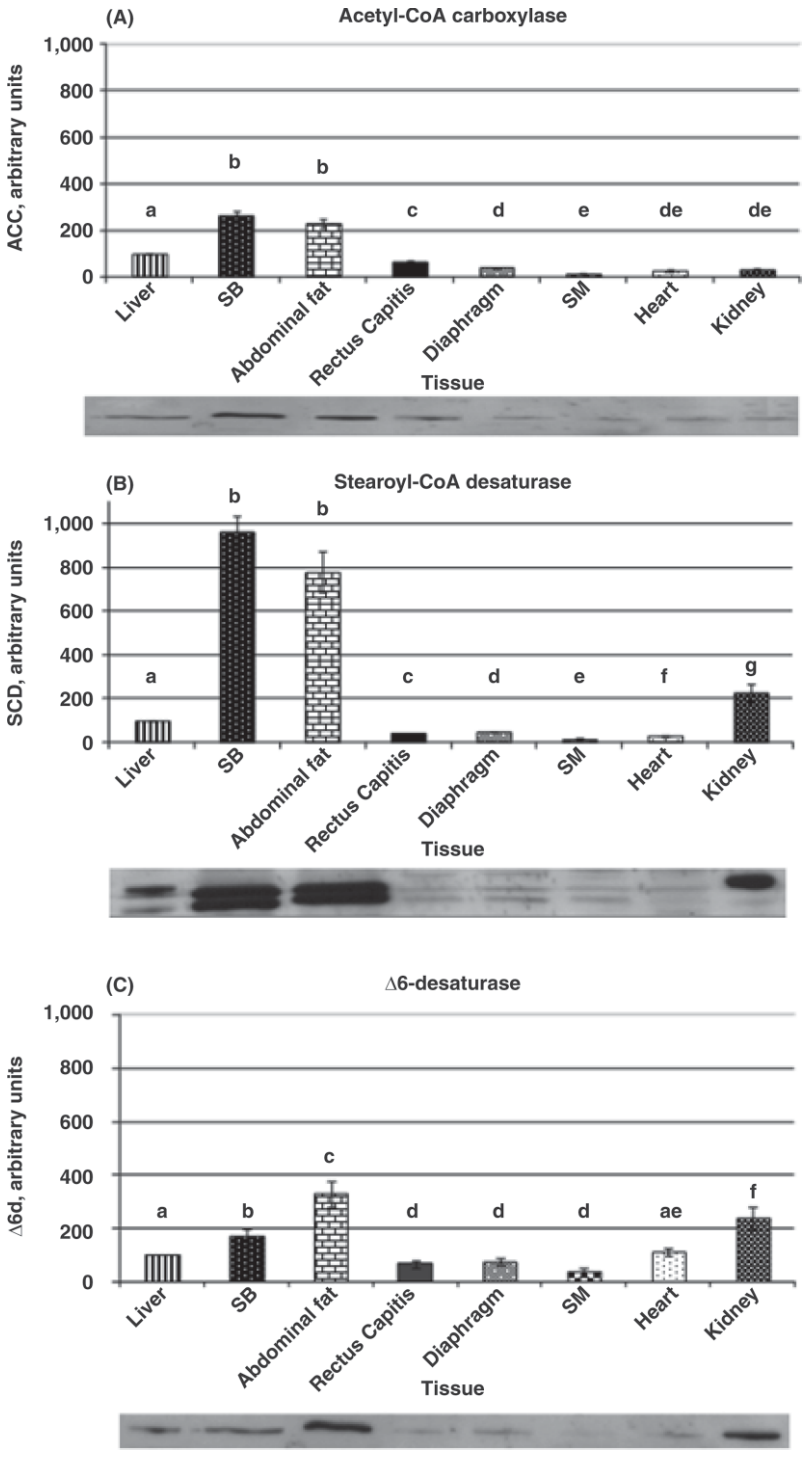


Figure 5. Expression profile of acetyl-CoA carboxylase (ACC), stearoyl-CoA desaturase (SCD), and Δ^6-desaturase (Δ6d) proteins in pig tissues. Each bar represents means of 4 pigs. All measurements were done in duplicate. Error bars represent SEM. [a–g]Means without a common letter differ ($P < 0.05$). Preparations from liver were used as reference samples in all the cases. The intensity of the signal of the reference sample for each enzyme was taken as 100 arbitrary units. The intensity of the signals of other samples on the same blot has been calculated as a fraction of the reference sample. Representative blots of tissue-specific distribution of ACC, SCD, and Δ6d proteins are given under corresponding graphs. SB = subcutaneous adipose tissue; SM = semimembranosus muscle.

that take place in adipose tissue (Ntambi and Miyazaki, 2004). Interestingly, we have also observed greater levels of SCD and Δ6d proteins in kidney (when compared with the liver, rectus capitis muscle, diaphragm, SM, and heart). Moreover, a greater level of ACC, SCD, and Δ6d proteins was also observed in the liver (when compared with the diaphragm, rectus capitis muscle, SM, and heart). Tissue-specific distribution of lipogenic enzymes is well known in other species and might be related to tissue-specific distribution of particular transcription factors and gene-specific promoter signals (Kim and Tae, 1994; Raclot and Oudart, 1999). Results of this study contribute to our understanding of the mechanisms regulating whole body fatty acid metabolism and partitioning in pigs.

Favorable scenarios for IMF content increase during genetic selection are expected as long as predicted breeding values based on IMF records are available (Solanes et al., 2009). At the present time, IMF evaluation in live animals is mainly conducted by ultrasound, which is expensive and not very accurate. Evaluation of IMF in carcasses can be performed by gas chromatography or similar techniques that are time-consuming and also expensive. Several DNA polymorphisms have been considered for developing of genetic tests, but the known polymorphisms only explain a small percentage of variations in IMF. Therefore, it would be beneficial to identify reliable biomarkers for rapid IMF evaluation. From the results of the present study, we conclude that SCD might be effective potential biomarker for fat deposition and partition in pigs. Further validation of the strength of the relationship between the lipogenic enzyme expression and fat content in a larger pig population is required.

LITERATURE CITED

AOAC. 2000. Official Methods of Analysis. 17th ed. Assoc. Off. Anal. Chem., Gaithersburg, MD.

Bosch, L., M. Tor, J. Reixach, and J. Estany. 2009. Estimating intramuscular fat content and fatty acid composition in live and post-mortem samples in pigs. Meat Sci. 82:432–437.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.

Cho, H. P., M. T. Nakamura, and S. D. Clarke. 1999. Cloning, expression, and nutritional regulation of the mammalian delta-6 desaturase. J. Biol. Chem. 274:471–477.

Da Costa, N., C. McGillivray, Q. Bai, J. D. Wood, G. Evans, and K. C. Chang. 2004. Restriction of dietary energy and protein induces molecular changes in young porcine skeletal muscles. J. Nutr. 134:2191–2199.

Decision 2001/775/CE. 2001. Methods for grading pig carcasses in Spain. Off. J. Eur. Commun. 291:50–51.

Doran, O., S. K. Moule, G. A. Teye, F. M. Whittington, K. G. Hallett, and J. D. Wood. 2006. A reduced protein diet induces stearoyl-CoA desaturase protein expression in pig muscle but not in subcutaneous adipose tissue: Relationship with intramuscular lipid formation. Br. J. Nutr. 95:609–617.

Eikelenboom, G., A. H. Hoving-Bolink, and P. G. Van der Wal. 1996. The eating quality of pork. 2. The influence of intramuscular fat. Fleischwirtschaft 76:517–518.

Ellis, M., A. J. Webb, P. J. Avery, and I. Brown. 1996. The influence of terminal sire genotype, sex, slaughter weight, feeding regimen and slaughter hose on growth performance, and carcass and meat quality in pigs and on the organoleptic properties of fresh pork. J. Anim. Sci. 74:521–530.

Enoch, H. G., A. Catala, and P. Strittmatter. 1976. Mechanisms of rat microsomal stearoyl CoA desaturase. J. Biol. Chem. 251:5095–5103.

Fernandez, X., G. Monin, A. Talmant, J. Mourot, and B. Lebret. 1999. Influence of intramuscular fat content on the quality of pig meat—1. Composition of the lipid fraction and sensory characteristics of m. longissimus lumborum. Meat Sci. 53:59–65.

Gardan, D., F. Gondret, and I. Louveau. 2006. Lipid metabolism and secretary function of porcine intramuscular adipocytes compared with subcutaneous and perirenal adipocytes. Am. J. Physiol. Endocrinol. Metab. 291:E372–E380.

Girard, J., D. Perdereau, F. Foufelle, C. Prip-Buus, and P. Ferre. 1994. Regulation of lipogenic enzyme gene expression by nutrients and hormones. FASEB J. 8:36–42.

Gondret, F., N. Guitton, C. Guillerm-Regost, and I. Louveau. 2008. Regional differences in porcine adipocytes isolated from skeletal muscle and adipose tissues as identified by a proteomic approach. J. Anim. Sci. 86:2115–2125.

Goransson, A., G. von Seth, and E. Tornberg. 1992. The influence of intramuscular fat content on the eating quality of pork. Page 245 in 38th Int. Congr. Meat Sci. Technol., 23–28 August, Clermont-Ferrand, France.

Hanson, S. W. F., and J. Olley. 1963. Application of the method of lipid extraction to tissue homogenate. Biochem. J. 89:101–102.

Hermesch, S., B. G. Luxford, and H.-U. Graser. 2000. Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs: 2. Genetic relationships between production, carcase and meat quality traits. Livest. Prod. Sci. 65:249–259.

Kaestner, K. H., J. M. Ntambi, T. J. Kelly Jr., and M. D. Lane. 1989. Differentiation-induced gene expression in 3T3–L1 preadipocytes. A second differentially expressed gene encoding stearoyl-CoA desaturase. J. Biol. Chem. 264:14755–14761.

Kim, H. J., M. Miyazaki, and J. M. Ntambi. 2002. Dietary cholesterol opposes PUFA-mediated repression of the stearoyl-CoA desaturase-1 gene by SREBP-1 independent mechanism. J. Lipid Res. 43:1750–1757.

Kim, K. H., and H. J. Tae. 1994. Pattern and regulation of acetyl-CoA carboxylase gene expression. J. Nutr. 124:1273S–1283S.

Lengi, A. J., and B. A. Corl. 2008. Comparison of pig, sheep and chiken SCD5 homologs: Evidence for an early gene duplication event. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 150:440–446.

McPhee, C. P., and G. R. Trout. 1995. The effects of selection for lean growth and the halothane allele on carcass and meat quality of pigs transported long and short distances to slaughter. Livest. Prod. Sci. 42:55–62.

Miyazaki, M., M. J. Jacobson, W. C. Man, P. Cohen, E. Asilmaz, J. M. Friedman, and J. M. Ntambi. 2003. Identification and characterization of murine SCD4, a novel hart-specific stearoyl-CoA desaturase isoform regulated by leptin and dietary factors. J. Biol. Chem. 278:33904–33911.

Miyazaki, M., and M. Ntambi. 2003. Role of stearoyl-conenzyme A desaturase in lipid metabolism. Prostaglandins Leukot. Essent. Fatty Acids 68:113–121.

Nicolau-Solano, S. I., J. D. McGivan, F. M. Whittington, G. J. Nieuwhof, J. D. Wood, and O. Doran. 2006. Relationship between the expression of hepatic but not testicular 3β-hydroxysteroid dehydrogenase with androstenone deposition in pig adipose tissue. J. Anim. Sci. 84:2809–2817.

Ntambi, J. M., S. A. Buhrow, K. H. Kaestner, R. J. Christy, E. Sibley, T. J. Kelly Jr., and M. D. Lane. 1988. Differentiation-induced gene expression in 3T3–L1 preadipocytes. Characterization of a differentially expressed gene encoding stearoyl-CoA desaturase. J. Biol. Chem. 263:17291–17300.

Ntambi, J. M., and M. Miyazaki. 2004. Regulation of stearoyl-CoA desaturases and role in metabolism. Prog. Lipid Res. 43:91–104.

O'Hea, E. K., and G. A. Leveille. 1969. Significance of adipose tissue and liver as sites of fatty acid synthesis in the pig and the efficiency of utilization of various substrates for lipogenesis. J. Nutr. 99:338–344.

Raclot, T., and H. Oudart. 1999. Selectivity of fatty acids on lipid metabolism and gene expression. Proc. Nutr. Soc. 58:633–646.

Reixach, J., M. Tor, C. Díaz, and J. Estany. 2008. Between and within breed variation for lean growth and intramuscular fat content and fatty acid composition in pigs. Page 9.13 in 54th Int. Congr. Meat Sci. Technol., Cape Town, South Africa. (Abstr.)

Reixach, J., M. Tor, and J. Estany. 2009. Respuesta a la seleccion contra grasa dorsal a grasa intramuscular constante en cerdos Duroc. Pages 120–122 in XIII Jornada sobre Produccion Animal. M. Joy, J. H. Calvo, C. Calvete, M. A. Latorre, I. Casasus, A. Bernues, B. Panea, A. Sanz, and J. Balcells, ed. AIDA, Zaragoza, Spain.

Ren, J., C. Knorr, L. Huang, and B. Brenig. 2004. Isolation and molecular characterisation of the porcine stearoyl-CoA desaturase gene. Gene 340:19–30.

Ruiz, J., J. Ventanas, R. Cava, A. Andres, and C. Garcia. 2000. Texture and appearance of dry-cured ham as affected by fat content and fatty acid composition. Food Res. Int. 33:91–95.

Rule, D. C., M. D. MacNeil, and R. E. Short. 1997. Influence of sire growth potential, time on feed, and growing-finishing strategy on cholesterol and fatty acids of the ground carcass and longissimus muscle of beef steers. J. Anim. Sci. 75:1525–1533.

Schenkman, J. B., and D. L. Cinti. 1978. Preparation of microsomes with calcium. Methods Enzymol. 52:83–89.

Scott, A. R., S. G. Gornelius, and H. J. Mersmann. 1981. Effect of age on lipogenesis and lipolysis in lean and obese swine. J. Anim. Sci. 53:505–511.

Solanes, F. X., J. Reixach, M. Tor, J. Tibau, and J. Estany. 2009. Genetic correlations and expected response for intramuscular fat content in a Duroc pig line. Livest. Sci. 123:63–69.

Thiede, M. A., J. Ozols, and P. Strittmatter. 1986. Construction and sequence of cDNA for rat liver stearoyl coenzyme A desaturase. J. Biol. Chem. 261:13230–13235.

Tor, M., J. Estany, F. Francesch, and M. D. Cubillo. 2005. Comparison of fatty acid profiles of edible meat, adipose tissues and muscles between cocks and capons. Anim. Res. 54:413–424.

Ventanas, S., J. Ruiz, C. Garcia, and J. Ventanas. 2007. Preference and juiciness of Iberian dry-cured loin as affected by intramuscular fat content, crossbreeding and rearing system. Meat Sci. 77:324–330.

Verbeke, W., M. J. Van Oeckel, N. Warnants, J. Viaene, and Ch. V. Boucque. 1999. Consumer perception, facts and possibilities to improve acceptability of health and sensory characteristics of pork. Meat Sci. 53:77–90.

Zheng, Y., S. M. Prouty, A. Harmon, J. P. Sundberg, K. S. Stenn, and S. Parimoo. 2001. Scd3—A novel gene of the stearoyl-CoA desaturase family with restricted expression in skin. Genomics 71:182–191.

IV. Discusión General

En los últimos años, el mercado se ha desplazado hacia un incremento de la demanda de carne de cerdo de alta calidad debido a un mayor interés del consumidor en los aspectos relacionados con la calidad sensorial, nutricional y la salud de los alimentos en general. Así, los productores buscan producir carne de cerdo con unos niveles más bajos de colesterol y triglicéridos sin alterar otros parámetros de la calidad del producto. Por este motivo, en la última década, una parte de la investigación en genética porcina se ha centrado en el estudio de caracteres relacionados con la calidad de la carne desde múltiples perspectivas que incluyen la genética molecular, la nutrigenómica y la genética cuantitativa, entre otras.

El carácter calidad de carne es un carácter complejo ya que engloba varios aspectos y propiedades sensoriales, nutricionales y tecnológicas que influencian el valor del producto para el consumidor y la industria cárnica (López-Bote, 1998). El contenido de lípidos y la composición de fibras musculares son causantes de la variación de una gran diversidad de caracteres relacionados con la calidad de carne de enorme importancia para el sector porcino (Wood et al., 2008) (Tabla 9).

Tabla 9: Correlaciones entre medidas bioquímicas y de la canal del cerdo y características sensoriales en el músculo *longissimus dorsi* (adaptado de Huff-Lonergan *et al*., 2002).

Carácter	GD 10ª costilla		% Lípidos		Colesterol	
	r	p-*value*	r	p-*value*	r	p-*value*
Color	-0.10	0.02	-0.15	0.0001	-0.07	0.03
Textura	0.24	0.0001	0.31	0.0001	0.02	0.72
Terneza	0.19	0.0001	0.19	0.0001	-0.12	0.01
Jugosidad	0.01	0.83	0.25	0.03	-0.09	0.04
Aroma y Sabor	0.24	0.001	0.23	0.0001	-0.04	0.41

Respecto a los lípidos, el porcentaje de GIM se asocia de forma favorable con la textura de la carne, la terneza, el sabor y la jugosidad mientras que la proporción de AG están relacionados con la consistencia de la grasa y el gusto de la carne (Huff-Lonergan et al., 2002). Sin embargo, la industria del porcino lleva muchos años mejorando el contenido de magro de la canal mediante selección contra el espesor de GD (Toro y Silió, 1992), provocando una reducción de la GIM debido a la correlación genética positiva que existe entre GD y GIM (Solanes *et al*., 2009). Por tanto, un objetivo del sector es disponer de metodologías de selección que permitan manipular GD y GIM de forma independiente, así como el contenido y composición de AG muscular.

Por este motivo, en esta tesis doctoral nos hemos centrado en estudiar los factores genéticos que afectan a la GIM y la composición de AG debido al marcado efecto que estos dos caracteres tienen sobre varios atributos como la terneza, la capacidad de retención de agua, la jugosidad, el gusto y el aroma (Olsson y Pickova, 2005 y Candek-Potokar *et al*., 2002).

Hoy en día, con las nuevas tecnologías adaptadas en la especie porcina es posible realizar un estudio más exhaustivo de genes candidatos a ser responsables de la variación genética de caracteres complejos relacionados con la deposición de la grasa. De este modo, en esta tesis se han utilizado y combinado diferentes métodos complementarios como el análisis de la expresión génica global mediante la tecnología de *microarrays*, la detección de eQTL combinado los datos de expresión génica con los estudios de ligamiento de QTL, y estudios más exhaustivos de genes candidatos (incluyendo análisis de la expresión proteica y estudios funcionales en células cultivadas *in vitro*).

Para llevar a cabo el experimento de expresión génica, de una población comercial de 385 machos Duroc distribuidos en cinco familas de medios hermanos, se seleccionaron un total de 70 animales extremos (35 por grupo de ALTO y BAJO nivel de engorde) para una combinación lineal de parámetros de lípidos sanguíneos y engorde utilizando un análisis de componentes principales con el fin de detectar cuáles de las 110 variables fenotípicas registradas explicaban mejor la variabilidad entre individuos. Se utilizaron finalmente 13 caracteres incluidos en la primera componente principal para elaborar un índice de selección con el que clasificar los animales. La variables GIM, AGS y AGMI se agruparon conjuntamente en el mismo eje en la componente principal 1, mientras que los AGPI se agruparon de una manera opuesta, concordando con las correlaciones fenotípicas y las relaciones fisiológicas existentes entre estos caracteres (Solanes *et al*., 2009). El análisis de componentes principales ofrece ventajas frente a la utilización de un solo carácter a la hora de seleccionar aquellos animales extremos para un conjunto de características ya que tiene en cuenta la suma de la variabilidad global de todos aquellos caracteres y variables con una fuerte relación con el metabolismo lipídico y la deposición de la grasa.

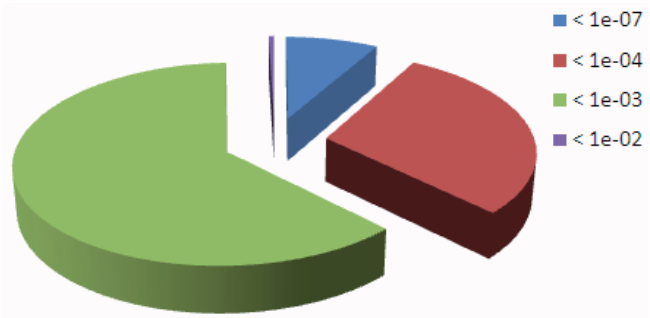


Figura 12: Distribución de los valores de significación de los genes DE entre los grupos de animales de ALTO y BAJO nivel de engorde.

Esta gran capacidad para segregar y seleccionar los animales ha resultado en la detección de un gran número de genes diferencialmente expresados (DE) con valores de significación comprendidos entre $p<10^{-2}$ y $p<10^{-7}$ tras utilizar el método del *false discovery rate* (FDR) para corregir el *multiple testing* (Figura 12).

En este sentido, se ha confirmado diferencias en el patrón de expresión entre los dos grupos de animales en el músculo *gluteus medius*, obteniendo 1060 sondas diferencialmente expresadas (correspondientes a 956 genes), de las cuales, la mayoría de ellas (n=839) estaban sobreexpresadas en el grupo de animales con ALTO nivel de engorde. Además, los valores de tasa de cambio entre ambos grupos estaban comprendidos entre 0.38 y 5.58, aunque únicamente 388 genes obtuvieron valores mayores a 1.5. En este contexto, los valores de tasa de cambio dependen, en su gran mayoría, del pre-procesamiento de los datos y de la normalización de éstos ya que según el algoritmo empleado para analizar los datos de expresión, los valores de tasa de cambio variarán (Murie et al., 2009). Por este motivo, en el análisis de los datos de expresión de esta tesis doctoral se optó por el uso del algoritmo gcRMA, ya que trata los datos de una manera moderada (menos restrictivo que el algoritmo RMA, pero más que el algoritmo MAS5). No obstante, se considera que existen diferencias biológicas con valores de tasa de cambio superiores a 1.5.

Otro aspecto que ha facilitado la comparación de expresión génica entre grupos es la plataforma de *microarray* utilizada. En este contexto, para realizar el experimento de *microarrays*, se utilizó *GeneChip Porcine Genome Arrays*® (*Affymetrix*) ya que ofrecía una mayor representación de sondas (20.201 genes) de 25 bases frente a los 8.000 genes que ofrecía en su momento la plataforma de *QIAgen* (actualmente representa a 20.000 genes). Además, en el caso de la plataforma de *Affymetrix*, se realiza una hibridación no

competitiva, permitiendo hibridar una sola muestra por chip, lo que representa una ventaja frente a la hibridación competitiva a la hora de analizar los datos ya que permite comparar los resultados de cada gen en cada chip sin necesidad de muestras de referencia o de comparación con muestras intermedias. Este punto es especialmente relevante cuando se tiene un gran número de muestras a analizar como ocurre en nuestro caso (n=70).

La heterogeneidad entre plataformas es un problema a la hora de integrar los datos obtenidos de los diferentes grupos, a fin de poder comparar y corroborar resultados (Tuggle *et al*., 2007). Por este motivo, se creó una base de datos de expresión dentro del *National Center for Biotechnology Information* (NCBI) denominada *Gene Expression Omnibus* (GEO) que utiliza un formato específico para introducir y exportar los datos provenientes de experimentos de *microarrays* teniendo en cuenta las diferentes plataformas utilizadas. En este contexto, un estudio comparativo entre estos dos chips de porcino concluyo que el 35% de las sondas concordaban entre ambas plataformas. Las diferencias existentes entre las sondas integrantes de una plataforma u otra se debe, probablemente, a que las sondas provenientes de la plataforma de *QIAgen* fueron diseñadas en el año 2002, mientras que las de *Affymetrix* se diseñaron dos años más tarde (2004), teniendo en cuenta que ésta pudo integrar nueva información sobre la identificación de transcritos alternativos, nuevos exones y procesamientos alternativos (Tuggle *et al*., 2007). Además, mencionar que aunque la concordancia entre las versiones más actuales de ambas plataformas es mayor, en el estudio comparativo del año 2007 (Tuggle *et al*., 2007) se realizó entre los chips de *Affymetrix* de 23K y *QIAgen* de 13K, explicando que solo se obtuviera el 35% de concordancia entre ambas plataformas al haber tanta diferencia en el número de sondas integrantes en cada plataforma.

Tabla 10: Número de sondas anotadas en la especie porcina según la base de datos de *Affymetrix* y la desarrollada por el Dr. Couture (Wang *et al*., 2008).

Base de datos	Nª genes LOC	Secuencias EST/ORF	*Nº genes anotados	*% genes anotados
Affymetrix (Sep. 2008)	1.101	869	2.300	9.5%
Affymetrix (Nov. 2010)	2.752	2.172	5.804	24%
Dr. Couture (2008)	1.213	1.957	17.609	73%

*Sin tener en cuenta las secuencias EST y ORF.

La incompleta anotación de las sondas de los chips de porcino limita el análisis e interpretación biológica de la lista de sondas diferencialmente expresadas derivadas del análisis de expresión génica. En nuestro caso, para anotar, identificar y comprobar los genes diferencialmente expresados derivados del experimento de *microarrays*, se utilizó la base de datos de *Affymetrix* (con fecha septiembre 2008; 9.5% de sondas anotadas) y la desarrollada por el Dr. Couture (Wang et al., 2008; 73% de sondas anotadas) (Tabla 10). Eso nos permitió combinar y comprobar la anotación de los genes con ambas bases de datos. Actualmente, desde que se finalizó la secuenciación del genoma porcino a finales del 2009, se ha notado una mejora significativa en la anotación de los genes como se puede observar notablemente en la base de datos de *Affymetrix* que pasa de tener el 9.5% de genes anotados en el año 2008 a tener el 24% de genes anotados en el año 2010 (Tabla 10). Además, gracias a la integración de diferentes bases de datos dentro de la herramienta *Biomart* de *Ensembl* es posible utilizar protocolos de anotación mucho más eficientes (Ballester et al., 2010). Esta nueva anotación del genoma, cada vez creciente, ha mejorado muchísimo la calidad de la anotación de las sondas de *microarrays*, disminuyendo de esta manera los errores y "vacios" provenientes de la anotación de éstos.

Para dar una interpretación biológica a los resultados obtenidos tras la comparación de los perfiles de expresión de los grupos ALTO y BAJO, utilizamos herramientas de ontología génica. El estudio ontológico tiene como objetivo unificar de un modo jerárquico la información genómica disponible clasificando cada uno de los genes en categorías de tipo biológico, molecular o celular, dividiéndose a su vez en subcategorías interrelacionadas entre sí por un identificador único (Hu *et al*., 2008). Utilizando la lista de 1.060 sondas diferencialmente expresadas, se identificaron cuatro términos ontológicos relacionados con el metabolismo lipídico en la categoría de proceso biológico, mostrando evidencias de la relación existente entre una fracción de genes diferencialmente expresados y procesos biológicos relacionados con el metabolismo lipídico y la deposición de la grasa. Los resultados obtenidos relacionados con el metabolismo lipídico sugieren que los genes relacionados con la activación del catabolismo de los AG como *APOE*, *ACACA*, *DGAT2*, *ELOV5*, *ELOV6*, *FASN*, *LIPIN1*, *LIPE*, *PPARG* y *SCD* se encuentran sobreexpresados en el grupo de cerdos con valores altos en parámetros relacionados con el metabolismo lipídico y deposición de grasa aunque no se ha encontrado una evidencia con una expresión diferencial entre los grupos de ALTO y BAJO nivel de engorde de los genes integrantes de la vía de la β-oxidación probablemente debido a que los genes *ADIPOQ* y *PPARD* podrían estar actuando mediante modificaciones post-traduccionales en las enzimas involucradas en la vía de la β-oxidación en lugar de estar actuando modificando los niveles de expresión en los cerdos con ALTO y BAJO nivel de engorde. Con los resultados obtenidos, aportamos una nueva hipótesis sobre los genes que están regulando el metabolismo lipídico y la deposición de la grasa en músculo porcino obteniendo resultados totalmente opuestos a los resultados observados hasta el momento en humanos y que

sugieren que la oxidación de los AG es más reducida en los individuos obesos (Kiens, 2006).

Aunque los estudios ontológicos nos dan una idea de la clasificación funcional de los genes diferencialmente expresados, es posible realizar un estudio más exhaustivo de las rutas metabólicas donde intervienen estos genes con el objetivo de saber dónde y cómo están actuando los genes diferencialmente expresados. Además, la utilización de herramientas que integren la anotación de las sondas con los valores de *fold-change* y *p-value* asociados a cada una de ellas permite reflejar en los resultados la importancia relativa que tiene cada gen diferencialmente regulado.

En este sentido, se han detectado varios cambios significativos en la expresión de genes de vías metabólicas relacionadas con el metabolismo lipídico, metabolismo de la glucosa e insulina, inmunidad y crecimiento y diferenciación muscular.

En el caso de las vías metabólicas relacionadas con la desaturación de la grasa, se observa una sobre regulación de los genes *ACACA*, *APOE*, *DGAT2*, *FABP4*, *FASN* y *SCD* en los animales pertenecientes al grupo de ALTO nivel de engorde. Este incremento de la expresión de las moléculas responsables de la captación de AG podría representar una respuesta compensatoria a los elevados niveles de lípidos de los animales con ALTO nivel de engorde. Por otro lado, una sobreexpresión de los genes *ACACA* y *FASN*, podría conducir a una acumulación de las reservas de grasa en el músculo debido a que el sustrato de *FASN* (*malonyl-CoA*) es un potente inhibidor de la β-oxidación de los AG, sintetizado a su vez por el gen *ACACA* (Wakil y Abu-Elheiga, 2009). Del mismo modo, una sobreexpresión de los genes *DAGT2* y *SCD* en el músculo podría decantar los AG hacia el almacenamiento en lugar de hacia la oxidación (Flowers y Ntambi, 2008).

Estos resultados están relacionados directamente con la calidad de la carne, ya que las diferencias de expresión observadas en estos genes tienen un marcado efecto sobre el porcentaje y composición de la GIM, asociándose favorablemente con la textura de la carne, la terneza, el sabor y la jugosidad (Huff-Lonergan, *et al*., 2002). Además, los genes *ACACA*, *FABP4*, *FASN* y *SCD* desempeñan un papel muy importante en la síntesis de AG, afectando a la consistencia de la grasa y al gusto de la carne, dos atributos muy importantes en la calidad de la carne. Por esta razón, dos de estos genes (*ACACA* y *SCD*) han estado analizados como genes candidatos confirmando su relación con la composición de lípidos de la grasa subcutánea y muscular al estar directamente asociados con la síntesis de AGS y AGMI respectivamente.

En relación a los genes implicados en el metabolismo de la glucosa y la insulina, se ha obtenido una expresión diferencial de los genes y receptores más importantes relacionados con la regulación de la glucosa y la insulina como son *GLUT4* e *IRS2* (ambos validados mediante qPCR). Estos genes están sobrexpresados en los individuos con altos niveles de colesterol y mayor engorde concordando con la hipótesis de que en la lipogénesis es posible la utilización de glucosa como sustrato.

El estudio de expresión génica muscular se ha complementado mediante la caracterización de eQTL con el objetivo de identificar las posibles regiones genómicas que pueden estar controlando la expresión génica en el músculo esquelético porcino. Este trabajo se ha realizado combinando los datos de expresión génica derivados del experimento de *microarrays* con los genotipados de marcadores moleculares de un trabajo anterior en la misma población (Gallardo et al., 2009). En este sentido, se utilizan todas aquellas sondas que mostraron una mínima variabilidad en la expresión de los genes (6.139 sondas) teniendo en cuenta la expresión génica en el *gluteus medius* de

un mayor número de animales (n=105), detectando 613 eQTL ($p<0.5$) distribuidos de forma no uniforme a lo largo de los 18 autosomas.

Debido a la incompleta anotación del genoma porcino disponible hasta el momento, únicamente 478 fueron mapeados (63 en *cis*-, es decir, eQTL físicamente cerca del gen/transcrito diana y 415 en *trans*- localizados lejos del gen/transcrito diana, normalmente en otro cromosoma).

Tabla 11: Comparación de resultados en términos de número de eQTL entre esta Tesis Doctoral y los trabajos de eQTL publicados en la especie porcina (diciembre 2010).

Total eQTL mapeadas	*cis*-eQTL	*trans*-eQTL	Tejido	Carácter	Referencia
104	8	96	*Longissimus dorsi*	Capacidad retención agua	Ponsuksili *et al.*, 2008
[a)]25	13	12			
[b)]139	35	104	*Longissimus dorsi*	Conductividad, pH y textura	Ponsuksili *et al.*, 2010*
[c)]955	125	830			
21	7	14	*Longissimus dorsi*	Capacidad retención agua	Ponsuksili *et al.*, 2010
478	63	415	*Gluteus medius*	Contenido/composición grasa	Datos presente tesis

*Resultados según el umbral de significación aplicado: a) $3.52*10^{-6}$; b) $4.50*10^{-5}$; c) $4.59*10^{-4}$.

Si comparamos estos resultados con otros trabajos de eQTL publicados en diferentes organismos, se pueden observar trabajos donde obtienen una mayor proporción de *cis*-eQTL respecto al número obtenido de *trans*-eQTL como es el caso de Goring *et al*. (2007) o Schadt *et al*. (2008) mientras que otros autores detectan un número más elevado de *trans*-eQTL (Morley *et al*., 2004; Myers *et al*., 2007; Ponsuksili *et al*., 2008; Ponsuksili *et al*., 2010) coincidiendo con los resultados presentados en esta tesis doctoral donde se han identificado una mayor proporción de *trans*-eQTL respecto a *cis*-eQTL (Tabla 11). Estas diferencias pueden ser debidas, probablemente, al tamaño muestral utilizado en los diferentes experimentos así como los lindares y el nivel de significación establecidos (Cheung y Spielman, 2009). Además, el alto número de

eQTL detectados (613), podría ser debido por una parte, al diseño experimental utilizado. El diseño de medios hermanos utilizado en nuestra población comercial Duroc es una estructura idónea para la detección de eQTL ya que analiza la presencia de eQTL a partir de las diferencias fenotípicas entre las familias de descendientes (Giovambattista *et al*., 1998).

A fin de identificar aquellos eQTL directamente relacionados con los caracteres bajo estudio, se realizó un análisis de concordancia posicional entre los eQTLs significativos a nivel genómico y los QTLs (para los caracteres de GIM, GD, composición de AG (AGS, AGI y AGPI) y colesterol en músculo) identificados previamente en la misma población Duroc analizada (Gallardo 2009 y resultados en proceso de publicación), resultando 80 coincidencias posicionales entre eQTL y QTL para GIM, porcentaje de GD, contenido de colesterol en músculo y contenido de varios AG (66 *trans*-eQTL y 14 *cis*-eQTL). El porcentaje de concordancia posicional resultante entre los eQTL y QTL (16.7%) avala la idoneidad del diseño experimental para detectar eQTLs directamente relacionados con el metabolismo lipídico y la deposición de la grasa. Con el propósito de corroborar la relación de los posibles genes regulados por los *cis*- y *trans*-eQTLs, con los caracteres relacionados con la calidad de la carne, el metabolismo lipídico y la deposición de grasa, se realizó un análisis de correlación entre la expresión de los genes diana y varios fenotipos relacionados con los caracteres a estudiar. De esta manera, se ha confirmado la relación de los 29 posibles genes candidatos seleccionados con caracteres relacionados con la calidad de la carne.

Estos resultados permiten elaborar una lista de posibles genes responsables de la variación de caracteres relacionados con la deposición de la grasa y de gran interés económico y productivo para el sector porcino. Es necesario un estudio más exhaustivo

(caracterización estructural y funcional, estudios de asociación, etc) de estos posibles genes candidatos a fin de validar el efecto que tienen sobre los diferentes caracteres como la GIM y el porcentaje de AG con el objetivo de poderlos introducir, en un futuro, dentro de los esquemas de selección porcina.

Dentro de los 25 genes seleccionados para validar la expresión diferencial entre los grupos de ALTO y BAJO nivel de engorde derivados del experimento de *microarrays* se encontraron los genes *ACACA* y *SCD* relacionados con la síntesis de AG, resultando el gen *SCD* uno de los 6 genes que no se validaron al no obtener una expresión diferencial significativa mediante qPCR. Dentro del marco de esta tesis doctoral, se decidió realizar un estudio más exhaustivo de estos dos genes por su relación funcional y fisiológica con las rutas metabólicas relacionadas con la elongación y desaturación de grasas y con la síntesis de colesterol (Figura 13), seleccionándolos como genes candidatos junto con los genes *HMGCR* y *Δ6D.*

El gen *HMGCR* está directamente relacionado con la síntesis del colesterol (Friesen y Rodwell, 2004) y con la expresión de los receptores *LDL* en el hígado (Kajinami et al., 2004). Durante la caracterización de este gen se estudió el perfil de expresión en seis tejidos implicados en la biosíntesis y el transporte del colesterol y observamos que el gen *HMGCR* porcino mostraba la expresión máxima en el duodeno y en la grasa, mientras que el hígado fue el tejido que mostró una menor expresión, concordando con el hecho que, en la especie porcina, el órgano lipogénico más importante es el tejido adiposo, mientras que en mamíferos la biosíntesis del colesterol se produce sobretodo en hígado, duodeno y grasa (Champe *et al*., 2008).

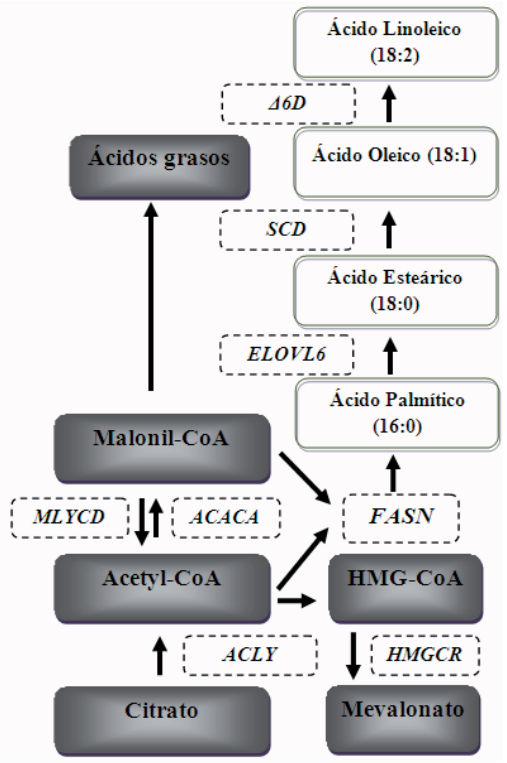


Figura 13: Representación gráfica de la actuación de los genes candidatos *HMGCR*, *ACACA*, *SCD* y *Δ6D*.

Además, hemos detectado correlaciones entre la expresión del gen *HMGCR* en músculo y el contenido de GIM, niveles de colesterol y diferentes AG como el ácido oleico, linoleico y palmítico. Estos resultados sugieren que elevadas concentraciones de lípidos en el plasma sanguíneo podrían estar contribuyendo a favorecer la deposición de la grasa en el tejido muscular mediante la regulación de la expresión del gen *HMGCR* (Nagao y Yanagita, 2008). Los resultados totalmente opuestos en la respuesta *in vitro* del promotor proximal a diferentes estimuladores e inhibidores de su actividad transcripcional entre las células musculares y hepáticas sugiriere la existencia de diferencias tejido-específicas en la regulación de la expresión por factores de

transcripción en cada tejido. Esta hipótesis concuerda con los diferentes patrones de metilación en muestras de hígado y músculo observado en el promotor *HMGCR* porcino relacionándose con las diferencias en la expresión de este gen así como en las funciones celulares (Illingworth *et al*., 2008).

Por tanto, los resultados obtenidos en estos estudios permiten concluir que el gen *HMGCR* porcino está asociado no sólo con los niveles de lípidos plasmáticos, sino que también está afectando a diferentes caracteres productivos relacionados con la calidad de la carne como es el contenido y composición de la GIM sugiriéndolo como gen candidato para tenerlo en cuenta dentro de los esquemas de selección porcina.

Para el estudio de los genes candidatos *ACACA*, *SCD* y *Δ6D* relacionados con la síntesis de AGS, AGMI y AGPI respectivamente, se utilizaron animales de la misma población base Duroc, seleccionando en este experimento animales con menor GD a GIM constante.

La importancia de estas enzimas en relación con la biosíntesis de lípidos y la deposición de grasas se ve reflejada en la máxima expresión de las tres enzimas en la grasa subcutánea y la grasa abdominal respecto a tejidos con otras funciones fisiológicas (hígado, musculo *rectus capitis*, musculo *semimembranoso*, diafragma, corazón y riñón),

No se observaron diferencias en los niveles proteicos de ninguna de las tres enzimas lipogénicas en el músculo *semimembranoso*, pero sí en la grasa subcutánea, mostrando una disminución de la expresión de *ACACA* y *SCD* respecto a un grupo control. La expresión proteica de *SCD* se correlacionaba también con el contenido de AGMI en la grasa subcutánea. La regulación tejido-específica de esta enzima lipogénica podría ser

debida a una presencia diferencial de las diferentes isoformas en estos tejidos; de hecho, se han descrito diferencias morfológicas y metabólicas entre los diferentes depósitos de grasa del cerdo (Gondret et al., 2008). Sin embargo, los niveles de *Δ*6D no difirieron entre los dos grupos analizados en ninguno de los dos tejidos.

Tanto el gen *ACACA* como el gen *SCD* resultaron ser genes sobreexpresados en los animales del grupo ALTO del experimento de *microarrays*. Por otra parte, los resultados de los análisis de eQTL, revelaron que el gen *ACACA* podría ser susceptible a ser regulado por polimorfismos localizados en el gen *ALPK3* (*alpha-kinase 3*), relacionado con la regulación del canal K-ATP de la actividad de las células-β del páncreas (Stanojevic et al., 2008), ya que se detectó un *trans*-eQTL en la misma posición donde está localizado en gen *ALPK3*. Asimismo, en el estudio posicional entre los eQTL detectados y los QTL previamente identificados (Gallardo et al 2008), se obtuvo una concordancia posicional entre el *trans*-eQTL correspondiente al gen *ACACA* y un QTL para GIM identificado en el SSC7 porcino. Además se ha observado una correlación positiva entre la expresión del gen *ACACA* en el musculo *gluteus medius* porcino y el contenido de GIM lo cual confirma las observaciones de Gallardo *et al*., 2008 que vinculan polimorfismos de este gen y GIM.

Estos resultados, y todos los expuestos anteriormente confirman el interés de *HMGCR*, *SCD* y *ACACA* como verdaderos genes candidatos para caracteres relacionados con la calidad de la carne tan importante como la GIM y el contenido de AG. Estos genes desempeñan, además, un papel clave en la salud humana, ya que los niveles de AGMI y AGPI en la dieta se relacionan con la susceptibilidad a padecer enfermedades cardiovasculares al alterar los niveles de LDL y colesterol sin reducir los niveles de HDL (Ventanas et al., 2007).

Finalmente, concluir que gracias a la combinación de las estrategias, aproximaciones y tecnologías utilizadas hoy en día en los estudios genómicos nos está permitiendo ampliar los conocimientos sobre la base genética de los caracteres complejos pudiendo conocer mejor los mecanismos de regulación de la expresión génica subyacentes. Además, recientemente se han desarrollado chips de genotipado masivo de alta densidad en la especie porcina (PorcineSNP60; Illumina) abriendo las puertas de la selección genómica (GS) en la especie porcina en el futuro inmediato para el estudio de la variabilidad estructural a nivel genómico que complementaria los estudios funcionales aquí presentados. En este contexto, la utilización de datos de genotipados masivos provenientes de los chips de SNP de alta densidad permitiría una mejor caracterización de estos eQTL, incluyendo una mejora muy significativa en la resolución de los mapas de eQTL, así como en la potencia estadística de este tipo de trabajos.

V. Conclusiones

1.-Existen cambios significativos en la expresión de genes en el músculo *gluteus medius* porcino entre animales con fenotipos extremos para nivel de engorde de la población Duroc analizada. Estos genes están implicados en rutas metabólicas relacionadas con el metabolismo lipídico, el crecimiento y la diferenciación muscular, la inmunidad y el metabolismo de la insulina y la captación de glucosa.

2.- Se han descrito numerosas regiones genómicas responsables de la variación de expresión génica en el músculo *gluteus medius* porcino. Un análisis de QTL de expresión (eQTL) ha detectando 613 regiones, correspondientes a 63 *cis*- y 415 *trans*-eQTL, distribuidos de forma no uniforme a lo largo de los 18 autosomas.

3.- Once de los eQTL detectados están afectando los niveles de expresión de un elevado número de genes (eQTL *hotspots*) incluyendo los funcionalmente relacionados con el metabolismo lipídico y la deposición de las grasas.

4.- El gen *HMGCR* porcino muestra variación estructural tanto a nivel de región codificante como a nivel de promotor. El polimorfismo de la región codificante está asociado con la concentración plasmática de colesterol LDL y caracteres relacionados con la calidad de la carne como la grasa intramuscular y varios ácidos grasos.

5.- Dos polimorfismos en las posiciones -225 y -1 del promotor del gen *HMGCR* porcino, podrían estar afectando a sitios de unión de los factores de transcripción bHLH y Ets-2 respectivamente.

6.- La secuenciación del promotor *HMGCR* porcino reveló la existencia de varias secuencias *consensus* para varios factores de transcripción conservados en otros mamíferos. La actividad de este promotor responde a la acción de diferentes activadores e inhibidores como la insulina, el colesterol y el forskolin. El polimorfismo localizado en la posición -225 afecta la actividad transcripcional del promotor *HMGCR* porcino de manera diferente en las células hepáticas HepG2 y en las musculares C2C12. Esta diferencia de activación podría estar ligada a las diferencias en el patrón de metilación de los motivos CpG del promotor del gen *HMGCR* porcino entre los tejidos de músculo e hígado.

7.- Se ha estudiado la expresión proteica de tres genes relacionados con la elongación y desaturación de las grasas. Los niveles de ACACA y SCD en la grasa subcutánea fue menor en los animales seleccionados por tener menor grasa dorsal a grasa intramuscular constante, mientras que los niveles de Δ6D, no mostraron diferencias significativas entre los grupos de animales seleccionados y control.

8.- Además, la expresión proteica del gen *SCD* se correlaciona positivamente con el contenido de grasa intramuscular y ácido oleico en el musculo semimembranoso, apuntando el gen *SCD* como un potencial biomarcador efectivo para la deposición de grasa intramuscular.

9.- Los genes *ACACA*, *HMGCR* y *SCD* están diferencialmente expresados entre los animales del grupo ALTO y BAJO nivel de engorde de la población Duroc analizada. Además, el gen *ACACA* está regulado por un *trans*-eQTL localizado en el cromosoma 7

porcino. Se reafirma así, el papel relevante de estos tres genes en los mecanismos moleculares implicados en la deposición de la grasa en el músculo porcino.

10.- A partir de la combinación de los resultados derivados del análisis de expresión génica, mapas de eQTL y genes candidatos se propone una lista de genes candidatos potenciales para estudiar más a fondo en un futuro mediante técnicas complementarias como los chips de SNP de alta densidad y metodologías de modelización de las redes génicas que permitan elucidar los mecanismos moleculares implicados en el metabolismo de los lípidos y los caracteres relacionados con la calidad de la carne.

VI. Referencias

Abu-Elheiga, L., Jayakumar, A., Baldini, A., Chirala, S.S., Wakil, S.J. 1995. Human acetyl-CoA carboxylase: characterization molecular cloning, and evidence for two isoforms. *Proc Natl Acad Sci U S A,* 92**,** 4011-4015.

Aguado, M. 2007. Microarrays de DNA en microbiología. *Revista complutense de ciencias veterinarias.,* 1.

Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet,* 2**,** 130-8.

Bai, Q., McGillivray, C., da Costa, N., Dornan, S., Evans, G., Stear, M. J. & Chang, K. C. 2003. Development of a porcine skeletal muscle cDNA microarray: analysis of differential transcript expression in phenotypically distinct muscles. *BMC Genomics,* 4**,** 8.

Ballester, B., Johnson, N., Proctor, G., Flicek, P. 2010. Consistent annotation of gene expression arrays. *BMC Genomics,* 11**,** 294.

Barber, M. C., Pooley, L., Travers, M. T. 2001. Developmental regulation of alternatively spliced acetyl-CoA carboxylase-alpha mRNAs encoding isozymes with or without an eight amino acid domain upstream of the Ser-1200 phosphorylation motif in the mammary gland. *J Mol Endocrinol,* 27**,** 349-56.

Barber, M. C., Price, N. T., Travers, M. T. 2005. Structure and regulation of acetyl-CoA carboxylase genes of metazoa. *Biochim Biophys Acta,* 1733**,** 1-28.

Berger, A., Roberts, M.A. 2005. Unraveling Lipid Metabolism with Microarrays. *E. Berger & M.A. Roberts Eds. Marcel Dekker, New York, USA.*

Bernard, C., Cassar-Malek, I., Martine, L. C., Dubroeucq, H., Renand, G., Hocquette, J.F. 2007. New indicators of beef sensory quality revealed by expression of specific genes. *Journal of agricultural and food chemistry,* 55**,** 5229-5237.

Bonneau, M., Lebret, B. 2010. Production systems and influence on eating quality of pork. *Meat Science,* 84**,** 293-300.

Bucharles, C., Girard J, Yuan, C, Bonet, M. 1985. Les lipides animaux dans la filière viande.

Burkhardt, R., Kenny, E. E., Lowe, J. K., Birkeland, A., Josowitz, R., Noel, M., Salit, J., Maller, J. B., Pe'er, I., Daly, M. J., Altshuler, D., Stoffel, M., Friedman, J. M., Breslow, J. L. 2008. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. *Arterioscler Thromb Vasc Biol,* 28**,** 2078-84.

Cagnazzo, M., te Pas, M. F., Priem, J., de Wit, A. A., Pool, M. H., Davoli, R., Russo, V. 2006. Comparison of prenatal muscle tissue expression profiles of two pig breeds differing in muscle characteristics. *J Anim Sci,* 84**,** 1-10.

Calvo, J. H., Lopez-Corrales, N. L., Anderson, S. I., Skinner, T. M., Marcos, S., Osta, R., Archibald, A. L., Zaragoza, P. 2000. Assignment of acetyl-coenzyme A carboxylase alpha (ACACA) to pig chromosome 12 (12p13-->p12) by fluorescence in situ hybridization and confirmation by genetic mapping. *Cytogenet Cell Genet,* 90**,** 238-9.

Candek-Potokar, M., Monin, G., Zlender, B. 2002. Pork quality, processing, and sensory characteristics of dry-cured hams as influenced by Duroc crossing and sex. *J Anim Sci,* 80**,** 988-96.

Champe, P. C., Harvey, R.A., Ferrier, D.R. 2008. Cholesterol and Steroid Metabolism. *Biochemistry, Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia*.

Chasman, D. I., Posada, D., Subrahmanyan, L., Cook, N. R., Stanton, V. P., Jr. & Ridker, P. M. 2004. Pharmacogenetic study of statin therapy and cholesterol reduction. *JAMA,* 291**,** 2821-7.

Cheung, V. G., Spielman, R.S. 2009. Genetics of human gene expression:mapping DNA variants that influence gene expression. *Nature Reviews Genetics*.

Chmurzynska, A., Mackowski, M., Szydlowski, M., Melonek, J., Kamyczek, M., Eckert, R. 2004. Polymorphism of intronic microsatellites in the A-FABP and LEPR genes and its association with productive traits in the pig. *Journal of Animal Feed Science,* 13**,** 615-624.

Chmurzynska, S., Szydlowski, M., Stachowiak, M., Stankiewicz, M., Switonski, M. 2007. Association of a new SNP in promoter region of the porcine FABP3 gene with fatness traits in a Polish synthetic line. *Animal Biotechnology,* 18**,** 37-44.

Cieslak, J., Nowacka-Woszuk, J., Bartz, M., Fijak-Nowak, H., Grzes, M., Szydlowski, M., Switonski, M. 2009. Association studies on the porcine RETN, UCP1, UCP3 and ADRB3 genes polymorphism with fatness traits. *Meat Science,* 83**,** 551-554.

Ciobanu, D. C., Bastiaansen, J. W., Lonergan, S. M., Thomsen, H., Dekkers, J. C., Plastow, G. S., Rothschild, M. F. 2004. New alleles in calpastatin gene are associated with meat quality traits in pigs. *J Anim Sci,* 82**,** 2829-39.

Clop, A., Ovilo, C., Perez-Enciso, M., Cercos, A., Tomas, A., Fernandez, A., Coll, A., Folch, J. M., Barragan, C., Diaz, I., Oliver, M. A., Varona, L., Silio, L., Sanchez, A., Noguera, J. L. 2003. Detection of QTL affecting fatty acid composition in the pig. *Mamm Genome,* 14**,** 650-6.

Coma, J., Piquer, J. 1999. Calidad de carne en porcino: efecto de la nutrición. *XV Curso de Especialización. Avances en Nutrición y Alimentación Animal*, 1-22.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., Lathrop, M. 2009. Mapping complex disease traits with global gene expression. *Nat Rev Genet,* 10, 184-94.

da Costa, N., McGillivray, C., Bai, Q., Wood, J. D., Evans, G., Chang, K. C. 2004. Restriction of dietary energy and protein induces molecular changes in young porcine skeletal muscles. *J Nutr,* 134, 2191-9.

Davis, A. M., White, B. A., Wheeler, M. B. 1995. Rapid communication: a DraI restriction fragment length polymorphism at the porcine 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCoAR) locus. *J Anim Sci,* 73, 310.

Davoli, R., Braglia, S. 2007. Molecular approaches in pig breeding to improve meat quality. *Brief Funct Genomic Proteomic,* 6, 313-21.

de Koning, D. J., Cabrera, C. P., Haley, C. S. 2007. Genetical genomics: combining gene expression with marker genotypes in poultry. *Poult Sci,* 86, 1501-9.

de Koning, D. J., Janss, L. L., Rattink, A. P., van Oers, P. A., de Vries, B. J., Groenen, M. A., van der Poel, J. J., de Groot, P. N., Brascamp, E. W., van Arendonk, J. A. 1999. Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa). *Genetics,* 152, 1679-90.

Directorate-General for Agriculture and Rural Development. 2007. Prospects for agricultural markets and income in the european union 2007-2014. *European Comission*, 1-55.

Doran, O., Moule, S. K., Teye, G. A., Whittington, F. M., Hallett, K. G., Wood, J. D. 2006. A reduced protein diet induces stearoyl-CoA desaturase protein expression in pig muscle but not in subcutaneous adipose tissue: relationship with intramuscular lipid formation. *Br J Nutr,* 95, 609-17.

Estany, J., Tor, M., Villalba, D., Bosche, L., Gallardo, D., Jiménez, N., Altet, L., Noguera, J.L., Reixach, J., Amills, M., Sánchez, A. 2007. Association of CA repeat polymorphism at intron 1 of insulin-like growth facto (IGF-1) gene with circulating IGF-1 concentration, growth, and fatness in swine. *Physiological Genomics,* 31, 236-243.

Fan, B., Onteru, S.K., Nikkilä, M.T., Stalder, K.J., Rothschild, M. 2009a. Identification of genetic markers associated with fatness and leg weakness traits in the pig. *Animal Genetics,* 40, 967-970.

Fan, B., Onteru, S.K., Plastow, G.S., Rothschild, M. 2009b. Detail characterization of the porcine MC4R gene in relation to fatness and growth. *Animal Genetics,* 40, 401-409.

Ferraz, A. L., Ojeda, A., Lopez-Bejar, M., Fernandes, L. T., Castello, A., Folch, J. M., Perez-Enciso, M. 2008. Transcriptome architecture across tissues in the pig. *BMC Genomics,* 9**,** 173.

Flowers, M.T., Ntambi, J.M. 2008. Role of stearoyl-coenzyme A desaturase in regulating lipid metabolism. *Curr Opin Lipidol*, 19:248-256.

Fontanesi, L., Scotti, E., Buttazzoni, L., Dall'Olio, S., Davoli, R., Russo, V. 2010. A single nucleotide polymorphism in the porcine cathepsin K (CTSK) gene is associated with back fat thickness and production traits in Italian Duroc pigs. . *Mol Biol Rep,* 37**,** 491-495.

Franke, L., Jansen, R. C.2009. eQTL analysis in humans.*Methods Mol Biol,*573**,**311-28.

Friesen, J. A., Rodwell, V. W. 2004. The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol,* 5**,** 248.

Fujii, J., Otsu, K., Zorzato, F., de Leon, S., Khanna, V. K., Weiler, J. E., O'Brien, P. J., MacLennan, D. H. 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science,* 253**,** 448-51.

Gallardo, D., Canovas, E., Lopez-Bejar, M., Ramirez, O., Pena, R., Quintanilla, R., Amills, M. 2008. Alternative splicing at exon 28 of the acetyl-coenzyme A carboxylase alpha gene in adult pigs and embryos. *Anim Genet,* 39**,** 205-6.

Gallardo, D., Quintanilla, R., Varona, L., Diaz, I., Ramirez, O., Pena, R. N., Amills, M. 2009. Polymorphism of the pig acetyl-coenzyme A carboxylase alpha gene is associated with fatty acid composition in a Duroc commercial line. *Anim Genet,* 40**,** 410-7.

Germán, C., Camacho, J, Gallegos, J. 2005. Producción de cerdos. *Institución de enseñanza e investigación en ciencias agrícolas***,** 1-21.

Giovambattista, G., Ripoli, M.V., Peral-García, P., De Lúea, J.C., Dulout, F.N. 1998. Mapping of QTLs associated with milf production: present status. *Agro Sur,* 26**,** 78-96.

Girard, J., Perdereau, D., Founfelle, F., Prip-Buus, C., Ferre, P. 1994. Regulation of lipogenic enzyme gene expression by nutrients and hormones. *FASEB J,* 8**,** 36-42.

Gondret, F., Guitton, N., Guillerm-Regost, C., Louveau, I. 2008. Regional differences in porcine adipocytes isolated from skeletal muscle and adipose tissues as identified by a proteomic approach. *J Anim Sci,* 86**,** 2115-25.

Göring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JB, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, MacCluer

JW, Kissebah AH, Collier GR, Moses EK, Blangero J. 2007. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet,* 39**,** 1208-1216.

Guillou, H., Zadravec, D., Martin, P.G., Jacobsson, A. 2010. The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. *Progress in Lipid Research,* 49**,** 186-199.

Guo, T., Ren, J., Yang, K., Ma, J., Zhang, Z., Huang, L. 2009. Quantitative trait loci for fatty acid composition in longissimus dorsi and abdominal fat: results from a White Duroc x Erhualian intercross F2 population. *Anim Genet,* 40**,** 185-91.

Harshman, K. 2002. Detección y ácidos nucleicos y nanotecnología: El gran potencial de los ensayos a pequeña escala. *Encuentros multidisciplinares nº12. ,* http://www.encuentros-multidisciplinares.org/.

Hausman, G. J., Poulos, S.P., Richarson, R.L., Barb, C.R., Andacht, T., Kirk, H.C., Mynatt, R.L. 2006. Secreted proteins and genes in fetal and neonatal pig adipose tissue and stromal-vascular cells. *Journal of Animal Science,* 84**,** 1666-1681.

Hofmann, K. 1994. What is quality. *Meat Focus***,** 73-82.

Hovatta, I., Zapala, M. A., Broide, R. S., Schadt, E. E., Libiger, O., Schork, N. J., Lockhart, D. J., Barlow, C. 2007. DNA variation and brain region-specific expression profiles exhibit different relationships between inbred mouse strains: implications for eQTL mapping studies. *Genome Biol,* 8**,** R25.

Hu, Z. L., Bao, J, Reecy, J.M. 2008. CateGOrizer: a web-based program to batch analyze gene ontology classification categories. *Online Journal of Bioinformatics,* 9**,** 108-112.

Hu, Z. L., Reecy, J. M. 2007. Animal QTLdb: beyond a repository. A public platform for QTL comparisons and integration with diverse types of structural genomic information. *Mamm Genome,* 18**,** 1-4.

Huff-Lonergan, E., Baas, T. J., Malek, M., Dekkers, J. C., Prusa, K., Rothschild, M. F. 2002. Correlations among selected pork quality traits. *J Anim Sci,* 80**,** 617-27.

Ibeagha-Awemu, E. M., Kgwatalala, P., Zhao, X. 2008. A critical analysis of production-associated DNA polymorphisms in the genes of cattle, goat, sheep, and pig. *Mamm Genome,* 19**,** 591-617.

Illingworth, R., Kerr, A., Desousa D, Jørgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A. 2008. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol,* 6**,** e22.

Jansen, R. C., Nap, J.P. 2001. Genetical genomics: the added value from segregation *Trends Genet,* 17**,** 388-91.

Jiang, Z., Michal, J.J., Tobey, D.J., Daniels, T.F., Rule, D.C., MacNeil, M.D. 2008. Significant associations of stearoyl-CoA desaturase (SCD1) gene with fat deposition and composition in skeletal muscle. *International journal of biological sciences,* 4**,** 345-351.

Jiang, Z. H., Gibson, J.P. 1999. Genetic polymorphisms in the leptin gene and their association with fatness in four pig breeds. *Mamm Genome,* 10**,** 191-193.

Kadarmideen, H. N., Janss, L. L. 2007. Population and systems genetics analyses of cortisol in pigs divergently selected for stress. *Physiol Genomics,* 29**,** 57-65.

Kajinami, K., Takekoshi, N., Brousseau, M. E., Schaefer, E. J. 2004. Pharmacogenetics of HMG-CoA reductase inhibitors: exploring the potential for genotype-based individualization of coronary heart disease management. *Atherosclerosis,* 177**,** 219-34.

Kiens, B. 2006. Skeletal muscle lipid metabolism in exercise and insulin resistance. *Physiol Rev,* 86**,** 205-243.

Kim, K. S., Larsen, N., Short, T., Plastow, G., Rothschild, M. F. 2000. A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm Genome,* 11**,** 131-5.

Le Dividich, J., Esnault, T., Lynch, B., Hoo-Paris, R., Castex, C., Peiniau, J. 1991. Effect of colostral fat level on fat deposition and plasma metabolites in the newborn pig. *J Anim Sci,* 69**,** 2480-8.

Li, M., Zhu, L., Shuai, S., Teng, X., Xiao, H., Li, Q., Chen, L., Guo, Y., Wang, J. 2008. Expression profiling analysis for genes related to meat quality and carcass traits during postnatal development of backfat in two pig breeds. *Science in China Series C: Life Sciences,* 51**,** 718-733.

Li, Y., Xu, Z., Li, H., Xiong, Y, Zuo, B. 2010. Differential transcriptional analysis between red and white skeletal muscle of chinese Meishan pigs. *International journal of biological sciences,* 6**,** 350-360.

Lin, C. S., Hsu, C.W. 2005. Differential transcribed genes in skeletal muscle of Duroc and Taoyuan pigs. *Journal of Animal Science,* 83**,** 2075-2086.

Liu, K., Wang, G., Zhao, S.H., Liu, B., Huang, J.N., Bai, X., Yu, M. 2010. Molecular characterization, chromosomal location, alternative splicing and polymorphism of porcine GFAT1 gene. *Mol Biol Rep*.

Liu, Z., Jeppesen, P. B., Gregersen, S., Chen, X., Hermansen, K. 2008. Dose- and Glucose-Dependent Effects of Amino Acids on Insulin Secretion from Isolated Mouse Islets and Clonal INS-1E Beta-Cells. *Rev Diabet Stud,* 5**,** 232-44.

Lobjois, V., Liaubet, L., SanCristobal, M., Glénisson, J., Rallières, J., Le Roy, P., Milan, D., Cherel, P., Hatey, F. 2008. A muscle transcriptome analysis identifies positional candidate genes for a complex trait in pig. *Animal Genetics,* 39**,** 147-162.

López-Bote, C. 1998. Sustained utilization of the Iberian pig breed. *Meat Sci,* 49 Suppl. 1**,** 17-27.

Lunney, J. K. 2007. Advances in swine biomedical model genomics. *Int J Biol Sci,* 3**,** 179-84.

Lynch, M., Walsh, B. 1998. Genetics and analysis of Quantitative Traits. *Sinauer Associates Inc. Publishers. MA, USA*.

Mackowski, M., Szymoniak, K., Szydlowski, M., Kamyczek, M., Eckert, R., Switonski, M. 2005. Missense mutations in exon 4 of the porcine LEPR gene encoding extracellular domain and their association with fatness traits. *Animal Genetics,* 36**,** 135-137.

Malek, M., Dekkers, J. C., Lee, H. K., Baas, T. J., Prusa, K., Huff-Lonergan, E., Rothschild, M. F. 2001. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. II. Meat and muscle composition. *Mamm Genome,* 12**,** 637-45.

Milan, D., Jeon, J. T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannuccelli, N., Rask, L., Ronne, H., Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P. L., Chardon, P., Andersson, L. 2000. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science,* 288**,** 1248-51.

Moe, M., Lien, S., Bendixen, C., Hedegaard, J., Hornshoj, H., Berget, I., Meuwissen, T. H., Grindflek, E. 2008. Gene expression profiles in liver of pigs with extreme high and low levels of androstenone. *BMC Vet Res,* 4**,** 29.

Morales, J. 2002. Efecto de la fermentación microbiana en el intestino grueso sobre la digestión, absorción y utilización de nutrientes: Comparación entre el cerdo Landrace y el Ibérico. *Tesis doctoral*.

Morley M, Molony, CM., Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG. 2004. Genetic analysis of genome-wide variation in human gene expression. *nature,* 430**,** 743-747.

Muñoz, G., Alves, E., Fernandez, A., Ovilo, C., Barragan, C., Estelle, J., Quintanilla, R., Folch, J. M., Silio, L., Rodriguez, M. C., Fernandez, A. I. 2007. QTL detection on porcine chromosome 12 for fatty-acid composition and association analyses of the fatty acid synthase, gastric inhibitory polypeptide and acetyl-coenzyme A carboxylase alpha genes. *Anim Genet,* 38**,** 639-46.

Murie, C., Woody, O., Lee, A.Y., Nadon, R. 2009. Comparison of small n statistical tests of diffeential expression applied to microarrays. *BMC Bioinformatics*, 10(45): 41-45.

Myers AJ, Gibbs, JR, Webster JA, Rohrer K, Zhao A, Marlowe L, Kaleem M, Leung D, Bryden L, Nath P, Zismann VL, Joshipura K, Huentelman MJ, Hu-Lince D, Coon KD, Craig DW, Pearson JV, Holmans P, Heward CB, Reiman EM, Stephan D, Hardy J. 2007. A survey of genetic human cortical gene expression. *Nature Genet,* 39**,** 1494-1499.

Nagao K, Yanagita T. 2008. Bioactive lipids in metabolic syndrome. *Prog Lipid Res,* 47**,** 127-146.

Nii, M., Hayashi, T., Tani, F., Niki, A., Mori, N., Fujishima-Kanaya, N., Komatsu, M., Aikawa, K., Awata, T., Mikawa, S. 2006. Quantitative trait loci mapping for fatty acid composition traits in perirenal and back fat using a Japanese wild boar x Large White intercross. *Anim Genet,* 37**,** 342-7.

Ntambi, J. M., Miyazaki, M. 2004. Regulation of stearoyl-CoA desaturases and role in metabolism. *Prog Lipid Res,* 43**,** 91-104.

O'Hea, E. K., Leveille, G. A. 1969. Significance of adipose tissue and liver as sites of fatty acid synthesis in the pig and the efficiency of utilization of various substrates for lipogenesis. *J Nutr,* 99**,** 338-44.

Oczkowicz, M., Tyra, M., Walinowicz, K., Rózycki, M., Rejduch, B. 2009. Known mutation (A3072G) in intron 3 of the IGF2 gene is associated with growth and carcass composition in Polish pig breeds. *J Appl Genet,* 50**,** 257-259.

Oliver, M. A., Gou, P, Gispert, M, Distre, A, Arnau, J, Noguera, J.L, Blasco, A. 1994. Comparison of five types of pig crosses. II. Fresh meat quality and sensory characteristics of dry cured ham. *Livestock Production Science,* 40**,** 179-185.

Olsson, V., Pickova, J. 2005. The influence of production systems on meat quality, with emphasis on porl. *Ambio,* 34**,** 338-343.

Ordovas, J. M. 2006. Genetic interactions with diet influence the risk of cardiovascular disease. *Am J Clin Nutr,* 83**,** 443S-446S.

Ovilo, C., Clop, A., Noguera, J. L., Oliver, M. A., Barragan, C., Rodriguez, C., Silio, L., Toro, M. A., Coll, A., Folch, J. M., Sanchez, A., Babot, D., Varona, L., Perez-

Enciso, M. 2002. Quantitative trait locus mapping for meat quality traits in an Iberian x Landrace F2 pig population. *J Anim Sci,* 80**,** 2801-8.

Paszek, A. A., Wilkie, P. J., Flickinger, G. H., Miller, L. M., Louis, C. F., Rohrer, G. A., Alexander, L. J., Beattie, C.W., Schook, L. B. 2001. Interval mapping of carcass and meat quality traits in a divergent swine cross. *Anim Biotechnol,* 12**,** 155-65.

Perez-Enciso, M., Clop, A., Noguera, J.L., Ovilo, C., Coll, A., Folch, J.M., Babot, D., Estany, J., Oliver, M. A., Diaz, I., Sanchez, A. 2000. A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *J Anim Sci,* 78**,** 2525-31.

Pond, W.G., Mersmann, H.J. 1996. Genetically diverse pig models in nutrition research related to lipoprotein and cholesterol metabolism. *In advances in swine in biomedical research (ed. ME Tumbleson and LB Schook.,* Plenum Press, New York.**,** 843-861.

Ponsuksili, S., Jonas, E., Murani, E., Phatsara, C., Srikanchai, T., Walz, C., Schwerin, M., Schellander, K., Wimmers, K. 2008. Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. *BMC Genomics,* 9**,** 367.

Ponsuksili, S., Murani, E., Phatsara, C., Schwerin, M., Schellander, K., Wimmers, K. 2009. Porcine muscle sensory attributes associate with major changes in gene networks involving CAPZB, ANKRD1, and CTBP2. *Funct Integr Genomics,* 9**,** 455-71.

Ponsuksili, S., Murani, E., Walz, C., Schwerin, M., Wimmers, K. 2007. Pre- and postnatal hepatic gene expression profiles of two pig breeds differing in body composition: insight into pathways of metabolic regulation. *Physiol Genomics,* 29**,** 267-79.

Ponsuksili, S., Murani, E., Schwerin, M., Schellander, K., Wimmers, K. 2010. Identification of expression QTL (eQTL) of genes expressed in porcine M. longissimus dorsi and associated with meat quality traits. *BMC Genomics,* 11**,** 572.

Rohrer, G.A., Thallman, R.M., Shackelford, S., Wheeler, T., Koohmaraie, M. 2006. A genome scan for loci affecting pork quality in a Duroc-Landrace F population. *Anim Genet,* 37**,** 17-27.

Rosthschild, M.F. 2003. Advances in pig genomics and functional gene discovery. *Comparative and Functional Genomics,* 4**,** 266-270.

Rothschild, M.F., Hu, Z. L., Jiang, Z. 2007. Advances in QTL mapping in pigs. *Int J Biol Sci,* 3**,** 192-7.

Sanchez, M.P., Iannuccelli, N., Basso, B., Bidanel, J.P., Billon, Y., Gandemer, G., Gilbert, H., Larzul, C., Legault, C., Riquet, J., Milan, D., Le Roy, P. 2007. Identification of QTL with effects on intramuscular fat content and fatty acid composition in a Duroc x Large White cross. *BMC Genet,* 8**,** 55.

Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich R. 2008. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol,* 6**,** e107.

Schliekelman, P. 2008. Statistical power of expression quantitative trait loci for mapping of complex trait loci in natural populations. *Genetics,* 178**,** 2201-16.

Simopoulos, A.P. 2008. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp Biol Med (Maywood),* 233**,** 674-88.

Sjogren, P., Sierra-Johnson, J., Gertow, K., Rosell, M., Vessby, B., de Faire, U., Hamsten, A., Hellenius, M.L., Fisher, R.M. 2008. Fatty acid desaturases in human adipose tissue: relationships between gene expression, desaturation indexes and insulin resistance. *Diabetologia,* 51**,** 328-35.

Solanes, F. X., Reixach, J., Tor, M., Tibau, J., Estany, J. 2009. Genetic correlations and expected response for intramuscular fat content in Duroc pig line. *Livestock Science,* 123**,** 63-69.

Stachowiak, M., Szydlowski, M., Cieslak, J., Switonski, M. 2007. SNPs in the porcine PPARGC1a gene: interbreed differences and their phenotypic effects. *Cell Mol Biol (Noisy-le-grand),* 12**,** 231-239.

Stanojevic V., Habener J.F., Holz G.G., Leech C.A. 2008. Cytosolic adenylate kinases regulate K-ATP channel activity in human beta-cells. *Biochem Biophys Res Commun*, 368(3):614-9.

Suzuki, K., Ishida, M., Kadowaki, H., Shibata, T., Uchida, H., Nishida, A. 2006. Genetic correlations among fatty acid compositions in different sites of fat tissues, meat production, and meat quality traits in Duroc pigs. *J Anim Sci,* 84**,** 2026-34.

Switonski, M., Stachowiak, M., Cieslak, J., Bartz, M., Grzes, M. 2010. Genetics of fat tissue accumulation in pigs: a comparative approach. *Journal Applied Genetics,* 51**,** 153-168.

Te Pas, M.F., De Wit, A.A., Priem, J., Cagnazzo, M., Davoli, R., Russo, V., Pool, M.H. 2005. Transcriptome expression profiles in prenatal pigs in relation to myogenesis. *Journal of Muscle Research and Cell Motility,* 26**,** 157-165.

Tesson, B.M., Jansen, R.C. 2009. eQTL analysis in mice and rats. *Methods Mol Biol,* 573**,** 285-309.

Tibau, J. 2004. La genética y la producción porcina. *Lecciones Máster "Mejora genética animal y Biotecnología de la reproducción". IRTA*.

Tibau, J., Puigvert, X, Solar, J, Trilla, N, Diestre, A, Gispert, M, Fernández, J, Manteca, X. 1997. Incidéncias de factores genéticos y de comportamiento en la eficiencia del crecimiento, la composición, la calidad de la canal y de la carne en distintas razas porcinas. *Anaporc,* 171**,** 74-91.

Tong, Y., Zhang, S., Li, H., Su, Z., Kong, X., Liu, H., Xiao, C., Sun, Y., Shi, J. J. 2004. 8302A/C and (TTA)n polymorphisms in the HMG-CoA reductase gene may be associated with some plasma lipid metabolic phenotypes in patients with coronary heart disease. *Lipids,* 39**,** 239-41.

Toro, M. A., Silió, L. 1992. La mejora de la selección porcina. *Porci,* 7**,** 69-78.

Tuggle, C.K., Wang, Y., Couture, O. 2007. Advances in swine transcriptomics. *Int J Biol Sci,* 3**,** 132-52.

Van der Steen, H., Prall G, Plastow G. 2005. Application of genomics to the pork industry. *Journal of Animal Science,* 83**,** 1-8.

Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C. S., Buys, N., Tally, M., Andersson, G., Georges, M., Andersson, L. 2003. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature,* 425**,** 832-6.

van Nas, A., Ingram-Drake, L., Sinsheimer, J.S., Wang, S.S., Schadt, E.E., Drake, T., Lusis, A.J. 2010. Expression Quantitative Trait Loci: replication, tissue- and sex-specificity in mice. *Genetics,* 185**,** 1059-1068.

Varona, L. 2008. Genética molecular y calidad de carne. *Sitio argentino de producción animal,* www.produccion-animal.com.ar.

Varona, L., Gomez-Raya, L., Rauw, W.M., Noguera, J.L. 2005. A simulation study on the detection of causal mutations from F2 experiments. *J Anim Breed Genet,* 122**,** 30-6.

Ventanas, S., Ruiz, J, Garcia, C, Ventanas, J. 2007. Preference and juiciness of Iberian dry-cured loin as effected by intramuscular fat content, crossbreeding and rearing system. *Meat Science,* 77**,** 324-330.

Vessby, B. 2003. Dietary fat, fatty acid composition in plasma and the metabolic syndrome. *curr Opin Lipidol,* 14**,** 15-9.

Wakil, S., Abu-Elheiga, L. 2009. Fatty acid metabolism: target for metabolic syndrome. *J Lipid Res*, 50(suppl): S138-143.

Wang, X., Xue, C., Liu, H., Xu, Y., Zhao, R., Jiang, Z., Dodson, M.V., Chen, J. 2009. Differential display of expressed genes reveals a novel function of SFRS18 in regulation of intramuscular fat deposition. *Int J Biol Sci,* 5**,** 28-33.

Wang, Y., Couture, O.P., Qu, L., Uthe, J.J., Bearson, S.M., Kuhar, D., Lunney, J.K., Nettleton, D., Dekkers, J.C., Tuggle, C.K. 2008. Analysis of porcine transcriptional response to Salmonella enterica serovar Choleraesuis suggest novel targets of NFkappaB are activated in the mesenteric lymph node. *BMC Genomics*, 9:437.

Wheeler, H.E., Metter, E.J., Tanaka, T., Absher, D., Higgins, J., Zahn, J.M., Wilhelmy, J., Davis, R.W., Singleton, A., Myers, R.M., Ferrucci, L., Kim, S.K. 2009. Sequential use of transcriptional profiling, expression quantitative trait mapping, and gene association implicates MMP20 in human kidney aging. *PLoS Genet,* 5**,** e1000685.

Wilding, J.P. 2007. The importance of free fatty acids in the development of Type 2 diabetes. *Diabet Med,* 24**,** 934-45.

Wimmers, K., Murani, E., Ponsuksili, S. 2010. Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. *Brief Funct Genomics,* 9**,** 251-8.

Wood, J., Enser, M, Whittington, F, Moncrieff, C, Kempster, A. 1989. Backfat composition in pigs: Differences between fat thickness groups and sexes. *Livestock Production Science,* 22**,** 351-362.

Wood, J. D., Enser, M, Fisher A.V, Nute, G.R, Sheard, P.R, Richardson, R.I, Hughes, S.I, Whittington, F.M. 2008. Fat deposition, fatty acid composition and meat quality: A review. *Meat Science,* 78**,** 343-358.

Xu, Z. Y., Yang, H., Xiong, Y.Z., Deng, C.Y., Li, F.E., Lei, M.G., Zuo, B. 2010. Identification of three novel SNPs and association with carcass traits in porcine TNNI1 and TNNI2. *Mol Biol Rep*.

Zhang, J., He, Q., Liu, Q.Y., Guo, W., Deng, X.M., Zhang, W.W., Hu, X.X., Li, N. 2007. Differential gene expression profile in pig adipose tissue treated with/without clenbuterol. *BMC Genomics,* 8**,** 433.

VII. Anexos

7.1. - SNP Discovery in Bovine Milk Transcriptome using RNA-Seq Technology

Cánovas, A., Rincón, G., Islas-Trejo, A., Wickramasinghe, S., Medrano, J.F

***Mammalian Genome*. 21: 592-598. 2010**

Mamm Genome (2010) 21:592–598
DOI 10.1007/s00335-010-9297-z

SNP discovery in the bovine milk transcriptome using RNA-Seq technology

Angela Cánovas · Gonzalo Rincon · Alma Islas-Trejo · Saumya Wickramasinghe · Juan F. Medrano

Received: 18 August 2010 / Accepted: 14 October 2010 / Published online: 6 November 2010

Abstract High-throughput sequencing of RNA (RNA-Seq) was developed primarily to analyze global gene expression in different tissues. However, it also is an efficient way to discover coding SNPs. The objective of this study was to perform a SNP discovery analysis in the milk transcriptome using RNA-Seq. Seven milk samples from Holstein cows were analyzed by sequencing cDNAs using the Illumina Genome Analyzer system. We detected 19,175 genes expressed in milk samples corresponding to approximately 70% of the total number of genes analyzed. The SNP detection analysis revealed 100,734 SNPs in Holstein samples, and a large number of those corresponded to differences between the Holstein breed and the Hereford bovine genome assembly Btau4.0. The number of polymorphic SNPs within Holstein cows was 33,045. The accuracy of RNA-Seq SNP discovery was tested by comparing SNPs detected in a set of 42 candidate genes expressed in milk that had been resequenced earlier using Sanger sequencing technology. Seventy of 86 SNPs were detected using both RNA-Seq and Sanger sequencing technologies. The KASPar Genotyping System was used to validate unique SNPs found by RNA-Seq but not observed by Sanger technology. Our results confirm that analyzing the transcriptome using RNA-Seq technology is an efficient and cost-effective method to identify SNPs in transcribed regions. This study creates guidelines to maximize the accuracy of SNP discovery and prevention of false-positive SNP detection, and provides more than 33,000 SNPs located in coding regions of genes expressed during lactation that can be used to develop genotyping platforms to perform marker-trait association studies in Holstein cattle.

A. Cánovas
IRTA, Genètica i Millora Animal, 191 Alcalde Rovira Roure Av, 25198 Lleida, Spain

G. Rincon · A. Islas-Trejo · S. Wickramasinghe · J. F. Medrano (✉)
Department of Animal Science, University of California-Davis, One Shields Ave, Davis 95616, CA, USA
e-mail: jfmedrano@ucdavis.edu

Introduction

Next-generation sequencing technologies have provided unprecedented opportunities for high-throughput functional genomic research, including gene expression profiling, genome annotation, small ncRNA discovery, and profiling and detection of aberrant transcription (Bentley 2006; Morozova and Marra 2008). Among these approaches, RNA sequencing (RNA-Seq) is a powerful new method for mapping and quantifying transcriptomes developed to analyze global gene expression in different tissues. Recently, this technique has also been used as an efficient and cost-effective method to systematically identify SNPs in transcribed regions in different species (Chepelev et al. 2009; Cirulli et al. 2010; Cloonan et al. 2008; Morin et al. 2008).

RNA-Seq generates sequences on a very large scale at a fraction of the cost required for traditional Sanger sequencing, allowing the application of sequencing approaches to biological questions that would not have been economically or logistically practical before (Marguerat et al. 2008). Taking this into account, we applied this novel approach to identify SNPs in the expressed coding regions of the bovine milk transcriptome.

The majority of gene expression analyses in the bovine mammary gland have been developed using a biopsy sample (Boutinaud and Jammes 2002; Finucane et al. 2008). An alternative sampling procedure has been proposed by

isolating mRNA directly from somatic cells that are naturally released into milk during lactation (Boutinaud et al. 2002). Recently, Medrano et al. (2010), using the RNA-Seq technique, compared the milk and mammary gland transcriptomes and showed extensive similarities of gene expression in both tissues.

In the present study we performed a SNP discovery analysis in milk transcriptome using RNA-Seq technology. For this purpose, seven milk samples from Holstein cows at different stages of lactation were analyzed by sequencing cDNA libraries using an Illumina GAII analyzer (Illumina, San Diego, CA) system. To evaluate the accuracy of SNPs detected with RNA-Seq, a comparison was made with SNPs detected in a set of 42 candidate genes expressed in milk that had been resequenced previously using Sanger sequencing technology. SNPs that were observed with only one technique were validated by the KASPar SNP Genotyping System.

Materials and methods

RNA-Seq library preparation

Seven milk samples were obtained from Holstein cows at two stages of lactation (day 15 and day 250). Milk samples were collected in 50-ml tubes 3 h after milking, kept on ice, and processed immediately for RNA extraction. Samples were centrifuged at $2000\times g$ for 10 min to obtain a pellet of cells. Total RNA was purified following a Trizol protocol (Invitrogen, Carlsbad, CA), and mRNA was isolated and purified using an RNA-Seq sample preparation kit (Illumina). mRNA was fragmented and first- and second-strand cDNA were synthesized. After adapters were ligated to the ends of double-stranded cDNA, a 300-bp fragment size was selected by gel excision and each sample was individually sequenced on an Illumina GAII analyzer.

RNA-Seq analysis and SNP detection

Short sequence reads (36-40 bp) were assembled and mapped to the annotated bovine reference genome Btau4.0 (http://www.ncbi.nlm.nih.gov/genome/guide/cow/index.html) using CLC Genomics Workbench software (CLC Bio, Aarhus, Denmark). Sequencing reads for each of the seven samples were pooled to perform the RNA-Seq and SNP discovery analyses. We applied stringent criteria in order to reduce the rate of detection of false-positive SNPs. For the assembly procedure, the sequences were mapped to the consensus genome accounting for a maximum of two gaps or mismatches in each sequence. Reads were then classified as uniquely mapped reads and nonspecifically mapped reads as shown in Table 1. SNP detection was performed using the following quality and significance filters: (1) the minimum average quality of surrounding bases and minimum quality of the central base were set as 15 and 20 quality score units, respectively; (2) minimum coverage was set at ten reads; (3) minimum variant frequency or count was set at 20% or two read counts per SNP; and (4) SNPs located in read ends (last three bases) were not considered in the analysis due to possible sequencing errors.

Sanger resequencing of target genes and SNP detection

Resequencing was performed in a DNA resource population specifically developed for SNP discovery as described by Rincon et al. (2007). This population consisted of eight Holstein animals that were unrelated at least three generations back in their pedigrees. Genomic sequences for 42 candidate genes that were expressed in milk samples were obtained from the Btau4.0 assembly and resequenced using Sanger sequencing technology. Exons and conserved noncoding regions were identified using multiple-species genome alignments with Genome VISTA (Couronne et al. 2003). Coding regions and the conserved noncoding regions of each gene were resequenced at SeqWright DNA Technology Services (Houston, TX) using Sanger sequencing technology. SNPs were analyzed using CodonCode aligner software (http://www.codoncode.com); gene sequences and SNPs were assembled and annotated in Vector NTI advance 10.1.1 software (Invitrogen, Carlsbad, CA).

SNP validation by the KASPar SNP genotying system

The KASPar SNP Genotyping System (KBiociences, Herts, UK) was used to validate SNPs detected by RNA-Seq and not detected by Sanger sequencing. For this purpose, 15 bovine DNA samples (8 cows used for Sanger resequencing and 7 cows used for RNA-Seq) were selected. Genomic DNA was extracted from 5 ml of cow's blood following the protocol of the Gentra Puregene blood kit (Qiagen, Valencia, CA). KASPar assay primers (Table 2) were designed using the Primer Picker software available at http://www.kbioscience.co.uk/primer-picker.htm (KBiociences). Genotyping assays were carried out with a 7500 Fast Real Time instrument (Applied Biosystems, Foster City, CA) in a final volume of 8 µl containing 4× Reaction Mix (KBiociences), 120 nM of each allele-specific primers and 300 nM of common primer, 2.2 mM of $MgCl_2$, and 2 mM KTaq polymerase (KBiociences). The following thermal profile was used for all reactions: 15 min at 94°C; 20 cycles of 10 s at 94°C, 5 s at 57°C, and 10 s at 72°C; and 18 cycles of 10 s at 94°C, 20 s at 57°C, and 40 s at 72°C.

Table 1 Summary of mapping all the RNA-Seq reads to the reference genome (Btau4.0) obtained from seven pooled milk samples

	Uniquely mapped reads		Nonspecifically mapped reads		Mapped reads	
	No. of reads	%	No. of reads	%	No. of reads	%
Total exon reads	59888107	83	12480476	17	72368583	87.5
Exon-exon reads[a]	10961868	89	1313183	11	12275051	
Total intron reads	9100877	88	1271041	12	10371918	12.5
Exon-intron reads[b]	1475160	90	166238	10	1641398	
Total gene reads	68988984	83	13751517	17	82740501	100

[a] *Exon-exon reads* reads mapping to two contiguous exons. Number is included in total exon reads

[b] *Exon-intron reads* reads mapping an exon and a contiguous intron. Number is included in total intron reads

Table 2 KASPar primers used to validate SNP detected by RNA-Seq

Primer name	Sequence 5′ → 3′	Position[a]
DDIT3_60417924_A	GAAGGTCGGAGTCAACGGATTGGACTTCAGCCTTTAATATTGGAGAAA	I2
DDIT3_60417924_T	GAAGGTGACCAAGTTCATGCTGGACTTCAGCCTTTAATATTGGAGAAT	
DDIT3_60417924	CCATGGGATTTTCCAGGCAAGAGTA	I2
INSIG2_73132468_C	GAAGGTCGGAGTCAACGGATTAAGCACTCTTATAGTCTGCATGACG	I8
INSIG2_73132468_T	GAAGGTGACCAAGTTCATGCTAAAGCACTCTTATAGTCTGCATGACA	
INSIG2_73132468	ATATCGTATCACAGTGTTGATGTGCCAAA	I8
STAT5A_43749704_G	GAAGGTGACCAAGTTCATGCTCGAGCACCGGGTCAGGGC	I20
STAT5A_43749704_A	GAAGGTCGGAGTCAACGGATTCGAGCACCGGGTCAGGGT	
STAT5A_43749704	GCAGGCCAGCTCCCTCTGATA	E20
STAT5A_43746587_C	GAAGGTCGGAGTCAACGGATTCGCCTGGAAGTTTGACTCTCC	E15
STAT5A_43746587_G	GAAGGTGACCAAGTTCATGCTCGCCTGGAAGTTTGACTCTCG	
STAT5A_43746587	GGAGTGTGGGCAATGCAGGGAA	I16
STAT5A_43741732_T	GAAGGTGACCAAGTTCATGCTCTCCGCCAACTTCTCACACCT	I18
STAT5A_43741732_A	GAAGGTCGGAGTCAACGGATTCTCCGCCAACTTCTCACACCA	
STAT5A_43741732	GGCCCTGGGGCTCGGGTT	E18

[a] Position according to exon/intron distribution in bovine gene

Results and discussion

Detecting genetic variation in pooled milk transcriptome reads by RNA-Seq

RNA-Seq analysis included 118 million reads, ranging from 36 to 40 bp in size, that were assembled and mapped to the annotated NCBI bovine whole-genome assembly (27,368 genes). An average of 17 million short-sequence reads was obtained for each individual sample. The median coverage for the exons was 38×. The analysis revealed that 82.7 million reads (~70%) were categorized as mapped reads (68.9 million were uniquely mapped reads and 13.7 million were nonspecifically mapped reads), while 35 million were unmapped reads (Table 1). Most of the uniquely mapped reads corresponded to total exon reads (87.5%), whereas a small fraction corresponded to total intron reads (12.5%; Table 1). Intron reads are expressed regions that are not annotated as exons in Btau4.0.

RPKM (reads per kilobase per million mapped reads) (Mortazavi et al. 2008) values were used to identify the total number of genes that were expressed in the milk transcriptome. A RPKM threshold value of 0.3 was established in order to balance the number of false positives and false negatives as described in Bentley et al. (2008) and Ramskold et al. (2009). A total of 13,807 genes were selected with RPKM threshold values greater than 0.3. For those genes with RPKM < 0.3, a detailed analysis was performed to determine the number of unique reads falling outside exon regions that can be representing either annotation errors or new exons not included in the current Btau4.0 genome. Using this strategy, 5368 expressed genes/regions were found with more than ten unique reads. We detected 19,175 expressed genes in milk samples

(70.06%) of the 27,368 total bovine annotated genes in Btau4.0 genome assembly. The SNP detection analysis revealed 100,734 SNPs in the seven Holstein samples. Of these SNPs, 67,689 (67.2%) were homozygous, corresponding to differences between Holsteins and the Hereford bovine whole-genome assembly Btau4.0. This is a large number of SNPs that are fixed in Holstein for a different allele to that found in the Hereford genome reference and requires further investigation. In some cases these SNPs may represent artifacts due to errors in the reference sequence or due to misalignment of the short reads to the reference (see subsection "Validation of unique SNP detected by RNA-Seq" below). It may also be possible that some of the Holstein fixed SNPs in fact correspond to variants with a very low frequency and a large number of cows will be needed to detect the common Hereford variant. A total of 33,045 (32.8%) SNPs were polymorphic within Holsteins. Allele frequencies for these heterozygous SNPs were obtained for the pooled samples by counting the number of reads representing each allele. In summary, 1,849 SNPs had an allele frequency of 80/20, 5,511 SNPs had an allele frequency of 70/30, 15,411 SNPs had an allele frequency of 60/40, and 10,274 SNPs had an allele frequency of 50/50. Figure 1 represents the total number of SNPs per gene mapped to the bovine Btau4.0 genome assembly. SNPs that are different between the Hereford consensus sequence and that of Holstein are shown in red, and SNPs that are polymorphic in Holstein samples are in blue. We observed that SNPs in expressed regions are distributed along the entire genome, but there are an increasing number of polymorphisms located in the extremes of the chromosomes' centromeric and telomeric regions. The pattern of SNP distribution in each chromosome is very similar between those that differentiate Holstein and Hereford and those SNPs that are polymorphic in Holsteins, suggesting that there are genomic regions that tend to accumulate a large number of SNPs. Interestingly, the collagen family genes in BTA1, BTA4, BTA12, and BTA19 showed the highest SNP count difference between the Holstein and the Hereford consensus sequences. A large amount of data was generated in this study; a detailed description of the SNPs is available from the authors upon request.

Accuracy of RNA-Seq technology for SNP detection

To analyze the accuracy of RNA-Seq technology for SNP detection, 42 genes highly expressed in milk and related to fatty acid synthesis and the growth hormone GH/IGF axis were resequenced using Sanger methodology. Nine genes did not show polymorphisms in exons by Sanger resequencing and were excluded from the SNP discovery and validation analyses: *IGF1*, *IGFBP3*, *IGFBP4*, *MBTPS1*, *MBTPS2*, *NR3C1*, *PIAS1*, *STAT2*, and *STAT4*. Eighty-six SNPs were detected in the remaining 33 candidate genes that exhibited variation in Holsteins. Seventy of 86 SNPs were also detected by RNA-Seq in 18 genes (Table 3). From the 16 SNPs that were not detected by RNA-Seq, 6 were located in exons that were not expressed in milk samples and therefore no sequencing reads were found.

It is important to note that the samples used for RNA-Seq were different from those sequenced by the Sanger method, so we were not expecting a 100% concordance of the results. However, it is noteworthy that despite the difference in sample composition in the analysis, only ten SNPs observed in Sanger were not detected in RNA-Seq.

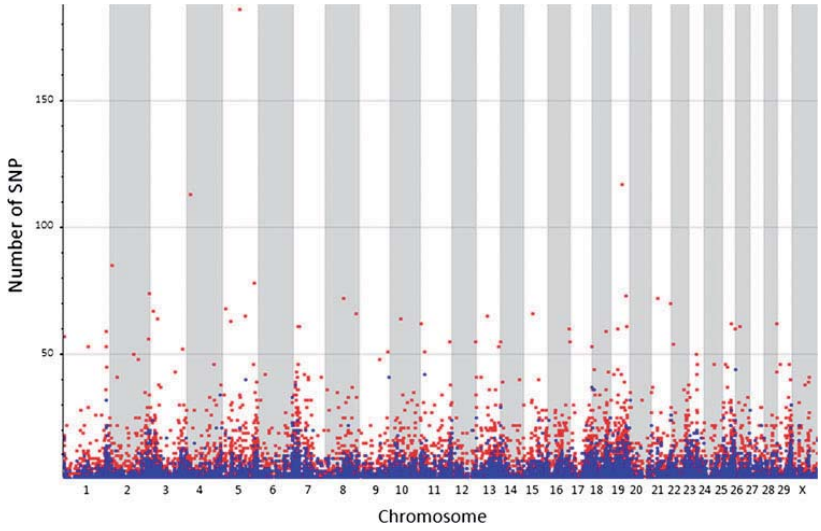


Fig. 1 Total number of SNPs per gene expressed in milk cells mapped to the Btau4.0 genome assembly. *Red dots* represent the number of SNPs per gene in coding regions that are different between Hereford and Holstein. *Blue dots* represent the SNPs per gene that are polymorphic in Holsteins

Table 3 List of 70 SNPs in 18 genes validated in coding regions using RNA-Seq and Sanger sequencing

Gene	BTA	SNP location[a]	Allele variation	Frequency
ADCY4 (Adenylate cyclase 4) ENSBTAG00000018419	10	21091039	A/T	50/50
		21093914	A/T	60/40
		21095171	A/T	60/40
		21096245	C/G	60/40
		21096645	T/A	60/40
CISH (Cytokine-inducible SH2-containing protein) ENSBTAG00000022622	22	50657856	G/T	50/50
		50659783	T/C	60/40
		50659810	C/T	60/40
DDIT3 (DNA damage-inducible transcript 3) ENSBTAG00000031544	5	60415461	C/G	60/40
		60416148	G/T	60/40
		60418030	G/A	60/40
FURIN (Trans Golgi network protease furin) ENSBTAG00000002939	21	21527200	C/G	50/50
		21528000	A/T	60/40
		21528001	T/G	60/40
		21528082	C/T	50/50
		21532215	C/A	60/40
IGF1R (Insulin-like growth factor 1 receptor) ENSBTAG00000021527	21	6862322	T/C	60/40
		6866036	T/C	60/40
		6868728	C/A	60/40
		6869964	A/C	50/50
IGFBP6 (Insulin-like growth factor-binding protein 6) ENSBTAG00000021467	5	29836177	A/T	50/50
INSIG1 (Insulin-induced gene 1) ENSBTAG00000001592	4	121362562	T/A	60/40
		121363743	C/A	60/40
		121363744	T/G	60/40
		121364402	A/C	50/50
		121364403	T/C	70/30
		121364574	T/G	60/40
		121364685	G/A	50/50
		121364911	A/G	60/40
		121365407	G/A	60/40
		121365607	T/C	60/40
		121365773	G/A	60/40
		121365941	G/A	60/40
		121366597	A/G	50/50
		121369136	G/A	60/40
		121369282	A/G	60/40
		121369843	G/A	60/40
		121370020	T/A	60/40
		121370021	T/A	50/50
		121371117	C/G	50/50
INSIG2 (Insulin-induced gene 2) ENSBTAG00000002112	2	73130467	C/T	50/50
NMI (N-myc (and STAT) interactor) ENSBTAG00000016219	2	47179315	C/G	60/40
PAPPA (Pregnancy-associated plasma protein-A) ENSBTAG00000004010	8	110792768	G/A	60/40
		110968532	C/T	60/40

Table 3 continued

Gene	BTA	SNP location[a]	Allele variation	Frequency
SCAP (Sterol regulatory element-binding protein cleavage- activating protein) ENSBTAG00000015782	22	53546833	C/T	50/50
		53549091	C/G	50/50
		53552713	A/T	60/40
SOCS5 (Suppressor of cytokine signaling 5) ENSBTAG00000008987	11	30359073	T/A	60/40
		30360874	T/C	60/40
SREBF1 (Sterol regulatory element binding protein-1) ENSBTAG00000007884	19	35680267	G/A	60/40
		35680574	A/G	60/40
		35682842	A/G	60/40
		35683588	G/C	60/40
		35685082	A/T	60/40
SRPR (Signal recognition particle receptor subunit alpha) ENSBTAG00000014105	29	31200612	A/C	50/50
STAT1 (Signal transducer and activator of transcription 1) ENSBTAG00000007867	2	83382392	A/T	60/40
STAT3 (Signal transducer and activator of transcription 3) ENSBTAG00000021523	19	43780445	A/G	70/30
		43780740	C/A	60/40
STAT5A (Signal transducer and activator of transcription 5A) ENSBTAG00000009496	19	43729581	C/G	50/50
		43730210	G/A	60/40
		43730211	T/A	60/40
		43741509	G/A	60/40
		43743914	C/G	50/50
		43745596	C/G	60/40
		43748702	C/G	70/30
STAT5B (Signal transducer and activator of transcription 5B) ENSBTAG00000010125	19	43655236	A/C	50/50
STAT6 (Signal transducer and activator of transcription 6) ENSBTAG00000006335	5	60837392	A/G	60/40
		60837393	C/T	50/50
		60845709	G/T	60/40
		60845948	G/T	60/40

[a] SNP location is based on the bovine genome assembly Btau4.0

On the other hand, five SNPs were observed in three genes, *DDIT3* (DNA-damage-inducible transcript 3), *INSIG2* (insulin-induced gene 2), and *STAT5A* (signal transducer and activator of transcription 5A), in RNA-Seq that were not detected by Sanger (Table 4). These SNPs were further validated using the KASPar SNP Genotyping System.

Validation of unique SNPs detected by RNA-Seq

In order to confirm the presence of the five SNPs uniquely found by RNA-Seq, they were genotyped using the KASPar SNP Genotyping System. Three out the five SNPs were validated, as shown in Table 4. Two SNPs in the *STAT5A* gene that failed with the KASPar assay were further examined with a detailed analysis of the sequence reads containing the putative SNPs. We observed that the corresponding 40-bp sequence that mapped to a *STAT5A* region had a 99% homology with the *STAT5B* gene. In the SNP discovery analysis we set up a threshold of a maximum number of mismatches to two. With this mismatch

Table 4 Unique SNPs detected by RNA-Seq that were validated using the KASPar Genotyping System

Gene	Chromosome	SNP	Position	Frequency (%)	Confirmed[a]
DDIT3	5	T/A	60417924	50.0/50.0	Yes
INSIG2	2	T/C	73132468	50.0/50.0	Yes
STAT5A	19	G/A	43749704	54.5/45.5	Yes
STAT5A	19	G/C	43746587	70.6/29.4	No
STAT5A	19	T/A	43741732	66.7/33.3	No

DDIT3 DNA damage inducible transcript 3, *INSIG2* insulin-induced gene 2, *STAT5A* signal transducer and activator of transcription 5A

[a] SNP confirmed by KASPar SNP Genotyping System and Sanger resequencing

rate, reads that correspond to a given gene, like *STAT5B*, can be assigned to *STAT5A*. This was not a common situation for most of the genes studied in this analysis, but it could represent a problem in gene families with highly conserved domains when using short sequence reads. In a similar study, Cirulli et al. (2010) observed that some false-positive SNPs identified in cDNA arose from alignment of a read to the wrong gene and that in these cases the correct gene and the gene chosen for the alignment always had very similar sequences. This situation has also been observed in regions associated with sequence repeats (Morozova and Marra 2008). Although the short-read structure of next-generation sequencers has some potential problems with respect to sequence assembly, the result is a system that generates accurate data and large coverage of consensus sequence and SNP calling at very high throughput and low cost (Thomas et al. 2006).

Conclusion

We have demonstrated that analyzing the transcriptome using RNA-Seq technology is an efficient and cost-effective method to identify SNPs in transcribed regions. Stringent criteria have to be applied to maximize the accuracy and prevent false-positive SNP detection. This study provides a valuable resource of more than 33,000 SNPs located in coding regions of genes expressed during lactation that can be used for further gene variation analysis and association studies in Holstein cattle.

Acknowledgments This project was supported by a grant from Dairy Management Inc. and the California Dairy Research Foundation. We thank Charlie Nicolet of the UC Davis Genome Center for his excellent technical expertise to perform Illumina GAII sequencing. A. Cánovas received a predoctoral scholarship from INIA, Spain.

References

Bentley DR (2006) Whole-genome re-sequencing. Curr Opin Genet Dev 16:545–552

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59

Boutinaud M, Jammes H (2002) Potential uses of milk epithelial cells: a review. Reprod Nutr Dev 42:133–147

Boutinaud M, Rulquin H, Keisler DH, Djiane J, Jammes H (2002) Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. J Anim Sci 80:1258–1269

Chepelev I, Wei G, Tang Q, Zhao K (2009) Detection of single nucleotide variations in expressed exons of the human genome using RNA-Seq. Nucl Acids Res 37:e106

Cirulli ET, Singh A, Shianna KV, Ge D, Smith JP et al (2010) Screening the human exome: a comparison of whole genome and whole transcriptome sequencing. Genome Biol 11:R57

Cloonan N, Forrest A, Kolle G, Gardiner B, Faulkner G et al (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods 5:613–619

Couronne O, Poliakov A, Bray N, Ishkhanov T, Ryaboy D et al (2003) Strategies and tools for whole-genome alignments. Genome Res 13:73–80

Finucane KA, McFadden TB, Bond JP, Kennelly JJ, Zhao FQ (2008) Onset of lactation in the bovine mammary gland: gene expression profiling indicates a strong inhibition of gene expression in cell proliferation. Funct Integr Genomics 8:251–264

Marguerat S, Wilhelm BT, Bahler J (2008) Next-generation sequencing: applications beyond genomes. Biochem Soc Trans 36:1091–1096

Medrano JF, Rincon G, Islas-Trejo A (2010) Comparative analysis of bovine milk and mammary gland transcriptome using RNA-Seq. In: 9th World congress on genetics applied to livestock production, Leipzig, Germany, August 1-6, 2010, Paper no 0852

Morin R, O'Connor M, Griffith M, Kuchenbauer F, Delaney A et al (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 18:610–621

Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92:255–264

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628

Ramskold D, Wang ET, Burge CB, Sandberg R (2009) An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. PLoS Comput Biol 5:e1000598

Rincon G, Thomas M, Medrano JF (2007) SNP identification in genes involved in GH-IGF1 signaling on BTA5. In: Plant & animal genomes XV conference, San Diego, CA, January 13-17, 2007, Abstract no P536

Thomas RK, Nickerson E, Simons JF, Janne PA, Tengs T et al (2006) Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. Nat Med 12:852–855

7.2. - Technical Note: Efficient protocol for isolation of good-quality total RNA from lyophilised fat and muscle pig samples

Pena, R.N., **Cánovas, A.,** Estany, J

***Journal of Animal Science*. 88:442-445. 2010**

Technical note: Efficient protocol for isolation of total ribonucleic acid from lyophilized fat and muscle pig samples[1]

R. N. Pena,*[2] A. Cánovas,* and J. Estany†

*Institut de Recerca i Tecnologia Agroalimentàries, Genètica i Millora Animal, 191 Av Rovira Roure, 25198 Lleida, Spain; and †Departament de Producció Animal, Universitat de Lleida, 191 Av Rovira Roure, 25198 Lleida, Spain

ABSTRACT: Isolation of total RNA from frozen muscle and fat samples typically results in small yields due to the presence of connective tissue between muscle fibers, which impairs complete tissue homogenization, and the excess of fat and relatively small cellularity of adipose tissue. Meat quality studies involve determination of fatty acid composition and content from muscle and subcutaneous fat samples, a process that may produce an excess of lyophilized tissue samples. The purpose of this work was to investigate the stability of total RNA in lyophilized tissue samples generated during the routine detection of fatty acid content of pig muscle and fat tissues, stored at room temperature or at −20°C. The protocol described here results in increased yields of total RNA from freeze-dried samples stored at −20°C, which facilitates the homogenization step. The isolated RNA is suitable for common gene expression techniques such as final point and quantitative reverse transcription-PCR.

Key words: freeze-dry, gene expression, lyophilization, meat quality, pig

J. Anim. Sci. 2010. 88:442–445
doi:10.2527/jas.2009-2298

INTRODUCTION

Fat composition and content are important traits influencing sensory and nutritional meat quality properties. Thus, there is a growing interest in the study of the molecular mechanisms underlying subcutaneous and intramuscular fat deposition events. Many of these studies focus on global or candidate gene expression characterization in muscle and fat tissues. Isolation of RNA from muscle and fat samples is largely inefficient because connective tissue within the muscle samples is difficult to homogenize, even with mechanical rotors, and the large fat content of adipose tissue interferes with the isolation process, whereas its relatively small cellularity results in decreased RNA concentrations.

Meat quality studies often rely also on muscle and fat chemical composition analysis, some of which involve previous sample lyophilization. Freeze-drying or lyophilization is a widely used technique for sample preservation commonly applied to store vaccines, microorganisms, and plant (Jaiprakash et al., 2003) and virus (Vaughan et al., 2006) samples. Although RNA is an unstable molecule, easily degraded by enzymes and at alkaline environments, viral RNA has been detected from a wide range of freeze-dried biological samples. The RNA is rapidly degraded in shelf-preserved freeze-dried biopsies, but it can still be used for some reverse transcription-PCR (**RT-PCR**) analyses (Matsuo et al., 1999). Despite these reports, the use of lyophilized animal samples is nearly anecdotic in molecular assays and particularly restricted to protein detection by Western blotting (Salomon et al., 1994) and tissue immunostaining (Louis et al., 2000).

The protocol presented here results in large yields of isolated total RNA from freeze-dried samples generated during the ordinary detection of fatty acid content of pig muscle and fat tissues. Additionally, the stability of total RNA in lyophilized tissue samples, stored at room temperature or at −20°C, and the suitability of this RNA as a start-up material for commonly used molecular techniques have been investigated.

MATERIALS AND METHODS

All the experimental procedures were approved by the Ethics Committee for Animal Experimentation of the University of Lleida.

[1]We thank Teresa Giró (Departament de Producció Animal, Universitat de Lleida) for her technical help. A. Cánovas received a PhD scholarship from Instituto Nacional de Investigaciones Agrarias (Spain). This work was partially funded by the Spanish Ministry of Education and Science (CICYT grant AGL2006-01243).

[2]Corresponding author: romi.pena@irta.cat

Received July 13, 2009.
Accepted October 2, 2009.

Sample Collection and Processing

Samples of semimembranosus muscle and subcutaneous fat from a Duroc commercial line were collected at slaughter, snap frozen in liquid nitrogen and stored at −80°C until analyzed. Once defrosted, muscle and adipose tissue samples were freeze-dried at −20°C, 0.01 hectopascals for 72 h using a HETO FD8 system (Heto-Holten A/S, Allerod, Denmark) and thoroughly homogenized with sand using a glass stirring rod. Lyophilized samples were stored at room temperature for 5 d or at −20°C before being used for RNA isolation.

RNA Isolation

Isolation of RNA was performed in parallel from frozen and freeze-dried samples. Muscle samples (60 mg if frozen or 30 mg if lyophilized) were homogenized in 1 mL of TRI Reagent (Sigma-Aldrich, Madrid, Spain) using a mechanical rotor (IKA Werke, Staufen, Germany) following the manufacturer's instructions. For adipose samples, the protocol was modified as follows: fat tissue (0.3 g) was homogenized in 3 mL of TRI Reagent and centrifuged at 3,000 × *g*, 4°C, 20 min. The lower (aquose) layer was vigorously mixed with 1 mL of chloroform for 30 s, incubated at room temperature for 10 min, and spun at 3,000 × *g*, 4°C, 20 min. From the aqueous layer, RNA was precipitated with ice-cold isopropanol and resuspended in 100 μL of diethyl-pyrocarbonate-treated water and quantified in a Nanodrop-1000 spectrophotometer (Nanodrop, Wilmington, DE). Integrity of the isolated RNA was checked by electrophoresis in a 1.2% formaldehyde-agarose gel using standard methods (Sambrook and Russell, 2001) and in a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) where RNA quality was evaluated by the RNA integrity number calculated with the Agilent 2100 expert software (Schroeder et al., 2006).

First-Strand cDNA Synthesis

Total RNA (1 μg) was treated with Turbo DNA-free DNase (Ambion, Austin, TX) according to the manufacturer's protocol and retrotranscribed with 0.5 pmol of random hexamers using 100 U of MuMLV reverse transcriptase (Fermentas, St. Leon-Rot, Germany) at 37°C for 1 h.

Final Point RT-PCR

The PCR reactions for the porcine low-density lipoprotein receptor (**LDLR**) mRNA were carried out in a PTC-100 thermocycler (MJ Research, Waltham, MA) as in Pena et al. (2009). Primer sequences are described in Supplemental Table 1 (http://jas.fass.org/content/vol88/issue2/).

Real-Time Quantitative RT-PCR

Single-strand cDNA was diluted 1:10 in diethyl-pyrocarbonate-treated water before real-time quantitative RT-PCR (**qRT-PCR**) analysis of 3 porcine genes [stearoyl-CoA desaturase (***SCD***), acetyl-CoA carboxylase α (***ACACA***) and hypoxanthine phosphoribosyltransferase 1 (***HPRT1***)]. Primers and TaqMan probes are described in Supplemental Table 1 (http://jas.fass.org/content/vol88/issue2/). For the *SCD* and *ACACA* genes, real-time qRT-PCR assays were carried out in triplicate in an ABI-7500 device (Applied Biosystems, Foster City, CA) in a final volume of 5 μL containing 1× SYBRgreen Master mix (Applied Biosystems) and 200 n*M* of each primer. For the endogenous control (*HPRT1*), reactions contained 300 n*M* primers, 200 n*M* TaqMan probe, and 1× Universal Taqman Master Mix (Applied Biosystems). The following thermal profile was used for all reactions: 10 min at 95°C, 40 cycles of 15 s at 93°C, and 1 min at 60°C. Mean expression measures from the endogenous control were used as an internal reference to normalize and quantify the expression data from the other 2 genes as in Yuan et al. (2006).

RESULTS AND DISCUSSION

In the present work, the use of freeze-dried pig muscle and adipose samples has been investigated as a more efficient source of total RNA for gene expression studies. For this, total RNA was isolated from tissues frozen at −80°C and freeze-dried tissues that were then stored at −20°C (**Lyoph.+Frozen**) or kept at room temperature for 5 d (**Lyoph.+RT**). The isolation protocol was slightly modified for the adipose samples to eliminate the excess of fat previous to RNA separation by the phenol acid method. Lyophilization of pig muscle and fat samples resulted in double the yield of isolated RNA per milligram of tissue without the need to scale up the TRI Reagent volume (Supplemental Table 2; http://jas.fass.org/content/vol88/issue2/). Although RT-PCR and many qRT-PCR studies do not require large amounts of RNA as a starting material, other gene expression applications such as microarrays require greater amount of total RNA at a concentration of >500 μg/μL. Isolating high yields from small samples would also be advantageous for the analysis of muscle and fat biopsies from time-course experiments (Bosch et al., 2009). Yields did not differ between lyophilized samples stored at −20°C or at room temperature for 5 d. However, RNA integrity was lost in the latter (Figure 1 and Supplemental Figure 1; http://jas.fass.org/content/vol88/issue2/). Electrophoresis of total RNA gave comparable integrity results, as measured by the RNA integrity number, for frozen and Lyoph.+Frozen samples (Figure 1 and Supplemental Figure 1; http://jas.fass.org/content/vol88/issue2/), indicating that RNA remained stable during the freeze-drying process.

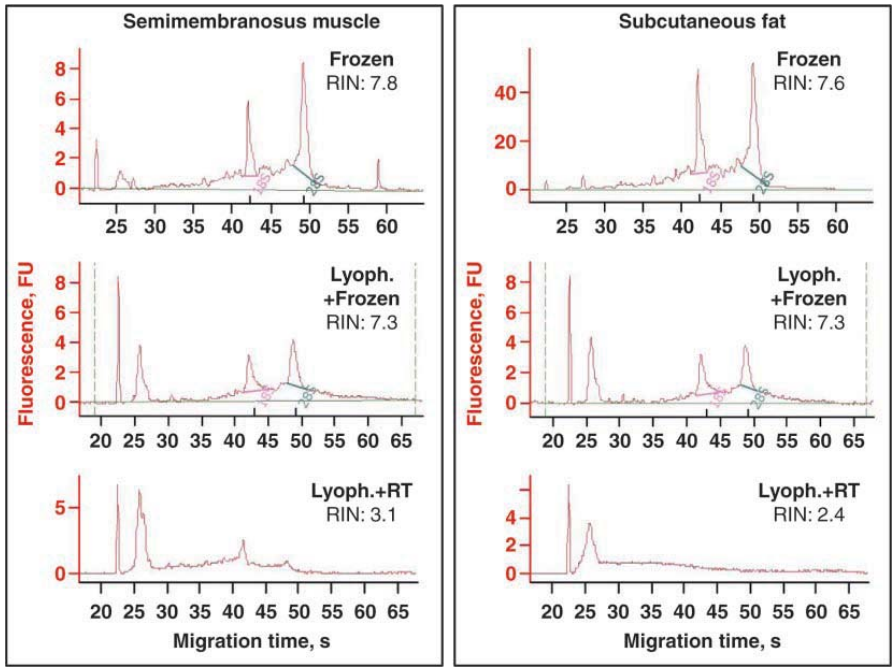


Figure 1. Analysis on a 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA) of RNA integrity from muscle and adipose samples that were frozen or lyophilized and then frozen at −20°C (Lyoph.+Frozen) or kept at room temperature for 5 d (Lyoph.+RT). The y-axis of the electropherograms reports the fluorescence in fluorescence units (FU), whereas the x-axis reports the migration time in seconds. For each sample, the RNA integrity number (RIN) is indicated.

The suitability of RNA isolated from freeze-dried samples as a starting material for common gene expression techniques was tested by final point RT-PCR and by real-time qRT-PCR. Two fragments of 1,150 and 600 bp of the coding region of the pig *LDLR* gene were successfully amplified from frozen and Lyoph.+Frozen samples in both tissues (Figure 2A). In contrast, only the smaller fragment could be amplified at a reduced efficiency from the Lyoph.+RT degraded RNA. Some reports indicate that RNA integrity does not affect amplification efficiency for many PCR-based applications on the range of 70 to 250 bp (Fleige and Pfaffl, 2006). Thus, qRT-PCR assays, which are usually based on very small amplicons (range of 75 to 150 bp), should not be influenced by degradation of starting RNA. Two genes related to the metabolism of fatty acids (*SCD* and *ACACA*) and 1 gene commonly used as a reference gene (*HPRT1*) were analyzed by qRT-PCR. Identical amplification curves were obtained for the 3 genes in frozen and Lyoph.+Frozen samples (Supplemental Figure 2; http://jas.fass.org/content/vol88/issue2/). Thus, relative quantification of *SCD* and *ACACA* expression gave similar results in samples treated either way (Figure 2B). In contrast, a delayed amplification curve was obtained with fat Lyoph.+RT samples, indicating a less efficient amplification (Supplemental Figure 2B; http://jas.fass.org/content/vol88/issue2/). Consequently, quantification values differed from the 2 other sample types (Figure 2B), indicating that RNA obtained after storing of freeze-dried samples at room temperature might not always be suitable for quantitative methods such as qRT-PCR. It is important to remark that, in both tissues, amplification plots and quantification measures were undistinguishable between frozen and Lyoph.+Frozen samples, validating the use of freeze-dried samples for gene expression analysis.

In conclusion, the method presented here represents an alternative to isolate large yields of total RNA from freeze-dried muscle and fat samples such as those obtained after processing meat and subcutaneous fat for fatty acid content determination. Freeze-drying facilitates RNA isolation from tissues difficult to homogenize (Tsuka et al., 1997), whereas integrity of RNA, end point, and quantitative RT-PCR results are not impaired. Although results for a limited number of samples are presented here, we routinely use lyophilized samples in gene expression assays with consistent results. In the practicalities of meat quality projects, this new protocol eliminates the need to collect additional tubes for DNA/RNA/protein studies, unburdening the tissue-collecting protocol at the abattoir.

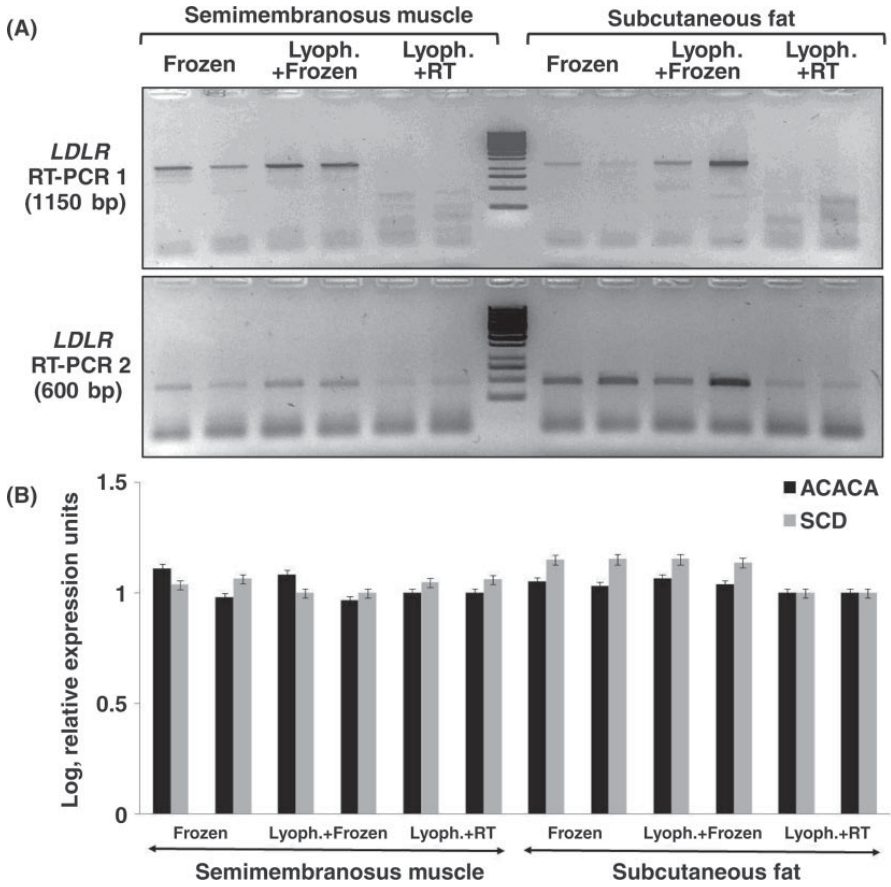


Figure 2. Isolated RNA was used in 2 common gene expression assays. A) Two fragments of 1,150 and 600 bp from the coding region of pig low-density lipoprotein receptor (*LDLR*) gene were successfully amplified by reverse transcription-PCR (RT-PCR) from frozen samples and samples lyophilized and then frozen at −20°C (Lyoph.+Frozen), whereas samples lyophilized and then stored at room temperature (Lyoph.+RT) amplified only the smaller fragment; B) quantitative RT-PCR analysis of 2 genes [stearoyl-CoA desaturase (*SCD*) and acetyl-CoA carboxylase α (*ACACA*)] gave identical results in frozen and Lyoph.+Frozen samples, whereas Lyoph.+RT estimations differed in adipose fat. Bars represent mean of 3 replicates. Error bars indicate SE.

LITERATURE CITED

Bosch, L., M. Tor, J. Reixach, and J. Estany. 2009. Estimating intramuscular fat content and fatty acid composition in live and post-mortem samples in pigs. Meat Sci. 82:432–437.

Fleige, S., and M. W. Pfaffl. 2006. RNA integrity and the effect on the real-time qRT-PCR performance. Mol. Aspects Med. 27:126–139.

Jaiprakash, M. R., P. Beena, V. Purna, N. Subramanian, V. P. Sinkar, and P. P. Sadhale. 2003. RNA isolation from high-phenolic freeze-dried tea (*Camellia sinensis*) leaves. Plant Mol. Biol. Rep. 21:465a–465g.

Louis, H., J. Lavie, P. Lacolley, D. Daret, J. Bonnet, and J. M. Lamaziere. 2000. Freeze-drying allows double nonradioactive ISH and antigenic labeling. J. Histochem. Cytochem. 48:499–508.

Matsuo, S., T. Sugiyama, T. Okuyama, K. Yoshikawa, K. Honda, R. Takahashi, and S. Maeda. 1999. Preservation of pathological tissue specimens by freeze-drying for immunohistochemical staining and various molecular biological analyses. Pathol. Int. 49:383–390.

Pena, R., A. Cánovas, L. Varona, I. Díaz, D. Gallardo, O. Ramírez, J. Noguera, and R. Quintanilla. 2009. Nucleotide sequence and association analysis of pig apolipoprotein-B and LDL-receptor genes. Anim. Biotechnol. 20:110–123.

Salomon, D., E. Masgrau, S. Vischer, S. Ullrich, E. Dupont, P. Sappino, J. H. Saurat, and P. Meda. 1994. Topography of mammalian connexins in human skin. J. Invest. Dermatol. 103:240–247.

Sambrook, J., and D. Russell. 2001. Molecular Cloning: A Laboratory Manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schroeder, A., O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow, and T. Ragg. 2006. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. BMC Mol. Biol. 7:3.

Tsuka, H., H. Mori, K. Okada, and S. Matsukawa. 1997. Utilization of the freeze-drying method in the preparation of biologically active, intact RNA from hard frozen tissues of human prostate. Anal. Biochem. 247:458–461.

Vaughan, H., V. J. Chalker, Z. Mee, A. Rossouw, and V. James. 2006. Stability of lyophilised specimens for the molecular detection of viral DNA/RNA. J. Clin. Virol. 35:135–140.

Yuan, J. S., A. Reed, F. Chen, and C. N. Stewart Jr. 2006. Statistical analysis of real-time PCR data. BMC Bioinformatics 7:85.

7.3. - Nucleotide sequence and association analysis of pig Apolipoprotein-B and LDL-receptor genes

Pena, R.N., **Cánovas, A.**, Varona, L., Díaz, I., Gallardo, D., Ramírez, O., Noguera, J.L., Quintanilla, R

***Animal Biotechnology*. 20(3):110-23. 2009**

Animal Biotechnology, 20: 110–123, 2009

ISSN: 1049-5398 print/1532-2378 online
DOI: 10.1080/10495390902892518

Taylor & Francis
Taylor & Francis Group

NUCLEOTIDE SEQUENCE AND ASSOCIATION ANALYSIS OF PIG APOLIPOPROTEIN-B AND LDL-RECEPTOR GENES

R. N. Pena[1], A. Cánovas[1], L. Varona[1,*], I. Díaz[2], D. Gallardo[3], O. Ramírez[3], J. L. Noguera[1], and R. Quintanilla[1]

[1]*IRTA, Genètica i Millora Animal, Lleida, Spain*

[2]*IRTA, Tecnologia dels Aliments, Finca Camps i Armet, Monells, Spain*

[3]*Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra, Spain*

Three genes are the major determinants of heritable hypercholesterolemia diseases in humans: APOB, LDLR and LDLRAP1, which encode for proteins that physically interact to promote cholesterol uptake in the cell. We have carried out association analyses of these variants with serum cholesterol and triglycerides concentrations in a half-sib Duroc pig population. Given the structure of the population (six paternal half-sib families), we have used a statistical model that considers separately the allele transmission through dams (at population level) and through sires (within-families from heterozygous sire). Only polymorphisms showing a relevant substitution effect for both male- and female-transmitted alleles are likely to be causal mutations. Thus, although we have found statistical association between genotypes for LDLR and APOB polymorphisms and serum lipid levels (mean allele substitution effects ranging from 15 to 40% of the standard deviation of these traits), none of them seem to be the causal mutation but probably represent closely linked polymorphisms. We have shown here that these three genes also contribute to genetic variability in pigs, with the description of new polymorphisms in their coding regions. Moreover, we have demonstrated that variants on two of these three genes are segregating in a number of commercial breeds. Finally, we report here the coding region for the porcine LDLRAP1 gene and describe a polymorphism in the last exon of this gene.

Keywords: APOB; Cholesterol; LDLR; Pig; Triglycerides

Multiple factors are known to affect serum cholesterol levels including body mass index, age, gender, menopausal status, and baseline LDL cholesterol levels. In concert with these, genotypic differences are also known to play a very relevant role (reviewed in [1]). Hence, familial hypercholesterolemia is a syndrome of genetic

*Current Affiliation: Departamento de Anatomía, Embriología y Genética. Universidad de Zaragoza, Zaragoza, Spain.

We are thankful to *Selección Batallé S.A.* for providing the animal material and for their cooperation in the experimental protocol and to D Almuzara for technical support. RN Pena received a partial contractual grant from INIA (Spain). A Canovas holds a predoctoral scholarship from INIA (Spain). The work was funded by grants AGL2002-04271-C03-02 and AGL2007-66707-C02-01 (MEC, Spain).

Address correspondence to R. N. Pena, IRTA, Genètica i Millora Animal, 191 Av Alcalde Rovira Roure, E-25198 Lleida, Spain. E-mail: romi.pena@irta.cat

origin characterized by increased total and LDL-bound cholesterol serum levels, and it is linked to coronary heart disease in humans, mice, and pigs (2, 3). The mutation of the LDL-receptor (*LDLR*) gene is the most frequent genetic cause of this disease in humans (4), although polymorphisms in the Apolipoprotein-B (*APOB*) gene and in the LDLR adaptor protein (*LDLRAP1*) gene have also been reported (1). APOB is the most abundant surface protein in LDL particles. In fact, LDL uptaking by endocytosis is triggered by APOB binding onto LDLR. The function of LDLRAP1 is not completely established, although a role in stabilizing the activated LDL-LDLR complex during clustering and internalization has been suggested (5). Thus, proteins encoded by these three genes physically interact during the LDLR-mediated endocytosis of LDL particles in the cell, which is the main step for cholesterol uptake in cells.

The existence of a genetic component affecting variability of pig serum cholesterol concentrations has been proved in several divergent selection experiments (6, 7), but reports on genetic variation for porcine cholesterol-related genes are scarce. In this regard, polymorphisms associated with variation on serum cholesterol levels in pigs include a missense mutation at residue 94 of the porcine *LDLR* gene (8, 3) and the *APOB* variant-5 (*APOB-5*), which is associated with hypercholesterolemia (9) and atherosclerotic lesions (10). The main limitation of these association studies is the inability to discern causative mutations from other (closely linked) polymorphisms (11). In contrast with the two previous genes, no studies for *LDLRAP1* have been reported in pigs.

In order to better define the contribution of these three genes to the pig cholesterol metabolism, we have analyzed the role that variability in *LDLR, APOB* and *LDLRAP1* genes play in the serum levels for total (TC), HDL- and LDL-bound cholesterol and triglycerides (TG) in a half-sib population of Duroc pigs. In the case of causal mutations, or polymorphisms very closely linked to causal mutation at population level, similar associations with phenotypic data are expected for alleles transmitted both through sire and dam. Thus, to investigate further the nature of associations found with phenotypic data, we considered separately whole-population (*via* female) and within paternal family (*via* male) allele substitution effects.

MATERIALS AND METHODS

Animal Material

Animals came from a high intramuscular fat commercial Duroc line used in the production of fine quality cured ham. The experiment was based on a half-sib design generated by mating six parental boars with 385 females and taking only one male offspring per litter to a total of 385 animals (12). These animals were castrated and controlled in four batches at IRTA-CCP experimental station under normal intensive conditions being all subjected to the same rearing and feeding regime from weaning to the end of the fattening period. During fattening, pigs were fed *ad libitum* on two standard pig diets, one up to 30–40 days before slaughter (around 90 kg of live weight) and the second one up to slaughter (mean 122 kg of live weight, 190 d of age), containing 2,450 and 2,375 $kcal \cdot kg^{-1}$ of energy concentration, respectively. The experimental procedures, traits recording, and blood sampling were approved by the Ethical Committee of the Institution (IRTA - *Institut de Recerca i Tecnologia Agroalimentàries*).

In addition, a panel of 186 animals from Pietrain (n = 44 pigs), Landrace (n = 50), Large White (n = 20), Meishan (n = 17), and Iberian (n = 7) breeds and from an additional Duroc (n = 48) population was used to study allele frequency distribution across breeds.

Phenotypic Data

Blood samples were taken at 45 and 190 d of age (the day before slaughter) into EDTA-containing tubes to measure total serum cholesterol, LDL, HDL, and triglyceride concentrations.

We used a cholesterol oxidase based method, where cholesterol esterase (CHE), cholesterol oxidase (CO) and peroxidase enzymes (POD) are employed, to determine TC levels (13). Serum HDL levels were measured using a direct immuno-inhibition method after precipitation with anti human β-lipoprotein antibody and with the previously reported enzymatic reactions mediated by CHE, CO, and POD (14). Triglyceride concentrations were estimated by using the glycerol kinase reaction with a method reported by Fossati et al. (15). Finally, serum LDL concentration was calculated according to the equation of Friedewald et al. (16).

After a data filtering process, where some measures from extremely young animals, strongly hemolyzed blood samples, or those considered to be the result of inaccurate measurements were discarded from the dataset, analyzed phenotypes belonged to 355 individuals of the experimental Duroc population. Table 1 shows the structure and size of data analyzed in this study. Before carrying out the association analyses, several exploratory analyses were done in order to investigate the

Table 1 Structure of data on the Duroc pig population analyzed. Number of available records (N) for serum lipid concentrations[a] at two ages and mean (standard deviation) of the analyzed measures (in $mg \cdot dl^{-1}$), by half-sib family

	TOTAL POPULATION	Families					
		♂1	♂2	♂3	♂4	♂5	♂6
Post weaning (45.7 ± 0.41 d)	N = 355	N = 60	N = 79	N = 82	N = 72	N = 40	N = 22
TC[a]	76.85 (13.49)	76.82 (13.92)	75.16 (11.01)	77.55 (13.86)	78.30 (14.25)	79.90 (15.48)	70.14 (12.28)
HDL[b]	30.44 (6.83)	30.65 (7.53)	30.29 (6.41)	29.61 (6.02)	30.62 (7.40)	31.69 (7.79)	30.69 (6.12)
LDL[c]	37.64 (8.78)	38.66 (9.69)	35.81 (7.80)	37.79 (8.54)	39.75 (9.01)	38.87 (7.71)	31.65 (8.30)
TG[d]	43.68 (19.13)	37.28 (11.45)	45.23 (22.69)	50.61 (24.39)	39.62 (13.68)	46.65 (14.19)	37.64 (9.50)
At slaughter (190.4 ± 0.45 d)	N = 333	N = 56	N = 71	N = 81	N = 69	N = 38	N = 18
TC[a]	124.15 (25.75)	133.62 (27.33)	115.82 (20.02)	129.37 (24.90)	122.25 (30.08)	125.05 (23.38)	109.39 (13.99)
HDL[b]	51.53 (9.94)	52.94 (7.56)	49.21 (8.39)	53.67 (11.85)	51.51 (9.23)	50.39 (12.60)	49.17 (7.00)
LDL[c]	62.30 (19.99)	71.78 (23.14)	56.35 (15.81)	63.60 (16.31)	61.58 (25.27)	63.55 (16.08)	50.54 (8.33)
TG[d]	51.18 (23.13)	44.93 (19.63)	51.46 (21.59)	59.32 (25.22)	44.85 (20.32)	54.95 (27.31)	49.22 (19.18)

[a]TC = total cholesterol.
[b]HDL = high density lipoproteins.
[c]LDL = low density lipoproteins.
[d]TG = triglycerides.

distribution of raw data and the model of fit. Data in raw form were quite skewed to the right. The Box-Whisker plots highlighted the presence of outlier phenotypes for TG levels at both ages and for LDL at 190 d, but the log-transformation of data generally corrected these departures from normality.

The influence of non-genetic environmental effects (farm of origin, batch, and box of fattening) and covariates (age at blood extraction, live weight, or backfat thickness) over these traits was tested by means of the GLM procedure of SAS (SAS Inst. Inc., Cary, NC). Normal probability plots and Kolmogorov-Smirnov tests were performed to investigate the goodness of fit of the residuals with the normal distribution. In order to avoid departures from Gaussian distribution, log-transformed data were employed to perform statistical analyses.

Isolation of DNA and RNA

Liver, pituitary gland, and fat and muscle at two anatomic points (*gluteus medius* and *longissimus dorsi*) were collected at slaughter, snap frozen at −198°C and stored at −80°C for subsequent RNA analyses. To isolate total RNA, liver samples were ground with mortar and pestle in liquid nitrogen and homogenized with a mechanical rotor. RNA was isolated by the acid phenol method (17) using the RiboPure kit (Ambion Inc., Austin, TX). DNA was isolated from blood samples using standard procedures (18).

Gene Characterization and Polymorphism Discovery

Primer design was carried out using the *Primer Express* software (Applied Biosystems, Foster City, CA) using the pig *LDLR* sequence (AF065990) as a template. Two regions comprising most of the coding region were amplified: the first fragment included the coding region for the first two extracytoplasmatic domains of this receptor (LDL-binding and EGF-like) which are highly polymorphic in humans; the second fragment included the coding regions for the third extracytoplasmatic domain, the intramembranal, and part of the intracytoplasmatic domains. As no sequence was available from the public genomic data bases for the porcine *LDLRAP1* gene, we aligned cDNA sequences from human, cow, chicken, and mouse (NM_015627, BT.880, XM_417736 and NM_145554, respectively) in order to identify conserved regions to design primers that would amplify the complete coding region of the porcine mRNA (Table 2).

Total RNA (0.5–1 μg) was retrotranscribed to cDNA using an Oligo(dT) primer and MuMLV RT enzyme (Fermentas, Glen Burnie, MD) at 37°C for 1 h. PCR reactions were carried out in a PTC-100 PCR device (MJ Research, Waltham, MA) in a volume of 25 μl with 1x buffer, 200 mM dNTP mix, 2.0 mM $MgCl_2$, 400 nM of each primer, 1U of Taq polymerase (ECOGEN SL, Barcelona, Spain), and 0.5 μl of cDNA. Annealing temperature for each primer set is indicated in Table 2.

PCR products were directly sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit v3.1 (Applied Biosystems) in an ABI-3100 capillary electrophoresis system. Sequences obtained were edited using the *Sequencing* software (Applied Biosystems) and aligned with the *ClustalW* program (19). Sequencing

Table 2 Primers used for the amplification of *LDLR* and *LDLRAP1* cDNA

Primer	Primer Sequence 5′ → 3′	Localization*	T°ann
LDLR_F1	5′-TGTGATGGGAACACCGAGTG-3′	E3	62
LDLR_R1	5′-CAACCTGAAGAACGTGGTCG-3′	E11	
LDLR_F2	5′-CTGACTCCATCCTGGGCACTG-3′	E12	60
LDLR_R2	5′-GATCAACCCACGCTCGCC-3′	E16	
LDLRAP1_F4	5′-CAGCTCATTGAGAACGTGTC-3′	E2*	
LDLRAP1_F5	5′-CTGATCCGGAGCCCCAGC-3′	E3*	64
LDLRAP1_R3	5′-AGTCTAGCAGGTTCCCAGTG-3′	E4*	
LDLRAP1_F3	5′-GGAGAAAGCCAGCCAAGAG-3′	E4*	58
LDLRAP1_R2	5′-GTCGCTTTGGTCCGTCATG-3′	E5*	

*Putative exon localization, based on human structure of this gene.

was performed on a total of 20 animals which included Duroc, Meishan, Iberian, Landrace, and Large White origins.

Polymorphism Genotyping

Primer design was carried out using the *Primer Express* software (Applied Biosystems). For the genotyping of *APOB* alleles, a PCR was designed around the *APOB-5* specific insertion of 230 bp at intron 28 (Table 3). Genotype was assigned based on the length of the amplified product (500 bp for *APOB-5* and 270 bp for other *APOB* alleles) as resolved by electrophoresis in a 1.5% agarose gel.

For the *LDLR:c.1653C>T* polymorphism, a 500 bp-long genomic sequence around position +1653 was amplified using primers LDLR_c.1653C>T_F on exon 13 and LDLR_c.1653C>T_R on exon 14 (Table 3). Genotyping was performed by restricting the PCR products with *Hin*1I (Fermentas), which cut the *C* but not the *T* allele. An allelic discrimination assay based on real time PCR was designed to genotype *LDLR:c.1661C>T* using allele discriminating probes. Primers and probes were placed on exon 13 (Table 3) and the assay was performed on an AB-7500 real time thermocycler (Applied Biosystems). To genotype the *LDLRAP1:c.909C>T* mutation, we performed a PCR-RFLP assay with primers located on exon 8

Table 3 Primers used for the genotyping of polymorphisms on pig *APOB, LDLR* and *LDLRAP1* genes

Primer name	Primer sequence 5′ → 3′	Localization*	T°ann
LDLR_c.1653C>T_F	GGCTTCATGTACTGGCTGATTG	E13	58°C
LDLR_c.1653C>T_R	GCTTGGAGTCGACCCAGTAAAG	E14	
LDLR_c.1661C>T_F	AGAAGGGCGGCCTGAAC	E13	
LDLR_c.1661C>T_R	CCATTGGGCCACTGGATGT	E13	60°C
MGB-probe_T	FAM-CCACTCCATTCAGGC-Quencher	E13	
MGB-probe_C	JOE-CCACTCCGTTCAGGC-Quencher	E13	
LDLRAP1_c.909C>T_F	CTTGCCCAGTCTCGGACGA	E8	58°C
LDLRAP1_c.909C>T_R	CTCCCTCAAGAGAGGGCAG	E8	
APOB-I28_F	AACCTCTACTACCGCCCTCA	I28	54°C
APOB-I28_R	CATGTGTGCATCCAATAAAATCA	I28	

*Putative exon localization, based on human structure of this gene.

(Table 3). The 410 bp-long PCR fragment was restricted with *Bse*GI, which cut the *T* but not the *C* allele.

Population-Wide and Within-Family Association Analysis

Association studies were performed with the following phenotypes: serum levels of TC, HDL- and LDL- bound cholesterol, and TG at 45 and 190 d of age. The model of analysis considered the allele substitution effects transmitted both through the male (within each family of half-sibs) and through the females (at population level). If the analyzed mutation is causal or is in total linkage disequilibrium at population level, both effects are expected to be coincident. Otherwise, the mutation must be discarded as causal mutation.

The common model of analysis for all traits was:

$$\mathrm{y} = \mathrm{Xb} + \mathrm{Zu} + \mathrm{T}_S\lambda_S + \mathrm{T}_D\lambda_D + \mathrm{e}$$

where:

y is the vector of phenotypic observations (logarithm of serum TC, LDL, HDL, or TG concentrations at 45 and 190 d)
b is a vector with all systematic effects (farm of origin, batch of fattening, age, or backfat thickness)
u is a vector of polygenic effects
X and Z are the incidence matrices for, respectively, the systematic and polygenic effects
λ_S is the vector of allele substitution effects transmitted through the sires (six levels)
λ_D is the allele substitution effect transmitted through the dam (a unique level)
T_S and T_D are the random incidence matrix and vector relating the phenotypes with the substitution effects *via* sire and *via* dam, respectively
e is the vector of residual effects.

The model was solved by Bayesian inference. Assuming normality of the transformed phenotypes, the likelihood of data is:

$$f\left(\mathrm{y}|\mathrm{b},\mathrm{u},\mathrm{T}_S,\mathrm{T}_D,\lambda_S,\lambda_D,\sigma_e^2\right) = N\left(\mathrm{Xb} + \mathrm{Zu} + \mathrm{T}_S\lambda_S + \mathrm{T}_D\lambda_D, \mathrm{I}\sigma_e^2\right)$$

The prior distribution for the polygenic effect was $f\left(\mathrm{u}|\sigma_u^2\right) = N\left(0, \mathrm{A}\sigma_u^2\right)$, where A is the additive numerator relationship matrix, which included pedigree information for up to five generations. The prior distributions for b, λ_S, λ_D, σ_e^2 and σ_u^2 were bounded uniform distributions, while for T_S and T_D they were discrete distributions defined by the information of the genotypes and the population gene frequency (p), i.e., $f(\mathrm{T}_S, \mathrm{T}_D|\mathrm{G}, p)$, where G is the matrix of genotypes for each polymorphism. The prior distribution of the gene frequency (p) was assumed uniform between 0 and 1.

From the likelihood and prior distributions, the joint posterior distribution was developed for the unknown parameters of the model. The marginal posterior distributions of the unknown parameters were obtained using a Gibbs Sampler (20). The conditional distributions needed on the Gibbs sampler were Gaussian

(b, u, λ_S, λ_D), inverted Chi Square (σ_e^2 and σ_u^2) and discrete distributions for the elements of T_S and T_D. Samples from the conditional posterior distribution of p were obtained using the Metropolis – Hastings algorithm (21). For each analysis, a total of 25,000 Gibbs iterations were performed after discarding the first 1,000. Convergence was checked using the Raftery and Lewis algorithm (22). The statistical relevance of the allelic substitution effects (λ) was calculated from their posterior distribution, by computing the probability (P) of these effects to be different from zero. Thus, we consider an effect as relevant when P over (when $\lambda > 0$) or below (when $\lambda < 0$) zero was bigger than 0.90.

RESULTS

Polymorphisms Detection on the Porcine *LDLR* Gene and Genotyping in the Analyzed Duroc Population

We identified three polymorphic sites on nucleotide positions +1086, +1653 and +1661 of *LDLR* cDNA, with respect to the ATG start codon. The first polymorphism (*LDLR:c.1086C>T*) is positioned on exon 10 and is characterized by a C/T substitution. We identified heterozygous animals in all breeds sequenced except in Duroc. The other two polymorphisms are C/T substitutions located on exon 13 (*LDLR:c.1653C>T* and *LDLR:c.1661C>T*), where *LDLR:c.1653C>T* is a synonymous change while *LDLR:c.1661C>T* is a non-synonymous substitution which entails a non-conservative Ser to Leu change in protein amino acid +554, affecting the second domain of the mature protein (EGF-like intracellular domain). In contrast with the former *LDLR:c.1086C>T* mutation, both polymorphisms on exon 13 segregated in our Duroc population with collected data on lipid serum levels, so we carried out the genotyping of the 385 individuals plus the six parental sires. For *LDLR:c.1653C>T*, allele *C* was the most frequent, with an allelic frequency of 0.83 (Fig. 1). Only two of the six parental males were heterozygous for this mutation. The genotyping of *LDLR:c.1661C>T* also gave allele *C* as the most prevalent with an allelic frequency of 0.87 (Fig. 1). Three of the parental sires were heterozygous for this polymorphism.

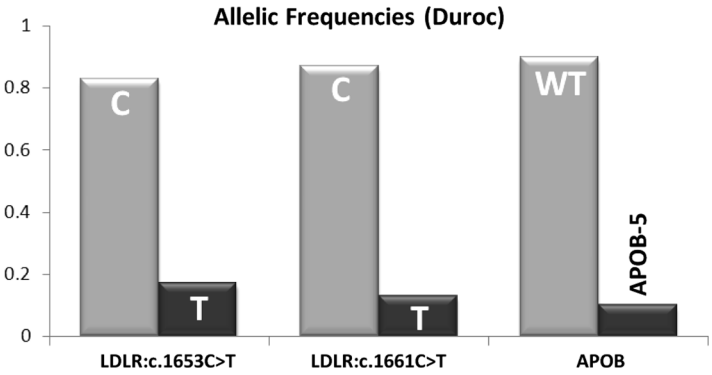


Figure 1 Allelic frequencies for the *LDLR:c.1653C>T*, *LDLR:c.1661C>T*, and *APOB-5* allele in the experimental Duroc pig population.

Genotyping the *APOB-5* Variant in the Analyzed Duroc Population

The *APOB-5* allele, although present in our population, had a low prevalence with an allelic frequency of 0.1 (Fig. 1). The rest of the alleles (designated as *WT* alleles) represented the remaining 0.9 and were pooled together as they represent other alleles not associated with hypercholesterolemia (10). Only one of the six half-sib families descended from a heterozygous sire.

Characterization and Polymorphism Detection of the Porcine *LDLRAP1* Gene

The amplified fragment for the porcine *LDLRAP1* cDNA included the complete *LDLRAP1* coding region of 857 bp and was submitted to GenBank with accession number EU072455. One single synonymous polymorphism on exon 9 was detected in a Meishan animal at position +909 (*LDLRAP1:c.909T>C*). However, this polymorphism did not segregate in our Duroc population, being all animals of *TT* genotype.

Genotyping of *LDLR* and *APOB* Polymorphisms in Other Breeds

To investigate the presence of these polymorphisms in other pig breeds, we genotyped the two *LDLR* polymorphisms and the *APOB-5* allele on a panel of 186 animals from Duroc (additional population), Meishan, Landrace, Iberian, Pietrain, and Large White breeds (Fig. 2). *LDLR:c.1653C>T* and the *APOB-5* allele were segregating in most populations. While the frequency of the *APOB-5* allele did not exceed 0.15, *LDLR:c.1653C>T* frequencies varied considerably among breeds (ranging from 0 to 0.88; Fig. 2). In clear contrast, animals of all breeds were monomorphic for *LDLR:c.1661C>T*, suggesting that this SNP is probably exclusive of our Duroc experimental population (Fig. 2).

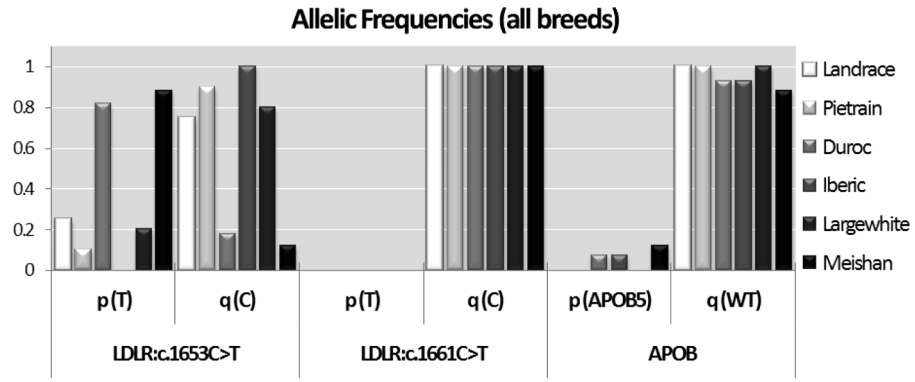


Figure 2 Distribution of allelic frequencies for the *LDLR:c.1653C>T*, *LDLR:c.1661C>T*, and *APOB* genotypes in pigs from six different breeds. Duroc animals are from a different population than those used to describe the polymorphism.

Association Studies with Phenotypic Data

Relevant results regarding the *APOB* gene association analyses are shown in Table 4. Consistent with its low allelic frequency (0.1), only one (♂1) of the six founder males was heterozygous for the *APOB-5* variant. Therefore, we could estimate the substitution effect *via* male for this mutation in only one family, along with the population-wise substitution effect *via* female. At both levels (*via* male and *via* female), associations between *APOB* genotype and serum lipid concentrations were found. Overall, the *APOB-5* allele exhibited an unfavorable effect over cholesterol metabolism at population level, since it was strongly associated to a lower ratio of HDL- *vs* LDL-bound cholesterol (Table 4). This was achieved by lowering the HDL-cholesterol and, consequently, the TG serum levels, which was more evident at 45 than at 190 d of age. Going further on this unfavorable effect, the presence of the *APOB-5* allele was also associated with higher LDL levels to the point of increasing total cholesterol concentration at 190 d, but only in ♂1 family, suggesting different phases of linkage disequilibrium of *APOB-5* with the causative mutation in this sire. According to our results, the posterior mean estimates of allele effects (in logarithmic scale) of *APOB* genotype on blood lipid parameters represented between the 16 and the 33% of the trait phenotypic standard deviation.

As for the two *LDLR* polymorphisms, *LDLR:c.1653C>T* had a very consistent association across families with serum TG levels on blood samples taken at 45 d of age (Table 5). Hence, the *T* allele was associated with higher TG content at population level (transmission *via* female) and also *via* male in the two families with a heterozygous sire (♂2 and ♂3). Within the ♂3 family, this mutation was also linked to lower cholesterol levels, in particular lower LDL-bound cholesterol, in blood samples taken both at weaning and at maturity. In fact, the allelic substitution effect over plasma LDL levels transmitted via ♂3 was the highest substitution effect detected in this study, representing 33 and 40% of the standard deviation of LDL levels at 45 and 190 d of age, respectively. These results were partially consistent with those observed at population level, where a favorable association with serum cholesterol *via* lowering both the HDL and LDL concentrations was shown for the T allele at

Table 4 Results from association analysis of *APOB* genotype with blood lipid parameters[a] at 45 and 190 d. Mean represents the *APOB-5* allele mean substitution effect in a logarithmic scale ($\ln(\text{mg} \cdot \text{dl}^{-1})$) and *P* is the posterior probability below or above 0

	Trait[a]	λ^b	Mean	*P*
45 d	HDL	$\lambda_{♀}$	−0.0622	0.96
	HDL/LDL	$\lambda_{♀}$	−0.0973	0.99
	TG	$\lambda_{♀}$	−0.0645	0.89
190 d	TC	$\lambda_{♂1}$	0.0641	0.93
	HDL	$\lambda_{♀}$	−0.0505	0.93
	LDL	$\lambda_{♂1}$	0.0981	0.93
	HDL/LDL	$\lambda_{♀}$	−0.0598	0.87

[a]TC = total cholesterol; HDL = high density lipoproteins; LDL = low density lipoproteins; TG = triglycerides.

[b]Allele substitution effect in the transmission *via* dam (♀) in the whole population or *via* sire (♂) in the half-sib families with heterozygous sire ($♂_1$).

Table 5 Results from association analysis of *LDLR_c.1653C>T* genotype with blood lipid parameters[a] at 45 and 190 days. Mean represents the *T* allele mean substitution effect in a logarithmic scale ($\ln(mg \cdot dl^{-1})$) and *P* is the posterior probability below or above 0

	Trait[a]	λ^b	Mean	*P*
45 d	TC	$\lambda_{♂3}$	−0.0442	0.89
	HDL	$\lambda_{♀}$	−0.0467	0.90
	LDL	$\lambda_{♂3}$	−0.0896	0.96
	LDL	$\lambda_{♀}$	−0.0616	0.94
	TG	$\lambda_{♂2}$	0.0817	0.90
	TG	$\lambda_{♂3}$	0.0942	0.93
	TG	$\lambda_{♀}$	0.1001	0.97
190 d	TC	$\lambda_{♂3}$	−0.0708	0.97
	HDL	$\lambda_{♂3}$	−0.0699	0.95
	LDL	$\lambda_{♂3}$	−0.1098	0.97
	HDL/LDL	$\lambda_{♀}$	0.0693	0.89

[a]TC = total cholesterol; HDL = high density lipoproteins; LDL = low density lipoproteins; TG = triglycerides.

[b]Allele substitution effect in the transmission *via* dam (♀) in the whole population or *via* sire (♂) in the half-sib families with heterozygous sire ($♂_2$ and $♂_3$).

45 d, representing 20 and 25% of the population phenotypic variation, respectively. As regards the *LDLR:c.1661C>T* polymorphism, overall, the *T* allele was associated with increased total cholesterol and with a negative effect over the ratio of HDL- *vs* LDL-bound cholesterol (Table 6). Three males were heterozygous for this polymorphism: ♂1, ♂4, and ♂5. The ♂4 sire within-family and whole population analyses gave consistent results from blood samples at 190 d, showing an increase of LDL-cholesterol associated to female- and ♂4-transmitted *T* alleles, which resulted

Table 6 Results from association analysis of *LDLR_c.1661C>T* genotype with blood lipid parameters[a] at 45 and 190 days. Mean represents the *T* allele mean substitution effect in a logarithmic scale ($\ln(mg \cdot dl^{-1})$) and *P* is the posterior probability below or above 0

	Trait[a]	λ^b	Mean	*P*
45 d	TC	$\lambda_{♀}$	0.0720	0.95
	HDL	$\lambda_{♂1}$	−0.0812	0.93
	HDL	$\lambda_{♂4}$	−0.0835	0.96
	LDL	$\lambda_{♀}$	0.0959	0.97
	HDL/LDL	$\lambda_{♂1}$	−0.0851	0.95
190 d	TC	$\lambda_{♀}$	0.0543	0.92
	HDL	$\lambda_{♂5}$	0.0729	0.90
	LDL	$\lambda_{♂4}$	0.0931	0.94
	LDL	$\lambda_{♀}$	0.1162	0.98
	HDL/LDL	$\lambda_{♂4}$	−0.1063	0.94
	HDL/LDL	$\lambda_{♀}$	−0.1298	0.98

[a]TC = total cholesterol; HDL = high density lipoproteins; LDL = low density lipoproteins; TG = triglycerides.

[b]Allele substitution effect in the transmission *via* dam (♀) in the whole population or *via* sire (♂) in the half-sib families with heterozygous sire ($♂_1$, $♂_4$ and $♂_5$).

in a decrease of the HDL/LDL ratio. The *T* allele transmitted *via* ♂1 also decreased the HDL/LDL ratio at 45 d, due to lower levels of serum HDL.

DISCUSSION

In humans, polymorphisms in a number of genes have been associated with an effect on serum lipids levels (23). Out of all these, three genes involved in cholesterol uptake and regulation are, however, the major determinants of heritable hypercholesterolemic diseases: *APOB*, *LDLR* and *LDLRAP1*. In contrast, little is known regarding the genetic basis of serum cholesterol levels in pigs and the porcine cholesterol-related genes. We show here that, like in humans, these three genes represent a source of DNA variability in pigs, which exhibit statistical association with serum lipid parameters.

We have described three new polymorphisms on the porcine *LDLR* gene. Conversely, the non-synonymous *LDLR:c.343C>T* mutation on exon 4 previously described by Grunwald and co-workers (3) and associated with high serum cholesterol levels was identified in none of the animals sequenced from a variety of breeds. This might be due to either low allelic frequency or the non existence of this SNP in these populations. From the three new polymorphisms, the synonymous *LDLR:c.1653C>T* segregated in all populations genotyped but in Iberian. In clear contrast, the non-synonymous *LDLR:c.1661C>T* could only be detected in the phenotypically studied Duroc population and in no other breed.

The *LDLR:c.1661C>T* SNP is located on exon 13 and promotes a non-conservative Ser → Leu change on residue +554. This mutation lies on the second domain of the mature protein, an EGF precursor-like domain, which participates in the displacement of LDL from the binding domain during the internalization step. In humans, mutations on this second domain diminish LDLR recycling after ligand endocytosis (24). Also, the degradation rate of the mature protein is largely affected resulting in familial hypercholesterolemia with elevated levels of LDL lipoproteins (25). Accordingly, the statistical analysis we performed showed a marked association between the pig *LDLR:c.1661C>T* allele transmitted *via* female and LDL-cholesterol and consequently the HDL/LDL ratio. Nevertheless, when broken down into paternal families, only in one family the allele transmitted *via* male exhibited the same results. This could either indicate a linkage disequilibrium of the *LDLR:c.1661C>T* mutation with the causative mutation, with different linkage phases in the paternal sire of this family, or, due to the low number of *T*-allele carriers, the lack of statistical power in the rest of families with heterozygous sire but fewer descendants.

On the other hand, a clear effect of *LDLR:c.1653C>T* on TG at 45 d of age was evident both for female- and male-transmitted alleles. Published reports on *LDLR* mutations give complete analyses on serum cholesterol fractions while serum TG levels are rarely assessed. Since this is a synonymous mutation, we shall expect the causative polymorphism to be in linkage disequilibrium within the same or a closely linked gene. While missense mutations of the coding regions are a well-reported cause of protein functional diversity, polymorphisms in regulatory elements are currently emerging as a very important source of phenotype variation (26) through alteration of the gene expression levels. Moreover, an increasing number of silent

polymorphisms are reported to influence the mRNA stability and subsequently its translation rate (27, 28). Overall, both SNPs on the porcine *LDLR* gene segregating in the analyzed Duroc population were associated with deleterious combinations of cholesterol composition, supporting the idea that a mutation in linkage disequilibrium is influencing the serum cholesterol parameters.

As previously reported by Purtell et al. (9), the *APOB-5* allele was associated with opposite changes on LDL and HDL levels which led to higher serum cholesterol at maturity and shifted the HDL/LDL coefficient to more unfavorable ratios at two ages, particularly at 45 d. In addition, we have shown that the *APOB-5* allele is present, although at very low frequency, in Duroc, Iberian, and Meishan animals; three breeds characterized by a high proportion of fat (23, 29). Estimates of genetic correlations between pig serum lipid levels and fatness are inexistent in the literature. However, although the lack of response for backfat thickness in some divergent selection experiments for serum cholesterol (6, 7) would suggest the absence of relevant genetic relationship between fatness and serum cholesterol, we have found a significant positive phenotypic correlation between backfat thickness and serum cholesterol content in our Duroc population (phenotypic correlation 0.33, $p < 0.0001$). Nevertheless, an environmental, rather than a genetic, correlation cannot be rule out at this stage. Finally, it is worth mentioning the lack of global concordance in the associations found for the three polymorphisms at the two analyzed ages. This is totally consistent with the weak or lack of phenotypic correlations between serum lipid concentrations at 45 and 190 d previously shown in the analyzed Duroc population (12).

Overall, results obtained in the present study for the pig *LDLR* and *APOB* mutations analyzed in a Duroc population are in agreement with results shown in human genetics. In humans, mutations on LDLR and its ligand APOB are the most common causes of autosomal dominant familial hypercholesterolemia. While over 800 different mutations, all rare, have been reported for the human *LDLR* gene, an Arg3500Gln point mutation at the *APOB* gene is very frequent in Europe (up to 1/500 carrier individuals) (30). Consistently with our results, the most prominent effects of *LDLR* and *APOB* mutations on human lipid serum levels are on raising TC and LDL levels while decreasing HDL concentrations.

Concerning discrepancies between substitution effects estimated in whole population and within-family analyses, it must be taken into account that the performance of gene association analyses in a half-sibs design of an outbred population suffers from several drawbacks that explain why associations are not always consistent across families and at population level. First, parents are only informative if they harbor heterozygous genotypes for the analyzed polymorphism and for the causal mutation. Second, in the case of a polymorphism in linkage disequilibrium with the causal mutation, the phase of linkage between the marker and the causal mutation alleles may vary amongst families. Third, the limited family size of some families might also limit the statistical power to estimate the within-family effects. And lastly, low allele frequencies further affect the power of the experiment. In summary, the aforementioned discrepancies between population and within family allele substitution effects should be interpreted with special caution because it is difficult to disentangle if they have been produced by genetic or experimental factors. In addition, the complex interaction between genes might play an important role in

phenotypic variability associated to these genes. Hence, most human *LDLR* mutations associated with hypercholesterolemia have a recessive or a semi-dominant effect. This is also shown with the +84 Arg → Cys missense mutation previously described in pigs (20). In this sense, we performed a conjoint analysis of the three mutations segregating in the analyzed Duroc population, in order to test the interaction between genotypes. However, the relative low number of samples in each category seriously compromised the statistical power of the test and impeded obtaining conclusive results to this respect.

As a general conclusion, results from this study show the presence of genetic variability in several porcine genes related to uptaking of LDL particles, and highlight the association between individual polymorphism on two of these genes and serum cholesterol and triglycerides levels in pigs. Further studies with more powerful experimental designs, larger number of animals and combined analysis of multiple genetic variants would be worthy to disentangle the complete role of these genes in pig cholesterol metabolism and the complex interaction between them.

REFERENCES

1. Corella D, Ordovas JM. Single nucleotide polymorphisms that influence lipid metabolism: Interaction with dietary factors. Annu Rev Nutr 2005; 25:341–390.
2. Kuller LH. Nutrition, lipids and cardiovascular disease. Nutr Rev 2006; 64:S15–S26.
3. Grunwald KA, Schueler K, Uelmen PJ, Lipton BA, Kaiser M, Buhman K, Attie AD. Identification of a novel Arg → Cys mutation in the LDL receptor that contributes to spontaneous hypercholesterolemia in pigs. J Lipid Res 1999; 40:475–485.
4. Varret M, Abifadel M, Rabès JP, Boileau C. Genetic heterogeneity of autosomal dominant hypercholesterolemia. Clin Genet 2008; 73:1–13.
5. Sirinian MI, Belleudi F, Campagna F, Ceridono M, Garofalo T, Quagliarini F, Verna R, Calandra S, Bertolini S, Sorice M, Torrisi MR, Arca M. Adaptor protein ARH is recruited to the serum membrane by low density lipoprotein LDL binding and modulates endocytosis of the LDL/LDL receptor complex in hepatocytes. J Biol Chem 2005; 280:38416–38423.
6. Lu CD, Pond WG, Mersmann HJ, Su DR, Krook J, Harris JJ, Savell JW. Dietary-fat and cholesterol in young-adult boars genetically selected for high or low serum-cholesterol response. J Anim Sci 1995; 73:2043–2049.
7. Harris KB, Pond WG, Mersmann HJ, Smith EO, Cross HR, Savell JW. Evaluation of fat sources on cholesterol and lipoproteins using pigs selected for high or low serum cholesterol. Meat Sci 2004; 66:55–61.
8. Hasler-Rapacz J, Ellegren H, Fridolfsson AK, Kirkpatrick B, Kirk S, Andersson L, Rapacz J. Identification of a mutation in the low density lipoprotein receptor gene associated with recessive familial hypercholesterolemia in swine. Am J Med Genet 1998; 76:379–386.
9. Purtell C, Maeda N, Ebert DL, Kaiser M, Lund-Katz S, Sturley SL, Kodoyianni V, Grunwald K, Nevin DN, Aiello RJ, Attie AD. Nucleotide sequence encoding the carboxyl-terminal half of apolipoprotein B from spontaneously hypercholesterolemic pigs. J Lipid Res 1993; 34:1323–1335.
10. Prescott MF, Hasler-Rapacz J, von Linden-Reed J, Rapacz J. Familial hypercholesterolemia associated with coronary atherosclerosis in swine bearing different alleles for apolipoprotein B. Ann NY Acad Sci 1995; 748:283–292.
11. Varona L, Gomez-Raya L, Rauw WM, Noguera JL. A simulation study on the detection of causal mutations from F-2 experiments. J Anim Breed Genet 2005; 122:30–36.

12. Gallardo D, Pena RN, Amills M, Varona L, Ramirez O, Reixach J, Diaz I, Tibau J, Soler J, Prat-Cuffi JM, Noguera JL, Quintanilla R. Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs. Physiol Genomics 2008; 35:199–209.
13. Richmond W. Analytical reviews in clinical biochemistry: The quantitative analysis of cholesterol. Ann Clin Biochem 1992; 29:577–597.
14. Rafai N, Warnick GR, Eds. Laboratory Measurement of Lipids, Lipoproteins and Agolipoproteins; AACC Press Laboratory, Washington DC, 1994.
15. Fossati P, Prencipe L. Serum triglycerides determined colonmetrically with an enzyme that produces hydrogen peroxide. Clin Chem 1982; 28:2077–2080.
16. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in serum, without use of the preparative ultracentrifuge. Clin Chem 1972; 18:499–502.
17. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987; 162:156–159.
18. Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual, 3rd edn.; Cold Spring Harbor Laboratory Press, New York, 2001.
19. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. Multiple sequence alignment with the Clustal series of programs. Nucl Ac Res 2003; 31:3497–3500.
20. Gelfand AE, Smith AFM. Sampling based approaches to calculating marginal densities. J. Am Stat Assoc 1990; 85:398–409.
21. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970; 57:97–109.
22. Raftery AE, Lewis SM. How many iterations in the Gibbs Sampler? In Bayesian Statistics IV; Bernardo JM, Berger JO, Dawid AP, Smith AFM, Eds.; New York, Oxford University Press, 1992:763–774.
23. Ordovas JM. Genetic interactions with diet influence the risk of cardiovascular disease. Am J Clin Nutr 2006; 83:443S–446S.
24. Van Hoof D, Rodenburg KW, Van der Horst DJ. Intracellular fate of LDL receptor family members depends on the cooperation between their ligand-binding and EGF domains. J Cell Sci 2005; 118:1309–1320.
25. Martin de Llano JJ, Fuertes G, Andreu EJ, Puig O, Chaves FJ, Soutar AK, Armengod ME, Knecht E. A single point mutation in the low-density lipoprotein receptor switches the degradation of its mature protein from the proteasome to the lysosome. Int J Biochem Cell Biol 2006; 38:1340–1351.
26. Pampin S, Rodriguez-Rey JC. Functional analysis of regulatory single-nucleotide polymorphisms. Curr Opin Lipidol 2007; 18:194–198.
27. Chen JM, Ferec C, Cooper DN. A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes I: General principles and overview. Hum Genet 2006; 120:1–21.
28. Chen JM, Ferec C, Cooper DN. A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes II: The importance of mRNA secondary structure in assessing the functionality of 3′ UTR variants. Hum Genet 2006; 120:301–333.
29. Daza A, Lopez-Bote C, Rey A, Olivares A. Effect of age at the beginning of the free-range fattening period on growth and carcass and fat quality in Iberian pigs. Arch Anim Nutr 2006; 60:317–324.
30. Soutar AK, Naoumova RP. Mechanisms of disease: Genetic causes of familial hypercholesterolemia. Nat Clin Pract Cardiovasc Med 2007; 4:214–225.