UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Ciència Animal i dels Aliments

Facultat de Veterinària

CENTRE DE RECERCA EN AGRIGENÒMICA

Departament de genètica animal

Functional genomics and candidate genes for meat quality traits in pigs

JORDI COROMINAS GALBANY

Ph.D Thesis in Animal Production

Bellaterra, September 2013

Supervisors:

Dr. Josep Maria Folch Albareda **Dra. Maria Ballester Devis**

“If we all did the things we are capable of,

we would astound ourselves.”

— Thomas A. Edison

El Dr. **Josep Maria Folch Albareda**, professor titular del Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de Barcelona,

i

la Dra. **Maria Ballester Devis**, investigadora del Departament de Genètica Animals del Centre de Recerca en Agrigenòmica.

CERTIFIQUEN:

Que en **Jordi Corominas Galbany** ha realitzat sota la seva direcció el treball de recerca "Functional genomics and candidate genes for meat quality traits in pigs" per obtenir el grau de Doctor per la Universitat Autònoma de Barcelona.

Que aquest treball s'ha dut a terme al Departament de Ciència Animal i dels Aliments de la Facultat de Veterinària de la Universitat Autònoma de Barcelona i a la unitat de Genètica Animal del Centre de Recerca en Agrigenòmica.

Bellaterra, a 27 de setembre de 2013

Dr. Josep Maria Folch Albareda Dra. Maria Ballester Devis

CONTENT

SUMMARY

Pig is one of the most economically important domestic species, since they have been extensively used not only as the major human meat source, but also for biomedical research. A good understanding of the molecular mechanisms affecting meat quality traits in pigs is essential for obtaining meat with a fatty acids profile more in line with public health recommendations without affecting the production yield. Food fatty acid composition is highly relevant for determining meat quality traits, but its metabolism is composed by a complex network of processes and pathways that are not fully understood. Elucidating these molecular processes is the general objective of this thesis, in order to better understand the genetic architecture of meat quality traits in pigs.

We evaluated the functional implication of the genetic variant *DQ144454:c.2645G<A* on *ACSL4* gene expression, observing in liver a higher expression in animals with the G allele than animals with the A allele. A SNP genome-wide association study with *ACSL4* gene expression was performed and two significant regions on SSC6 and SSC12 with candidate genes associated with *ACSL4* gene expression were identified.

On the other hand, *ELOVL6* gene was analyzed as a candidate gene for the QTL on SSC8 affecting palmitic and palmitoleic acid content. The complete genetic architecture of *ELOVL6* gene was described, showing two different isoforms in liver and adipose tissue. The *ELOVL6:c.-533C>T* SNP, identified in the promoter region, was highly associated with its own gene expression and the percentages of palmitic and palmitoleic fatty acids in muscle and adipose tissue. Functional analyses of *ELOVL6* promoter showed the occupancy of SREBF1 and ERα, suggesting an important role of both transcription factors on the regulation of *ELOVL6* gene expression. Interestingly, the *ELOVL6:c.-394G>A* SNP is in linkage disequilibrium with *ELOVL6:c.-533C>T* SNP and also is located within the only predicted ERE on *ELOVL6* promoter. In addition, ChIP assays showed that ERα binding depends on *ELOVL6:c.-394G>A* genotype, showing union only in animals carrying the G allele. It has been described that ERα binding causes the recruitment of Dnmts and, consequently, an increase on the methylation levels of the surrounding CpG motifs. Therefore, differences on ERα binding and the methylation pattern may be the causal factors of the lower *ELOVL6* expression in animals with the *ELOVL6:c.-394G* allele. Hence, we proposed *ELOVL6:c.-394G>A* SNP as the major putative causal mutation for explaining the phenotypic variation of the QTL analyzed.

Finally, the adipose tissue transcriptomes of two groups of phenotypically extreme pigs for intramuscular fatty acid composition were sequenced using RNA-Seq. Differential expression analysis identified 396 genes differentially expressed between groups. The major metabolic pathway differentially modulated between groups was lipogenesis, probably caused by the differences on PUFA content between the two groups. Therefore, the linked effect of fatty acid composition on lipid-related gene expression suggested that differentially-expressed genes may play an important role in lipid and fatty acid metabolism.

RESUMEN

El cerdo es la especie doméstica más importante económicamente, no solo por ser la principal fuente de carne en la dieta humana, sino también por su papel como modelo animal para la investigación biomédica. La buena compresión de los mecanismos moleculares que afectan a la calidad de la carne en cerdos es esencial para la obtención de carnes con un patrón saludable de ácidos grasos sin afectar el rendimiento productivo. La composición de ácidos grasos es un carácter muy importante para la calidad de la carne; no obstante, nuestro conocimiento sobre la compleja red de procesos y vías metabólicas que forman el metabolismo de los ácidos grasos sigue siendo limitado. Esta tesis tiene como objetivo general mejorar el entendimiento de estos procesos moleculares, con el fin de comprender mejor la arquitectura genética de la calidad de la carne en el cerdo.

Se ha evaluado la implicación funcional de la variante *DQ144454:c.2645G<A* en la expresión del *ACSL4*, observando en hígado una mayor expresión en los animales con el alelo G en comparación con los que presentan el alelo A. Se realizó un estudio de asociación (eGWAS) con la expresión génica de *ACSL4* identificando dos regiones en los cromosomas 6 y 12 con genes que pueden estar relacionados con la expresión del gen.

El gen *ELOVL6* ha sido analizado como gen candidato para el QTL del cromosoma 8, relacionado con el contenido de ácidos palmítico y palmitoleico. Se ha determinado la arquitectura genética del gen ELOVL6, mostrando dos isoformas expresadas en hígado y tejido adiposo. El SNP ELOVL6:c.-533C>T presentó una alta asociación con la expresión del propio gen y el contenido de los ácidos palmítico y palmitoleico en músculo y tejido adiposo. Estudios funcionales del promotor del gen validaron la unión de SREBF1 y ERα, sugiriendo un papel importante de ambos factores de transcripción en la regulación del gen *ELOVL6*. Curiosamente, el SNP *ELOVL6:c.-394G>A* se encuentra en desequilibrio de ligamiento con el SNP *ELOVL6:c.-533C>T* y también se encuentra localizado en el único lugar de unión de ERα en el promotor. Además, los estudios realizados permitieron observar una unión diferencial de ERα en función del genotipo del *ELOVL6:c.-394G>A*, mostrando su unión solo en los animales portadores del alelo G. La unión de ERα ha sido asociada al reclutamiento de metilasas y, consecuentemente, al aumento de los niveles de metilación de los motivos CpG adyacentes. Por lo tanto, las diferencias en la unión de ERα y en el grado de metilación del promotor pueden ser los factores causales de la menor expresión del gen en animales con el alelo G. Por este motivo proponemos el SNP *ELOVL6:c.-394G>A* como el principal SNP causal de las variaciones fenotípicas del QTL.

Finalmente, el transcriptoma del tejido adiposo de cerdos fenotípicamente extremos para la composición de ácidos grasos en músculo fue secuenciado mediante *RNA-Seq*. Se realizó un análisis de expresión diferencial que permitió identificar 396 genes diferencialmente expresados que modulaban principalmente la lipogénesis, probablemente debido a las diferencias en el contenido de ácidos grasos poliinsaturados entre grupos. Este efecto de la composición de ácidos grasos sobre la expresión de genes lipídicos sugiere que los genes diferencialmente expresados pueden desempeñar un papel importante en el metabolismo de los lípidos y de los ácidos grasos.

LIST OF TABLES

LIST OF FIGURES

General Introduction

General Discussion

Annexes

Paper I:

LIST OF PUBLICATIONS

This thesis is based on the work contained in the following published papers:

Paper I **Corominas J.**, Ramayo-Caldas Y., Puig-Oliveras A., Pérez-Montarelo D., Noguera J.L., Folch J.M. & Ballester M. (2013) Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE* **8**, e53687.

Paper II **Corominas J.**, Ramayo-Caldas Y., Castelló A., Muñoz M., Ibáñez-Escriche N., Folch J.M. & Ballester M. (2012) Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs. *Animal Genetics* **43**, 714-20.

Paper III **Corominas J.**, Ramayo-Caldas Y., Puig-Oliveras A., Estellé J., Castelló A., Alves E., Pena R.N., Ballester M. & Folch J.M. (2013) Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics*. In revision.

and the preliminary manuscript:

Paper IV **Corominas J.** et al. Different patterns of methylation on ELOVL6 promoter caused by a promoter polymorphism is associated with a major QTL effect on fatty acid composition in pigs.

RELATED PUBLICATIONS BY THE AUTHOR

(Not included in the thesis)

- Ramayo-Caldas Y., Mach N., Esteve-Codina A., **Corominas J.**, Castello A., Ballester M., Estelle J., Ibanez-Escriche N., Fernandez A., Perez-Enciso M. & Folch J.M (2012). Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics* 13, 547.

- Muñoz M, Alves E, **Corominas J**, Folch JM, Casellas J, Noguera JL, Silió L, Fernández AI. (2012). Survey of SSC12 Regions Affecting Fatty Acid Composition of Intramuscular Fat Using High-Density SNP Data. *Front Genet* 2,101. doi: 10.3389/fgene.2011.00101.

ABBREVIATIONS

IMF → intramuscular fat

FA → fatty acids

PUFA → polyunsaturated fatty acids

SFA → saturated fatty acids

MUFA → monounsaturated fatty acids

QTL → quantitative trait loci

TAS → trait-associated SNP

RNA-Seq→ RNA sequencing

GWAS → genome-wide association studies

BC1_LD → IBMAP backcross between Iberian and Landrace breeds

SNP → Single nucleotide polymorphism

ChIP→ Chromatin-immunoprecipitation

UTR → untranslated region

1. GENERAL INTRODUCTION

1.1.- Porcine meat production

Worldwide meat production of livestock species has increased in the last decade due to the application of different breeding programmes. This upward trend was clearly observed in the porcine meat production with a 21% increase from 2000 to 2010 (Figure 1.1). Porcine is the major source of meat, representing the 39% of the meat produced worldwide, followed by chicken (31%) and bovine (22.5%). Asia is so far the major producer, accounting for the 57% of the produced porcine meat, followed by Europe (25%) and North and Central America (12%) (FAO 2011: http://faostat.fao.org/). In the European community, the two main producers are Germany and Spain and both increased their porcine meat production between 2009 and 2011: 6.1% in Germany and 5.7 % in Spain. Catalonia is leading the porcine production in Spain with the 40% of the Spanish porcine meat production, according to the Annual report from the *Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural* (Gaspa *et al.* 2011).

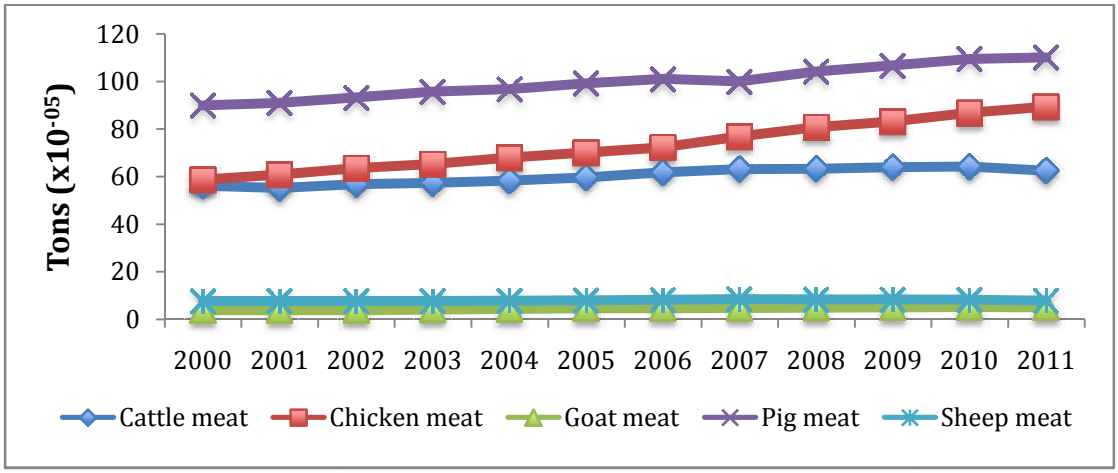

Figure 1.1- Evolution of meat production worldwide in the different livestock animals from the period 2000-2010 (FAOSTAT, 2012).

Pig breeding programmes in the early twenty-century have focused on increasing the percentage of lean meat in carcass, producing a strong negative selection on fat in pig commercial breeds. The dramatic reduction of intramuscular fat (IMF) content caused detrimental changes on sensory meat quality, by affecting negatively to taste and tenderness (Wood & Whittemore 2007). As consequence, during the last decade the genetics of pig meat quality has became a subject of increased research and the improvement of this trait is essential to satisfy the consumer's demands for excellent eating, healthy and nutritional quality. Hence, current genetics and genomics works on porcine meat quality traits and the increasing consumer's awareness of food quality have driven pig breeding companies to include these traits as an integral part of pig selection programmes.

1.2.- Pork meat quality

In general, meat quality is a composite concept difficult to measure in a simple and unique manner (Davoli & Braglia 2008) and also its definition varies between countries and meat industries. Perhaps the most extensive definition of meat quality is that of Hoffman (1994) who suggested that meat quality is the global result of the sensory properties, technological factors, nutritive values and hygienic and toxicological or food safety aspects that influence the product value for the consumer and meat industries (Hoffman 1994). The importance of all these qualitative aspects depends on which kind of meat we are analyzing. In fresh meat, appearance (colour and marbling), tenderness, juiciness, texture, flavour are essential to satisfy the consumer's demand (Rothschild & Ruvinsky 2011). In processed pork products, quality is determined by factors such as firmness, water-holding capacity, cooking loss, conductivity (Rothschild & Ruvinsky 2011). Table 1 lists technological, sensory and nutritional quality characteristics that are frequently considered in determining the quality of pork fresh and processed meats.

Table 1.1- Sensory, technological and nutritional factors related with pork quality traits. Asterisks indicate which final pork product are affected for each trait. Adapted from (Sellier 1998).

Perception	Traits	Fresh meat	Processed meat
Sensory quality	Post morten proteolysis	*	
	Texture	*	
	Flavour	*	
	Tenderness	*	
	Juiciness	*	
	Appearance (Colour and marbling)	*	
Technological quality	Firmness	*	*
	pH	*	*
	Water-holding capacity	*	*
	Cooking loss		*
	Conductivity		*
	Processing yield		*
Nutritional quality	Protein		*
	Lipid		*
	Vitamins		*
	Minerals		*
	Digestibility		*

Meat quality traits show low to moderate heritabilities (average ranking from 0.10-0.30) in sensorial and technological quality traits (Rothschild & Ruvinsky 2011), while nutritional quality traits, such as FA composition, show moderate to high heritabilities with average values into the range 0.15-0.55 (Casellas *et al.* 2010; Ntawubizi *et al.* 2010). Most of meat quality traits are dependent on muscle composition, which is determined by diet, animal handling, animal welfare and genetic factors. It is well know the importance of these factors in determining the IMF content and FA composition in muscle, which are two key traits affecting firmness, colour, tenderness, juiciness, flavour and texture of meat (Huff-Lonergan *et al.* 2002; Wood *et al.* 2004; Wood *et al.* 2008; Ramayo-Caldas *et al.* 2012b).

The interest of meat FA composition stems from the need to find ways to produce healthier meat, presenting a higher ratio of polyunsaturated fatty acids (PUFA) to

saturated fatty acids (SFA) and decreasing the ratio n-6/n-3 PUFA (Simopoulos 2001; Wood *et al.* 2004; Guillevic *et al.* 2009). PUFA, mainly n-3, have been considered beneficial for human health due to their effect in the reduction of total cholesterol concentration and the modulation of immune functions and inflammatory processes (Rudel *et al.* 1995; Poudyal *et al.* 2011). On the other hand, the high consumption of SFA is associated with diseases associated with modern life, such as various cancers, obesity and especially coronary heart diseases (Chizzolini *et al.* 1999; Wood *et al.* 2004). Obtaining optimal FA ratios for healthy meat is important for human consumption, but the impact on meat quality traits must be considered. High levels of PUFA have negative effects on the oxidative stability of meat, producing undesirable colour and flavour in pork meat, decreasing the consumer's acceptance (Wood *et al.* 2008). In contrast, SFA and monounsaturated fatty acids (MUFA) tend to be associated with desirable sensorial characteristics (Carrapiso *et al.* 2003; Wood *et al.* 2008). Additionally, diets rich in MUFA produced meat with a favourable nutritional profile in comparison with diets with SFA (Wood & Enser 1997) and have also been associated with meat flavour (Cameron & Enser 1991) and a lower susceptibility to oxidation compared with animals fed with diets rich in PUFA (Lopez-Bote *et al.* 1997). Therefore, elucidating the genetic basis of the biological processes controlling FA metabolism is essential to improve pork meat quality and also to determine the role of FA metabolism in metabolic human diseases.

1.3.-Fatty acid metabolism

Lipids are molecules that include a broad variety of compounds indisputable for life. FA do not only serve as a major source of energy, but are also crucial structural components of membranes. In addition, several studies have described the role of FA as signalling molecules, thus exerting key biological functions such as regulating FA metabolism (Duplus & Forest 2002; Cao *et al.* 2008).

Dietary FA are modulated in the digestive tract before reaching target tissues. Free FA released by pancreatic lipases diffuse into the intestinal epithelial cells, where triglycerides are synthesized from the monoglycerol and free FA. Lipoprotein particles

called chylomicrons are formed from triglycerides together with cholesterol, phospholipids and proteins. Via the lymphatic system, the chylomicrons reach the circulation, delivering fatty acid to various tissues such as liver to obtain energy or adipose tissue to store energy. The direction of FA metabolism depends on the nutritional status. In the fed state, carbohydrates are converted to FA which are stored in triglycerides as energy reservoirs, while in the fasting state, triglycerides breakdown and FA oxidation predominates.

1.3.1- FA β-oxidation

In situations where energy expenditure exceeds energy intake, the breakdown of FA provides the organism with energy in a process called β-oxidation (Neely & Morgan 1974; Felig & Wahren 1975). Hepatic β-oxidation also provides ketone bodies (acetoacetate and β-hydroxybutyrate) to brain, when blood glucose levels are low (Eaton *et al.* 1996). The breakdown of FA up to 18 carbons in length takes place in the mitochondria while longer FA need to be shortened in peroxisomes before further oxidation in mitochondria occurs.

The enzymes of β-oxidation all act on CoA esters, so a preliminary step to β-oxidation is the ATP-dependent formation of fatty acyl-CoA esters, catalysed by acyl-CoA synthetases (ACS). Secondly, carnitine palmitoyltransferases 1 (CPT1) transfers the acyl-CoA ester to carnitine and transports the carnitine derivates through the inner mitochondrial membrane. Once there, the opposite reaction is performed with the carnitine palmitoyltransferases 2 (CPT2), separating the carnitine from the acyl-CoA ester (Eaton *et al.* 1996) (Figure 1.2). The β-oxidation process is produced by the consecutive actions of several enzymes, starting with the formation of a double bound in the β carbon by acyl-CoA dehydrogenases, followed by a hydration of the double bound by 2-enoyl-CoA hydratases and ending with a second dehydrogenation by 3-hydroxyacyl-CoA dehydrogenases, generating a second keto group. Finally, the 3-oxoacyl-CoA thiolases cleaves between the two continuous keto groups, in a process called thiolysis (Eaton *et al.* 1996). Products generated are one molecule of acetyl-CoA and one acyl-CoA with two carbon atoms less (n-2 acyl-CoA), which starts again the β-oxidation process until all acyl-CoA is degradated to acetyl-CoA molecules (Figure 1.2).

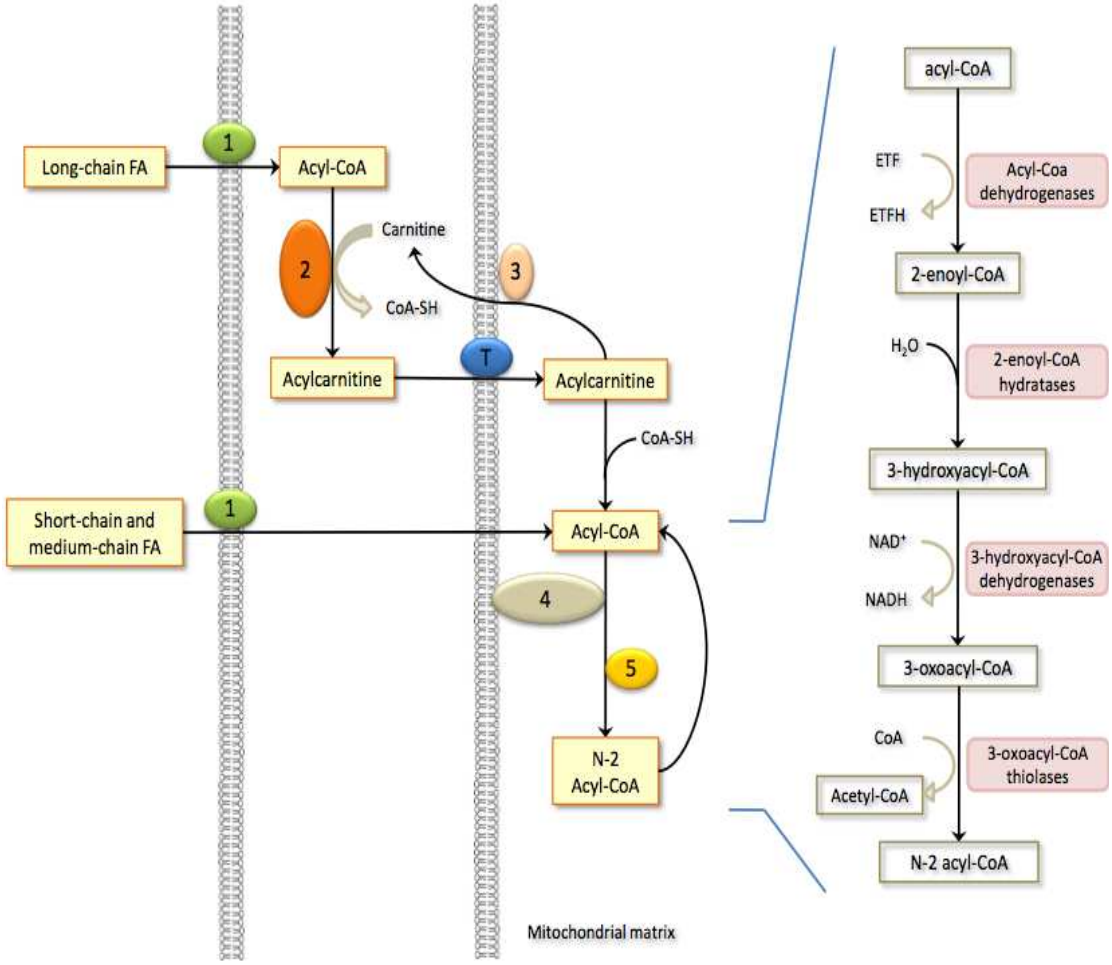

Figure 1.2- FA mitochondrial β-oxidation. Enzymes: 1: Acyl-CoA synthetase; 2: carnitine palmitoyltransferase 1; 3: carnitine palmitoyltransferase 2; 4: β-oxidation enzymatic system associated to inner membrane; 5: soluble β-oxidation enzymatic system; T: transporter.

1.3.2- De novo FA synthesis

FA can either be derived directly from the diet or they can be *de novo* synthesized through lipogenesis, a key event in the energy storage system. Lipids, carbohydrates and amino acids can be metabolized into acetyl-CoA, thus serving as substrate for lipogenesis.

De novo FA synthesis starts with the carboxylation of the acetyl-CoA by acetyl-CoA carboxylase alpha (ACACA) to produce malonyl-CoA. Besides being used for FA synthesis, malonyl-CoA also acts as an inhibitor for CPT1, which transports acyl-CoA into the mitochondria for β-oxidation (McGarry *et al.* 1977). FA up to 16 carbons in

length are synthesized in the cytosol by the multifunctional protein fatty acid synthase (FAS), which utilizes acetyl-CoA as the priming substrate and malonyl-CoA as the two-carbon donor in the sequential synthesis of palmitic acid (C16:0) (Wakil *et al.* 1983; Jayakumar *et al.* 1995). The reaction steps are condensation, reduction, dehydration and one further reduction (Figure 1.3).

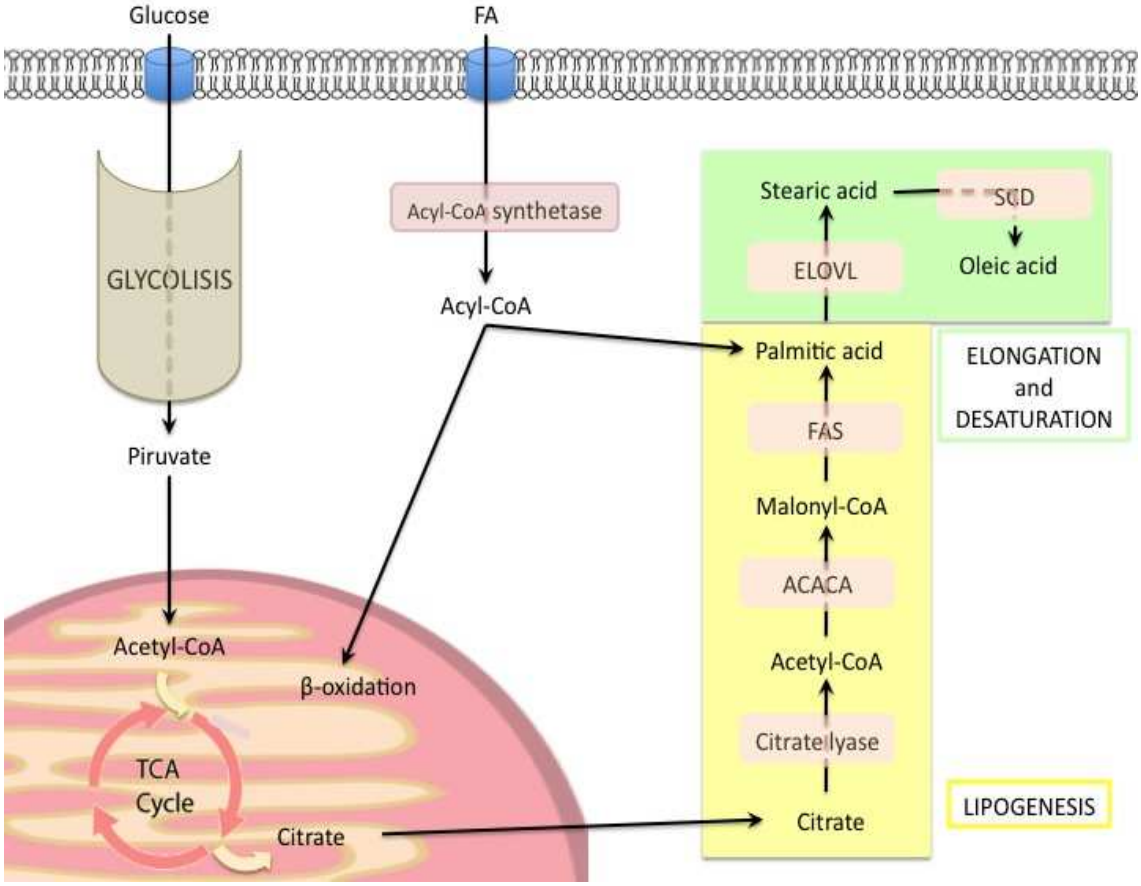

Figure 1.3- Metabolic pathways in the synthesis of FA. Inside the cell, glucose is converted to pyruvate via glycolysis. Pyruvate enters the Krebs cycle in the mitochondria and subsequently leaves the mitochondria as citrate, which is used as substrate for lipogenesis initialization. On the other hand, transformed FA are also incorporated to lipogenesis for elongation and desaturation. Adapted from (Postic & Girard 2008).

FA taken up from the diet, as well as, a significant amount of FA produced by FAS, undergo further elongation into long chain FA (>C18) and very long chain FA (>C20). FA elongation involves also the addition of two-carbon units to a fatty acyl-CoA, employing malonyl-CoA as the donor and NADPH as the reducing agent. The process of

elongation requires four separate enzymatic reactions: 1) the first and rate-limiting step is performed by elongation of very long chain fatty acids (ELOVL) enzymes, condensing fatty acyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA, 2) reduction of the 3-ketoacyl-CoA using NADPH to form a 3-hydroxyacyl-CoA, 3) dehydration of the 3-hydroxyacyl-CoA to trans-2-enoyl-CoA, and 4) reduction of the trans-2-enoyl-CoA to the final elongated fatty acyl-CoA (Moon *et al.* 2001; Guillou *et al.* 2010). Besides the variation in chain length, the acyl chain of a fatty acid can be desaturated by the introduction of a position specific double bond performed by acyl-CoA desaturases (Figure 1.3).

1.4.- Pig genomics

1.4.1- Pig genome chronology

The pig (*Sus scrofa*) was the first livestock species whose scientific community decided to map their genome in the early 1990s (Haley *et al.* 1990; Rothschild & Ruvinsky 2011). The EU-funded Pig Gene Mapping Project (PiGMaP) was the first internationally coordinated effort to map the porcine genome, focused on the development of genetic markers in pigs (Davies *et al.* 1994; Coppieters *et al.* 1995; Groenen *et al.* 1995) and the establishment of genetic linkage (Archibald *et al.* 1995) and cytogenetic (Echard *et al.* 1992) maps. Genetic markers of the pig genome identified from PiGMaP project data, allowed the implementation of linkage analysis for the identification of quantitative trait loci (QTLs) (Andersson *et al.* 1994; Pérez-Enciso *et al.* 2000; Clop *et al.* 2003). In September of 2003, the Swine Genome Sequencing Consortium (SGSC) was originated to provide international coordination for sequencing the pig genome. The main goal of this consortium was to advance biomedical research for animal production and health by the development of DNA based tools and products resulting from the sequencing of the swine genome (Schook *et al.* 2005a). The sequencing strategy followed a hybrid approach combining hierarchical shotgun sequencing of BAC clones and whole genome shotgun sequencing. The annotated genome assembly Sscrofa9 was released with Ensembl 56 in September of 2009 and later revised to Sscrofa10 by incorporating the whole genome shotgun sequence data, providing >30x

genome coverage (Archibald *et al.* 2010). More recently, a high quality draft of the pig genome sequence (Sscrofa10.2) has been published by the SGSC (Groenen *et al.* 2012). This new assembly comprises 2.6 Gb assigned to chromosomes and more than 212 Mb in unplaced scaffolds. The genome annotation allowed the identification of 21,640 coding genes, 380 pseudogenes and 2,965 non coding RNAs (ncRNAs). Furthermore, a *de novo* repeat discovery and annotation strategy revealed a total of 95 novel repeat families. The repetitive elements represent the 40%, being the LINE1 and PRE (porcine repetitive element) the most abundant elements (Groenen *et al.* 2012).

The development of the annotated porcine genome represents an incredibly rich source of information that allows the identification of gene markers for meat quality traits, providing significant insights into the molecular basis of phenotypic variation of production traits and assisting breeders in pig selection. The availability of pig genome sequence together with the development of a high-density single nucleotide polymorphism (SNP) panel for pig genotyping (Porcine 60K SNP Beadchip, Illumina) (Ramos *et al.* 2009) caused an important increase on the number of identified QTLs (Duijvesteijn *et al.* 2010; Fernandez *et al.* 2012; Gregersen *et al.* 2012; Ramayo-Caldas *et al.* 2012b; Yang *et al.* 2013). Additionally, the SNP chip has also been employed for population genetic studies (Burgos-Paz *et al.* 2013) and for porcine structural variant identification (Ramayo-Caldas *et al.* 2010; Chen *et al.* 2012b; Wang *et al.* 2012). Moreover, the continuous improvement of the high-throughput sequencing technologies may be used to explore the genetic architecture of complex traits and also to generate transcriptional profiles (RNA-Seq) that can be subject to expression QTLs (eQTL) mapping procedures. New developed fields related with functional genomics and protemics are very useful tools for the study of complex traits. Considering many genes and proteins at the same time is important not only for increasing the knowledge of gene functions and gene regulation but also to know how these genes are participating in complex networks controlling the phenotypic characteristics of traits.

1.4.2- Genome-wide transcriptome profiling using RNA-Seq

Development of microarrays in porcine species allowed the production of an important number of studies related with transcriptome profiling and its mechanisms of regulation (Ponsuksili *et al.* 2008; Pérez-Enciso *et al.* 2009; Freeman *et al.* 2012; Pérez-Montarelo *et al.* 2012). However, the rapid evolution of high-throughput sequencing has impacted heavily on the methodology of transcriptome analysis. The RNA-Seq approach involves the conversion of isolated transcripts into the complementary DNA (cDNA), which is then directly sequenced in a massively parallel sequencing-based approach. Beyond gene expression analysis, RNA-Seq can identify novel transcripts, novel isoforms, alternative splice sites, allele-specific expression and rare transcripts in a single experiment. Hence, RNA-Seq offers several advantages in comparison with microarrays (Table 1.5).

Table 1.5- Comparison of RNA-Seq technology with expression microarrays (Adapted from Illumina white paper: see http://eh.uc.edu/genomics/files/ Illumina_Whitepaper_RNASeq_to_arrays_comparison.pdf).

Application	RNA-Seq	Microarray
Transcript discovery	Yes	No
High run-to-run reproducibility	Yes	Yes
Dynamic range comparable to actual Transcript abundances within cells	Yes	No
Able to detect allele-specific expression	Yes	No
De novo analysis of sample without a Reference genome	Yes	No
Re-analyzable data	Yes	No

Due to the novelty of the technique, the number of RNA-Seq studies in pigs is still scarce (Chen *et al.* 2011; Esteve-Codina *et al.* 2011; Jung *et al.* 2012; Ramayo-Caldas *et al.* 2012a), but in the coming years there will be a significant increase of published works using this technology. In general, one of the main purposes of RNA-Seq studies

in livestock species is the identification of candidate functional genes for the analyzed traits and also the identification of gene networks and tissue-specific expression clusters. In this sense, recent studies in the IBMAP project are focused on the transcriptome analysis of the most important tissues in lipid metabolism: liver (Ramayo-Caldas *et al.* 2012a), adipose tissue (Corominas *et al.* 2013a) and muscle (Puig-Oliveras *et al.*). The main goal of these studies is to determine the effects of each transcriptome on FA composition traits in *longissimus dorsi* muscle. These analyses will allow us to have an overall picture of all tissues transcriptomes affecting these traits, what is essential for a better understanding of the complexity and the interconnectivity of lipid metabolic processes.

1.5.- Genomic studies of porcine meat quality traits

1.5.1- Mapping of meat quality QTLs

Most economically important traits in pigs are complex or quantitative and, thus, are influenced by multiple genes and environmental factors. The stretches of DNA containing or linked to the genes that affect the variation of a quantitative trait are known as QTL (Andersson 2001). The identification of QTLs affecting characters of high economic interest and their causal mutations is one of the major goals of animal genetics.

The hunt for QTLs in pigs has been ongoing for nearly two decades, beginning with the first publication of a QTL for fatness on pig chromosome 4 (Andersson *et al.* 1994). Since then, hundreds of publications have documented thousands of QTLs for a wide variety of traits. The current resource to summarize the QTL identified in pigs is the Pig QTL database (PigQTLdb, http://www.animalgenome.org/cgi-bin/QTLdb/SS/index), which is an on-line accessible resource that archives all the porcine QTL published (Hu *et al.* 2013). The PigQTLdb includes information for 8,935 QTLs from 371 publications representing 644 different traits (until September 2013). There exist QTLs identified in all porcine chromosomes, being the biggest chromosome (SSC1) the one that contains more described QTLs (1,629 QTLs).

As is observed in the Figure 1.4, meat and carcass quality traits have been extensively studied in the last decades for their economically interest and their relevance in human health. Other groups of traits have also been studied, including health traits (blood parameters, immune capacity, disease susceptibility or pathogens), reproduction traits (litter size, reproductive organs, reproductive traits and endocrine), exterior traits (conformation, defects or coat colour) or production traits (growth, feed intake, feed conversion or digestive organs).

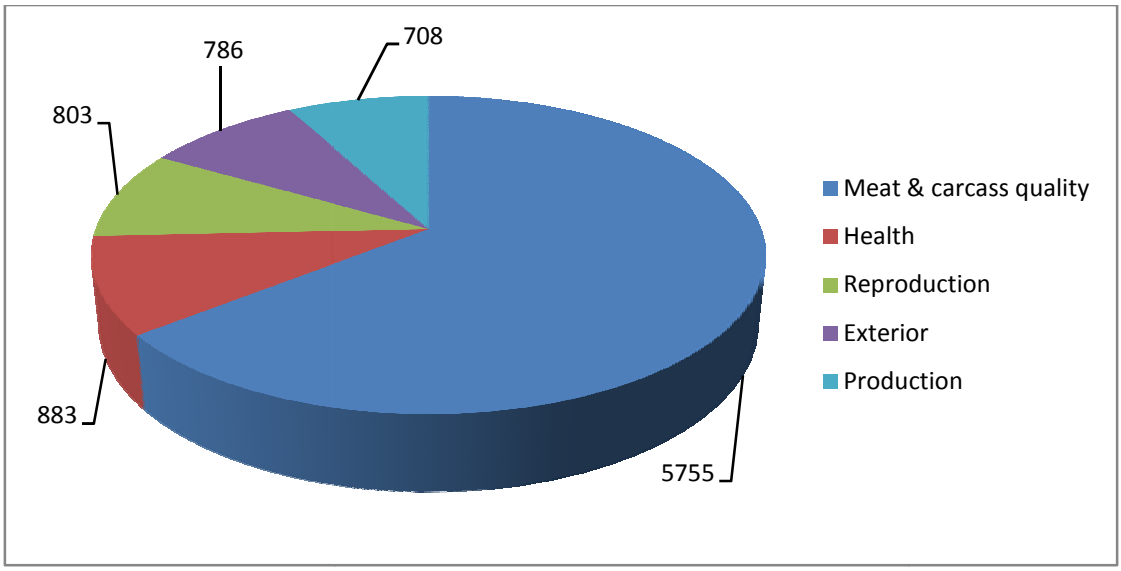

Figure 1.4-Distribution of porcine QTLs among the different trait classes. Source: PigQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index).

Focusing on meat and carcass quality traits, we can observe that the most important factors affecting meat quality explained above (see section 1.2) are present in the list of QTLs related with meat and carcass quality traits (Table 1.2). Despite that, the traits showing a higher number of QTLs are fatness, texture and anatomy. In this thesis, we focused on the study of fat composition, which is critical for meat quality traits.

Table 1.2- Overview of QTLs currently deposited in the PigQTLdb related with meat and carcass quality (as accessed in September 2013).

Trait type	Number of QTL
Fatness	1,461
Texture	1,366
Anatomy	1,361
Fat composition	423
Meat colour	323
pH	334
Chemical	208
Conductivity	151
Flavour	81
Enzyme activity	31
Odour	13
Stiffening	3

The first QTL study for fatty acid composition traits was performed by Clop and collaborators (2003) using the IBMAP cross (see section 1.6). Since this work, several groups have identified QTLs for fatty acid composition in pigs using microsatellite as molecular markers in different animal populations. Nii *et al*. (2006) found 25 significant effects for 17 backfat fatty acid composition traits at 13 chromosomal positions in a Japanese Wild Boar x Large White cross. SFA QTLs were mapped to porcine chromosomes SSC1, SSC9 and SSC15, while QTLs related with unsaturated fatty acids were detected on chromosomes SSC1, SSC4, SSC5, SSC9, SSC15 and SSC17 (Nii *et al.* 2006). A similar study was performed with a White Duroc x Erhualian intercross population, where two regions on SSC4 and SSC7 showed significant pleiotropic effects on MUFA and PUFA in *longissimus dorsi* muscle and abdominal fat (Guo *et al.* 2009). Additionally, two QTLs with significant multi-faceted effects on MUFA and PUFA in *longissimus dorsi* muscle were found on SSC8 and SSCX, respectively (Guo *et al.* 2009). In recent years, QTL identification efforts have greatly been enhanced with the development of a high-density SNP panel for pig genotyping (Ramos *et al.* 2009). Recently, this chip was used for trait-associated SNPs (TASs) identification using GWAS (Genome-Wide Association Studies) approaches. Association testing exploits population-wide linkage disequilibrium, as well as within-family linkage, and provides

more resolution to map QTLs compared with using only within-family linkage information (Ernst & Steibel 2013). Few GWAS has recently been reported for fatty acid composition in pork; the first one was performed last year within the IBMAP project (Ramayo-Caldas *et al.* 2012b), which is explained in the next section. More recently, GWAS for fatty acid composition in *longissimus dorsi* muscle and abdominal fat tissues of 591 White Duroc x Erhualian F_2 animals and muscle samples of 282 Chinese Sutai pigs were performed, and new knowledge of the complex genetic architecture of these traits was provided by the identification of 46 QTLs on 15 pig chromosomes for 12 FA (Yang *et al.* 2013).

1.5.2- Causal mutations and functional validation

Following the QTL identification, efforts are focused on the more difficult step of identifying the causal gene and/or the causal polymorphism. The identification of causal mutations is one of the greatest challenges in genetics for understanding how phenotypes are built. Difficulties on this step are caused by the large number of genes included in the QTL regions and also by the linkage disequilibrium generated in experimental populations (Varona *et al.* 2005). Once the candidate gene is selected, characterization, amplification and sequencing of the regulatory and coding regions of the gene are usually done to detect polymorphisms. Further association studies can be performed with these new polymorphisms and phenotypic data. Moreover, functional analysis of the gene can be achieved by gene expression studies with quantitative PCR (qPCR), obtaining valuable expression data for the analysis of association between a hypothetical mutation, gene expression values and/or phenotypic data. In the literature, several procedures for studying and validating hypothetical causal mutations have been proposed. The location of the mutation is crucial for determining which type of approach should be used. For polymorphisms located in the promoter region of the gene, the first efforts may be focused on *in silico* analyses based on literature and transcription factor (TF) database search. Following *in silico* analyses, DNA-TF binding approaches (Table 1.3) are very useful to investigate the effect of the promoter polymorphisms on TF binding (Galas & Schmitz 1978; Carey & Smale 2000;

Knight 2003). Furthermore, functional promoter approaches can be performed with transfection or transgenic expression assays, with the main goal of investigating the alteration on promoter activity caused by the mutation (Knight 2003; Nobrega *et al.* 2003; Maston *et al.* 2006). The most common assays performed for each approach are described in table 1.3.

Table 1.3- Assays involved in determining the effects of polymorphisms on complex traits.

SNP region	Approach	Aim	Assay
Promoter	DNA-TF binding analysis	Investigate protein binding to promoter in vitro	EMSA
			DNase I-footprinting
			ChIP
	Functional promoter analysis	Investigate if promoter mutation alters promoter activity in vivo	Transient transfection
			Stable transfection
			Transgenic expression
Coding region	Protein activity analysis	Investigate if a non-synonymous mutation alters protein activity	Several commercial kits are available for some proteins
3'UTR	Functional 3'UTR analysis	Investigate if 3'UTR mutations alters gene expression	mRNA stability
	DNA-microRNA binding analysis	Investigate if 3'UTR mutations alters a hypothetical regulation by microRNAs	Dual-luciferase reporter assay

Table 1.3 also summarizes different approaches to analyze the effect of polymorphisms located in the coding region and in the 3'UTR region. In general, the most relevant polymorphisms located in the coding region are those that cause an aminoacidic change: non-synonymous mutation. The analyses of these mutations are mainly based on determining the effect of the aminoacidic change on protein activity.

Nowadays, several commercial kits are available for measure protein activity: e.g., Estellé *et al*. (2009a). On the other hand, polymorphisms located in the 3'UTR regions are also clear candidate polymorphisms to be causal mutations, due to the implication of this region on gene expression (Lee *et al.* 2012). There are two main approaches based on functional studies for determining the 3'UTR mutation effect on: i) mRNA stability (Wang *et al.* 2006), and ii) gene expression regulation by microRNA binding (Clop *et al.* 2006). MicroRNAs are single-stranded non-coding RNAs involved in post-transcriptional regulation mechanisms acting mainly through down-regulation of target messenger RNAs (mRNAs) in a wide range of biological and pathological processes (Bartel 2009). The development of high-throughput sequencing techniques allows the identification of an important number of microRNAs (Li *et al.* 2011; Chen *et al.* 2012a; Li *et al.* 2012; Timoneda *et al.* 2012), generating valuable data for the analysis of putative 3'UTR causal mutations of complex traits such as meat quality. A clear example is the microRNAs miR-33a/b, which are located within *SREBP* gene sequence and have been associated with the regulation of FA metabolism and insulin signaling (Dávalos *et al.* 2011). Finally, despite the promoter, coding region and 3'UTR are the main regions where we expected to found the causal mutation, we should not discard other genomic regions. For instance, there are cases in which a causal mutation was detected in introns (Van Laere *et al.* 2003; Georges 2007).

Nowadays, the scientific community has focused their interest in epigenetics, as is demonstrated by the increasing number of publications on this topic. Epigenetic modifications play a role in the regulation of several physiological and pathological processes such as lipid related diseases: obesity, insulin resistance, diabetes or cardiovascular diseases (Ferrari *et al.* 2012). Chromatin modifications such as methylation, phosphorylation or acetylation are crucial to determine the genome accessibility to transcriptional machinery, which is essential in the regulation of gene expression. Additionally, some studies have demonstrated that genetic polymorphisms in promoter regions can influence an epigenetic state of the promoter (Boumber *et al.* 2008). Hence, for a complete characterization of a candidate gene is important to know its epigenetic situation and also if this situation can be related with a putative causal mutation.

Few studies have succeeded in identifying the causal mutations of relevant traits for the pork industry. One example is the non-synonymous mutation C1843T located in the *RYR1* (*Ryanodine receptor 1*) gene, which is related with the porcine stress syndrome in animals homozygous for the recessive allele. In addition, these animals have a reduced quality of meat, known as PSE from pale, soft and exhudative (Fujii *et al.* 1991). Other mutations with clear effects on meat quality traits have been identified in the following genes: *CAST* (*Calpastatin*) (Ciobanu *et al.* 2004), *IGF2* (*Insulin-like growth factor 2*) (Jeon *et al.* 1999), *MC4R* (*Melanocortin receptor 4*) (Kim *et al.* 2000) and *PRKAG3* (*Protein kinase, AMP-activated, gamma 3 non-catalytic subunit*) (Milan *et al.* 2000). These mutations have been used by pig breeding companies with the aim of including meat quality traits as an integral part of selection programs (van der Steen *et al.* 2005; Davoli & Braglia 2008).

1.6.- IBMAP Consortium

1.6.1- IBMAP cross

The classical approach for QTL detection in livestock is to perform an experimental cross between two divergent lines. In collaboration with INIA (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria) and IRTA (Insitut de Recerca i Tecnologia Agroalimentàries), our group created and Iberian x Landrace experimental cross (IBMAP cross). The Iberian pig is the native breed of the Iberian Peninsula and is characterized by an excellent meat quality, a high body-fat composition and a very low prolificacy (Serra *et al.* 1998). Specifically, the Guadyerbas line is one of the few original Iberian pig strains, that has remained isolated on an experimental farm in the Castilla-La Mancha since 1945 (Odriozola 1976; Béjar *et al.* 1992). The Guadyerbas pigs are voracious eaters with slow-growing, very low prolificity and high fat (Serra *et al.* 1998). In comparison with Iberian pigs (Guadyerbas line), the Landrace population has lower levels of fat, less IMF, higher content of PUFA and lower content of MUFA and SFA (Serra *et al.* 1998).

As shown in Figure 1.5, the IBMAP population was created by crossing 3 Iberian (Guadyerbas line) boars with 31 Landrace (Nova Genètica, Lleida, Spain) sows, generating 63 F_1 individuals (5 males and 58 females). These animals were used to create different generations, including F_2, F_3 and different Iberian backcrosses with Landrace, Pietrain and Duroc breeds. Analyses performed in this thesis are focused on the 144 animals from the Iberian backcross with Landrace (BC1_LD), which was created by crossing five F_1 boars with 26 Landrace sows.

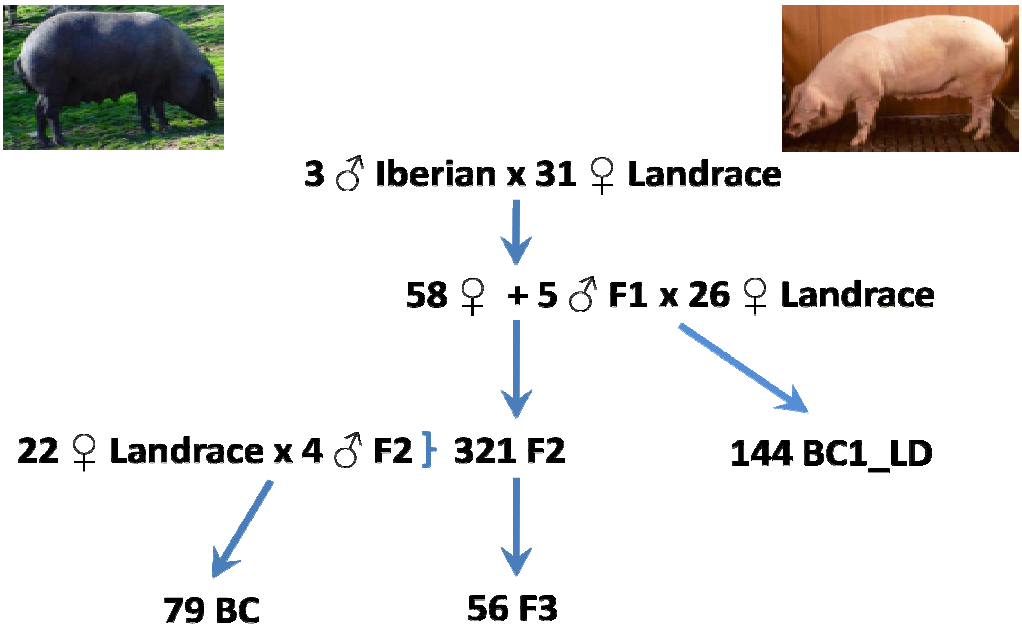

Figure 1.5-Schematic representation of the Iberian by Landrace cross (IBMAP).

1.6.2- QTLs identified in the IBMAP cross

The IBMAP project was created with the goal of detecting QTLs related with carcass and growth traits, meat quality traits and fatty acid composition traits. Initially, the first QTLs were identified by linkage analysis using microsatellites as molecular markers. The study of growth, carcass and fatness traits resulted in the identification of significant related QTLs in chromosomes SSC2, SSC4, SSC6 and SSCX (Óvilo *et al.* 2000; Pérez-Enciso *et al.* 2000; Varona *et al.* 2002; Mercadé *et al.* 2005a; Óvilo *et al.* 2005; Pérez-Enciso *et al.* 2005). On the other hand, QTLs related with meat quality traits were detected in chromosomes SSC3, SSC4, SSC6, SSC7, SSC8 and SSCX (Óvilo *et*

al. 2000; Óvilo *et al.* 2002; Pérez-Enciso *et al.* 2002; Mercadé *et al.* 2005a). Finally, several QTLs were identified in chromosomes SSC4, SSC6, SSC8, SSC10, SSC12 and SSCX, associated with fatty acid composition traits (Pérez-Enciso *et al.* 2000; Clop *et al.* 2003; Mercadé *et al.* 2006a; Muñoz *et al.* 2007). More recently, all animals of BC1_LD have been genotyped with the Illumina Porcine 60K SNP Beadchip, allowing the QTL identification by linkage analysis (Fernández *et al.* 2012) and GWAS (Ramayo-Caldas *et al.* 2012b). The following two sections are centred on QTLs for FA composition on SSC8 and SSCX, the chromosomes that were analysis in this work.

1.6.2.1- SSC8

The first report of a genome scan for QTLs affecting directly FA composition in pigs was performed by Clop and collaborators (2003) in the IBMAP cross. In this study, FA composition of adipose tissue from backfat samples of F_2 generation animals were analysed. Figure 1.6 highlights the detection of QTLs affecting palmitic, palmitoleic acid (C16:1n-7) and average chain lentgh (ACL) of FA in the porcine SSC8 (Clop *et al.* 2003).

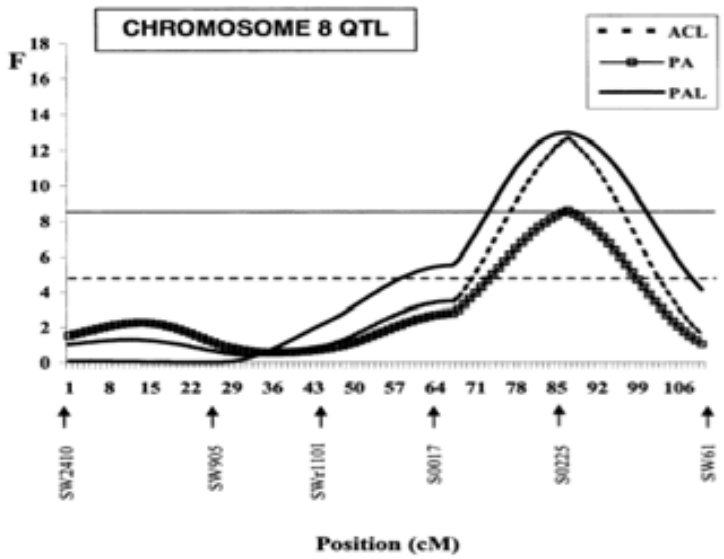

Figure 1.6- F-value profiles of the FA related QTLs in chromosome 8 identified in Clop *et al*. (2003). The horizontal solid line indicates the 95% genome-wise threshold and the dashed line indicates the 95% chromosome-wise threshold. ACL: average chain length; PA: palmitic acid; PAL: palmitoleic acid.

A recent GWAS study identified several chromosomal regions and positional candidate genes that may be associated with the profile of IMF FA composition of *longissimus dorsi* muscle in the IBMAP cross (Ramayo-Caldas *et al.* 2012b). Interestingly, this study confirmed SSC8 QTLs affecting palmitic and palmitoleic acids in IMF FA composition of BC1_LD animals. Additionally, a direct effect of these QTLs on palmitic and palmitoleic acids content was suggested in a Duroc x Erhualian cross (Yang *et al.* 2011) using a causal phenotype network. Combining the QTLs described above, a pleiotropic effect of palmitic and palmitoleic QTLs in IMF and backfat was suggested (Ramayo-Caldas *et al.* 2012b) and later confirmed in the BC1_LD animals (Muñoz *et al.* 2013). SSC8 QTLs affecting palmitic and palmitoleic acids were among the strongest signals obtained using GWAS in the IBMAP cross (Ramayo-Caldas *et al.* 2012b). Hence, an important part of this thesis was focused on analyzing the causes and consequences of this QTL.

Previous study in our group evaluated the *Microsomal triglyceride transfer protein* (*MTTP*) gene as a positional candidate gene for the FA QTLs of SSC8 (Estellé *et al.* 2009a). This study identified a non-synonymous polymorphism in a conserved residue of the lipid transfer domain of *MTTP*, which was associated with the palmitic and palmitoleic composition of porcine fat and the MTTP lipid transfer activity (Estellé *et al.* 2009a). Additionally, an interaction between the *MTTP* genotype and the type of fat source provided in the pig's diet was also described (Estellé *et al.* 2009a).

1.6.2.2- SSCX

Pérez-Enciso *et al.* (2002) described a flexible method for analyzing the X-linked QTL in crosses between outbred lines. The method was applied in the IBMAP F_2 population, detecting QTLs related with IMF and colour traits (Pérez-Enciso *et al.* 2002; Pérez-Enciso *et al.* 2005). An additional study, including FA composition data in the X-linked QTL analysis, revealed the presence of two QTLs affecting the percentages of oleic fatty acid and MUFA at position 73 cM, and the percentage of gadoleic acid at position 52 cM (Figure 1.7) (Mercadé *et al.* 2006a).

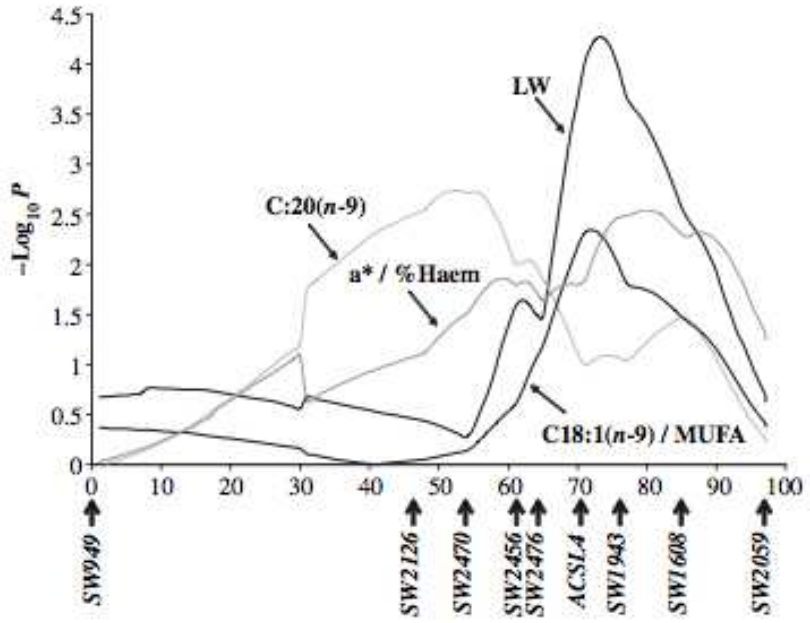

Figure 1.7- QTL profiles for meat quality and FA composition traits performed in the study of Mercadé *et al*. (2006a). LW, live weight; C20:1(n-9), gadoleic acid; C18:1(n-9), oleic acid; % Haem, pigment content; a*, minolta *a* colour component.

1.6.3- Candidate genes analyzed in the IBMAP cross and in this work

Since the beginning of the IBMAP project, numerous candidate genes have been analyzed (Table 1.4). These genes were related with the most interesting QTLs affecting growth, carcass, fat metabolism and meat quality, which were located in chromosomes SSC2, SSC4, SSC6, SSC8, SSC12 and SSCX (Table 1.4).

In this thesis, the analysis of putative causal factor of a QTL was performed by the positional candidate gene approach. The availability of the porcine genome allowed us to use this approach, in which the selection of candidate genes is done directly by sequence exploring. Therefore, this approach is based on the identification of genes biologically involved with the trait of interest in the QTL region, followed by molecular and association studies analyses of these genes. The following sections describe the

positional candidate genes analyzed in this work and their physiological relationship with the QTLs described in porcine SSC8 and SSCX of the IBMAP cross.

Table 1.4- Principal candidate genes analyzed in the IBMAP cross (genes analyzed in this work are shown in bold characters).

Chr	QTL associated traits	Candidate genes	References
SSC2	Growth and Fatness	*IGF2*	(Estellé *et al.* 2005a)
SSC4	Growth, Fatness and Form	*DECR*	(Clop *et al.* 2002)
		DGAT1	(Mercadé *et al.* 2005b)
		FABP4	(Mercadé *et al.* 2006b)
		FABP5	(Estellé *et al.* 2006)
SSC6	Fatness and IMF	*LEPR*	(Óvilo *et al.* 2005)
		FABP3	(Ovilo *et al.* 2002)
		ACADM	(Kim *et al.* 2006)
SSC8	FA composition	*CDS1*	(Mercadé *et al.* 2007)
		FABP2	(Estellé *et al.* 2009b)
		MTTP	(Estellé *et al.* 2005b) (Estellé *et al.* 2009a)
		ELOVL6	(Corominas *et al.* 2013b)
SSC12	FA composition	*FASN* *GIP* *ACACA*	(Muñoz *et al.* 2007)
SSCX	FA composition, Growth, Fatness and IMF	***ACSL4***	(Mercadé *et al.* 2006a) (Corominas *et al.* 2012)

1.6.3.1- *ELOVL6*

In mammals, the first and rate-limiting reaction of the elongation cycle is catalyzed by different condensing enzymes, which are members of *elongation of very long-chain*

fatty acids (*ELOVLs*) family (Moon *et al.* 2001; Jakobsson *et al.* 2006; Denic & Weissman 2007). Denic and Weissman (2007) showed that the catalysis of the condensation reaction takes place at the cytosolic face of the endoplasmic reticulum membrane. Studies performed in rodents and human elongases demonstrate that all ELOVL proteins contain several conserved amino acid consensus sequences (Leonard *et al.* 2004) such as the HxxHH motif, which is a prerequisite for the formation of 3-ketoacyl-CoA (Denic & Weissman 2007). In mammals, *ELOVLs* family consists of at least seven members, being the *ELOVL1*, *ELOVL5* and *ELOVL6* ubiquitously expressed, while *ELOVL2*, *ELOVL3*, *ELOVL4* and *ELOVL7* display a more distinct tissue-specific expression pattern. Moreover, ELOVL enzymes could also be classified by their substrate preferences for FA of different lengths and degrees of unsaturation: ELOVL1, ELOVL3, ELOVL6 and ELOVL7 metabolize SFA and MUFA; and ELOVL2, ELOVL4 and ELOVL5 are selective for PUFA (Figure 1.8) (Leonard *et al.* 2000; Wang *et al.* 2005; Kitazawa *et al.* 2009; Guillou *et al.* 2010).

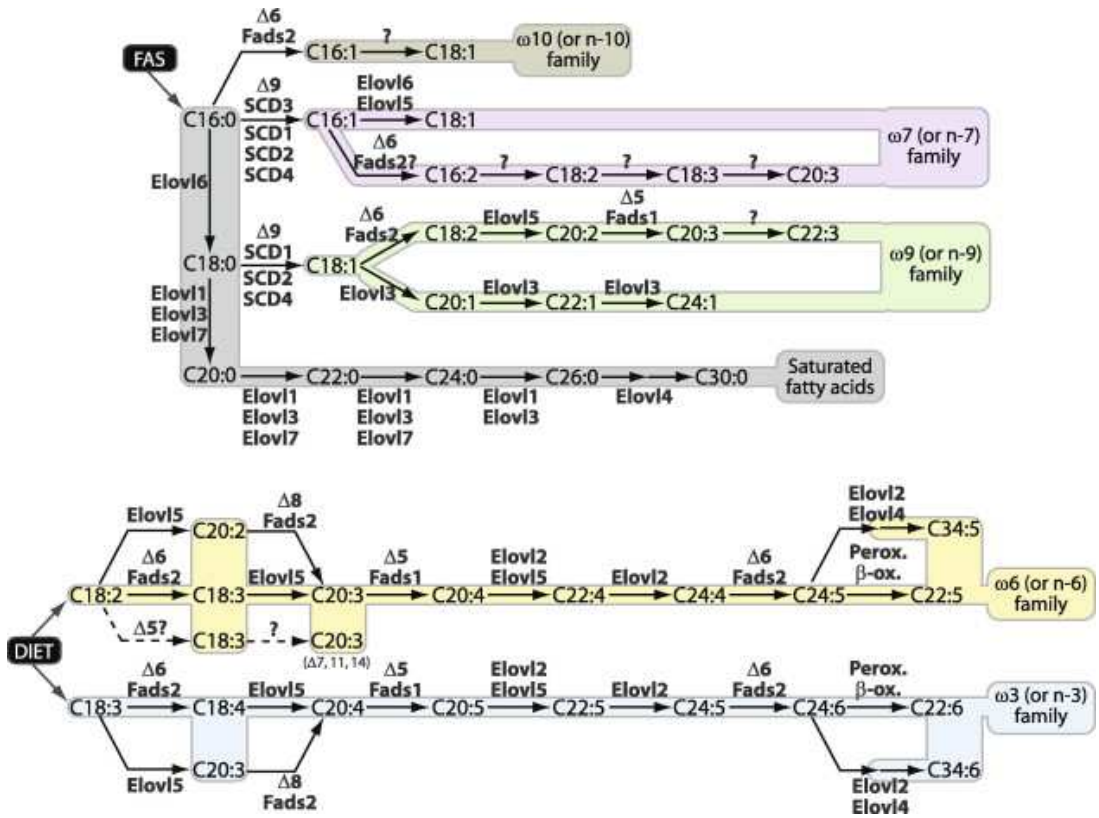

Figure 1.8- Long chain and very long-chain fatty acid biosynthesis in mammals. Long-chain FA and unsaturated FA can be synthesized from palmitic acid produced by FAS. On the other side, very long-chain FA of n-3 and n-6 families can only be synthesized from precursors obtained from the diet (Guillou *et al.* 2010).

The *ELOVL6* is the member of the family involved in the elongation of long-chain SFA and MUFA with 12-16 carbons, but it does not possess activity beyond 18 carbons (Jakobsson *et al.* 2006). Especially important is the elongation of the end product of FAS, palmitic acid and its desaturated product palmitoleic acid to yield stearic acid (C18:0) and vaccenic acid (C18:1n-7). Since the first identification of *ELOVL6* gene, its transcriptional regulation has been highly associated with sterol regulatory element binding transcription factors (SREBFs) (Moon *et al.* 2001; Matsuzaka *et al.* 2002; Kumadaki *et al.* 2008). Nevertheless, it is also well described that other factors such as circadian clock (Guillaumond *et al.* 2010) or dietary fatty acids (Schmitz & Ecker 2008) affect *ELOVL6* gene expression. In addition, studies performed with *ELOVL6*$^{-/-}$ mice demonstrated that this gene plays a critical role in the development of obesity-induced insulin resistance by modifying FA composition (Matsuzaka *et al.* 2007). In addition, *ELOVL6*$^{-/-}$ mice exhibited partial embryonic lethality, emphasizing the importance of endogenous C18 FA synthesis (Matsuzaka *et al.* 2007). Genetic variations in the *ELOVL6* gene have also been associated with insulin sensitivity in a human population (Morcillo *et al.* 2011). In humans, a major activity of ELOVL6 has been associated with the increase in the levels of n-7 and n-9 fatty acids, particularly palmitoleic acid and oleic acid, produced in cystic fibrosis (Thomsen *et al.* 2011).

The human *ELOVL6* gene presents two transcript variants differing in the number of exons and the 3'UTR length. Transcript variant 1 consists of 5 exons and has a total length of 6,454 bp (Ensembl transcript ID: ENST00000394607), while transcript variant 2 consists of 4 exons and has a total length of 3,090 bp (Ensembl transcript ID: ENST00000302274). The complete sequence of porcine *ELOVL6* gene and its correspondent mRNA is not known in the current databases. The porcine genome version Sscrofa 10.2 includes the *ELOVL6* gene in SSC8 within the confidence interval of the QTL affecting FA composition in the IBMAP cross (Clop *et al.* 2003; Ramayo-Caldas *et al.* 2012b).

1.6.3.2- *ACSL4*

The *long-chain acyl-CoA synthetase* (*ACSL*) family is comprised of several enzymes that have a key role in FA metabolism in mammals. Once FA have entered inside the cells, they are transformed into acyl-CoA esters to be used in different biochemical processes: synthesis of cellular lipids and degradation of fatty acid via β-oxidation (Suzuki *et al.* 1990; Mashek *et al.* 2004; Mashek *et al.* 2006; Cooke *et al.* 2011). In this reaction, ACSL ligates FA to CoA in a two-step reaction: 1) FA + ATP → fatty acyl-AMP + pyrophosphate; 2) fatty acyl-AMP + CoA→ fatty acyl-CoA + AMP (Suzuki *et al.* 1990; Mashek *et al.* 2004). Fatty acyl-CoA synthesized is used not only for anabolic and catabolic lipid processes, but also for the regulation of transmembranal transport, enzymatic activation and gene regulation (Piccini *et al.* 1998). The first ACSL member was originally described in 1953 (Kornberg & Pricer 1953), but it was not until 1990 when Suzuki cloned the cDNA encoding the ACSL (Suzuki *et al.* 1990). Until now, five genes and several isoforms of ACSL differing in their substrate preferences, enzyme kinetics, subcellular location and regulation have been identified and characterized (Mashek *et al.* 2006; Mercadé *et al.* 2006a; Cooke *et al.* 2011). Members of *ACSL* family have been classified in two groups: i) ACSL1, ACSL5 and ACSL6, which present a 60% of aminoacidic similarity between proteins and ii) *ACSL3* and *ACSL4*, with 30% of similarity (Van Horn *et al.* 2005).

The *ACSL4* gene encodes a protein that preferentially utilizes arachidonic acid and eicosapentaenoic acid (EPA) as substrates. This substrate preferences place ACSL4 as a potential enzyme in controlling the level of free arachidonic acid, which is a key regulatory point in the generation of lipid mediators derived from arachidonic acid, such as prostaglandins and leukotrienes (Cao *et al.* 1998). Despite ACSL enzymes act in both anabolic and catabolic lipid processes, some studies suggested that each member can act preferentially in one process. The location of *ACSL4* gene in the peroxisome and mitochondria suggests that acyl-CoA generated by ACSL4 is mainly used for β-oxidation (Coleman *et al.* 2002). Nevertheless, under situation of energy excess, ACSL4 also allocates acyl-CoA for lipid synthesis (Lewin *et al.* 2002). Polymorphisms and alterations in gene expression have been associated with neurological disorders (Covault *et al.* 2004; Kantojärvi *et al.* 2011; Modi *et al.* 2013), mental retardation

(Piccini *et al.* 1998; Meloni *et al.* 2002; Verot *et al.* 2003; Gazou *et al.* 2012) or cancers (Cao *et al.* 2001; Liang *et al.* 2005; Sung *et al.* 2007).

In pigs, this gene is located in SSCX (Mercadé *et al.* 2005a; Čepica *et al.* 2006), within or proximal to the IMF content QTL (Harlizius *et al.* 2000; Pérez-Enciso *et al.* 2002; Ma *et al.* 2009) and within QTLs affecting growth (Pérez-Enciso *et al.* 2005), live weight at slaughter and MUFA and oleic acid content (Mercadé *et al.* 2006a). Several groups have selected *ACSL4* gene as a candidate gene for these QTLs (Mercadé *et al.* 2006a; Čepica *et al.* 2007; Duthie *et al.* 2009; Ma *et al.* 2013). In the IBMAP cross study, an association between a polymorphisms in the 3'UTR of the porcine *ACSL4* gene and the percentages of oleic acids and MUFA was observed (Mercadé *et al.* 2006a). The observed effects of the *ACSL4* polymorphisms are in agreement with the phenotypic differences between Iberian and Landrace breeds. Iberian animals have a major content of oleic acid and MUFA (Serra *et al.* 1998).

1.7- Pig as a model animal for human diseases

Different animal species have been used to exploit their physiological uniqueness in addressing biomedical research issues, but in most cases biological differences between animal species and humans make impossible the experimental design. In general, rodents are the most common animals used as biological model, what is a good approximation in the first steps of biomedical research. Nevertheless, despite the important similarities between rodents and human, there are several barriers that must be exceed for applying data obtained with rodent models to the physiological/clinical settings in humans. At this point, models of large mammals are necessary to act as intermediary to humans.

The pig presents important similarities with humans in body size and other physiological/anatomical features, including organ development, disease progression and their innate tendency to overconsume food (Bergen & Mersmann 2005). For all these reasons, pigs are often mentioned as the preferred animal species for organ xenotransplantation (Cooper 2012) and also are an excellent model for studying

humans: since diseases such as cancer, diabetes, atherosclerosis or cardiovascular disease to lifestyles issues like nutrition or stress (Xi *et al.* 2004; Schook *et al.* 2005b; Lunney 2007). Interestingly, naturally occurring mutations in different animal populations may produce individuals with relevant characteristics for some disease studies (Groenen *et al.* 2012). For example, the MeLiM (Melanoblastoma-bearing Libechov minipig) are characterized by the appearance of tumors *in utero* or during the first 3 months of life, followed by the tumor progress until the onset of a spontaneous and complete tumor regression (Vincent-Naulleau *et al.* 2004; Bourneuf *et al.* 2011). This spontaneous regression converts this pig population in an optimal model for studying genetic events controlling melanoma development and regression. On the other hand, obesity, diabetes and cardiovascular diseases in humans are associated with excess energy consumption (chronic and high fat-diet). In these cases, the disease symptoms can be obtained by feeding animals with a high-fat cholesterol diet resulting in excellent models for disease progression and pathogenesis (Xi *et al.* 2004; Torres-Rovira *et al.* 2011).

2. OBJECTIVES

OBJECTIVES

This PhD thesis was done under the framework of the IBMAP Project funded by the projects AGL2008-04818-C03/GAN (MICINN) and AGL2011-29821-C02 (MINECO). The general goal of this thesis was to increase the knowledge of the genetic basis of intramuscular FA composition traits in pigs. To achieve this goal, we have utilized the Iberian x Landrace cross (IBMAP cross) created by the collaboration between INIA, IRTA and UAB.

More specifically, the objectives were:

1. To fine-map the QTL related with the percentages of palmitic and palmitoleic fatty acids in backfat and muscle on porcine chromosome 8.

2. To characterize the porcine *ELOVL6* gene as a candidate gene for fatty acid composition QTL on chromosome 8 by the identification of polymorphisms and the study of the polymorphisms effect on the analyzed traits.

3. To study the functional implication of an *ACSL4* genetic variant on its own gene expression and the QTL on porcine chromosome X affecting the percentages of oleic acid and MUFA.

4. To characterize the transcriptome architecture of porcine adipose tissue and to identify genes and pathways differentially-expressed in the backfat of animals with extreme phenotypes for intramuscular FA composition.

3. PAPERS AND STUDIES

PAPER I

Evaluation of the porcine *ACSL4* gene as a candidate gene for meat quality traits in pigs

J. Corominas, Y. Ramayo-Caldas, A. Castelló, M. Muñoz, N. Ibáñez-Escriche, J. M. Folch and M. Ballester

Animal Genetics 2012, Volume 43, Issue 6, pages 714–720. DOI: 10.1111/j.1365-2052.2012.02335.x. Article first published online: 11 Mar 2012

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/j.1365-2052.2012.02335.x

Evaluation of the porcine *ACSL4* gene as a candidate gene for meat quality traits in pigs

J. Corominas*,††, Y. Ramayo-Caldas*,††, A. Castelló*,††, M. Muñoz†, N. Ibáñez-Escriche‡, J. M. Folch*,†† and M. Ballester§,††

*Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, UAB, Bellaterra, 08193, Spain. †Departamento de Mejora Genética Animal, INIA, Ctra. De la Coruña km. 7, Madrid, 28040, Spain. ‡Genètica i Millora Animal, IRTA Lleida, Lleida, 25198, Spain. §Centre de Recerca en Sanitat Animal (CReSA), UAB-IRTA, Campus de la UAB, Bellaterra, 08193, Barcelona, Spain

Summary

Long-chain acyl-CoA synthetase (ACSL) family members catalyse the formation of long-chain acyl-CoA from fatty acid, ATP and CoA, thus playing an important role in both *de novo* lipid synthesis and fatty acid catabolism. Previous studies in our group evaluated *ACSL4* as a positional candidate gene for quantitative trait loci located on chromosome X in an Iberian × Landrace cross. A *DQ144454:c.2645G>A* SNP located in the 3′ untranslated region of the *ACSL4* gene was associated with the percentages of oleic and monounsaturated fatty acids. The aim of the present work was to evaluate the functional implication of this genetic variant. An expression analysis was performed for 120 individuals with different genotypes for the *DQ144454:c.2645G>A* polymorphism using real-time quantitative PCR. Differences between genotypes were identified in liver, with the *ACSL4* mRNA expression levels higher in animals with the G allele than in animals with the A allele. A SNP genome-wide association study with *ACSL4* relative expression levels showed significant positions on chromosomes 6 and 12. Description of positional candidate genes for *ACSL4* regulation on chromosomes 6 and 12 is provided.

Keywords *ACSL4*, fatty acid metabolism, genome-wide association study, *β*-oxidation.

Introduction

Long-chain acyl-CoA synthetases (ACSLs) are a family of enzymes, widely distributed in all eukaryotes, that play a key role in the synthesis of cellular lipids and degradation of fatty acids via *β*-oxidation (Singh *et al.* 1992). In mammals, ACSLs are essential for fatty acid degradation, phospholipid remodelling and the catalysis of long-chain acyl-CoA synthesis from fatty acid, ATP and CoA (Suzuki *et al.* 1990; Schneiter & Kohlwein 1997; Van Horn *et al.* 2005). In addition, long-chain acyl-CoAs are also involved in cell signal transduction by regulating membrane trafficking, ion fluxes, protein kinase C and gene expression (Piccini *et al.* 1998).

Since the first cDNA encoding ACSL was isolated from rat liver in 1990 (Suzuki *et al.* 1990), five genes and several isoforms of ACSL differing in their substrate preferences, enzyme kinetics, subcellular location and regulation have been identified and characterized. Until now, several spliced isoforms of the *ACSL4* gene, with a different expression pattern, have been characterized in mouse, rat, human and pig (Mercadé *et al.* 2006; Soupene & Kuypers 2008). The *ACSL4* gene encodes a protein that preferentially utilizes arachidonic acid and eicosapentaenoic acid as substrates. This substrate preference makes ACSL4 an important enzyme in controlling the level of free arachidonic acid, which is a key regulatory point in the generation of lipid mediators derived from this acid, such as prostaglandins and leukotrienes (Cao *et al.* 1998).

The porcine *ACSL4* gene was mapped on porcine chromosome X (SSCX; Mercade *et al.* 2005; Cepica *et al.* 2006). It has been described as a positional candidate gene in several quantitative trait locus (QTL) studies (Mercadé *et al.* 2006; Cepica *et al.* 2007; Duthie *et al.* 2009). Interestingly, the *ACSL4* gene has been located within or proximal to intramuscular fat QTL (Harlizius *et al.* 2000; Pérez-Enciso *et al.* 2002; Ma *et al.* 2009) and within a QTL affecting growth detected in an Iberian × Landrace F_2 intercross (IBMAP cross; Pérez-Enciso *et al.* 2005). More recently, new QTL affecting live weight at slaughter

Address for correspondence

J. Corominas, Consorci CSIC-IRTA-UAB (Centre de Recerca en Agrigenòmica), Edifici CRAG, Campus UAB, Bellaterra, Spain.
E-mail: Jordi.corominas@uab.cat

Current address: ††Consorci CSIC-IRTA-UAB (Centre de Recerca en Agrigenòmica), Edifici CRAG, Campus UAB, Bellaterra, 08193, Barcelona, Spain.

Accepted for publication 14 November 2011

and the percentages of oleic fatty acid and monounsaturated fatty acids (MUFA) were identified in the same genomic region. In this study, *ACSL4* was evaluated as a positional candidate gene for these QTL in the IBMAP cross (Mercadé *et al.* 2006). Sequence analysis revealed 10 polymorphisms in the 3′UTR region of the *ACSL4* gene segregating in only two haplotypes in the cross. The *DQ144454:c.2645G>A* polymorphism was genotyped in animals of the IBMAP cross, and QTL and association analysis showed that the SNP was more significantly associated with the percentages of oleic fatty acid and MUFA (Mercadé *et al.* 2006), reinforcing the interest of *ACSL4* as a positional candidate gene for fatty acid metabolism.

In the present study, we analysed the effect of the *DQ144454:c.2645G>A* polymorphism on *ACSL4* gene expression in liver and adipose tissue by real-time quantitative PCR (RT-qPCR). In addition, a genome-wide association study (GWAS) with genotypes from a subset of 54 988 SNPs allowed for the identification of genomic regions related to *ACSL4* gene expression.

Materials and methods

Animal samples

Animals used in this study belong to the IBMAP cross, a population generated by crossing three Iberian (Guadyerbas line) boars with 31 Landrace sows (Peréz-Enciso *et al.* 2000) and containing several generations and backcrosses. The *ACSL4* gene expression analysis was carried out in animals from a backcross (BC1_LD) generated by crossing five F1 (Iberian × Landrace) boars with 26 Landrace sows and producing 162 backcrossed animals. In this backcross material, only X chromosomes coming from Landrace (F0 and F1 sows) were segregating. At slaughter, samples of liver and adipose tissue (backfat) were collected, snap-frozen in liquid nitrogen and stored at −80 °C until analysed. Genomic DNA was obtained from blood samples of all animals. Genotyping of the *ACSL4 DQ144454:c.2645G>A* polymorphism in the BC1_LD animals was performed by a previously described pyrosequencing protocol (Mercadé *et al.* 2006).

Amplification and sequencing of the pig *ACSL4* promoter

Parental animals of the IBMAP cross (three Iberian boars and 10 Landrace sows) were used in this study. Two overlapping fragments of 600 and 584 bp of the proximal *ACSL4* gene promoter region were amplified by PCR. Primers (Table S1) were designed using the software PRIMER3 (Rozen & Skaletsky 2000) and the *ACSL4* gene sequence (ENSSSCG00000012583) available at the Sscrofa9 database of Ensembl (http://www.ensembl.org). The different designs were validated using the software PRIMER EXPRESS™ (Applied Biosystems).

PCRs were carried out in a total volume of 25 μl containing 0.6 units of AmpliTaq Gold (Applied Biosystems), 1.5–2 mM $MgCl_2$ (depending on the primers; Table S1), 0.2 mM of each dNTP, 0.5 μM of each primer and 50 ng of genomic DNA. Thermocycling was carried out under the following conditions: 94 °C for 10 min, 35 cycles of 94 °C for 1 min, 58–63 °C (depending on the primers; Table S1) for 1 min and 72 °C for 1 min, with a final extension of 72 °C for 5 min.

PCR products were purified using the ExoSAP-IT® method and sequenced with a Big Dye Terminator v.3.1 Cycle Sequencing Kit in an ABI 3730 analyser (Applied Biosystems).

To characterize the *ACSL4* promoter, a computer-assisted identification of putative promoter/enhancer elements was performed using the MATINSPECTOR application (set at a cut-off score >85%; Cartharius *et al.* 2005), a part of GENOMATIXSUITE software (Genomatix Software GmbH). Genomatix Matrix Library 8.3 was used with a core similarity threshold of 0.75 and a matrix similarity threshold of optimal −0.02. CpG island identification was performed with the EMBOSS CPGPLOT software (http://www.ebi.ac.uk/Tools/emboss/cpgplot) with the following parameters: observed/expected ratio > 0.80, per cent C + per cent G > 60.00, and length > 100.

Gene expression quantification

A total of 120 animals of the BC1_LD backcross with different genotypes for the *DQ144454:c.2645G>A* polymorphism were selected (Table 1). Total RNA was obtained from 120 liver samples and 47 backfat samples using the RiboPure™ Isolation of High Quality Total RNA (Ambion®), following the manufacturer's recommendations. RNA was quantified using the NanoDrop ND-1000 spectrophotometer (NanoDrop products) and checked for purity and integrity in a Bioanalyzer-2100 (Agilent Technologies). The isolated RNA was reverse-transcribed into cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) and random hexamers in a total volume of 20 μl containing 1 μg of total RNA, following the manufacturer's instructions.

PCR primers were designed using PRIMER EXPRESS™ software (Applied Biosystems) and are shown in Table S1. Primers for amplification of ACSL4 mRNA were designed

Table 1 Number of animals for each *DQ144454:c.2645G>A* genotype used in the *ACSL4* gene expression study. The number of animals in the initial liver study is indicated in brackets.

	ACSL4 DQ144454:c.2645G>A polymorphism			
Genotype	X^AX^A/X^AY	X^AX^G	X^GX^G/X^GY	Total
Liver	15 (13)	35 (11)	70 (23)	120 (47)
Backfat	13	11	23	47

from the available sequence (Mercadé *et al.* 2006) covering exons 8–9 to amplify a 124-bp-long fragment. Three genes frequently used as references in RT-qPCR experiments were analysed as endogenous controls: *β-2 microglobulin* (*B2M*), *hypoxanthine phosphoribosyltransferase 1* (*HPRT1*) and *glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*) (Erkens *et al.* 2006; Nygard *et al.* 2007). All reference genes were tested using the software GENORM (Overgard *et al.* 2010), and the best endogenous control for both tissues was *HPRT1.* PCR amplifications were performed in triplicate in a 20-μl final volume containing 2 μl of cDNA sample diluted 1:20 or 1:5 in DEPC-treated H_2O from liver or backfat samples respectively. For *HPRT1* amplification, FastStart Universal SYBR Green Master (Rox; Roche Applied Science) was used. SYBR Green PCR Core reagents (Applied Biosystems) with 2.5 mM $MgCl_2$ were used for *ACSL4* amplification. Primers were used at 900 nM each in all cases. PCR amplification was run on an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) using 96-well optical plates under the following conditions: 10 min at 95 °C, 40 cycles of 15 s at 95 °C and 1 min at 60 °C. A dissociation curve was drawn for each primer pair to assess that there was no primer–dimer formation.

To quantify and normalize the relative quantification (RQ) data, the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen 2001) was used. The sample of lowest expression level was selected as calibrator.

Expression data were adjusted using QXPAK5.0 software (print residual option; Pérez-Enciso & Misztal 2004, 2011) with a mixed model that included sex and batch as fixed effects and the infinitesimal effect as a random effect. Thus, the expression data for the family information were adjusted. Comparison of the mean residual values between genotypes was made using a linear procedure of R software. Differences were considered statistically significant at $P < 0.05$.

Genotyping

A total of 162 animals of the BC1_LD backcross were genotyped with the Porcine SNP60 BeadChip (Illumina) using the Infinium HD Assay Ultra protocol (Illumina). Raw data had a high genotyping quality (call rate > 0.99) and was visualized and analysed with the GENOMESTUDIO software (Illumina). For subsequent data analysis, a subset of 54 988 SNPs was selected by removing the SNPs with a minor allele frequency <5%, those with missing genotypes > 0.05% and the duplicated SNPs in the Sscrofa10 assembly.

Genome-wide association analysis

Association analyses of RT-qPCR expression data of ACSL4 mRNA in liver and whole-genome SNP genotypes were carried out with QXPAK 5.0 software. The position of the SNPs was based on build 10 of the *Sus scrofa* whole-genome sequence (contributed by Martien Groenen; http://www.animalgenome.org/repository/pig/). Two different models were used, and both included the fixed effect of sex and batch:

$$\text{Additive model: } y_i = \text{Sex}_i + \text{Batch}_i + \lambda_i a_k + \mu_i + e_i$$

$$\text{Additive and dominant model: } y_i = \text{Sex}_i + \text{Batch}_i + \lambda_i a_k + \lambda_i d_k + \mu_i + e_i,$$

where y_i is the individual record of data expression, sex (two levels) and batch (five levels) are fixed effects, λ_i is a -1, 0, $+1$ indicator variable depending on the individual genotype for the kth SNP, a_k is the additive effect of each SNP, d_k is the dominant effect of each SNP and μ_i represents the infinitesimal genetic effect distributed as N (0, $\mathbf{A}\sigma_u$), where $\mathbf{A}$ is a numerator of kinship matrix and e_i is the residual. The inclusion in the model of the infinitesimal effect allows us to adjust the data for the family information and, thus, correct the inter-chromosomal linkage disequilibrium effect. In this analysis, each SNP was tested individually to check the association. Chromosome X was analysed using the same models but including a dosage compensation parameter (Pérez-Enciso *et al.* 2002). The R package *q*-value (Storey & Tibshirani 2003) was used to calculate the false discovery rate-based *q*-value to measure the statistical significance at the genome-wide level for association studies. The cut-off of significant association at the whole-genome level was set at *q*-value ≤ 0.1. This significance threshold is likely too stringent because of the linkage association among SNP genotypes. Gene annotation for 2-Mb genomic intervals around the most significant SNPs was performed with BIOMART software in the Ensembl Sscrofa 9 data set (http://www.ensembl.org). The concordance between the position of the relevant genes and significant GWAS peaks was validated by using a BLAST search in the Ensembl Sscrofa 10 data set (http://www.animalgenome.org).

Results and discussion

Effect of the *ACSL4* gene polymorphism on its expression

The association of the *DQ144454:c.2645G>A ACSL4* polymorphism with the percentage of oleic fatty acid and MUFA (Mercadé *et al.* 2006) suggested a role of this mutation in the regulation of *ACSL4* gene expression and subsequently in the metabolism of fatty acids. To study this hypothesis, RT-qPCR analyses were performed on liver and adipose tissues, organs particularly important in the export and storage of lipids, of 47 BC1-LD animals. Although no differential expression was shown between *DQ144454:c.2645G>A* genotypes when backfat samples were analysed (Fig. 1a), different levels of expression

between *DQ144454:c.2645G>A* genotypes were obtained in liver samples ($P = 0.025$; Fig. 1b). No significant effect of sex on *ACSL4* expression was observed (data not shown); thus, dosage compensation was assumed and data from both sexes were joined for homozygous females and males (X^AX^A and X^AY; X^GX^G and X^GY). Even though a high variability in RQ data was shown in all genotype groups, animals with the X^AX^A/X^AY genotype showed a significant lower expression compared to animals with the X^GX^G/X^GY genotype. Such differential gene expression increased when the number of liver samples was enlarged to 120 animals, gaining in statistical significance (Fig. 1c; $P = 0.008$). Furthermore, a significant differential gene expression was also observed between animals with the

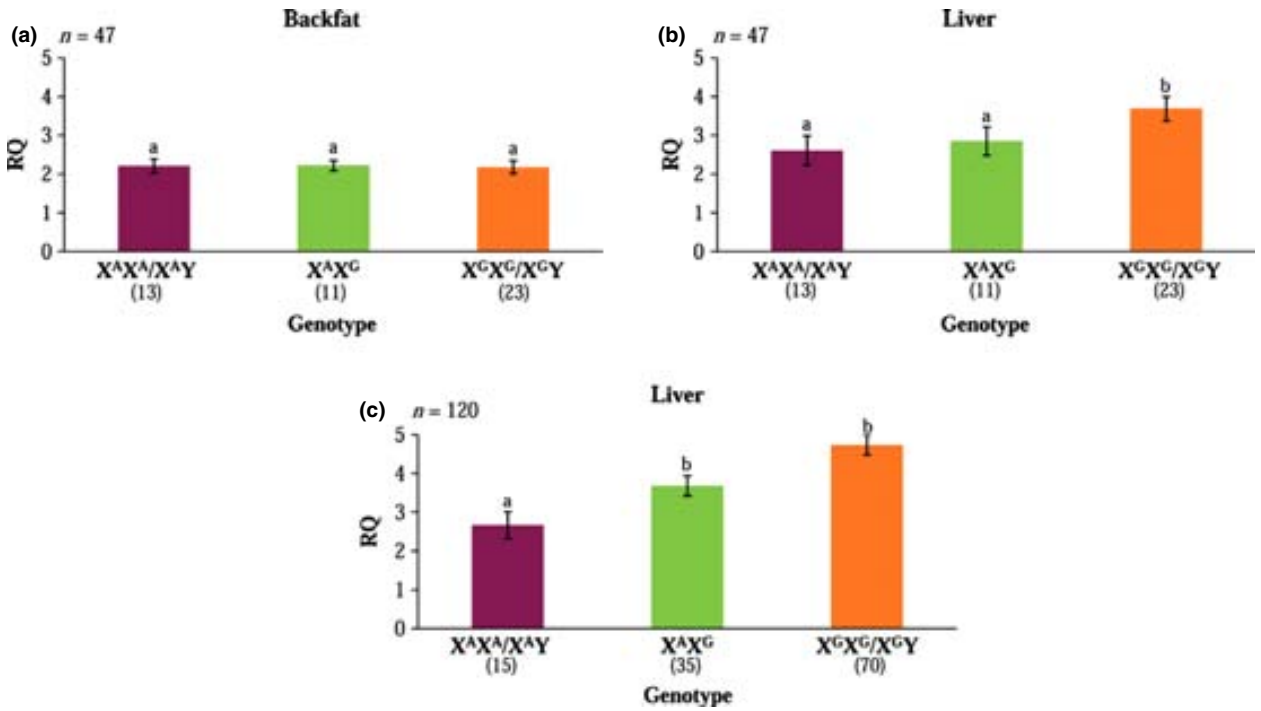

Figure 1 Relative quantification of mRNA levels of different genotypes of the *DQ144454:c.2645G>A* polymorphism, using a, backfat samples or b, c, liver samples. Results from liver are shown for the first 47 samples (b) and also for the total 120 samples (c). The number of animals analysed for each genotype is indicated in brackets. Data represent means ± SEM. Values with different superscript letters (a and b) indicate significant differences between groups ($P < 0.05$) as determined by the linear procedure of R software.

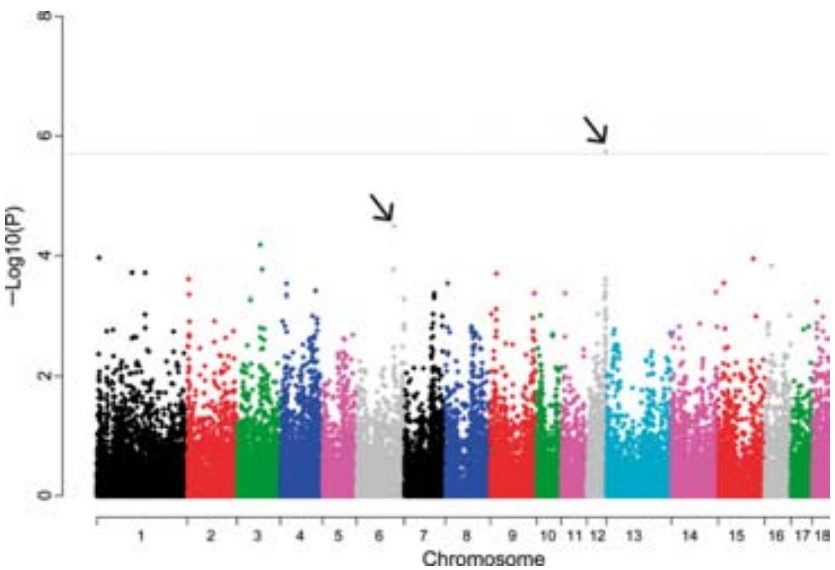

Figure 2 Association analysis between liver *ACSL4* expression levels and genotypes of autosomic SNPs. Results of the additive-dominant model are shown. Positions in Mb are relative to the *Sus scrofa* Assembly 10. The horizontal dashed line indicates the genome-wide significance level (FDR-based *q*-value ≤ 0.1). Arrows show the studied regions.

X^AX^A/X^AY and the heterozygous X^AX^G genotypes (Fig. 1c; $P = 0.008$).

These results support previous experiments in which differences in the regulation of *ACSL4* expression observed between rat tissues (Mashek *et al.* 2006) suggested that the mechanisms controlling *ACSL4* expression are different in liver and adipose tissues.

Characterization of *ACSL4* proximal promoter

Although different expression levels were obtained between *DQ144454:c.2645G>A* genotypes, high variabilty was observed within the genotyped groups, suggesting that other polymorphisms may explain these expression differences. One interesting region in which to find these putative polymorphisms was the *ACSL4* promoter. Hence, the amplification and sequencing of the proximal promoter region was performed for the *ACSL4* gene sequence (ENSSSCG00000012583) available at the Ensembl Sscrofa 9 database (http://www.ensembl.org) and assuming conservation with the human *ACSL4* promoter (AB061711) previously described by Minekura *et al.* (2001). A total of 801 bp of the *ACSL4* promoter were sequenced in two overlapping fragments of 600 and 584 bp. However, the comparative analysis of *ACSL4* promoter sequences did not reveal polymorphisms in animals from the IBMAP cross. Interestingly, the analysis with the EMBOSS CPGPLOT software predicted a 250-bp-long CpG island in the proximal promoter region of the *ACSL4* gene (at 88936576–88936825 bp of GenBank sequence NC_010461.2). This finding suggests that differences in the methylation pattern between liver and adipose tissue could explain the differential regulation of *ACSL4* expression between tissues. However, further functional studies are needed to characterize the degree of methylation in *ACSL4* promoter in different tissues.

Genome-wide association studies

To find new potential genomic regions associated with *ACSL4* gene expression, GWASs were carried out using the RQ expression data of the *ACSL4* gene in liver and the genotypes of 54 998 SNPs distributed across the pig genome and determined with the Porcine SNP60 BeadChip (Illumina). First, additive and additive-dominant models were applied for all the SNPs located in autosomes, and two candidate chromosomal regions, similar in both models, were identified in SSC6 and SSC12 (the additive-dominant plot is shown in Fig. 2). For SNPs located on chromosome X, a dosage compensation parameter was included in the models. Moreover, the *DQ144454:c.2645G>A ACSL4* polymorphism, which is not included in the porcine SNP60 BeadChip, was genotyped and incorporated into the SSCX SNP association analysis. The *ACSL4* polymorphism was the strongest

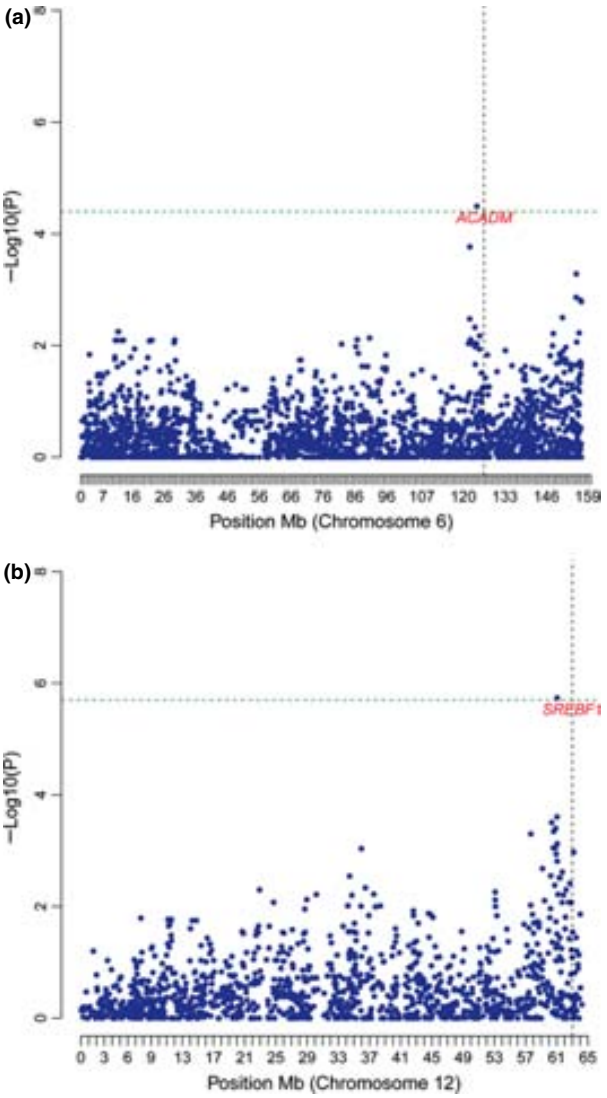

Figure 3 Association analysis between liver *ACSL4* expression level and SNP genotypes for (a) SSC6 and (b) SSC12. Results of the additive and dominant model are shown. Positions in Mb are relative to *Sus scrofa* Assembly 10. Vertical dashed lines indicate the location of positional candidate genes. Horizontal dashed lines mark the genome-wide significance level (FDR-based q-value ≤ 0.1).

signal on SSCX ($P = 0.00155$; Fig. S1), reinforcing the hypothesis of a cis-acting genetic variant modulating the *ACSL4* gene transcription. Conversely, the *ACSL4* polymorphism was less significant than other SNPs located in autosomes. This result suggests that genetic variants of other genomic regions are implicated in the regulation of *ACSL4* gene expression. Subsequently, additional plots were generated for individual chromosomes using only the additive-dominant model data, because there were no relevant differences with the additive model (data not shown). Chromosome-wise significant peaks were detected only on chromosomes SSC6 and SSC12 (Fig. 3). Gene annotation of these genomic regions was performed to

find polymorphic candidate genes regulating the *ACSL4* gene expression. However, the significance threshold was likely too stringent owing to the linkage dependence between the SNPs included in the analysis and, thus, other suggestive SNP peaks may also contain relevant genes. The most significant region associated with the RQ data was identified in the 59- to 63-Mb region of SSC12 (Fig. 3b). Gene annotation of this region showed an interesting gene near the significant SNP (*ASGA0100129*; $P = 0.00000184$): the *sterol regulatory element binding transcription factor 1* (*SREBF-1*), a family member of transcription factors that regulate lipid homoeostasis by promoting glycolysis, lipogenesis and adipogenesis (Horton *et al.* 2003). The mature protein translocates to the nucleus and activates transcription by binding to the sterol regulatory element-1 (SRE-1). Others have reported that *SREBF-1* and *peroxisome proliferator-activated receptor gamma* (*PPARG*) genes regulate ACSL4 in the liver of mice and rats (Horton *et al.* 2003; Mashek *et al.* 2006). Our results provide new evidence of this regulation because of the presence of SRE-1 in the *ACSL4* promoter (Fig. S2) and the localization of this gene near the significant SNPs in the GWAS analysis. However, our results did not allow us to validate the effect of the *PPARG* gene, which was mapped in assembly 10 but is located in SSC13, where no significant region was found. In SSC6, a significant region was found between positions 121 and 127 Mb (Fig. 3a; *ASGA0029488*; $P = 0.000032$). Gene annotation identified *medium-chain acyl coenzyme A dehydrogenase* (*ACADM*), which is a gene that plays a key role in the initial steps of fatty acid degradation via β-oxidation (Kelly *et al.* 1987), as well as *ACSL4*. *ACADM* is regulated by the *estrogen-related receptor alpha* (*ESRRA*) gene, which binds to an ERRα response element located in its promoter and contains a single consensus half-site 5′-TNAAGGTCA-3′ (Sladek *et al.* 1997; Luo *et al.* 2003). The porcine *ACSL4* gene promoter shares this cis-regulatory motif (Fig. S2), suggesting a possible co-regulation and co-expression of both genes. The *ESRRA* gene has been mapped to the SSC2 (position 4–9 Mb) in a non-significant SNP peak. However, this region was significant when we used a simpler statistical model without accounting for the infinitesimal genetic effect (data not shown).

In summary, the *DQ144454:c.2645G>A* polymorphism located in the 3′UTR of the *ACSL4* gene was associated with the expression of this gene in liver but not in adipose tissue. This result suggests the presence of a cis-acting regulatory element, coming from Landrace, on SSCX affecting the *ACSL4* gene expression in liver. In addition, stronger effects on *ACSL4* gene expression were found in other pig genomic regions. Different candidate genes that could act as trans-acting regulatory factors have been described, although further studies are necessary to evaluate their role in the *ACSL4* gene expression regulation and subsequently in fatty acid metabolism.

Acknowledgements

We would like to thank Dr. Jordi Estellé for help in the association studies for SSCX SNPs and Dr. Anna Mercadé for the genotyping of porcine chip. This study has been funded by MICINN project (AGL2008-04818-C03). J. Corominas was funded by a FPI PhD grant from the Spanish *Ministerio de Educación* (BES-2009-018223) and Y. Ramayo by a FPU PhD grant (AP2008-01450).

References

Cao Y., Traer E., Zimmerman G.A., McIntyre T.M. & Prescott S.M. (1998) Cloning, expression, and chromosomal localization of human *long-chain fatty acid-CoA ligase 4* (*FACL4*). *Genomics* **49**, 327–30.

Cartharius K., Frech K., Grote K., Klocke B., Haltmeier M., Klingenhoff A., Frisch M., Bayerlein M. & Werner T. (2005) MATINSPECTOR and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**, 2933–42.

Cepica S., Masopust M., Knoll A., Bartenschlager H., Yerle M., Rohrer G.A. & Geldermann H. (2006) Linkage and RH mapping of 10 genes to a QTL region for fatness and muscling traits on pig chromosome X. *Animal Genetics* **37**, 603–4.

Cepica S., Bartenschlager H. & Geldermann H. (2007) Mapping of QTL on chromosome X for fat deposition, muscling and growth traits in a wild boar x Meishan F2 family using a high-density gene map. *Animal Genetics* **38**, 634–8.

Duthie C.A., Simm G., Pérez-Enciso M., Doeschl-Wilson A., Kalm E., Knap P.W. & Roehe R. (2009) Genomic scan for quantitative trait loci of chemical and physical body composition and deposition on pig chromosome X including the pseudoautosomal region of males. *Genetics Selection Evolution* **41**, 27.

Erkens T., Van Poucke M., Vandesompele J., Goossens K., Van Zeveren A. & Peelman L.J. (2006) Development of a new set of reference genes for normalization of real-time RT-PCR data of porcine backfat and longissimus dorsi muscle, and evaluation with PPARGCIA. *BMC Biotechnology* **6**, 41.

Harlizius B., Rattink A.P., de Koning D.J., Faivre M., Joosten R.G., van Arendonk J.A. & Groenen M.A. (2000) The X chromosome harbors quantitative trait loci for backfat thickness and intramuscular fat content in pigs. *Mammalian Genome* **11**, 800–2.

Horton J.D., Shah N.A., Warrington J.A., Anderson N.N., Park S.W., Brown M.S. & Goldstein J.L. (2003) Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12027–32.

Kelly D.P., Kim J.J., Billadello J.J., Hainline B.E., Chu T.W. & Strauss A.W. (1987) Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and its expression in enzyme-deficient human tissue. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 4068–72.

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–8.

Luo J., Sladek R., Carrier J., Bader J.A., Richard D. & Giguère V. (2003) Reduced fat mass in mice lacking orphan nuclear recep-

tor *estrogen-related receptor alpha*. *Molecular and Cellular Biology* **23**, 7947–56.

Ma J., Ren J., Guo Y. *et al.* (2009) Genome-wide identification of quantitative trait loci for carcass composition and meat quality in a large-scale White Duroc x Chinese Erhualian resource population. *Animal Genetics* **40**, 637–47.

Mashek D.G.L., Li L.O. & Coleman R.A. (2006) Rat long-chain acyl-CoA synthetase mRNA, protein, and activity vary in tissue distribution and response to diet. *Journal of Lipid Research* **47**, 2004–10.

Mercadé A., Sanchez A. & Folch J.M. (2005) Assignment of the *acyl-CoA synthetase long-chain family member 4* (*ACSL4*) gene to porcine chromosome X. *Animal Genetics* **36**, 76.

Mercadé A., Estellé J., Pérez-Enciso M., Varona L., Silió L., Noguera J.L., Sánchez A. & Folch J.M. (2006) Characterization of the porcine *acyl-CoA synthetase long-chain 4* gene and its association with growth and meat quality traits. *Animal Genetics* **37**, 219–24.

Minekura H., Kang M.J., Inagaki Y., Cho Y.Y., Suzuki H., Fujino T. & Yamamoto T. (2001) Exon/intron organization and transcription units of the human *acyl-CoA synthetase 4* gene. *Biochemical and Biophysical Research Communications* **286**, 80–6.

Nygard A., Jorgensen C.B., Cirera S. & Fredholm M. (2007) Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Molecular Biology* **8**, 67.

Overgard A., Nerland A.H. & Patel S. (2010) Evaluation of potential reference genes for real time RT-PCR studies in Atlantic halibut (*Hippoglossus Hippoglossus L.*); during development, in tissues of healthy and NNV-injected fish, and in anterior kidney leucocytes. *BMC Molecular Biology* **11**, 36.

Pérez-Enciso M. & Misztal I. (2004) QXPAK: a versatile mixed model application for genetical genomics and QTL analyses. *Bioinformatics* **20**, 2792–8.

Pérez-Enciso M. & Misztal I. (2011) QXPAK.5: old mixed model solutions for new genomics problems. *BMC Bioinformatics* **12**, 202.

Peréz-Enciso M., Clop A., Noguera J.L. *et al.* (2000) A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *Journal of Animal Science* **78**, 2525–31.

Pérez-Enciso M., Clop A., Folch J.M., Sánchez A., Oliver M.A., Óvilo C., Barragán C., Varona L. & Noguera J.L. (2002) Exploring alternative models for sex-linked quantitative trait loci in outbred populations: application to an Iberian x Landrace pig intercross. *Genetics* **161**, 1625–32.

Pérez-Enciso M., Mercadé A., Bidanel J.P., Geldermann H., Cepica S., Bartenschlager H., Varona L., Milan D. & Folch J.M. (2005) Large-scale, multibreed, multitrait analyses of quantitative trait loci experiments: the case of porcine X chromosome. *Journal of Animal Science* **83**, 2289–96.

Piccini M., Vitelli F., Bruttini M., Pober B.R., Jonsson J.J., Villanova M., Zollo M., Borsani G., Ballabio A. & Renieri A. (1998) *FACL4*, a new gene encoding long-chain acyl-CoA synthetase 4, is deleted in a family with Alport syndrome, elliptocytosis and mental retardation. *Genomics* **47**, 350–8.

Rozen S. & Skaletsky H.J. (2000) PRIMER3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* **132**, 365–86.

Schneiter R. & Kohlwein S.D. (1997) Organelle structure, function, and inheritance in yeast: a role for fatty acid synthesis. *Cell* **88**, 431–4.

Singh I., Lazo O., Dhaunsi G.S. & Contreras M. (1992) Transport of fatty acids into human and rat peroxisomes. Differential transport of palmitic and lignoceric acids and its implication to X-adrenoleukodystrophy. *The Journal of Biological Chemistry* **267**, 13306–13.

Sladek R., Bader J. & Giuguère V. (1997) The orphan nuclear receptor *estrogen-related receptor* α is a transcriptional regulator of the human medium-chain *acyl coenzyme a dehydrogenase* gene. *Molecular and Cellular Biology* **17**, 5400–9.

Soupene E. & Kuypers F.A. (2008) Mammalian long-chain acyl-CoA synthetases. *Experimental Biology and Medicine (Maywood)* **233**, 507–21.

Storey J.D. & Tibshirani R. (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* **100**, 9440–5.

Suzuki H., Kawarabayasi Y., Kondo J., Abe T., Nishikawa K., Kimura S. & Hashimoto T. (1990) Structure and regulation of rat long-chain acyl-CoA synthetase. *The Journal of Biological Chemistry* **265**, 8681–5.

Van Horn C.G., Caviglia J.M., Li L.O., Wang S., Granger D.A. & Coleman R.A. (2005) Characterization of recombinant long-chain rat acyl-CoA synthetase isoforms 3 and 6: identification of a novel variant of isoform 6. *Biochemistry* **8**, 1635–42.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1. Association analysis of *ACSL4* expression level in liver with SNP genotypes for chromosome X including the *c.2645G>A ACSL4* polymorphism (labeled with a red circle).

Figure S2. Nucleotide sequences of the 5′-flanking region of porcine *ACSL4* gene.

Table S1. Primers for the *ACSL4* promoter sequencing (P) and RT-qPCR (RT) study.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

PAPER II

Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs

Jordi Corominas, Yuliaxis Ramayo-Caldas, Anna Puig-Oliveras, Dafne Pérez-Montarelo, Jose L. Noguera, Josep M. Folch, Maria Ballester

PLoS ONE 2013, 8(1): e53687. DOI:10.1371/journal.pone.0053687

OPEN ACCESS Freely available online

PLOS ONE

Polymorphism in the *ELOVL6* Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs

Jordi Corominas[1]*, Yuliaxis Ramayo-Caldas[1], Anna Puig-Oliveras[1], Dafne Pérez-Montarelo[2], Jose L. Noguera[3], Josep M. Folch[1,4], Maria Ballester[1]

1 Departament de genètica animal, Centre de Recerca en Agrigenòmica (CRAG), Bellaterra, Spain, **2** Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain, **3** Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Lleida, Spain, **4** Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain

Abstract

Background: The *ELOVL fatty acid elongase 6* (*ELOVL6*), the only elongase related to *de novo* lipogenesis, catalyzes the rate-limiting step in the elongation cycle by controlling the fatty acid balance in mammals. It is located on pig chromosome 8 (SSC8) in a region where a QTL affecting palmitic, and palmitoleic acid composition was previously detected, using an Iberian x Landrace intercross. The main goal of this work was to fine-map the QTL and to evaluate the *ELOVL6* gene as a positional candidate gene affecting the percentages of palmitic and palmitoleic fatty acids in pigs.

Methodology and Principal Findings: The combination of a haplotype-based approach and single-marker analysis allowed us to identify the main, associated interval for the QTL, in which the *ELOVL6* gene was identified and selected as a positional candidate gene. A polymorphism in the promoter region of *ELOVL6*, *ELOVL6:c.-533C>T*, was highly associated with the percentage of palmitic and palmitoleic acids in muscle and backfat. Significant differences in *ELOVL6* gene expression were observed in backfat when animals were classified by the *ELOVL6:c.-533C>T* genotype. Accordingly, animals carrying the allele associated with a decrease in *ELOVL6* gene expression presented an increase in C16:0 and C16:1(n-7) fatty acid content and a decrease of elongation activity ratios in muscle and backfat. Furthermore, a SNP genome-wide association study with *ELOVL6* relative expression levels in backfat showed the strongest effect on the SSC8 region in which the *ELOVL6* gene is located. Finally, different potential genomic regions associated with *ELOVL6* gene expression were also identified by GWAS in liver and muscle, suggesting a differential tissue regulation of the *ELOVL6* gene.

Conclusions and Significance: Our results suggest *ELOVL6* as a potential causal gene for the QTL analyzed and, subsequently, for controlling the overall balance of fatty acid composition in pigs.

Citation: Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, et al. (2013) Polymorphism in the *ELOVL6* Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. PLoS ONE 8(1): e53687. doi:10.1371/journal.pone.0053687

Editor: Stephen Moore, University of Queensland, Australia

Received August 3, 2012; **Accepted** December 3, 2012; **Published** January 14, 2013

Funding: This study has been funded by MICINN projects (AGL2008-04818-C03 and AGL2011-29821-CO2) and the Innovation Consolider-Ingenio 2012 Program (CSD2007-00036, Center for Research in Agrigenomics). J. Corominas was funded by a Formación de Personal Investigador (FPI) PhD grant from Spanish Ministerio de Educación (BES-2009-018223), Y. Ramayo by a Formación del Profesorado Universitario (FPU) PhD grant (AP2008-01450) and A. Puig was funded by a Personal Investigador en Formación (PIF) PhD grant from the Universitat Autónoma de Barcelona (458-01-1/2011). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jordi.corominas@uab.cat

Introduction

Food fatty acid (FA) composition is a critical aspect in human health and it is also relevant for meat quality. It determines important sensorial and technological aspects of meat due to the variability in the melting point of fatty acids. Thus, variation in fatty acids has an important effect on flavor, muscle color and firmness or softness of the fat in meat [1]. Meat fat is primarily composed of monounsaturated fatty acid (MUFA) and saturated fatty acid (SFA). Oleic acid is the most abundant and nutritionally relevant FA, followed by palmitic and stearic acids [2,3]. The highest rate of *de novo* synthesis of these FAs occurs in liver and adipose tissue, which converts the excess of glucose into FAs for storage and transport [4]. During *de novo* synthesis of FAs, palmitic acid (C16:0) produced by cytoplasmic acetyl-CoA carboxylase (ACC) and fatty acid synthase (FASN) is transferred to endoplasmic reticulum membranes, where FA elongase and desaturase enzymes catalyze the conversion of saturated FAs into monounsaturated FAs, such as palmitoleic acid (C16:1(n-7)) or oleic acid (C18:1(n-9)) [5,6]. Accordingly, FA elongase activity has an important role in regulating the synthesis of *de novo*-derived MUFAs and establishing the balance among C16:1(n-7), C18:1(n-7) and C18:1(n-9) [6].

In 2003, Clop et al. identified a QTL on porcine chromosome 8 (SSC8) with significant effects on C16:0 and C16:1(n-7) contents and a suggestive effect on C18:1(n-9) detected in backfat, using an Iberian x Landrace F_2 intercross (IBMAP) [7]. Previous studies in our group evaluated the *MTTP* gene as a positional candidate gene for this QTL fatty acid composition detected on SSC8 [8]. A mutation in the lipid transfer region of the *MTTP* protein (p.Phe840Leu) was associated with fatty acid composition of porcine fat and with the *MTTP* lipid transfer activity measured

with an *in vitro* assay. Furthermore, two QTL regions in 62 and 92 cM on SSC8, related with C16:0 and C16:1(n-7) fatty acid content in *Longissimus dorsi* muscle, respectively, were detected in a Chinese cross between Duroc and Erhualian [9]. More recently, a Genome-Wide Association Study (GWAS), performed on *Longissimus dorsi* muscle fatty acid composition from an Iberian x Landrace backcross population, detected this QTL between positions 92.1 Mb-96.7 Mb on SSC8 (according to Sscrofa 9.61 genome assembly) at 10 Mb from the *MTTP* gene [10]. This QTL was also identified using backfat fatty acid composition at positions 89 cM (C16:0) and 91 cM (C16:1(n-7) (Muñoz *et al.* (2012), manuscript in preparation). In this region, a relevant gene for fatty acid metabolism has been located: *ELOVL fatty acid elongase 6* (*ELOVL6*). The *ELOVL6* gene is a member of the elongation-of-very-long-chain-fatty-acid gene family (*ELOVLs*) of condensing enzymes that perform the first and rate-limiting step in the elongation cycle in mammals [11]. These enzymes use malonyl-CoA as the 2-carbon donor to initialize the elongation process. In pigs, the family of enzymes consists of at least seven members, differing in their substrate preferences for FAs of different lengths and degrees of unsaturation, and specific spatial and temporal expression. To generalize, FA elongases can be divided into two major groups: a) enzymes involved in the elongation of saturated and monounsaturated very-long-chain fatty acids (*ELOVL1*, *3*, *6* and *7*) and b) enzymes which are elongases of polyunsaturated fatty acids (*ELOVL2*, *4* and *5*) [12,13]. The *ELOVL fatty acid elongase 6* (*ELOVL6*) gene (also known as LCE and FACE) is the only elongase involved in *de novo* lipogenesis, which catalyzes the elongation of long-chain saturated and monounsaturated FAs with 12–16 carbons to C18, but it does not possess activity beyond C18 [11]. Analysis of *ELOVL6*-deficient mice demonstrated that *ELOVL6* plays a crucial role in the overall fatty acid composition balance [5], and alterations in this composition have important effects on *de novo* lipogenesis and fatty acid oxidation [5]. The clear relationship between *ELOVL6* function and the QTL phenotype makes this gene a promising positional and functional candidate gene for the traits analyzed.

In the present study, a refined localization of the QTL affecting C16:0 and C16:1(n-7) FA in muscle and the evaluation of the porcine *ELOVL6* gene as candidate gene for this QTL was carried out in an Iberian x Landrace backcross population. DNA sequencing, gene expression analyses and association studies were performed to evaluate the involvement of this gene in C16:0 and C16:1(n-7) FA contents. In this article, we present different evidence that supports the role of *ELOVL6* gene polymorphism in the determination of muscle fatty acid composition in pigs.

Materials and Methods

Animal samples

Animals used in this study belong to the IBMAP cross, a population generated by crossing three Iberian (Guadyerbas line) boars with 31 Landrace sows [14], and containing several generations and backcrosses. The *ELOVL6* sequencing and gene expression analyses were carried out in animals from a backcross (BC1_LD) generated by crossing five F1 (Iberian x Landrace) boars with 26 Landrace sows and producing 144 backcrossed animals. All animals were maintained under intensive conditions and feeding was *ad libitum* with a cereal-based commercial diet. Animal care and procedures were performed following national and institutional guidelines for the Good Experimental Practices and approved by the Ethical Committee of the Institution (IRTA-Institut de Recerca i Tecnologia Agroalimentàries). Animals were slaughtered at an average age of 179.8±2.3 days, and samples of liver, muscle (*Longissimus dorsi*) and adipose tissue (backfat) were collected, snap-frozen in liquid nitrogen and stored at −80°C until analyzed. Genomic DNA was obtained from blood samples of all animals by the phenol-chloroform method, as described elsewhere. Composition of fatty acid with 12 to 22 carbons was determined in muscle [10] and backfat (Muñoz *et al.* (2012), manuscript in preparation) using a protocol based on gas chromatography of methyl esters [15].

Linkage map and haplotype reconstruction

A total of 439 animals, including the founder populations, were genotyped with the Porcine SNP60K BeadChip [16]. CRI-MAP version 2.503, developed by Evans and Maddox [http://www.animalgenome.org/bioinfo/tools/share/crimap], was used to build the linkage map using the genotype information of SSC8. In addition, previously detected polymorphisms in the *MTTP* and *FABP2* genes were also included in the analysis [8,17]. Raw data had a high genotyping quality (call rate >0.99) and, after selecting SNPs with MAF >5%, markers with genotyping and mapping errors were excluded by using the "Chrompic" option of CRI-MAP and R scripts developed by our group. Finally, we recalculated the genetic distances, employing the "Fixed" option, and 2,565 SNPs were retained for subsequent analyses (Table S1). Haplotypes were reconstructed using DualPHASE software [18], which exploits population (linkage disequilibrium) and family information (Mendelian segregation and linkage) in a Hidden Markov Model setting.

Chromosome 8 association and fine-mapping analyses

GWAS for the intramuscular profile of palmitic and palmitoleic acids was performed with a mixed model [19,20] accounting for additive effects associated with each marker (see below) by using Qxpak 5.0 [21]:

$$y_{ijlkm} = Sex_i + Batch_j + \beta c_l + \lambda_l a_k + u_l + e_{ijlkm}, \qquad (1)$$

in which y_{ijlkm} is the l-th individual record, sex (two levels) and batch (five levels) are fixed effects, β is a covariate coefficient with *c* being carcass weight, λ_l is a −1, 0, +1 indicator variable depending on the l-th individual genotype for the k-th SNP, a_k represents the additive effect associated with SNP, u_l represents the infinitesimal genetic effect treated as random and distributed as $N(0, \mathbf{A}\sigma_u)$ where $\mathbf{A}$ is a numerator of the kinship matrix and e_{ijlkm} is the residual. The same model was carried out for studying the association of polymorphisms detected in the *ELOVL6* gene with palmitic and palmitoleic acid profiles in muscle and backfat.

QTL fine-mapping was performed by simultaneously exploiting linkage and linkage disequilibrium (LD) using a haplotype-based approach [18] and following the mixed model:

$$y = Xb + Z_h h + Z_u u + e, \qquad (2)$$

in which *b* is a vector of fixed effects (sex and batch), *h* is the vector of random QTL effects corresponding to the K cluster defined by the Hidden State (HS), *u* is the vector of random individual polygenic effects and *e* is the vector of individual error. The genome-wide significance was determined using the R-package q-value [22], and the cut-off of the significant association was set at q-value≤0.05.

In order to estimate the LD between the SNPs located within the candidate region, a LD analysis was performed using the genotype and phases information from DualPHASE software. The

LD estimated for each pair of SNPs was visualized using the "LDheatmap 0.9" R package [23].

RNA isolation and cDNA synthesis

Total RNA was obtained from liver, muscle and backfat tissues using the RiboPure^{TM} Isolation of High Quality Total RNA (Ambion®), following the manufacturer's recommendations. RNA was quantified using the NanoDrop ND-1000 spectrophotometer (NanoDrop products) and checked for purity and integrity in a Bioanalyzer-2100 (Agilent Technologies). The isolated RNA was reverse-transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) and random hexamers in a total volume of 20 µl containing 1 µg (liver and muscle) or 0.3 µg (backfat) of total RNA, following the manufacturer's instructions.

Amplification and sequencing of the pig *ELOVL6* coding region and proximal promoter

The proximal promoter and the entire coding region of the *ELOVL6* gene was amplified and sequenced in twenty samples from the BC1_LD. Primers (Table S2) to amplify two overlapping fragments of 688 bp and 499 bp, including the complete coding region, were designed from the human GenBank NM_024090.2 sequence, assuming conservation across species. The proximal promoter region was amplified for the Sus scrofa breed mixed chromosome 8 sequence (GenBank:NW_003610943) available at the Sscrofa10.2 database (primers in Table S2) and assuming conservation with the human and mouse *ELOVL6* promoters [24]. A total of 1046 bp of the *ELOVL6* promoter and exon 1 were sequenced in two overlapping fragments of 604 bp and 605 bp. Primers were designed using the software PRIMER3 [25] and were validated using the software PRIMER EXPRESS^{TM} (Applied Biosystems).

PCRs were carried out in a total volume of 25 µl containing 0.6 units of AmpliTaq Gold (Applied Biosystems), 1.5–2.5 mM $MgCl_2$ (depending on the primers; see Table S2), 0.2 mM of each dNTP, 0.5 µM of each primer and 50 ng of genomic DNA or 2 µl of cDNA. Thermocycling was carried out under the following conditions: 94°C for 10 min, 35 cycles of 94°C for 1 min, 58°C–62°C (depending on the primers; see Table S2) for 1 min and 72°C for 1 min, with a final extension of 72°C for 7 min.

PCR products were purified using the ExoSAP-IT® method and sequenced with a Big Dye Terminator v.1.1 Cycle Sequencing Kit in an ABI 3730 analyzer (Applied Biosystems).

To characterize the *ELOVL6* promoter, a computer-assisted identification of putative promoter/enhancer elements was performed using the TFSEARCH software [http://www.cbrc.jp/research/db/TFSEARCH.html] and MATINSPECTOR application (set at a cut-off score of >85%) [26], a part of GENOMATIXSUITE software (Genomatix Software GmbH). Genomatix Matrix Library 8.3 was used with a core similarity threshold of 0.85 and an optimized matrix similarity threshold.

Gene expression quantification

A total of 110 animals of the BC1_LD backcross were selected to perform gene expression quantification in liver, backfat and muscle. PCR primers were designed using PRIMER EXPRESS^{TM} software (Applied Biosystems) and are shown in Table S2. Primers for amplification of *ELOVL6* mRNA were designed from the available sequence (GenBank:XM_003357048) covering exons 3–4 to amplify a 103-bp-long fragment. Three genes frequently used as references in RT-qPCR experiments were analyzed as endogenous controls: *β-2 microglobulin* (*β2M*), *Hypoxanthine phosphoribosyltransferase1* (*HPRT1*) and *Glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*) [27,28]. All reference genes were tested using the software GeNorm [29], and the two best endogenous controls for all tissues were *β2M* and *HPRT1*. PCR amplification was performed in triplicate in a 20 µl final volume containing 2 µl of cDNA sample, diluted 1:20 in DEPC-treated H_2O from liver and muscle samples, and 1:5 from backfat samples. For gene amplification, FastStart Universal SYBR Green Master (Rox; Roche Applied Biosystems) was used. Primers were used at 900 nM for the *ELOVL6* gene and 600 nM for both references genes, except from *HPRT1* in the muscle study (900 nM). PCR amplification was run on an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) using 96-well optical plates under the following conditions: 10 min at 95°C, 40 cycles of 15 sec at 95°C and 1 min at 60°C. A dissociation curve was drawn for each primer pair to assess that there was no primer-dimer formation.

To quantify and normalize the relative quantification (RQ) data, the $2^{-\Delta\Delta CT}$ method [30] was applied using a sample with low expression as a calibrator. Comparison of mean values between genotypes was made using a linear procedure of R software, which employs a single stratum analysis of variance considering sex and batch as fixed effects. Differences were considered statistically significant at a p-value of < 0.05.

Genotyping

BC1_LD backcross animals (N = 144) were genotyped with the Porcine SNP60 BeadChip (Illumina) using the Infinium HD Assay Ultra protocol (Illumina). Raw data had a high genotyping quality (call rate >0.99) and was visualized and analyzed with the GenomeStudio software (Illumina). For subsequent data analysis, a subset of 54,998 SNPs was selected by removing the SNPs with a minor allele frequency <5%, those with missing genotypes >5% and the duplicated SNPs in the Sscrofa 10.2 assembly.

The SNPs *ELOVL6:c.-533C>T*, *ELOVL6:c.-480C>T* and *ELOVL6:c.416C>T* were genotyped using the KASP SNP genotyping system platform [http://www.kbioscience.co.uk/reagents/KASP/KASP.html]. A total of 160 animals were genotyped, 125 of those belonging to BC1_LD and the rest being parental animals of the IBMAP cross (F0 and F1).

GWAS of gene expression

Association analyses of RT-qPCR expression data of *ELOVL6* mRNA in liver, backfat and muscle, and whole-genome SNP genotypes, were carried out with Qxpak 5.0 software. The position of the SNPs was based on the *Sus scrofa* 10.2 genome assembly [http://www.animalgenome.org/repository/pig/]. For GWAS analysis, the previously described model (1), without correcting for carcass weight, was used. The infinitesimal effect allows us to adjust the data for family information and, thus, to correct the inter-chromosomal linkage disequilibrium effect. In this analysis, each SNP was tested individually to check the association. Chromosome X was analyzed using the same models, but including a dosage compensation parameter [31]. The R package q-value [22] was used to calculate the FDR-based q-value to measure the statistical significance at the genome-wide level for association studies. The cut-off of significant association at the whole genome level was set at q-value≤0.1. This significance threshold is likely too stringent due to the linkage association among SNP genotypes. Gene annotation for 2 Mb genomic intervals around the most significant SNPs was performed with *Biomart* software in the Ensembl Sscrofa 10.2 data set [www.ensembl.org]. For gene annotation, only those regions that showed

a cut-off at a chromosome-wise level lower than q-value < 0.05 were selected.

Results

Linkage and haplotype reconstruction

The length of the linkage map on SSC8 was 131.2 cM and the ratio between the genetic and the physical distance was 0.89 cM/Mb, similar to that previously reported [32]. Genotypes from a total of 2,565 SNPs of the Porcine SNP60 BeadChip (Illumina) were employed to reconstruct the haplotypes through DualPHASE software. Previous studies showed that the estimation of the phenotypic effect of haplotype clusters is a good approximation to identify the functionally relevant ones, as well as to reduce the confidence interval for the fine mapping QTL [18,33]. In this study, a method based on Hidden Markov Models that simultaneously phases and sorts haplotypes using linkage and LD information for haplotype reconstruction was employed. A total of twenty haplotype clusters (K = 20) were used for fine mapping, as described below.

Fine mapping and gene annotation

A combination of the haplotype-based approach and GWAS for the intramuscular profile of palmitic and palmitoleic acids was performed in 144 BC1_LD individuals and 2,565 SNPs. It is worth noting that, for both traits, the two strategies showed the highest association at the same position (Figure 1). For instance, the GWAS profile corresponding to palmitic acid was maximized at 119,727,822–119,887,525 bp (p-value = 6.19×10^{-09}) and the profile score from the haplotype-based analyses showed the maximum association signal at position 117,824,360–119,887,525 bp (p-value = 3.57×10^{-07}) (Figure 1A). For palmitoleic acid, the GWAS profile was maximized at 119,851,321–120,104,023 (p-value = 4.23×10^{-09}) and the profile scores from the haplotype-based analyses were maximized at position 117,824,360–119,727,822 bp (p-value = 1.09×10^{-06}) (Figure 1B). In general, the association signal obtained by GWAS was higher than were curves obtained with the haplotype-based approach. However, it should be noted that the haplotype-based approach allowed us to simultaneously exploit linkage analysis and LD (LDLA). In addition, although both strategies were modeled by a mixed model, a different parameterization was employed. Thus, in the LDLA approach, HS was treated as additive random effects, whereas in GWAS a single-marker regression analysis was performed and the SNP alleles were treated as additive fixed effects.

According to the fine mapping data, the region comprised between 117–121 Mb was annotated using *Biomart* software in the Ensembl Sscrofa10.2 dataset [www.ensembl.org]. A total of 21 genes were located in this region, but only two were clearly related to fatty acid metabolism: *ELOVL6* (at position 120,119,244 bp) and *PLA2G12A* (at position 120,566,787 bp). The coincidence between the biological function of *ELOVL6* and the observed QTL effect on fatty acid composition on SSC8 strengthens the interest of the *ELOVL6* gene as the positional candidate gene for this QTL.

Identification of polymorphisms in the porcine *ELOVL6* gene

To characterize the porcine *ELOVL6* gene, a 1,046-bp long fragment of the *ELOVL6* promoter and exon 1 was amplified from genomic DNA and sequenced, assuming conservation with the human and mouse genes. In addition, the entire coding region of the *ELOVL6* gene was amplified and sequenced. The alignment and analysis of these sequences allowed for the identification of eight polymorphisms (Table 1): one synonymous polymorphism in exon 4 and seven nucleotide substitutions in the promoter region. The SNPs located in the promoter were arranged in three haplotypes, which can be distinguished by genotyping the *ELOVL6:c.-533C>T* and *ELOVL6:c.-480C>T* polymorphisms (relative to the transcription start site, TSS, of the GenBank:NW_003610943). Hence, these two tag polymorphisms and the *ELOVL6:c.416C>T* SNP in exon 4 (GenBank:AB529461) were genotyped in parental and BC1_LD animals. Regarding the IBMAP founders, the *ELOVL6:c.-533C* allele and *ELOVL6:c.416T* allele were fixed in Iberian boars. The allele frequencies for these two SNPs were 0.25 for F1 Landrace sows and 0.78 and 0.72 for the BC1_LD Landrace sows, respectively. In contrast, *ELOVL6:c.-480C>T* SNP was not fixed in the Iberian founders, and therefore it was less informative. Both *ELOVL6:c.-533C>T* and *ELOVL6:c.416C>T* polymorphisms segregated in the BC1_LD animals with frequencies of 0.63 for allele C and 0.60 for allele T, respectively. Linkage disequilibrium analysis revealed that the three *ELOVL6* polymorphisms were in strong LD (D′ = 0.99) with three of the most significant SNPs (SIRI0000509, INRA0030422 and H3GA0025321) identified in both GWAS and fine mapping analyses (Figure S1).

To assess if polymorphisms in the promoter region could affect *ELOVL6* expression through the disruption of transcription factor-binding sites, a computer-assisted identification of potential cis-acting DNA-sequence motifs was carried out. As has been previously described in mouse liver, the *ELOVL6* gene is regulated by SREBP-1 [4,24,34,35]. SREBP-1 presents dual DNA sequence specificity, binding to both E-box and SRE motifs [36]. Four SREBP binding sites were identified in the pig *ELOVL6* promoter, three SRE elements in positions −27 to −17, −460 to −449 and −532 to −524 (Figure 2) and one E-box in position −341 to −330 (Figure 2A), relative to the TSS of the GenBank sequence NW_003610943, similar to those observed in the mouse promoter (Figure 2B) [24]. Also, other candidate transcription factors, with biological relevance, have elements in this promoter, such as MLX (at position −339 to −322), which belongs to the family of basic helix-loop-helix leucine zipper (bHLH-Zip) and induces *ELOVL6* gene expression by glucose in mice [6], HNF4γ (at position −719 to −694) or KLF10 (at position −377 to −372) (Figure 2). However, none of our polymorphisms changed these binding sites. Interestingly, two consecutive SNPs forming a haplotype at positions −533 (*ELOVL6:c.-533C>T)* and −*534 (ELOVL6:c.-534C>T)* were identified in the core binding site of the *estrogen-related receptor alpha* gene (*ESRR-α*), generating a multi-nucleotide polymorphism. Furthermore, the *ELOVL6:c.-480C>T* polymorphism was also located in a potential SP1 binding site (Figure 2).

Association of *ELOVL6* polymorphisms with C16:0 and C16:1(n-7) composition in muscle and backfat

An association analysis with the SSC8 genotypes from 2,565 SNPs of the Porcine SNP60 BeadChip (Illumina) and three *ELOVL6* SNPs in 125 BC1-LD animals was performed using the additive model (1). In this analysis, the *ELOVL6:c.-533C>T* polymorphism showed the highest association with the percentage of palmitic acid (p-value = 1.38×10^{-07}; â(estimated additive effect) = 0.742; see Figure S2A) and palmitoleic acid (p-value = 1.23×10^{-08}; â = 0.253; see Figure S2B) content in muscle. Also, a relevant association was observed between *ELOVL6:c.416C>T* polymorphism with palmitic, and palmitoleic acid content in muscle (p-value (C16:0) = 1.11×10^{-04}; â(C16:0) = 0530; p-value (C16:1(n-7)) = 6.98×10^{-07}; â(C16:1(n-7)) = 0.214; see Figure S2). In backfat, the *ELOVL6:c.-533C>T* polymorphism was the most significantly associated one with

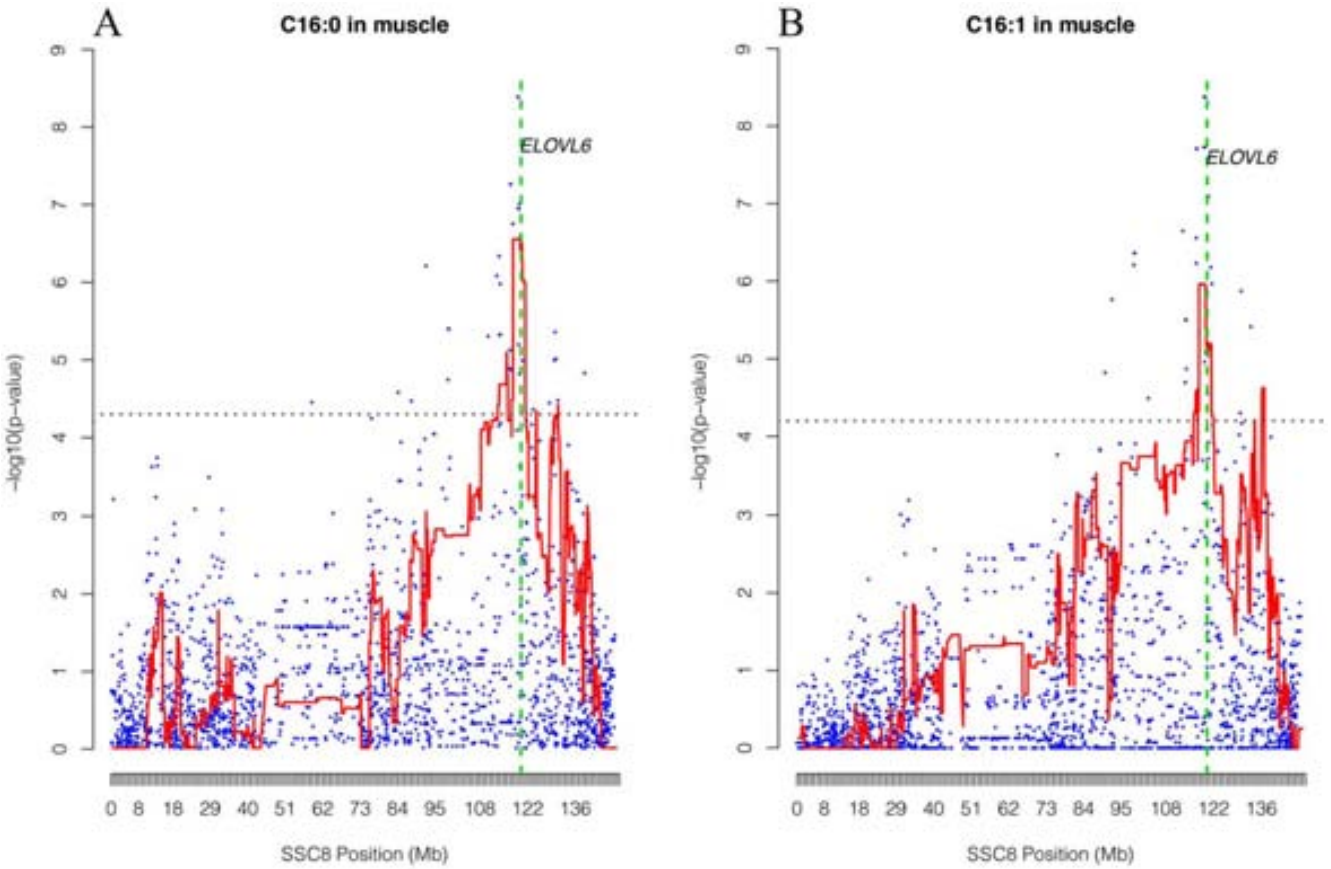

Figure 1. Reduction of the QTL interval by GWAS and LDLA analyses and gene mapping of *ELOVL6*. Plot of GWAS (blue points) and LDLA patterns (red line) for palmitic (A) and palmitoleic (B) acids. The X-axis represents chromosome 8 positions in Mb and the Y-axis shows the –log10 (p-value). The vertical green line represents the position of the *ELOVL6* gene on SSC8. Horizontal dashed lines mark the genome-wide significance level (FDR-based q-value≤0.05). Positions in Mb are relative to *Sscrofa10.2 assembly* of the pig genome.
doi:10.1371/journal.pone.0053687.g001

palmitic acid content (p-value = 2×10^{-15}; â = 0.976). In addition, *ELOVL6:c.416C>T* SNP showed a high association with palmitic acid content (p-value = 6.27×10^{-13}; â = 0.859) (data not shown). Analyzing the palmitoleic acid content in backfat, the most significantly associated SNP was H3GA0025290 (113,528,768 bp; â = 0.182, p-value = 8.54×10^{-10}). *ELOVL6* polymorphisms *ELOVL6:c.-533C>T* and *ELOVL6:c.416C>T* also showed a significant association with palmitoleic content: p-value = 6.14×10^{-09} (â = 0.168) and p-value = 6.95×10^{-08} (â = 0.151), respectively (data not shown). The clear association of the *ELOVL6:c.-533C>T* polymorphism with the percentage of both fatty acids in muscle and backfat yields new evidence to continue studying *ELOVL6* as a candidate gene for SSC8 QTL.

Table 1. Polymorphisms identified in the proximal promoter and coding regions of the *ELOVL6* gene.

Gene localization	Position (bp)	Polymorphism
Promoter[1]	−574	C/T
	−534	C/T
	−533[3]	C/T
	−492	G/A
	−480[3]	C/T
	−394	G/A
	−313	C/T
Exon 4[2]	+416[3]	C/T

[1]Positions relative to the transcription start site using, as reference, the GenBank NW_003610943 sequence.
[2]Referring to the coding region (GenBank:AB529461).
[3]SNPs genotyped in the BC1_LD population.
doi:10.1371/journal.pone.0053687.t001

Effect of the *ELOVL6: c.-533C>T* polymorphism on gene expression and fatty acid composition

The association of the *ELOVL6: c.-533C>T* polymorphism with the percentages of C16:0 and C16:1(n-7) suggests a role of this mutation in the regulation of *ELOVL6* gene expression and, subsequently, in fatty acid metabolism. Thus, the expression profile of the pig *ELOVL6* gene, in liver, backfat and muscle, organs particularly important in fatty acid metabolism, was studied by RT-qPCR in 110 BC1_LD animals. In accordance with previous results in mouse and rat, in which high *ELOVL6* expression was found in tissues with active lipogenesis [4,35,37], the highest expression was found in backfat tissue, followed by liver and muscle. Clear differences in *ELOVL6* expression were observed among samples in all tissues, with a highly significant effect of sex in liver (p-value = 6.5×10^{-03}), backfat (p-value = 3.4×10^{-04}) and muscle (p-value = 3.4×10^{-05}), where *ELOVL6* gene expression was higher in females than in males.

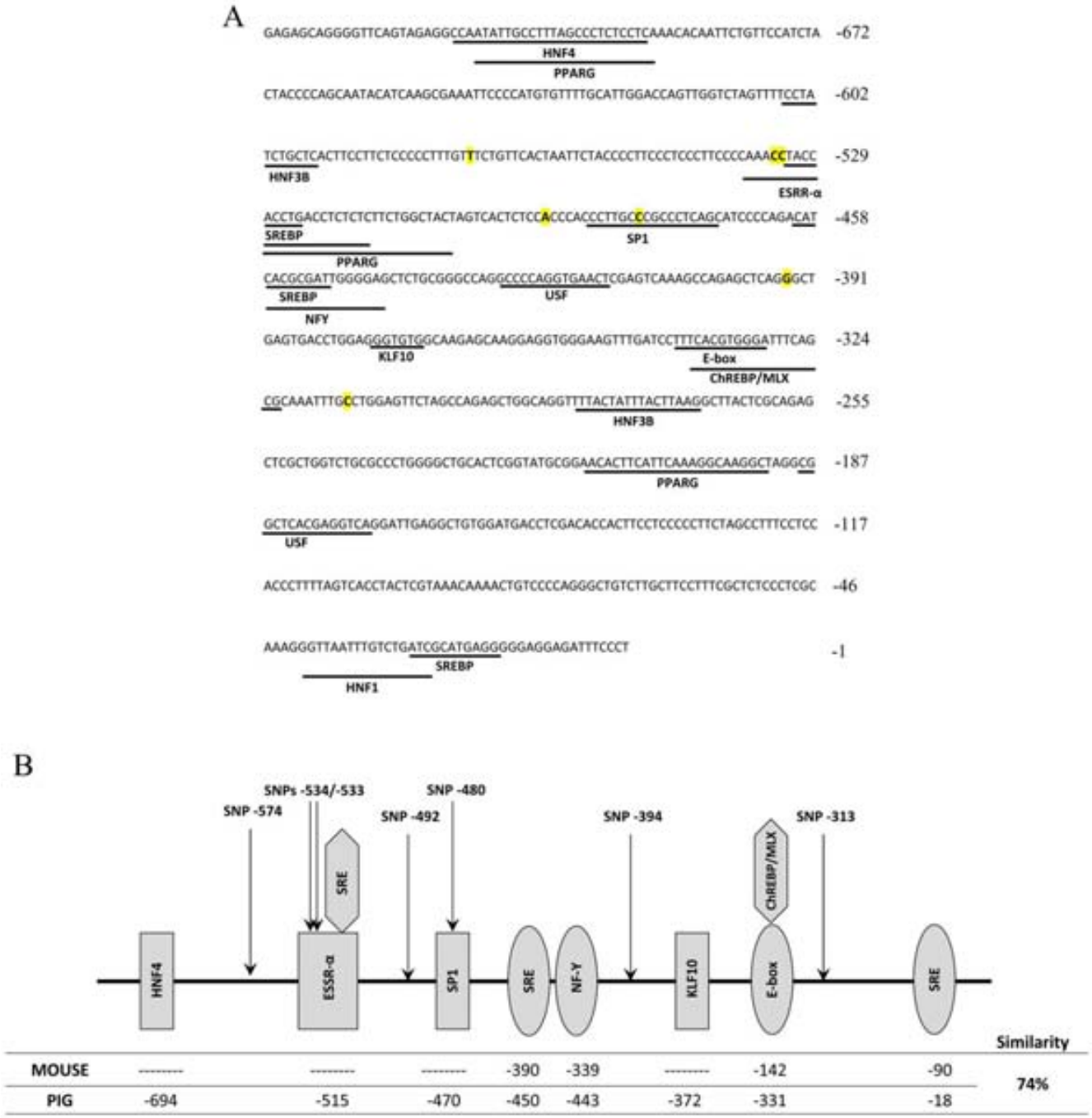

Figure 2. Genetic characterization of the *ELOVL6* pig promoter and identification of potential cis-acting DNA-sequence motifs. Summary of the *ELOVL6* pig promoter: A, nucleotide sequence of the 5′-flanking region of the porcine *ELOVL6* gene, where potential binding sites for transcription factors are underlined. Positions of *ELOVL6* promoter polymorphisms are labeled in yellow. B, comparison of transcription factor binding sites between mouse and the pig *ELOVL6* promoter, including *ELOVL6* SNPs localization.
doi:10.1371/journal.pone.0053687.g002

In addition, the correlation between the *ELOVL6* expression levels across the three tissues was analyzed, but no clear associations were observed among tissues. This result suggests that the mechanisms controlling *ELOLV6* expression are different in backfat, liver and muscle tissues [35].

When animals were classified according to the *ELOVL6:c.-533C>T* genotypes, no significant variations were found between genotypes when liver and muscle samples were analyzed (Figure 3C–D). Nevertheless, different levels of expression between genotypes were obtained in backfat samples (p-value = 8.7×10^{-05}; Figure 3B), where animals with the CC genotype showed a significantly lower expression, as compared to animals with the other two genotypes. Interestingly, when only individuals with a known allele origin for the *ELOVL6:c.-533C>T* polymorphism were analyzed, the Iberian allele C decreased the expression, in comparison with the Landrace allele T (p-value = 4.6×10^{-03}). Accordingly, CC homozygous individuals showed a higher percentage of C16:0 in muscle (p-value = 3.61×10^{-05}) and backfat

(p-value = 1.83×10^{-09}), in comparison with TT individuals (Figure 4A). Similar results were obtained for C16:1(n-7) in both tissues, with CC animals presenting a higher relative content of this fatty acid (p-value (muscle) = 7.1×10^{-06} and p-value (backfat) = 1.47×10^{-04}; Figure 4B). These data suggest a substrate accumulation in individuals with the C allele due to a hypothetical deficiency of the *ELOVL6* gene (Figure 3A). In agreement with these data, a decrease of C18:0 content was also observed in backfat (p-value = 5×10^{-02}) in animals with the C allele, but such differences were not present in muscle (data not shown). Although non-significant differences were observed in the 18-carbon fatty acid content in muscle and backfat (except for the C18:0 in backfat), a significant decrease in elongation activity ratios (C18:0/C16:0 and C18:1(n-7)+C18:1(n-9)/C16:1(n-7)) were observed in both tissues in animals with the CC genotype (Figure 4C–D).

Genome-wide association studies of *ELOVL6* gene expression

Taking into account that differences in *ELOVL6* gene expression were observed among tissues and animals, GWASs using the RQ expression data of the three tissues and the genotypes of 54,998 SNPs distributed across the pig genome were carried out to find new, potential genomic regions associated with *ELOVL6* gene expression. The promoter *ELOVL6* SNPs (*ELOVL6:c.-533C>T* and *ELOVL6:c.-480C>*) and the protein-coding region SNP (*ELOVL6:c.416C>T*) genotyped in this work, which are not included in the porcine SNP60 Bead-Chip, were incorporated into the study. First, backfat analysis of *ELOVL6* gene expression showed three relevant regions in chromosomes SSC2, SSC4 and SSC8 which were significant at a chromosome-wise level (Figure S3A). Interestingly, the most significant peak was localized in SSC8 inside the QTL region, very close to the *ELOVL6* gene (ALGA0049135; 117,548,144 bp; p-value = 2.74×10^{-06}) (Figure 5C). High association was also obtained with the *ELOVL6:c.-533C>T* polymorphism (p-value = 2.05×10^{-05}), whereas the other two *ELOVL6* polymorphisms were not significantly associated. Gene annotation of the other two regions was performed to find potential trans-acting genetic variants modulating *ELOVL6* gene expression. In SSC2, a significant region was found between positions 9.3 Mb and 9.8 Mb (DIAS0000337; 9,736,754 bp; p-value = 1.4×10^{-05}), in which several genes related to lipid metabolism were identified (Figure 5A). The most interesting ones were the *estrogen-related receptor alpha* (*ESRRα*), three genes which are members of the *fatty acid desaturase* family (*FADS1*, *FADS2* and *FADS3*), the *carnitine palmitoyltransferase 1A* (*CPT1A*) and the *nuclear receptor subfamily 1, group H, member 3* (*NR1H3*). Finally, the most significant region of SSC4 was found between positions 36 Mb and 44 Mb (ASGA0088888; 40,318,092 bp; p-value = 1.4×10^{-05}), where the *Kruppel-like factor 10* (*KLF10*) gene was annotated (Figure 5B). In liver, three candidate chromosomal regions were significantly associated with *ELOVL6* gene expression at a chromosomal level on SSC4, SSC5 and SSC9 (Figure S3B). The most significant region in SSC4 showed two peaks at the 30 Mb–35 Mb and 60 Mb–67 Mb regions (Figure 6). Gene annotation of both regions allowed us to identify several genes, which may be related to *ELOVL6* RQ, near the two most significant SNPs: ALGA0025162 (60,844,160 bp; p-value = 2.93×10^{-06}) and ALGA0024413 (34,206,333 bp; p-value = 3.58×10^{-06}). Proximal to

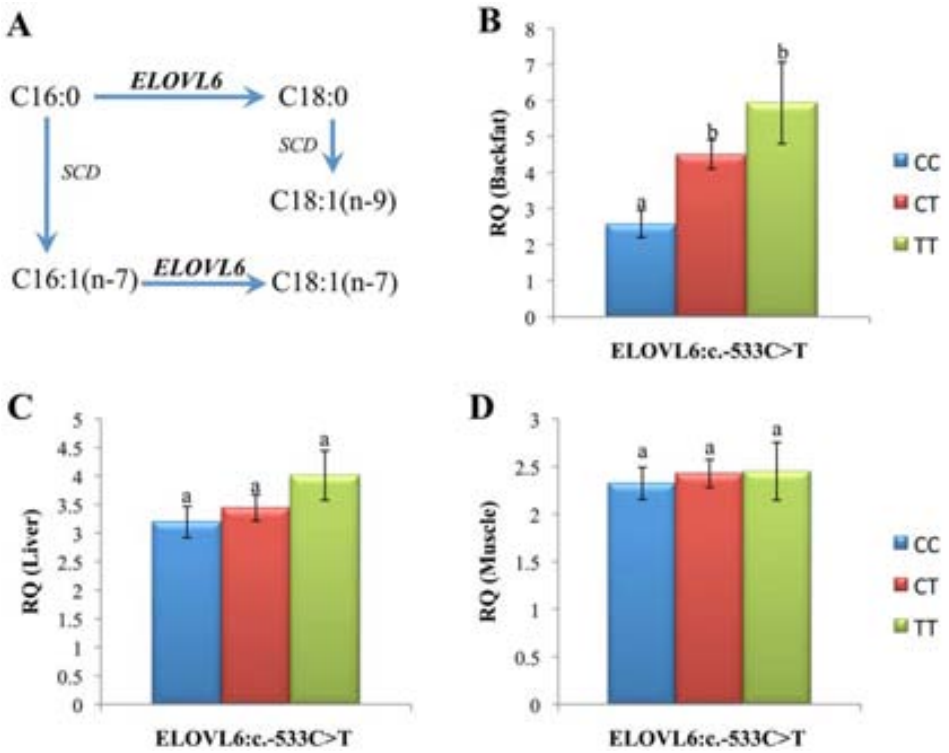

Figure 3. Association of *ELOVL6: c.-533C>T* genotypes on gene expression in backfat. A SNP genome-wide association study was performed with *ELOVL6* relative expression levels measured by RT-qPCR in 110 samples from backfat, liver and muscle. Data include: Schematic representation of the elongation pathway of 16-carbon fatty acid (A), *ELOVL6* expression levels in backfat (B), liver (C) and muscle (D). Data represent means ± SEM. Values with different superscript letters (a, b and c) indicate significant differences between groups (p-value < 0.05), as determined by a single stratum analysis of variance considering sex and batch as fixed effects.
doi:10.1371/journal.pone.0053687.g003

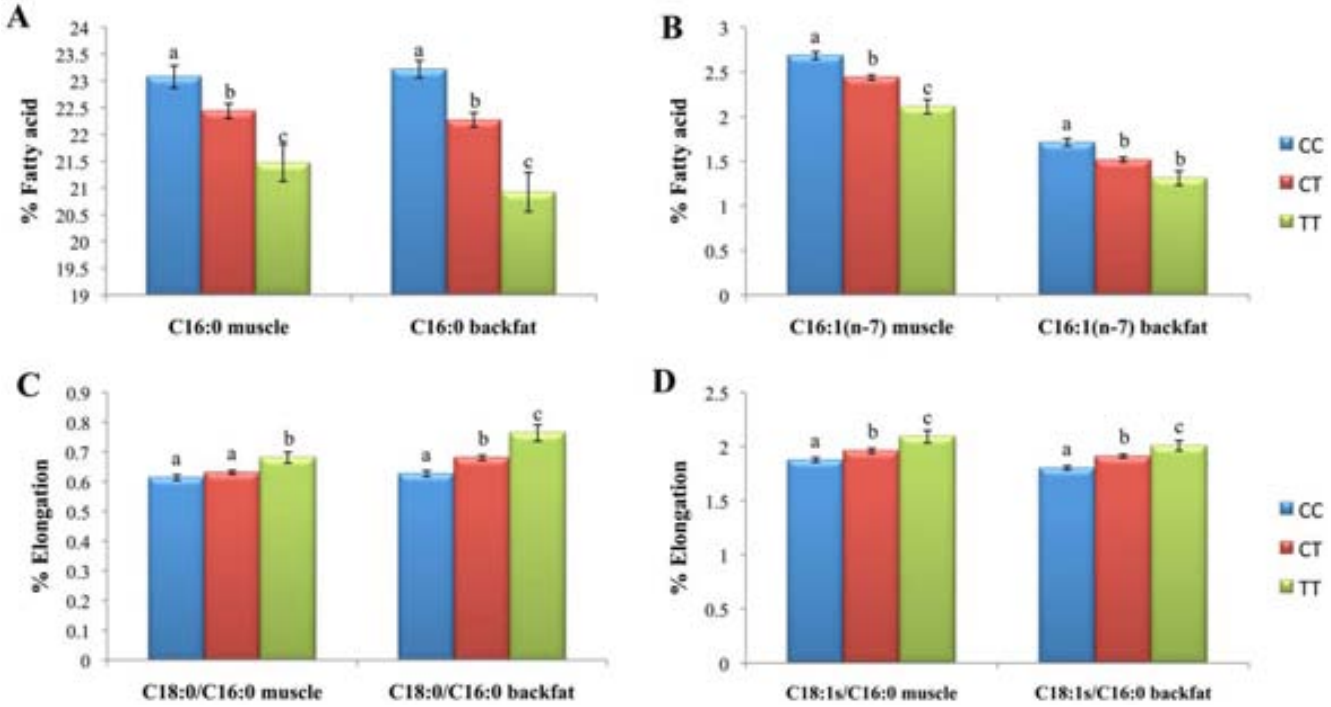

Figure 4. Fatty acid composition of different *ELOVL6:c.-533C>T* genotypes in muscle and backfat. Data include: percentage of C16:0 (A) and C16:1(n-7) fatty acids (B) in muscle and backfat; and the elongation ratios C18:0/C16:0 (C) and C18:1(n-7)+C18:1(n-9)/C16:0 (D) in muscle and backfat. Data represent mean ± SEM. Values with different superscript letters (a, b and c) indicate significant differences between groups (p-value < 0.05), as determined by a single stratum analysis of variance considering sex and batch as fixed effects.
doi:10.1371/journal.pone.0053687.g004

ALGA0025162 was located the *hepatocyte nuclear factor 4 gamma* (*HNF4γ*) and members 4 and 5 of the *fatty acid binding protein* family (*FABP4* and *FABP5*) (Figure 6). SNP ALGA0024413 was near the significant region detected in backfat analysis, suggesting a co-effect in both tissues by the porcine *KLF10* gene (Figure 6). No relevant genes were found in the 71 Mb to 79 Mb region of SSC5. In SSC9, the significant region was located in the 65 Mb–71 Mb interval, in which the *acyl-CoA dehydrogenase 8* (*ACAD8*) was located. Finally, muscle *ELOVL6* gene expression was asociated with three regions in SSC6, SSC8 and SSC12 (Figure S3C). The most significant one was situated in SSC6 between positions 18 Mb and 26 Mb (ALGA0034806; 19,862,636 bp; p-value = 4.02×10^{-06}), where the general transcription factor *SET domain containing 6* (*SETD6*) was identified (data not shown). In SSC8 and SSC12, a significant region was found in intervals 15 Mb–19 Mb and 10 Mb–14 Mb, respectively (data not shown). Nevertheless, no relevant genes were identified using the current porcine gene annotation information. The significance threshold was likely too stringent owing to the linkage dependence among the SNPs included in the analysis and, thus, other suggestive SNP peaks may also contain relevant genes.

Discussion

The QTL affecting palmitic and palmitoleic acid contents on SSC8 was previously identified and the porcine *MTTP* gene was analyzed as a positional candidate gene [8]. These studies were performed using a reduced number of microsatellite markers and, as a consequence, the confidence interval had several Mb. The improvements in the porcine genome and the use of the SNP data from the Illumina 60 K porcine chip allowed us to make a better estimation of the QTL position by GWAS and haplotype-based approaches. GWAS studies maximized the QTL peak at 10 Mb from the *MTTP* gene, in the region where the *ELOVL6* gene was located. Although the *ELOVL6* gene has been selected as a new functional and positional candidate gene, a lower effect of the *MTTP* gene cannot be ruled out.

Despite the crucial role of genes such as *ELOVL6* and members of the *SCD* family in determining the balance among C16:0, C16:1(n-7), C18:1(n-7) and C18:1(n-9) [5,6,38], the information regarding these genes in pigs is sparse. In this study, we characterized the porcine *ELOVL6* gene and we presented several pieces of evidence confirming that this lipogenic enzyme is highly associated with fatty acid composition in pigs. Among the eight polymorphisms found in the porcine *ELOVL6* gene, the *ELOVL6:c.-533C>T* polymorphism was clearly associated with C16:0 and C16:1(n-7) composition in muscle and backfat. An increase of C16s' fatty acid percentage in animals with the C allele, in comparison with animals carrying only the T alleles, was observed. In accordance with the function of *ELOVL6* (Figure 3A), which elongates C16 to C18 fatty acids [5], a lower *ELOVL6* gene expression was found in the backfat of animals with the Iberian allele. The lower *ELOVL6* gene expression was associated with the accumulation of C16:0 and C16:1(n-7) in muscle and backfat, as has been previously described in mammalian cells by modulating *ELOVL6* activity with siRNA [5,6]. Similar results were obtained using mice deficient for *ELOVL6* [5], where an increase of C16 fatty acids and a decrease of C18 fatty acids was observed in $ELOVL6^{-/-}$ mice. In agreement with these studies, the percentage of C18:0 showed a decrease in backfat, but no differences were observed in C18:1(n-7).

The relevance of adipose tissue in overall fatty acid synthesis in pigs must be considered for the interpretation of the present results. Liver and adipose tissue are the principal organs

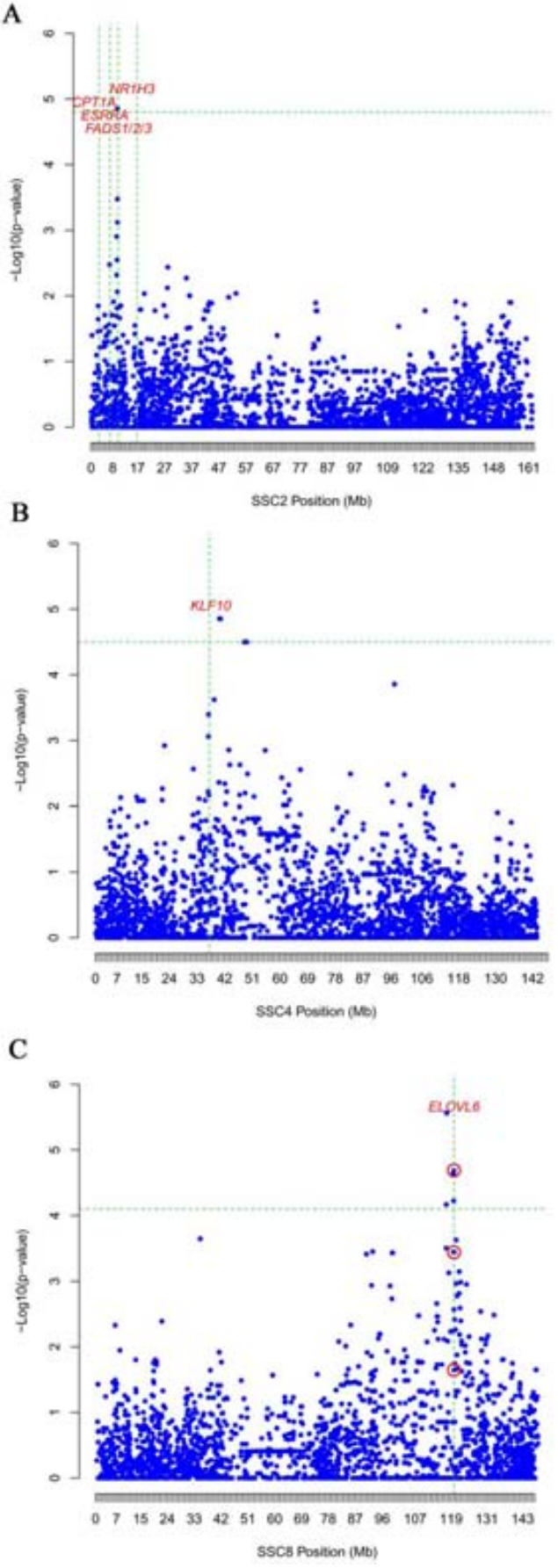

Figure 5. Significant region obtained in GWAS for backfat gene expression. Association analysis between the backfat *ELOVL6* expression level and SNP genotypes for SSC2 (A), SSC4 (B) and SSC8 (C). *ELOVL6* polymorphisms are included and labeled with a red circle. Positions in Mb are relative to *Sscrofa10.2 assembly* of the pig genome. Vertical, dashed lines indicate the location of positional candidate genes. Horizontal, dashed lines mark the genome-wide significance level (FDR-based q-value≤0.1).
doi:10.1371/journal.pone.0053687.g005

implicated in *de novo* lipogenesis, although their contribution differs across species. In ruminants, such as cow and sheep, both liver and adipose tissue appear to be important sites of synthesis [39], while in mouse and rat adipose tissue accounts for at least 50% of the newly synthesized fatty acids [40]. Pig adipose tissue seems to be responsible for a greater contribution to overall fatty acid synthesis than does liver [41], as has been similarly observed in humans [42,43]. In agreement with this, the expression of *ELOVL6*, a gene involved in *de novo* lipogenesis, was higher in adipose tissue than in the liver and muscle of 110 BC1_LD animals. Furthermore, the effect of SNP *ELOVL6:c.-533C>T* in *ELOVL6* expression was only significant in adipose tissue, suggesting that this polymorphism may have an influence in adipose fatty acid synthesis and, subsequently, in body fatty acid composition. In fact, adipose tissue is the major source of circulating free fatty acids (FFAs) and, together with the liver, supplies fatty acids to muscle [44]. In mice, the concentration of muscle palmitoleate is a direct reflection of adipose FFAs [44]. As in pig the contribution of adipose tissue in fatty acid synthesis is higher than in liver, we could hypothesize that the composition of fatty acids in muscle closely resembles that observed in adipose tissue [45,46]. High and moderate positive phenotypic correlations between backfat and muscle were found for C16 and C16:1(n-7) composition ($r_{C16:0} = 0.72$, p-value $= 2.2 \times 10^{-16}$ and $r_{C16:1(n-7)} = 0.43$, p-value $= 5.13 \times 10^{-06}$, respectively), supporting our hypothesis. Furthermore, a high correlation for palmitoleic fatty acid was not expected because another genomic region with a strong effect on this fatty acid in muscle, but not in backfat, was identified in SSC4 [10].

Despite the strong association (p-value $= 2.05 \times 10^{-05}$) between the *ELOVL6:c.-533C>T* polymorphism and backfat *ELOVL6* gene expression, SSC8 SNP ALGA0049135 (117,548,144 bp) was more significantly associated (p-value $= 2.74 \times 10^{-06}$). However, *ELOVL6:c.-533C>T* showed a higher additive effect ($\hat{a} = 0.174$), in comparison with SNPs ALGA0049135 ($\hat{a} = 0.154$). Hence, further investigation is required to validate the *ELOVL6:c.-533C>T* polymorphism as the causal mutation or to identify new genetic variants in this QTL region modulating *ELOVL6* gene transcription that could better explain the QTL underlying phenotypic variation in C16 and C16:7(n-1).

Apart from the significant effect of SSC8 in *ELOVL6* gene expression, other interesting genomic regions were identified as being directly associated with *ELOVL6* relative expression levels in backfat, muscle and liver. Among them, a common peak at 60 Mb of SSC4 was identified for backfat and liver, suggesting the presence of genes related to *ELOVL6* expression in both tissues. The porcine *KLF10* gene was identified in this chromosomal position. It is a circadian-clock-controlled transcription factor that regulates genes involved in glucose and lipid metabolism in liver, such as *SREBP* and *ELOVL6* [47]. The identification of a potential cis-acting DNA-sequence motif for *KLF10* in the proximal promoter region of porcine *ELOVL6* supports the involvement of this gene in the *ELOVL6* transcriptional regulation in both tissues. Another interesting region associated with *ELOVL6* expression in backfat was observed in SSC2, in which the *ESRR-α* gene was identified. *ESRR-α* codes for a transcriptional regulator which

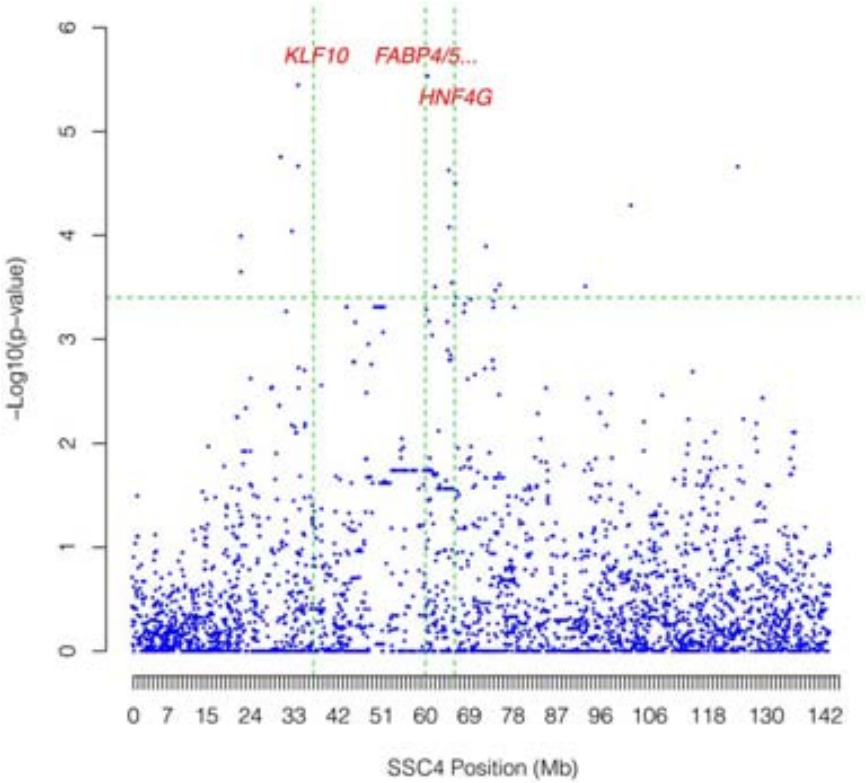

Figure 6. Significant region obtained in GWAS for liver gene expression. Association analysis between the liver *ELOVL6* expression level and SNP genotypes for SSC4. Positions in Mb are relative to *Sscrofa10.2 assembly* of the pig genome. Vertical, dashed lines indicate the location of positional candidate genes. Horizontal, dashed lines mark the chromosome-wide significance level (FDR-based q-value≤0.1).
doi:10.1371/journal.pone.0053687.g006

binds to an ERR-α response element (ERRE) containing a single-consensus half-site 5′-TNAAGGTCA-3′ and regulates a variety of genes related to fatty acid metabolism [48]. Interestingly, this transcription factor is regulated by estrogens, the primary female sex hormones. Thus, the higher *ELOVL6* gene expression observed in females may be explained by the increase of *ESRR-α* activity, due to the high levels of estrogens in females. Furthermore, ERRE was present in the *ELOVL6* promoter and included two polymorphic positions in the core binding element (Figure 2), one of which was *ELOVL6:c.-533C>T* SNP, reinforcing this polymorphism as a candidate mutation to explain the differences in *ELOVL6* mRNA observed among animals. Additionally, the *ESRR-α* binding site overlapped one SRE motif (−532 to −524 bp), suggesting that the two polymorphisms identified in this region (*ELOVL6:c.-533C>T* and *ELOVL6:c.-534C>T*) may have an important role in selecting which transcription factor (*ESRR-α* or *SREBP1*) binds to its corresponding element. Further studies are needed to determine both the effect of the two polymorphisms to *ESRR-α* or *SREBP1* binding and how the selection of the transcription factor can affect *ELOVL6* gene expression. In liver, a second significant peak was also obtained in SSC4, in which the *HNF4γ*, *FABP4* and *FABP5* genes were identified. The porcine *HNF4γ* gene is a member of the hepatocyte nuclear receptor superfamily, which is highly homologous to *HNF4α*, suggesting that it may have a similar function in the regulation of hepatic genes [49]. The protein structure of *HNF4γ* revealed that fatty acids bind to its ligand binding pocket, acting as a regulatory molecule of *HNF4γ* [50]. In spite of the presence of the *HNF4γ* binding site in the *ELOVL6* promoter, the main relationship described between both genes is that deficiencies in *ELOVL6* gene expression deplete the newly synthesized fatty acids, which are coactivators of the *HNF4γ* gene, producing a decrease in *HNF4γ* activity [5]. Interestingly, preliminary results in our lab indicate higher expression levels of *HNF4γ* in liver than in backfat (data not shown). These data point towards *HNF4γ* as a regulator of *ELOVL6* gene expression in liver and suggest that the polymorphism proximal to or within the *HNF4γ* gene partially determines the differences in liver *ELOVL6* gene expression. On the other hand, *FABP5* is a protein that binds and transports long-chain fatty acids into the nucleus [51], where they can act as transcription factors on lipogenic genes, such as elongases [44]. Association analysis with muscle *ELOVL6* gene expression data allowed us to identify significant regions, but only general transcription factors were found. Data obtained suggest a minimal elongation activity in muscle, and probably the difference in mRNA levels between animals was caused by the intramuscular adipocytes, as was observed in previous pig studies [52]. Apart from the potential relevance of all of these genes located in significant regions in regulating *ELOVL6* gene expression, we cannot discard the involvement of other genes located in non-significant regions but with biological relevance in the regulation of *ELOVL6* expression, such as *SP1* and *SREBP* genes. In the present study, we have identified four SREBP binding sites in the pig *ELOVL6* promoter, but none of these cis-acting motifs was affected by *ELOVL6* polymorphisms. However, *SREBP* has been described as a weak transcriptional activator that requires interaction with additional regulators like NF-Y and SP1 to activate the transcription of genes involved in fatty acid metabolism [4,53]. Interestingly, a SNP disrupting a potential SP1 binding site has been identified in the *ELOVL6* promoter, not

discarding the involvement of this mutation in the differences of gene expression observed among tissues. Taking into account the regulatory networks necessary for transcriptional activation, further investigation is required to determine the role of these mutations in the *ELOVL6* expression together with the implication of tissue-specific factors and epigenetic modifications.

Finally, the results provided in the present study are both helpful for the understanding of molecular mechanisms governing important economical traits like meat quality, but also to improve the knowledge of human diseases related to obesity, including diabetes and metabolic syndrome. Fatty acid composition has been highly associated with insulin sensitivity, especially the ratio of C18 to C16 fatty acids, which is controlled by *ELOVL6* activity [5]. The accumulation of C16 fatty acids, observed in our study, has been related to protection against hepatic lipotoxicity and insulin resistance [5]. Palmitoleic acid, segregated by adipose tissue, greatly strengthens the insulin-signaling pathway, avoiding tissue insulin resistance and obesity-related diseases [44].

Conclusions

In this work, the interval for the C16:0 and C16:1(n-7) QTL in SSC8 has been reduced, allowing for the identification of *ELOVL6* as a positional candidate gene. The characterization of the coding and proximal promoter regions of the porcine *ELOVL6* gene allowed for the identification of several mutations, especially the *ELOVL6:c.-533C>T* polymorphism strongly associated with muscle and backfat percentages of palmitic and palmitoleic acids. Interestingly, this SNP was also related to *ELOVL6* expression levels in backfat and fatty acid content and elongation activity ratios in muscle and backfat. Thus, the *ELOVL6:c.-533C>T* polymorphism is a candidate causal mutation to explain the variation in palmitic and palmitoleic acid content observed in an Iberian x Landrace cross. Hence, this work provides the first report of the importance of the porcine *ELOVL6* gene in the metabolism of fatty acids and, subsequently, in meat quality traits in pigs, but further functional studies in model organisms and validation in independent pig populations are required to confirm this causal mutation.

Supporting Information

Table S1 List of SNPs for SSC8 linkage map and haplotype reconstruction.
(XLSX)

Table S2 Primers for *ELOVL6* mRNA sequencing (R), promoter sequencing (P) and RT-qPCR (RT) study.
(DOC)

Figure S1 Linkage disequilibrium among *ELOVL6* polymorphisms. Pattern of linkage disequilibrium analysis between the three identified polymorphisms on the *ELOVL6* gene and the most significant SNP detected in both GWAS and fine mapping. Figure colored from blue to red according to LD strength between consecutive markers.
(TIF)

Figure S2 Association of SNPs from SSC8 and *ELOVL6* polymorphims with palmitic and palmitoleic acid content. Association analyses of C16:0 (A) and C16:1(n-7) (B) with genotypes of markers included in the Porcine SNP60 Bead-Chip (Illumina). *ELOVL6* polymorphisms are included and labeled with a red circle. Positions in Mb are relative to the *Sscrofa10.2 assembly* of the pig genome. The horizontal, dashed line indicates the genome-wide significance level (FDR-based q-value$\leq$0.05).
(TIF)

Figure S3 GWAS for *ELOVL6* gene expression in backfat, liver and muscle. Association analyses of *ELOVL6* expression levels in backfat (A), liver (B) and muscle (C) with genotypes of markers included in the Porcine SNP60 Bead-Chip (Illumina). Positions in Mb are relative to the *Sscrofa10.2 assembly* of the pig genome. The horizontal, dashed line indicates the genome-wide significance level (FDR-based q-value$\leq$0.1).
(TIF)

Acknowledgments

We wish to thank Anna Mercadé and Anna Castelló for genotyping the porcine chip. We are indebted to Martien Groenen (Wageningen, NL) for providing information about SNP positioning in Assembly 10.2. This manuscript has been proofread by Chuck Simons, a native English-speaking university instructor of English.

Author Contributions

Principal investigator of the project: JMF. Conceived and designed the experiments: JC JMF MB. Performed the experiments: JC YR-C AP-O MB. Analyzed the data: JC YR-C AP-O MB. Contributed reagents/materials/analysis tools: JC YR-C AP-O DP-M JLN JMF MB. Wrote the paper: JC YR-C AP-O JMF MB.

References

1. Wood JD, Richardson RI, Nute GR, Fisher AV, Campo MM, et al. (2004) Effects of fatty acids on meat quality: a review. Meat Science 66(1):21–32.
2. Valsta LM, Tapanainen H, Mannisto S (2005) Meat fats in nutrition. Meat Sci 70(3):525–530.
3. Lichtenstein AH (2006) Thematic review series: patient-oriented research. Dietary fat, carbohydrate, and protein: effects on plasma lipoprotein patterns. J Lipid Res 47(8):1661–1667.
4. Moon YA, Shah NA, Mohapatra S, Warrington JA, Horton JD (2001) Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins. Journal of Biological Chemistry 276(48):45358–45366.
5. Matsuzaka T, Shimano H, Yahagi N, Kato T, Atsumi A, et al. (2007) Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. Nat Med 13(10):1193–1202.
6. Green CD, Ozguden-Akkoc CG, Wang Y, Jump DB, Olson LK (2010) Role of fatty acid elongases in determination of de novo synthesized monounsaturated fatty acid species. Journal of Lipid Research 51(7):1871–1877.
7. Clop A, Ovilo C, Perez-Enciso M, Cercos A, Tomas A, et al. (2003) Detection of QTL affecting fatty acid composition in the pig. Mammalian Genome 14(9):650–656.
8. Estellé J, Fernández AI, Pérez-Enciso M, Fernández A, Rodríguez C, et al. (2009) A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. Animal Genetics 40(6):813–820.
9. Yang B, Navarro N, Noguera JL, Muñoz M, Guo TF, et al. (2011) Building phenotype networks to improve QTL detection: a comparative analysis of fatty acid and fat traits in pigs. Journal of Animal Breeding and Genetics 128(5):329–343.
10. Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, et al. (2012) Genome-wide association study for intramuscular fatty acid composition in an Iberian x Landrace cross. Journal of Animal Science 90(9):2883–2893.
11. Jakobsson A, Westerberg R, Jacobsson A (2006) Fatty acid elongases in mammals: Their regulation and roles in metabolism. Progress in Lipid Research 45(3):237–249.
12. Guillou H, Zadravec D, Martin PGP, Jacobsson A (2010) The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. Progress in Lipid Research 49(2):186–199.
13. Leonard AE, Bobik EG, Dorado J, Kroeger PE, Chuang LT, et al. (2000) Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. Biochem J 350(3):765–770.
14. Perez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, et al. (2000) A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. Journal of Animal Science 78(10):2525–2531.

15. Mach N, Devant M, Díaz I, Font-Furnols M, Oliver MA, et al. (2006) Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. Journal of Animal Science 84(11):3039–3048.
16. Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, et al. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by Next Generation Sequencing Technology. PLoS One 4(8):e6524.
17. Estellé J, Mercadé A, Pérez-Enciso M, Pena RN, Silió L, et al. (2009) Evaluation of FABP2 as candidate gene for a fatty acid composition QTL in porcine chromosome 8. Journal of Animal Breeding and Genetics 126(1):52–58.
18. Druet T, Georges M (2010) A hidden markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. Genetics 184(3):789–798.
19. Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31(2):423–447.
20. Henderson CR (1984) Applications of linear models in animal breeding. Guelph: University of Guelph. 462 p.
21. Perez-Enciso M, Misztal I (2011) Qxpak.5: Old mixed model solutions for new genomics problems. BMC Bioinformatics 12(1):202.
22. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences 100(16):9440–9445.
23. Shin JH, Blay S, McNeney B, Graham J (2006) LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. Journal of Statistical Software 16.
24. Kumadaki S, Matsuzaka T, Kato T, Yahagi N, Yamamoto T, et al. (2008) Mouse Elovl-6 promoter is an SREBP target. Biochemical and Biophysical Research Communications 368(2):261–266.
25. Rozen S, Skaletsky H, Misener S, Krawetz SA (1999) Primer3 on the WWW for general users and for biologist programmers. Bioinformatics Methods and Protocols. In. Edited by Walker JM, vol. 132: Humana Press; 1999: 365–386.
26. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, et al. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics, 21(13):2933–2942.
27. Nygard AB, Jorgensen C, Cirera S, Fredholm M (2007) Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. BMC Molecular Biology 8(1):67.
28. Erkens T, Van Poucke M, Vandesompele J, Goossens K, Van Zeveren A, et al. (2006) Development of a new set of reference genes for normalization of real-time RT-PCR data of porcine backfat and longissimus dorsi muscle, and evaluation with PPARGC1A. BMC Biotechnology 6(1):41.
29. Overgard AC, Nerland A, Patel S (2010) Evaluation of potential reference genes for real time RT-PCR studies in Atlantic halibut (Hippoglossus Hippoglossus L.); during development, in tissues of healthy and NNV-injected fish, and in anterior kidney leucocytes. BMC Molecular Biology 11(1):36.
30. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta CT) method. Methods 25(4):402–408.
31. Pérez-Enciso M, Clop A, Folch JM, Sánchez A, Oliver MA, et al. (2002) Exploring alternative models for sex-linked quantitative trait loci in outbred populations: Application to an Iberian x Landrace pig intercross. Genetics 161(4):1625–1632.
32. Muñoz M, Alves E, Ramayo-Caldas Y, Casellas J, Rodríguez C, et al. (2012) Recombination rates across porcine autosomes inferred from high-density linkage maps. Animal Genetics 43(5):620–3.
33. Karim L, Takeda H, Lin L, Druet T, Arias JAC, et al. (2011) Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat Genet 43(5):405–413.
34. Horton JD, Shah NA, Warrington JA, Anderson NN, Park SW, et al. (2003) Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. Proceedings of the National Academy of Sciences 100(21):12027–12032.
35. Matsuzaka T, Shimano H, Yahagi N, Yoshikawa T, Amemiya-Kudo M, et al. (2002) Cloning and characterization of a mammalian fatty acyl-CoA elongase as a lipogenic enzyme regulated by SREBPs. Journal of Lipid Research 43(6):911–920.
36. Kim JB, Spotts GD, Halvorsen YD, Shih HM, Ellenberger T, et al. (1995) Dual DNA binding specificity of ADD1/SREBP1 controlled by a single amino acid in the basic helix-loop-helix domain. Molecular and Cellular Biology 15(5):2582–2588.
37. Turyn J, Stojek M, Swierczynski J (2010) Up-regulation of stearoyl-CoA desaturase 1 and elongase 6 genes expression in rat lipogenic tissues by chronic food restriction and chronic food restriction/refeeding. Molecular and Cellular Biochemistry 345(1):181–188.
38. Morcillo S, Martín-Núñez GM, Rojo-Martínez G, Almaraz MC, García-Escobar E, et al. (2011) ELOVL6 genetic variation is related to insulin sensitivity: A new candidate gene in energy metabolism. PLoS ONE, 6(6):e21198.
39. Ballard FJ, Hanson RW, Kronfeld DS (1969) Gluconeogenesis and lipogenesis in tissue from ruminant and nonruminant animals. Federation proceedings 28(1):218–231.
40. Jansen GR, Hutchison CF, Zanetti ME (1966) Studies on lipogenesis in vivo: Effect of dietary fat or starvation on conversion of [^{14}C]Glucose into fat and turnover of newly synthesized fat. Biochem J 99:10.
41. O'Hea EK, Leveille GA (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. The Journal of Nutrition 99(3):338–344.
42. Aarsland A, Chinkes D, Wolfe RR (1997) Hepatic and whole-body fat synthesis in humans during carbohydrate overfeeding. The American Journal of Clinical Nutrition 65(6):1774–1782.
43. Swierczynski J, Goyke E, Wach L, Pankiewicz A, Kochan Z, et al. (2000) Comparative study of the lipogenic potential of human and rat adipose tissue. Metabolism 49(5):594–599.
44. Cao H, Gerhold K, Mayers JR, Wiest MM, Watkins SM, et al. (2008) Identification of a lipokine, a lipid hormone linking adipose tissue to systemic metabolism. Cell 134(6):933–944.
45. Ruiz J, Cava R, Antequera T, Martín L, Ventanas JS, et al. (1998) Prediction of the feeding background of Iberian pigs using the fatty acid profile of subcutaneous, muscle and hepatic fat. Meat Science 49(2):155–163.
46. Yang KX, Ma JW, Guo YM, Guo TF, Zhao YG (2010) Correlations between fat depot traits and fatty acid composition in abdominal subcutaneous adipose tissue and longissimus muscle: Results from a White Duroc x Erhualian intercross F2 population. Journal of Animal Science 88(11):3538–3545.
47. Guillaumond F, Gréchez-Cassiau A, Subramaniam M, Brangolo S, Peteri-Brünback B, et al. (2010) Kruppel-Like Factor KLF10 is a link between the circadian clock and metabolism in liver. Molecular and Cellular Biology 30(12):3059–3070.
48. Sladek R, Bader JA, Giguère V (1997) The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. Molecular and Cellular Biology 17(9):5400–5409.
49. Drewes T, Senkel S, Holewa B, Ryffel GU (1996) Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. Molecular and Cellular Biology 16(3):925–931.
50. Wisely GB, Miller AB, Davis RG, Thornquest AD Jr, Johnson R, et al. (2002) Hepatocyte Nuclear Factor 4 is a transcription factor that constitutively binds fatty acids. Structure 10(9):1225–1234.
51. Damcott CM, Moffett SP, Feingold E, Barmada MM, Marshall JA (2004) Genetic variation in fatty acid-binding protein-4 and peroxisome proliferator-activated receptor gamma interactively influence insulin sensitivity and body composition in males. Metabolism 53(3):303–309.
52. Canovas A, Quintanilla R, Amills M, Pena R (2010) Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. BMC Genomics, 11(1):372.
53. Reed BD, Charos AE, Szekely AM, Weissman SM, Snyder M (2008) Genome-wide occupancy of SREBP1 and its partners NFY and SP1 reveals novel functional roles and combinatorial regulation of distinct classes of genes. PLoS Genet 4(7):e1000133.

PAPER III

Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition

Jordi Corominas, Yuliaxis Ramayo-Caldas, Anna Puig-Oliveras, Jordi Estellé, Anna Castelló, Estefania Alves, Ramona N. Pena, Maria Ballester and Josep M. Folch.

BMC Genomics. In revisión.

Analysis of porcine adipose tissue transcriptome reveals differences in *de novo* fatty acid synthesis in pigs with divergent muscle fatty acid composition

Jordi Corominas[1,2]*, Yuliaxis Ramayo-Caldas[1,2], Anna Puig-Oliveras[1,2], Jordi Estellé[3,4,5], Anna Castelló[1], Estefania Alves[6], Ramona N. Pena[7], Maria Ballester[1,2] and Josep M. Folch[1,2].

[1] Centre de Recerca en Agrigenòmica (Consorci CSIC-IRTA-UAB-UB), Edifici CRAG, Campus UAB, Bellaterra, 08193, Barcelona, Spain.

[2] Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Campus UAB, Bellaterra, 08193, Barcelona, Spain.

[3] INRA, UMR 1313 Génétique Animale et Biologie Intégrative (GABI), Equipe Génétique Immunité Santé, Jouy-en-Josas F-78352, France.

[4] AgroParisTech, UMR 1313 GABI, Jouy-en-Josas F-78352, France.

[5]CEA, DSV/iRCM/SREIT/LREG, Jouy-en-Josas F-78352, France.

[6] Departamento de Mejora Genética Animal, INIA, Ctra. de la Coruña km. 7, 28040, Madrid, Spain.

[7]Genètica i Millora Animal, IRTA Lleida, 25198, Lleida, Spain.

Email addresses:

JC: jordi.corominas@uab.cat

YRC: yuliaxis.ramayo@uab.es

APO: anna.puig@cragenomica.es

JE: jordi.estelle@jouy.inra.fr

AC: anna.castello@uab.es

EA: esalves@inia.es

RP: romi.pena@prodan.udl.cat

MB: maria.ballester@cragenomica.es

JMF: josepmaria.folch@uab.cat

*__Corresponding author:__ Jordi Corominas Galbany. Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra 08193, Spain.

ABSTRACT

Background: In pigs, adipose tissue is one of the principal organs implicated in the regulation of lipid metabolism. It is particulary involved in the overall fatty acid synthesis with consequences in other lipid-target organ such as muscle and liver. With this in mind, we have used massive parallel high-throughput sequencing technologies to characterize the porcine adipose tissue transcriptome architecture in six Iberian x Landrace crossbred pigs showing extreme phenotypes for intramuscular fatty acid composition.

Results: High-throughput RNA sequencing was used to generate a whole characterization of adipose tissue (backfat) transcriptome. A total of 4,130 putative unannotated protein-coding sequences were identified in the 20% of reads which mapped in intergenic regions. Furthermore, the 36% of the unmapped reads were represented by interspersed repeats, being SINEs the most abundant elements. Differential expression analyses identified 396 candidate genes between divergent animals for intramuscular fatty acid composition. The 62% of these genes (247/396) were overexpressed in the group of pigs with higher content of intramuscular SFA and MUFA, while the remaining 149 showed higher expression in the group with higher content of PUFA. Pathway analysis related these genes to biological functions and canonical pathways controlling lipid and fatty acid metabolisms. In concordance with the phenotypic classification of animals, the major metabolic pathway differentially modulated between groups was *de novo* lipogenesis, being the group with more PUFA the one that down-expressed lipogenic genes.

Conclusions: These results will help in the identification of causal genetic variants at loci that affect fatty acid composition traits. The implications of these results range from the improvement of porcine meat quality traits to the application of pig as an animal model of human metabolic diseases.

Keywords: RNA-Seq, transcriptome, adipose tissue, pork, *de novo* lipogenesis.

BACKGROUND

The pig (*Sus scrofa*) is one of the most important livestock animals due to its economical importance in the alimentary industry, but it is also an interesting biomedical model for human diseases [1]. Over the last decades, the genetic selection in commercial pig breeds has greatly improved meat production efficiency at the expense of reducing the sensorial and technological properties of meat. These changes are mainly caused by the reduction in intramuscular fat (IMF) content and alterations in fatty acid (FA) composition, both critical for various meat quality attributes such as muscle color, firmness, water holding capacity and also important nutritional aspects [2]. In this regard, FA composition of food has also become a critical aspect in human nutrition: a high consumption of SFA has been associated with obesity, high plasma cholesterol and cardiovascular diseases [3, 4], while replacing SFA with MUFA or PUFA decreases serum LDL cholesterol and total cholesterol, reducing the risk of coronary heart disease [5, 6].

Factors other than dietary intake have been less characterized in relation with tissue lipid composition. These include the role of candidate gene genotypes on lipid and FA metabolism [7-12]. In this context, studies in pigs can find a dual purpose: first, to study the genetics of food, i.e., how the genotype of the animal influences the FA content and profile of meat; and secondly, as an animal model for nutrigenomic studies or human metabolic diseases.

Liver, adipose tissue and skeletal muscle are the principal organs implicated in the regulation of lipid metabolism. The adipose tissue is an organ responsible for energy storage in form of lipids and, in pigs, is the major source of circulating free FAs (FFA) [13]. It also acts as a major endocrine organ producing adipocytokines like TNFα, peptide hormones such as leptin, adiponectin, estrogen and resistin and lipid hormones (lipokines) such as palmitoleate, all of them involved in the maintenance of metabolic homeostasis [14, 15]. Furthermore, pig adipose tissue has a greater contribution to overall FA synthesis than liver [13]. Thus, the characterization of the transcriptome landscape of this organ may be relevant for the improvement of pork nutritional quality.

The development of next-generation sequencing (NGS) methods has provided new tools for both transcriptome characterization and gene expression profiling. RNA-Seq technique is based on sequencing the poly-A RNA fraction and allows whether to characterize isoforms from known genes or to discover novel predicted coding genes [16]. To date, the number of RNA-Seq analysis in livestock is still scarce, some recent reports have focused in the study of organs [17, 18], animal products such as milk [19, 20] or embryos [20]. Thus, in 2011, Esteve-Codina *et al.* compared the pig gonads of two individuals from two breeds (Iberian and Large White). The same year, Chen *et al.,* (2011) analyzed the transcriptome of three pig tissues (liver, *longissimus dorsi* muscle and abdominal fat) in two full-sib F_2 females with extreme phenotypes in growth and fatness (White Duroc x Erhualian). Furthermore, a liver RNA-Seq study was performed using four animals from genetically different porcine breeds (Berkshire, Duroc, Landrace and Yorkshire) [21].

In a previous work of our group, the liver of ten Iberian x Landrace backcrossed pigs classified in two phenotypically-extreme groups for intramuscular FA composition (five per group), were analyzed using RNA-Seq [22]. This study identified 55 genes differentially-expressed in liver that may play a crucial role in muscle lipid composition. Nevertheless, muscle lipids are derived not only from liver (mostly dietary lipids), but also from adipose tissue (mostly *de novo* lipogenesis) [15, 23]. Therefore, the aim of the present study was to investigate the contribution of backfat transcriptome to the FA content and profile of intramuscular fat in pigs. The two main goals of our study are: (i) the identification of genes and pathways differentially-expressed in the backfat of Iberian x Landrace crossbred pigs (BC1-LD) showing extreme phenotypes for intramuscular FA composition; and (ii) to describe transposable elements and new putative protein-coding genes in the transcriptome of pig backfat. Combining the new adipose tissue transcriptome data with the already available liver transcriptome information will allow us to study the expression of genes regulating the overall lipid metabolism in pigs.

RESULTS

Characterization of pig adipose tissue transcriptome

In a previous work by our group, animals from an Iberian x Landrace backcross (BC1-LD) were analyzed with a Principal Component Analysis (PCA) to describe the phenotypic variation of traits related to carcass quality and intramuscular FA composition [22]. The score information of the first principal component in PCA was used to classify the BC1-LD animals in two groups High (H) and Low (L) and the hepatic transcriptome from the most extreme females (five per group) was evaluated using an RNA-Seq approach [22]. In the present study, a total of six females (three per group) were selected for RNA sequencing of their backfat tissue. Pedigree information was used to avoid the selection of sibs in the same group. When phenotypic means between groups were compared, 54% of the traits showed significant statistical differences (14/26) (Table S1). In summary, the H group showed higher levels of intramuscular polyunsaturated fatty acids (PUFA) and group L showed higher levels of saturated fatty acids (SFA) and monounsaturated fatty acids (MUFA) (Table S1). The lack of statistical significance of palmitoleic (C16:1 n-7) and heptadecenoic (C17:1) acids content in comparison with Ramayo-Caldas *et al.* (2012) may be explained by the lower sample size.

Sequencing yield around 236M of 75 bp paired-end reads and approximately 84% of reads were mapped against the reference pig genome assembly Sscrofa10.2 using the Tophat software. Percentages of mapped reads observed per individual ranged from 80% to 87%, and were equally distributed in the H and L groups. These values were higher than the percentages reported in previous porcine transcriptome studies: 71.4-77.8% [22], 61.4-65.6% [24], 66.7 [25] and 54 % [21]. IntersectBed tool from BEDtools was employed to calculate the proportion of mapped reads annotated in exons, introns and intergenic regions. The highest percentages of reads were mapped to exons (66-71%), while 19-20% fell in intergenic regions and the lowest percentages were located in introns (10-14%) (Table 1). The proportion of reads mapped to exons was slightly higher than the values reported by Ramayo-Caldas *et al.,* (2012) and Chen *et al.,* (2012), due to the better pig genome assembly (Sscrofa10.2) and annotation

version employed in our work. Finally, transcripts generated from assembling the short reads by cufflinks fell mostly into annotated exons (42-47%). The remaining reads were classified into the following categories: intron retention events (7-12%), intergenic transcripts (17-20%), potentially novel isoforms of genes (19-21%), pre-mRNA molecules (3-4%) and polymerase run-on fragments (2-3%) (Table S2). The percentage of intergenic transcripts represented the third category in read abundance, which is relevant for detecting putative coding transcripts or new transposable elements not described in the current version of the pig genome.

Exploring for novel coding transcripts and transposable elements in the adipose tissue transcriptome

Transcript annotation performed with cufflinks showed a mean value of 10,862 total unknown intergenic transcripts (Table S2). This value doubles the number of intergenic transcripts detected in previous studies [22, 25]. To determine which of these transcripts encoded a protein, Augustus software [26] was used and, as it was expected, the total amount of predicted proteins was also higher compared to previous studies: 4,130 predicted proteins against 326 [22] and 714 [25]. Our analysis showed an improvement in the sequence length and sequence coverage. BLASTP analysis was performed to compare the putative proteins predicted by Augustus against the predicted proteins reported by Ramayo-Caldas *et al.,* (2012) in liver and Esteve-Codina *et al.,* (2011) in gonads. A total of 269 new putative proteins fitted in with the previously described ones: 93 putative proteins were expressed in liver transcriptome (34%) and 211 were expressed in gonad transcriptome (78%). Additionally, a functional annotation was performed using BLAST2GO software. BLASTP analysis of the 4,130 predicted proteins revealed that 2,100 proteins (50.8%) displayed significant similarity with existing protein sequences (the top hit species was *Sus scrofa*, 59.1%). These proteins corresponded to: 16 novel computationally predicted and 1,108 known human proteins, 361 novel and 598 known bovine proteins and 997 novel and 599 known porcine proteins. Hence, the number of novel predicted proteins was lower in better described genomes (human and bovine). From the 2,100

predicted proteins, only 1,226 were functionally annotated with at least one gene ontology (GO) term. At the third GO level, the most represented biological terms were 'primary metabolic process' (8% of predicted proteins), 'cellular metabolic process' (8%) and 'regulation of biological process' (7%). According to the molecular function, 'protein binding' (25%), 'ion binding' (17%) and 'nucleic acid binding' (16%) were the most represented categories. In the cellular compartments category, 'cell part' (29%) was the most represented term, followed by 'membrane-bounded organelle' (20%) and 'organelle part' (16%) (data not shown). Furthermore, the main metabolic pathways represented were purine metabolism (49 sequences), pyrimidine metabolism (11), phosphatidylinositol signaling system (11) and inositol phosphate metabolism (10); these and other pathways observed are shown in Table S3.

Repetitive elements (RE) were identified in the adipose tissue transcriptome using the RepeatMasker software. The total interspersed repeats represented the 36% of the intergenic transcripts (Table S4), a percentage higher than those observed in previous porcine transcriptome analyses: 5.8-7.3% [22] and 7.3% [25]. As described above, the differences obtained could be explained by the assembly used for each analysis. The description of these regions has been improved in the current assembly (Sscrofa10.2), as demonstrated in the results observed. The major part of the RE were classified as SINEs (n=92,007), representing the 13.96% of the transcriptome sequenced. The second group was LINEs (n=43,428), but these RE constitute a larger part of the transcriptome than SINEs due to the bigger size of LINEs (15.88% of adipose tissue transcriptome). The remaining RE were classified as: LTR elements (4.04%), DNA elements (2.06%) and unclassified RE (0.02%).

Gene expression analysis

Counting the reads mapped in each gene, around 15,747 annotated genes were expressed in adipose tissue with similar amounts between groups (L= 15,608-16,001, H=15,433-15,834). Taking into account only those genes with a minimum mean of 20 reads per gene in at least one of the extreme groups, 13,086 expressed genes were selected. Gene expression distribution was similar in both groups, classifying 1% of the

selected genes between 0-20 mapped reads; 27-28% among 20-200 mapped reads; most part of genes (46-47%) had between 200-2,000 mapped reads; 23% among 2,000-20,000 mapped reads, and the remaining genes (2%) more than 20,000 mapped reads (Figure S1). Mean gene expression levels were highly correlated between groups (r =0.98 between H and L groups) indicating that most genes were similarly expressed in both groups. Five of the 6 individuals were also assayed with the Gene-Chip® Porcine microarray (*Affymetrix*, Santa Clara, CA) to analyze the gene expression of 20,201 *Sus scrofa* genes. After signal normalization, correlation between the expression data obtained by microarrays and RNA-Seq was calculated. All animals showed a high Spearman correlation (r=0.65-0.68) (Figure S2) in accordance with previous studies of the porcine transcriptome [22, 25, 27]. Genes with intermediate expression values had higher correlation between technologies than genes with low or high expression values. This same pattern had already been observed in previous studies and it is explained by the higher dynamic range of RNA-Seq analyses [28, 29]. Finally, the top 100 expressed genes showed an overrepresentation in biological gene ontologies related to hormone-sensitive lipase-mediated triacylglycerol hydrolysis, peroxisome proliferator-activated receptors (PPAR) signaling pathway, lipid digestion, mobilization and transport, carbohydrate-responsive element-binding protein (ChREBP) activates metabolic gene expression, pyruvate metabolism and biosynthesis of unsaturated FAs.

Differential gene expression between animals with extreme phenotypes of intramuscular FA composition

Biological functions overrepresented in the differential expressed genes are candidate functions to explain the variation in intramuscular FA composition among the analyzed animals. Gene expression data from each group was compared using the DESeq software, which allowed the detection of genes differentially-expressed between groups. The data were filtered, discarding those genes with a mean of less than 20 reads mapped in one of the extreme groups. Finally, a total of 13,086 genes were used to perform the differential expression analysis using a standard cut-off of: |fold

change| ≥ 1.2, p-value ≤ 0.01 and a q-value ≤ 0.1 (Figure 1). DESeq software identified a total of 396 differentially-expressed genes (Figure 2), from which 247 were up-regulated in the L group and the remaining 149 genes were down-regulated in the same group (Table S5). It is noteworthy that 9 of the differential expressed genes in adipose tissue (*IVD*, *CRABP2*, *SLC2A12*, *AACS*, *RBP1*, *ACADL*, *APOB* and *THEM5*) have previously been reported to be associated with the profile of intramuscular FA composition in a GWAS study in the same animal population (Table 2) [30]. The gene overlapping between both studies gives new evidences about the relevance of lipid metabolism in adipose tissue in determining the intramuscular FA composition, therefore they should be considered as interesting candidate genes for pig meat quality traits in future studies.

To gain insight into the metabolic processes that differed between both groups, the list of 396 differentially-expressed genes was analyzed using the core analysis function included in Ingenuity Pathways Analysis (IPA). The main biological functions identified were related with cancer, lipid concentration, synthesis of lipids, homeostasis of blood, and FA metabolism (Table 3). Interestingly, within the general representation of lipid metabolism, it is remarkable *de novo* FA synthesis pathway where the most relevant genes were overexpressed in L group. The ATP citrate lyase (ACLY) (p-value = $2.47x10^{-04}$ and fold change = -2.07) is the primary enzyme responsible for the synthesis of cytosolic acetyl-CoA from citrate and CoA in many tissues [31]. Although citrate is the main substrate for initializing *de novo* lipogenesis in pigs, acetic acid is also used to produce cytosolic acetyl-CoA in species with extensive forestomach and hindgut fermentation such as rabbits, cattle, sheep and goats. The gene responsible for acetic acid conversion is the *acyl-CoA synthetase short-chain family member 2* (*ACSS2*) (p-value = $9.58x10^{-05}$ and fold change = -2.12) which was also overexpressed in L group suggesting that cytosolic acetate may also contribute to increase the lipogenesis in pig adipose tissue. Both genes provide the substrate necessary for acetyl-CoA carboxylase alpha (ACACA) to initialize *de novo* FA synthesis. The *ACACA* gene (p-value = $9.2x10^{-07}$ and fold change = -2.67) encodes the enzyme that catalyzes the carboxylation of acetyl-CoA to malonyl-CoA [32]. This gene was previously reported as a candidate gene for a porcine quantitative trait loci (QTL) affecting the percentages of palmitoleic,

stearic and vaccenic acids in chromosome 12 [32, 33]. The product of ACACA is catalyzed by the *fatty acid synthase* (*FASN*) gene (p-value = $3.28x10^{-16}$ and fold change = -9.3) to synthesize palmitate (C16:0) in the presence of NADPH [34, 35]. These steps are important for the conversion of intermediate metabolites into FAs contributing to the synthesis of cellular lipids and storage of fats. The overexpression of these genes may explain the higher levels of C16:0 observed in L group animals. The C16:0 fatty acid produced in the cytosol is transferred to endoplasmic reticulum (ER) membranes, where *ELOVL fatty acid elongase 6* (*ELOVL6*) gene (p-value = $7.2x10^{-04}$ and fold change = -1.93) and *stearoyl-CoA desaturase* (*SCD*) gene (p-value = $1.95x10^{-06}$ and fold change = -7.87) are sequentially involved to produce C18:1 [36]. Some studies performed in our group proposed the *ELOVL6* gene as a candidate gene explaining a QTL in porcine chromosome 8, related to C16:0 and C16:1n-7 in the same animal material [37]. Moreover, it is worth noting the overexpression in the L group of genes related to glucose metabolism, such as *glucose-6-phosphate dehydrogenase* (*G6PD*) or *malic enzyme 1* (*ME1*) whose functions in glucose metabolism contribute to the initial steps of lipogenesis. The porcine *G6PD* gene (p-value = $8.2x10^{-03}$ and fold change = -1.70) product is a cytosolic enzyme responsible for the first step of a chemical pathway that converts glucose to ribose-5-phosphate. On the other hand, the porcine *ME1* gene (p-value = $2.1x10^{-04}$ and fold change = -2.12) encodes a tetrameric NADP-dependent enzyme that catalyzes the reversible oxidative decarboxylation of L-malate to pyruvate [38]. This gene was previously reported as a candidate gene for a porcine QTL affecting fat deposition in chromosome 1 [38]. Both genes contribute to produce pyruvate, which is transported and metabolized in the mitochondria to produce citrate. This citrate is the substrate of the differentially-expressed gene *ACLY* and the initial molecule of pig lipogenesis. This complete metabolic pathway, from glucose metabolism to *de novo* lipogenesis, is well represented in the first IPA-generated network identified as "Lipid metabolism, Nucleic acid metabolism, Small molecule biochemistry" (score 36, focus molecules 23) (Figure 3) (Table S6). In clear consistency with the generated networks, the most representative canonical pathway significantly modulated between groups was the LXR/RXR activation (p-value = $6.19x10^{-10}$) (Table 4), which regulates the whole *de novo* FA synthesis pathway (Figure 3). Other canonical pathways modulated between groups were ethanol degradation II, noradrenaline and

adrenaline degradation, acute phase response signaling and super-pathway of serine and glycine biosynthesis I (Table 4).

Finally, the lower representation of *de novo* FA synthesis observed in H group was validated by the overexpression of the *thyroid hormone responsive* (*THRSP*) (p-value = $3.36x10^{-06}$ and fold change = 3.46). This gene is abundantly expressed in lipogenic tissues and plays an important role in the biosynthesis of triglycerides with medium-length FA chain and in modulating the lipogenesis [39]. The coexpression of this gene with the *MID1 interacting protein 1* (*MID1IP1*) leads to form an heterodimer between these proteins, which in turn inhibits the *MID1IP1* function of up-regulating the ACACA enzyme [39]. Moreover, several genes overexpressed in H group were related to lipid transport, such as albumin (*ALB*) (p-value = $1.23x10^{-07}$ and fold change = 3.05), which is a soluble protein responsible for carrying steroids, long-chain FAs and thyroid hormones. *ALB* also plays an important role controlling the level of cellular uptake of long-chain FAs, through interactions with the cell surface [40]. Furthermore, a low expression of *apolipoprotein B* (*APOB*) gene was observed in all animals, despite that a clear DE was observed. APOB is essential for the synthesis and secretion of chylomicrons and very low density lipoproteins (VLDL), but also acts as a ligand for the cellular binding and internalization of LDL particles [41]. Finally, the *apolipoprotein C-III* (*APOC3*) gene, which is also a lipid transporter, was overexpressed in H group (p-value = $3.4x10^{-04}$ and fold change = 4.28).

DISCUSSION

A comprehensive exploration of the pig backfat adipose tissue transcriptome was carried out in this study by using high-throughput RNA sequencing. In comparison to other technologies such as microarrays, RNA-Seq represents a relevant improvement in terms of lower background noise, larger dynamic range and higher technical reproducibility. In reference-based RNA-Seq studies performed in domestic species, such as the pig, the quality of the genome sequence annotation is critical to obtain an optimal representation of transcriptomes. The improved genome sequence length and sequence coverage of the current assembly (Sscrofa10.2) [42] caused a 10% increase in the number of mapped reads in comparison with studies that used previous versions. Nevertheless, the current gene annotation still needs to be improved in order to determine the function of approximately 20% of reads, which mapped in intergenic regions. The assembled transcripts in those regions allowed the prediction of 4,130 putative proteins; this amount higher than those found in previous studies may also be caused by the presence of new intergenic sequences that have not been described in the previous versions of pig genome assembly. Furthermore, the conservative approach used in Ramayo-Caldas *et al.,* (2012), in which only those transcripts expressed in at least four of the five animals of each group were considered, could aid to explain the differences obtained (4,130 vs. 326). The 2,100 proteins displaying significant similarity with existing protein sequences allowed the validation of a large number of novel computationally predicted proteins. As expected, the lowest number of novel predicted proteins were observed in the best annotated genomes (human and bovine). Finally, improvements in the Sscrofa10.2 version also allowed a better detection of new repetitive elements in the porcine genome.

The strong correlation observed between gene expression values obtained by RNA-Seq and Affymetrix microarrays allowed us to confirm the reproducibility of the data. As expected, the most represented functions into the top 100 expressed genes showed relevant functions closely related to adipose tissue metabolism such as storage and hydrolysis of triglycerides, metabolism of glucose and synthesis of unsaturated FAs. Additionally, key regulatory pathways were also detected: the PPAR signaling pathway, which is important for the induction of preadipocyte differentiation and FA storage

[43], or the ChREBP transcription factor, which has emerged as a major mediator of glucose action on lipogenic gene expression and as a key regulator of lipid synthesis [44].

On the other hand, the differential expression analysis identified an important number of genes implicated in determining the phenotypic differences observed in intramuscular FA composition between two groups of animals (H and L) from an Iberian x Landrace backcross. As described elsewhere, the most remarkable differences between Iberian and Landrace breeds were the strong development of backfat tissue and the higher content of IMF in Iberian pigs in comparison to Landrace pigs [45]. The breed effect also translates on the backfat FA composition, showing Iberian pigs higher percentages of C16:0 and MUFA (particularly C18:1) and lower content of PUFA than commercial breeds [45-47]. Serra *et al*. (1998) suggested that differences between Iberian and Landrace pigs may be an indirect consequence of differences in *de novo* lipid synthesis. In our animal material, pigs of the L group had a higher content of SFA and MUFA, similar to Iberian pigs. Meanwhile, animals of the H group had higher percentages of PUFA, as observed in Landrace animals. Taking our results into account, the main difference observed between the two groups was the differential expression of genes implicated in *de novo* FA synthesis pathway, confirming the relevance of this pathway in determining FA composition. In support of our results, several key genes involved in glucose metabolism and *de novo* FA synthesis had been previously reported as differential expressed between pigs with lean phenotype (Landrace) and pigs with obese phenotype (Rongchang pig) in adipose tissue [48]. Hence, differences in rates of tissue lipid accumulation between leaner pigs compared to fatter pigs are attributable to genetic factors resulting in different patterns of expression of anabolic and oxidative lipid metabolism genes.

FFAs derived from adipose tissue and VLDL-associated triglycerides derived from liver are important sources of FA supply to muscle, playing an important role in determining the intramuscular FA composition [10, 15, 23]. In fact, high positive phenotypic correlations between adipose tissue and muscle FA composition were found for C14:0 ($r_{C14:0}$=0.59, p-value = $1.14x10^{-14}$), C16:0 ($r_{C16:0}$=0.72, p-value = $2.2x10^{-16}$) and C17:0 ($r_{C17:0}$=0.65, p-value = $2.2x10^{-16}$) and moderate positive phenotypic correlations were

found for C16:1 n-7 ($r_{C16:1\ n-7}$=0.50, p-value = $3.3x10^{-10}$), C16:1 n-9 ($r_{C16:1\ n-9}$=0.47, p-value = $3.9x10^{-09}$), C18:0 ($r_{C16:0}$ =0.43, p-value = $9.8x10^{-08}$) and C18:1 n-9 ($r_{C18:1\ n-9}$=0.40, p-value = $9.2x10^{-07}$) in our animal material [49]. In a previous study [22] that analyzed the liver transcriptome on the same groups of animals, it was suggested that a higher PUFA bioavailability observed in H group induced a greater stimulation of both peroxisomal and mitochondrial β-oxidation and reduced triglyceride and cholesterol synthesis. This increase of FA oxidation observed in animals of H group, jointly with the ketone body production, is a "glucose sparing" mechanism of regulation in fasting conditions [43] in which the animals were at slaughter. In adipose tissue, fasting condition induces the lipolysis of triglycerides store and the blood transport of ALB-associated FFAs bound to organs such as heart and skeletal muscle to fulfill their energy requirements [50]. Previous studies performed in 3T3-L1 pre-adipocytes demonstrated that overexpression of *ALB* stimulates long-chain FAs uptake by direct interaction with adipose cells and suggested that this stimulatory effect may be a general phenomenon in other types of cells [40]. Hence, data obtained may explain the greater uptake of fatty acids into hepatocytes and their degradation in the β-oxidation pathway in liver of the H group animals. On the other hand, it is well known the negative effect that dietary PUFA causes on *de novo* FA synthesis [44, 51, 52] and this effect was also observed in our data. The down-regulation of this pathway in the group with higher content of PUFA (that is, the H group) may be caused by the inhibitory effect of *n-3* and *n-6* PUFA on the expression of *receptor subfamily 1, group H, member 3* (*NR1H3*) [52]. The *NR1H3* gene, also called *liver X receptor* (LXR), is a nuclear receptor which is highly expressed in adipose tissue. Studies performed in $NR1H3^{(-/-)}$ mice showed a decrease on *de novo* FA synthesis, due to the down-regulation of *SREBP1* and their target genes [53] (Figure 3). Other studies confirmed that *ChREBP* is a key transcriptional regulator for the coordinated inhibition of glycolytic and lipogenic genes by PUFA [54]. PUFA is also suppressing *ChREBP* gene function in a LXR-dependent manner, increasing its mRNA decay and altering ChREBP protein translocation from the cytosol to the nucleus [44]. In addition, we cannot rule out a direct inhibition of *SCD* expression by PUFA [55] in animals of H group. The repression of *SCD* increases the intracellular pool of saturated fatty acyl CoAs inhibiting the ACACA enzyme and *de novo* lipogenesis and activating the *carnitine*

palmitoyltransferase 1 (*CPT1*) gene, which is responsible of the rate-limiting step in the import and oxidation of FAs into the mitochondria [56].

Thus, altogether, our results are in agreement with a functional and anatomical separation of *de novo* lipid synthesis and -oxidation in the porcine adipose tissue and liver, respectively. This suggests a tightly coordinated process between different hormones (peptides and/or lipids), transcription factors and nuclear receptors to avoid the simultaneously activation of antagonistic pathways. However, there is a high controversy in explaining the relevance of β-oxidation in porcine adipose tissue. PPARA is considered the main transcription factor controlling fatty acid oxidation. There are some studies that described a greater expression of pig *PPARA* in adipose tissue than liver, suggesting that adipose tissue could oxidize fatty acids to any extend [57]. In contrast, other authors did not find *PPARA* expression in porcine adipose tissues [58]. Consistent with this, we found higher levels of *PPARA* expression in liver compared to adipose tissue which suggests an important role of liver in porcine β-oxidation.

Different studies have determined the importance of several adipose tissue-derived hormones in the regulation of systemic carbohydrate and lipid homeostasis [59, 60]. Communication between adipose tissue and distant organs has been previously described through the lipokine palmitoleate (C16:1 n-7) which strongly stimulates muscle insulin action while suppresses hepatosteatosis [15]. Studies performed *in vivo* in humans showed a clear increase of *SREBP1c* caused by insulin in muscle and consequently the induction of key lipogenic enzymes [61, 62]. The mean comparison of C16:1 n-7 FA composition between L and H groups showed suggestive differences in both muscle (Table S1) and adipose tissue (data not shown). Hence, different levels of C16:1 n-7 may determine a differential systemic regulation that may explain the phenotypic variations observed between groups. Nevertheless, further studies are needed to better understand the mechanisms affecting this regulation and its consequences in phenotypic traits. Finally, other adipose tissue-derived hormones such as leptin or adiponectin cannot be discarded, as well as other interesting genes not annotated in the current pig genome assembly. The identified differentially-expressed genes seem to be relevant in controlling the overall FA composition in adipose tissue and muscle, and should be considered as candidate genes for meat

quality traits in pigs. The knowledge of these genes and their regulatory networks may help in the design of new strategies for improving pork meat quality by increasing the ratios MUFA/SFA and n-3/n-6 PUFA [2]. The maintenance of these ratios is essential to reduce the imbalanced FA intake of today's consumers and to avoid several diseases, including cancers and coronary heart disease. The high similarities between pigs and human in body size and other physiological/anatomical features, converts the pig in an excellent biomedical model for human disease. Hence, results provided in the present study are also helpful to improve the knowledge of human diseases related to obesity, including diabetes and metabolic syndrome.

CONCLUSIONS

In this study, we provide a global view of adipose tissue (backfat) transcriptome of six pigs and extensive new knowledge about transposable elements, new putative protein-coding genes and the expression levels of known genes in adipose tissue. Animals were classified in two groups according to their intramuscular FA composition and 396 genes were found to be differentially-expressed between groups. These genes belong to molecular functions and gene networks related with lipid and FA metabolism. Pathway analysis showed a different modulation of lipogenesis between phenotypically extreme animals, probably caused by differences in PUFA levels (mainly linolenic and α-linolenic). Finally, it is well-known the crucial role of intramuscular FA composition in the technological and the nutritional and organoleptic quality of pork meat. Hence, this study will allow the identification of candidate genes and gene networks for FA composition traits which may help in the design of better selection strategies to improve porcine meat quality traits.

METHODS

Animal material

The IBMAP cross was originated by crossing three Iberian (Guadyerbas line) boars with 31 Landrace sows [49]. Animals used in this study belong to a backcross (BC1-LD) generated by crossing five F1 (Iberian x Landrace) boars with 26 Landrace sows and producing 144 backcrossed animals. All pigs were raised in a normal intensive system and feeding was *ad libitum* with a cereal-based commercial diet. Pigs were slaughtered at an average age of 179.8 ± 2.6 days following national and institutional guidelines for the ethical use and treatment of animals in experiments. Samples of adipose tissue (backfat) were collected at slaughterhouse, snap-frozen in liquid nitrogen and stored at -80ºC until analyzed. A total of 48 traits related with growth, carcass quality and intramuscular FA composition were measured. In Ramayo-Caldas *et al.* (2012), the phenotypic information from twenty-six of the total traits was used to classify BC1_LD animals in two groups (H and L) according to the first component of a PCA [22]. A total of six animals were selected for the study, considering pedigree information representing the parental genetic diversity and that only females were retained for RNA sequencing (three per group). Phenotypic mean comparison between groups was performed using R.

RNA isolation, library preparation and sequencing

Total RNA was isolated from backfat using the *RiboPure™ Isolation of High Quality Total RNA* (*Ambion®, Austin, TX*) following the manufacturer's recommendations. RNA was quantified using the *Nano-Drop ND-1000 spectrophotometer* (*NanoDrop products,* Wilmington, USA) and checked for purity and integrity in a *Bioanalyser-2100* (*Agilent Technologies, Inc.,* Santa Clara CA, USA). For each sample, one paired-end library with approximately 300 bp insert size was prepared using TruSeq RNA Sample Prep Kit v2 (*Illumina, Inc.*, San Diego CA, USA). Libraries were sequenced, in CNAG (Centro Nacional de Análisis Genómico), on an Illumina HiSeq2000 instrument (*Illumina, Inc.*,

San Diego CA, USA) that generated paired-end reads of 75 nucleotides. More than 462 million reads were generated in this study.

Mapping, assembling and annotation of reads

In order to map all reads generated, the software TopHat v2.0.1 [63, 64] was employed using as reference the version 10.2 of the pig genome (Sscrofa10.2) and the annotation database Ensembl Genes 67 [http://www.ensembl.org/info/data/ftp/index.html]. Tophat was used with an expected mean inner distance between mate pairs of 160. Quality control and reads statistics were determined with FASTQC [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]. Transcripts were assembled and quantified by Cufflinks v2.0.2 [64, 65] with a minimum alignment count per *locus* of 20. Additionally, for counting the number of reads mapping to exons, introns and intergenic positions the tool intersectBED from BEDtools was used [66].

Orthology detection and transposable element analysis

Intergenic expressed regions, according to the current pig genome assembly (Sscrofa10.2), were extracted with Cuffcompare [65]and custom Python and R scripts. Putative coding transcripts were identified with Augustus [26], providing exon boundaries and allowing complete protein translation. Functional annotation was performed by using BLASTP option from BLAST2GO, with the following parameters: *E*-value hit filter 1.00E-6, annotation cutoff 55, gene ontology (GO) weight 5 and HSP-hit coverage cutoff 0 [67]. Additionally, InterProScan specific tool implemented in the BLAST2GO software and the ANNEX data set were employed to refine the functional annotations. GO terms were summarized according to the three principal GO categories: cellular component, biological process and molecular functions. Enzyme mapping of annotated sequences was performed using direct GO to enzyme mapping and used to query the Kyoto Encyclopedia of Genes and Genomes (KEGG) to define the main metabolic pathways involved [68, 69].

Furthermore, the software RepeatMasker (http://www.repeatmasker.org/) version open-3.30 was employed with the 'quick search' option and 'pig' species, in order to indentify repetitive and transposable elements in the adipose tissue transcriptome.

The Search Engine used was NCBI/RMBLAST with the complete database rm-20120418.

Gene expression quantification and correlation analysis with expression microarrays

Qualimap v0.5 software was employed to count the number of reads mapped for each gene and the total number of counts were considered as expression values [70]. Correlations between mean expression values between groups were calculated. Five of the animals sequenced were also assayed with high-density oligonucleotide microarray chips (*GeneChip® Porcine*) from *Affymetrix* (Santa Clara, CA) containing a total of 23,937 probe sets (23,256 transcripts), representing 20,201 *Sus scrofa* genes. Microarrays were hybridized and scanned at the *Institut de Recerca Hospital Universitari Vall d'Hebron* (Barcelona, Spain) following *Affymetrix* standard protocols. The Gene-Chip Operating Software (GCOS) was used to generate expression data and probes were normalized and adjusted for background noises with the GCRMA R package [71]. All probes correspond to a total of 7,885 Ensembl gene IDs expressed in backfat and these genes were used to estimate the Spearman correlation between the log2 expression values of genes analysed by RNA-Seq and microarrays.

Differential gene expression analysis

The R package DESeq was employed to detect genes differentially-expressed between groups [72]. DESeq mediates a negative binomial distribution by modeling the biological and technical variance for testing DE genes in two experimental conditions. DESeq uses as input file the unambiguous table of counts per gene obtained from QualiMap software using the *comp-counts* option [70]. Before the analysis, some exploratory tests were performed to validate both the good data quality and the variance estimation. Per-gene estimates of the base variance against the base levels showed that the fit (red line) followed well the single-gene estimates (Figure S3). The *residualEcdfPlot* function used to check the uniformity of the cumulative probabilities revealed a similar curve pattern of the empirical cumulative density functions (ECDF) in both groups. Data was filtered by a minimum mean of 20 reads mapped per gene and

only those genes with a fold change between groups higher than 1.2 fold were retained. Then, the R package q-value [73] was employed to calculate the false-discovery rate and genes with a p-value ≤ 0.01 (which is equivalent to a q-value ≤ 0.1) were retained in both classifications.

Gene functional classification, network and canonical pathways analyses

A bioinformatics approach was used to elucidate the biological importance of differentially-expressed genes in adipose tissue transcriptome. Ingenuity Pathway Analysis software (IPA; Ingenuity Systems, www.ingenuity.com) was applied to identify functions and pathways represented and for generating biological networks. IPA program consists of the Ingenuity Pathway Knowledge Base (IPKB) which is derived from known functions and interactions of genes published in the literature. IPA presents the top canonical pathways associated with the uploaded data with a p-value calculated using right-tailed Fisher's exact test. Functional analysis was used to identify the biological functions that are differentially represented between both groups (H and L). Networks were algorithmically generated based on their connectivity, with a score representing the log probability of a particular network being found by random chance. Direct and indirect biological relationships between molecules (nodes) were represented as continuous and discontinuous lines, respectively. All lines are supported by at least one reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Pathways Knowledge Base. The intensity of the node color indicates the degree of up-(red) or down-(green) regulation of H versus L group.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

JC, JMF and MB conceived and designed the experiment. JMF was the principal investigator of the project. JC, YRC and APO performed the RNA-Seq data analysis. JC and JE performed the pathways analysis. JC, MB and AC performed the RNA isolation. EA, RP, JE and JMF collected the samples. JC, JMF and MB drafted the manuscript. All authors read and approved the final manuscript.

AKNOWLEDGMENTS

This work was funded by the *Ministerio de Economía y Competitividad* project AGL2011-29821-C02, and by the Innovation Consolider-Ingenio 2010 Program (CSD2007-00036, Centre for Research in Agrigenomics). J. Corominas was funded by a FPI PhD grant from the Spanish *Ministerio de Educación* (BES-2009-081223), Y. Ramayo-Caldas was funded by a FPU PhD grant from the Spanish *Ministerio de Educación* (AP2008-01450) and A. Puig-Oliveras was funded by a PIF PhD grant from the Universitat Autónoma de Barcelona (458-01-1/2011).

REFERENCES

1. Dodson MV HG, Guan L, Du M, Rasmussen TP, Poulos SP, Mir P, Bergen WG, Fernyhough ME, McFarland DC, Rhoads RP, Soret B, Reecy JM, Velleman SG, Jiang Z. : Lipid metabolism, adipocyte depot physiology and utilization of meat animals as experimental models for metabolic research. *Int J Biol Sci* 2010, 6(7):691-699.
2. Wood JD, Richardson RI, Nute GR, Fisher AV, Campo MM, Kasapidou E, Sheard PR, Enser M: Effects of fatty acids on meat quality: a review. *Meat Science* 2004, 66(1):21-32.
3. Chizzolini R, Zanardi E, Dorigoni V, Ghidini S: Calorific value and cholesterol content of normal and low-fat meat and meat products. *Trends in Food Science & Technology* 1999, 10(4-5):119-128.
4. Katan MB, Zock PL, Mensink RP: Effects of fats and fatty acids on blood lipids in humans: an overview. *The American Journal of Clinical Nutrition* 1994, 60(6):1017S-1022S.
5. Food, Agriculture O, Agriculture Organization of the United N: Fats and Fatty Acids in Human Nutrition: Report of an Expert Consultation : 10-14 November 2008, Geneva: Food and Agriculture Organization of the United Nations.
6. de Lorgeril M, Salen P: New insights into the health effects of dietary saturated and omega-6 and omega-3 polyunsaturated fatty acids. *BMC Medicine* 2012, 10(1):50.
7. Pérez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, Folch JM, Babot D, Estany J, Oliver MA, Díaz I *et al*: A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *Journal of Animal Science* 2000, 78(10):2525-2531.
8. Casellas J, Noguera JL, Reixach J, Díaz I, Amills M, Quintanilla R: Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs. *Journal of Animal Science* 2010, 88(7):2246-2254.
9. Ntawubizi M, Colman E, Janssens S, Raes K, Buys N, De Smet S: Genetic parameters for intramuscular fatty acid composition and metabolism in pigs. *Journal of Animal Science* 2010, 88(4):1286-1294.

10. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, Ballester M: Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE*, 8(1):e53687.
11. Estellé J, Fernández AI, Pérez-Enciso M, Fernández A, Rodríguez C, Sánchez A, Noguera JL, Folch JM: A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. *Animal Genetics* 2009, 40(6):813-820.
12. Corominas J, Ramayo-Caldas Y, Castelló A, Muñoz M, Ibáñez-Escriche N, Folch JM, Ballester M: Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs. *Animal Genetics* 2012, 43(6):714-720.
13. O'Hea EK, Leveille GA: 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. *The Journal of Nutrition* 1969, 99(3):338-344.
14. Kershaw EE, Flier JS: Adipose Tissue as an Endocrine Organ. *Journal of Clinical Endocrinology & Metabolism* 2004, 89(6):2548-2556.
15. Cao H, Gerhold K, Mayers JR, Wiest MM, Watkins SM, Hotamisligil GkS: Identification of a Lipokine, a Lipid Hormone Linking Adipose Tissue to Systemic Metabolism. *Cell* 2008, 134(6):933-944.
16. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth* 2008, 5(7):621-628.
17. Jager M, Ott C-E, Grunhagen J, Hecht J, Schell H, Mundlos S, Duda G, Robinson P, Lienau J: Composite transcriptome assembly of RNA-seq data in a sheep model for delayed bone healing. *BMC Genomics* 2011, 12(1):158.
18. Li R, Rinaldi M, Capuco A: Characterization of the abomasal transcriptome for mechanisms of resistance to gastrointestinal nematodes in cattle. *Veterinary Research* 2011, 42(1):114.
19. Cánovas A, Rincon G, Islas-Trejo A, Wickramasinghe S, Medrano J: SNP discovery in the bovine milk transcriptome using RNA-Seq technology. *Mammalian Genome* 2010, 21(11-12):592-598.
20. Huang W, Khatib H: Comparison of transcriptomic landscapes of bovine embryos using RNA-Seq. *BMC Genomics* 2010, 11(1):711.

21. Jung WY, Kwon SG, Son M, Cho ES, Lee Y, Kim JH, Kim B-W, Park DH, Hwang JH, Kim TW *et al*: RNA-Seq Approach for Genetic Improvement of Meat Quality in Pig and Evolutionary Insight into the Substrate Specificity of Animal Carbonyl Reductases. *PLoS ONE* 2012, 7(9):e42198.
22. Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castello A, Ballester M, Estelle J, Ibanez-Escriche N, Fernandez A, Perez-Enciso M *et al*: Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics* 2012, 13(1):547.
23. Frayn K, Arner P, Yki-Järvinen H: Fatty acid metabolism in adipose tissue, muscle and liver in health and disease. *Essays in Biochemistry* 2006, 042:89-103.
24. Chen C, Ai H, Ren J, Li W, Li P, Qiao R, Ouyang J, Yang M, Ma J, Huang L: A global view of porcine transcriptome in three tissues from a full-sib pair with extreme phenotypes in growth and fat deposition by paired-end RNA sequencing. *BMC Genomics* 2011, 12(1):448.
25. Esteve-Codina A, Kofler R, Palmieri N, Bussotti G, Notredame C, Perez-Enciso M: Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC Genomics* 2011, 12(1):552.
26. Stanke M, Diekhans M, Baertsch R, Haussler D: Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 2008, 24(5):637-644.
27. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y: RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research* 2008, 18(9):1509-1517.
28. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y: RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 2008, 18:1509-1517.
29. Wang Z, Gerstein M, Snyder M: RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009, 10(1):57-63.
30. Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, Díaz I, Ibánez-Escriche N, Noguera JL, Pérez-Enciso M *et al*: Genome-Wide Association

Study for Intramuscular Fatty Acid Composition in an Iberian x Landrace Cross. *J Animal Sci* 2012, 90:1-11.

31. Ren Z-Q, Wang Y, Xu Y-J, Wang L-J, Lei M-G, Zuo B, Li F-E, Xu D-Q, Zheng R, Deng C-Y *et al*: Identification of a differentially expressed gene, ACL, between Meishan x Large White and Large White x Meishan F1 hybrids and their parents. *Genetics Selection Evolution* 2008, 40(6):625-637.
32. Muñoz M, Alves E, Corominas J, Folch JM, Casellas J, Noguera JL, Silió L, Fernández AI: Survey of SSC12 regions affecting fatty acid composition of intramuscular fat using high density SNP data. *Frontiers in Genetics* 2011, 2:101.
33. Muñoz G, Alves E, Fernández A, Óvilo C, Barragán C, Estellé J, Quintanilla R, Folch JM, Silió L, Rodríguez MC *et al*: 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. *Animal Genetics* 2007, 38(6):639-646.
34. Jayakumar A, Tai MH, Huang WY, al-Feel W, Hsu M, Abu-Elheiga L, Chirala SS, Wakil SJ: Human fatty acid synthase: properties and molecular cloning. *Proceedings of the National Academy of Sciences* 1995, 92(19):8695-8699.
35. Wakil SJ, Stoops JK, Joshi VC: Fatty Acid Synthesis and its Regulation. *Annual Review of Biochemistry* 1983, 52(1):537-579.
36. Matsuzaka T, Shimano H, Yahagi N, Kato T, Atsumi A, Yamamoto T, Inoue N, Ishikawa M, Okada S, Ishigaki N *et al*: Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. *Nat Med* 2007, 13(10):1193-1202.
37. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, Ballester M: Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE* 2013, 8(1):e53687.
38. Vidal O, Varona L, Oliver MA, Noguera JL, Sànchez A, Amills M: Malic enzyme 1 genotype is associated with backfat thickness and meat quality traits in pigs. *Animal Genetics* 2006, 37(1):28-32.
39. Colbert CL, Kim C-W, Moon Y-A, Henry L, Palnitkar M, McKean WB, Fitzgerald K, Deisenhofer J, Horton JD, Kwon HJ: Crystal structure of Spot 14, a modulator of

fatty acid synthesis. *Proceedings of the National Academy of Sciences* 2010, 107(44):18820-18825.

40. Trigatti BL, Gerber GE: A direct role for serum albumin in the cellular uptake of long-chain fatty acids. *Biochemical Journal* 1995, 308(1):155-159.
41. Burnett JR, Wilcox LJ, Telford DE, Kleinstiver SJ, Barrett PHR, Newton RS, Huff MW: Inhibition of HMG-CoA Reductase by Atorvastatin Decreases Both VLDL and LDL Apolipoprotein B Production in Miniature Pigs. *Arteriosclerosis, Thrombosis, and Vascular Biology* 1997, 17(11):2589-2600.
42. Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J *et al*: Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 2012, 491(7424):393-398.
43. Ferré P: The Biology of Peroxisome Proliferator-Activated Receptors: Relationship With Lipid Metabolism and Insulin Sensitivity. *Diabetes* 2004, 53(suppl 1):S43-S50.
44. Postic C, Dentin R, Denechaud P-D, Girard J: ChREBP, a Transcriptional Regulator of Glucose and Lipid Metabolism. *Annual Review of Nutrition* 2007, 27(1):179-192.
45. Serra X, Gil F, Pérez-Enciso M, Oliver MA, Vázquez JM, Gispert M, Díaz I, Moreno F, Latorre R, Noguera JL: A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livestock Production Science* 1998, 56(3):215-223.
46. Webb EC, O'Neill HA: The animal fat paradox and meat quality. *Meat Science* 2008, 80(1):28-36.
47. 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 Sci* 2008, 78(4):343-358.
48. Li XJ, Yang H, Li GX, Zhang GH, Cheng J, Guan H, Yang GS: Transcriptome profile analysis of porcine adipose tissue by high-throughput sequencing. *Animal Genetics*, 43(2):144-152.
49. Perez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, Folch JM, Babot D, Estany J, Oliver MA, Díaz I *et al*: A QTL on pig chromosome 4 affects fatty acid

metabolism: evidence from an Iberian by Landrace intercross. *Journal of Animal Science* 2000, 78(10):2525-2531.

50. van der Vusse GJ: Albumin as Fatty Acid Transporter. *Drug Metabolism and Pharmacokinetics* 2009, 24(4):300-307.
51. Bergen WG, Mersmann HJ: Comparative Aspects of Lipid Metabolism: Impact on Contemporary Research and Use of Animal Models. *The Journal of Nutrition* 2005, 135(11):2499-2502.
52. Schmitz G, Ecker J: The opposing effects of n-3 and n-6 fatty acids. *Progress in Lipid Research* 2008, 47(2):147-155.
53. Peet DJ, Turley SD, Ma W, Janowski BA, Lobaccaro J-MA, Hammer RE, Mangelsdorf DJ: Cholesterol and Bile Acid Metabolism Are Impaired in Mice Lacking the Nuclear Oxysterol Receptor LXR alpha. *Cell* 1998, 93(5):693-704.
54. Dentin R, Benhamed F, xE, gorier J-P, Foufelle F, Viollet B, Vaulont S, Girard J, Postic C: Polyunsaturated fatty acids suppress glycolytic and lipogenic genes through the inhibition of ChREBP nuclear protein translocation. *The Journal of Clinical Investigation* 2005, 115(10):2843-2854.
55. Ntambi JM: Regulation of stearoyl-CoA desaturase by polyunsaturated fatty acids and cholesterol. *Journal of Lipid Research* 1999, 40(9):1549-1558.
56. Cohen P, Miyazaki M, Socci ND, Hagge-Greenberg A, Liedtke W, Soukas AA, Sharma R, Hudgins LC, Ntambi JM, Friedman JM: Role for Stearoyl-CoA Desaturase-1 in Leptin-Mediated Weight Loss. *Science* 2002, 297(5579):240-243.
57. Ding ST, Schinckel AP, Weber TE, Mersmann HJ: Expression of porcine transcription factors and genes related to fatty acid metabolism in different tissues and genetic populations. *Journal of Animal Science* 2000, 78(8):2127-2134.
58. Sundvold H, Grindflek E, Lien S: Tissue distribution of porcine peroxisome proliferator-activated receptor α: detection of an alternatively spliced mRNA. *Gene* 2001, 273(1):105-113.
59. Yamauchi T, Kamon J, Minokoshi Y, Ito Y, Waki H, Uchida S, Yamashita S, Noda M, Kita S, Ueki K *et al*: Adiponectin stimulates glucose utilization and fatty-acid

oxidation by activating AMP-activated protein kinase. *Nature Medicine* 2002, 8(11):1288-1295.

60. Baile CA, Della-Fera MA, Martin RJ: Regulation of metabolism and body fat mass by leptin. *Annual Review of Nutrition* 2000, 20(1):105-127.
61. Ducluzeau P-H, Perretti Nl, Laville M, Andreelli F, Vega N, Riou J-P, Vidal H: Regulation by Insulin of Gene Expression in Human Skeletal Muscle and Adipose Tissue: Evidence for Specific Defects in Type 2 Diabetes. *Diabetes* 2001, 50(5):1134-1142.
62. Guillet-Deniau I, Mieulet V, Le Lay S, Achouri Y, Carré D, Girard J, Foufelle F, Ferré P: Sterol Regulatory Element Binding Protein-1c Expression and Action in Rat Muscles: Insulin-Like Effects on the Control of Glycolytic and Lipogenic Enzymes and UCP3 Gene Expression. *Diabetes* 2002, 51(6):1722-1728.
63. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009, 25(9):1105-1111.
64. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L: Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protocols* 2012, 7(3):562-578.
65. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech* 2010, 28(5):511-515.
66. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010, 26(6):841-842.
67. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, 21(18):3674-3676.
68. Kanehisa M, Goto S: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 2000, 28(1):27-30.
69. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T *et al*: KEGG for linking genomes to life and the environment. *Nucleic Acids Research* 2008, 36(suppl 1):D480-D484.

70. García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A: Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 2012, 28(20):2678-2679.

71. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F: A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association* 2004, 99(468):909-917.

72. Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biology* 2010, 11(10):R106.

73. Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* 2003, 100(16):9440-9445.

TABLES

Table 1. Summary of mapped reads. L=low; H=high

Animal	Group	% Exons	% Introns	% Intergenic
BC1	L	71	10	19
BC2	L	71	10	19
BC3	L	67	13	20
BC4	H	69	12	20
BC5	H	66	14	20
BC6	H	69	12	19

Table 2. Differentially-expressed genes associated with intramuscular FA content in a genome-wide association study in the same population [30] L=low; H=high

Ensembl Gene ID	Gene Name	Counts L group	Counts H group	p-value	Fold change
ENSSSCG00000004172	SLC2A12	316.27	124.56	5.2E-03	-2.54
ENSSSCG00000004774	IVD	10,800.27	4,058.75	3.21E-06	-2.66
ENSSSCG00000008596	APOB	26.38	106.03	1.3E-03	4.02
ENSSSCG00000006472	CRABP2	1,215.36	463.39	1.9E-04	-2.62
ENSSSCG00000006614	THEM5	23.62	152.32	8.4E-05	6.45
ENSSSCG00000011664	RBP1	861.18	403.95	4.1E-03	-2.13
ENSSSCG00000009755	AACS	5,287.19	2,303.84	2.7E-05	-2.29
ENSSSCG00000016156	ACADL	4,533.01	2,366.33	1.6E-03	-1.92

Table 3. Top five biological functions significantly modulated in backfat adipose tissue when comparing H vs. L animals

Function	Genes	p-value
Cancer	*ACADL, ACAN, ACBD4, ACE2, ACLY, ACTG2, ACVR1C, ADH1A, AFAP1L1, AHSG, ALB, ALDH1A1, ALDOC, ANXA4, AP3M2, APAF1, APOB, AQPEP, ATP5J2, AZGP1, BCHE, BCL10, BMPER, BNIP2, C10orf116, C19orf53, C2orf40, C8A, CA3, CAPN6, CAPZA2, CCBP2, CCL21, CCT6A, CD1E, CD300LG, CENPF, CES1, CHST13, CLCA2, CLEC2D, CLIC5, CLK1, CMPK2, CNN1, COL11A1, COL15A1, COL5A1, COL8A2, COMT, CPXM2, CRABP1, CRABP2, CTCFL, CTNNAL1, CTSF, CXCL1…*	3.36E-07
Concentration of lipids	*ACACA, ACADL, ACLY, AHSG, ALB, ALDH1A1, APAF1, APOB, APOC3, CD4, CES1, CIDEC, COMT, CTDNEP1, CYP2E1, DHCR24, FASN, FFAR4, GC, GNAT1, HIF1AN, HP, MOGAT2, PHGDH, PLP1, PON1, PON2, PON3, RBP1, RDH16, RGS4, SCD, SNCA, STEAP4, TGFBR2, THEM5, THRSP, UGT8, VAV3*	1.22E-06
Synthesis of lipids	*ACACA, ACADL, ACLY, ACSS2, ALB, ALDH1A1, APOB, APOC3, C1QTNF3, CD4, CXCL1, CYP2E1, DHCR24, ESRRG, FASN, G6PD, KDR, KIT, LSS, MOGAT2, NFATC2, PLP1, PMVK, PON1, PON2, PRKG2, RBP1, RDH16, RDH5, SCD, SLC6A6, SNCA, THRSP, UGT8*	4.85E-06
Homeostasis of blood	*AHSG, APOC3, APOH, CIDEC, COMT, CYP2E1, ESRRG, MUT, VAV3*	9.46E-06
Fatty acid metabolism	*AACS, ACACA, ACADL, ACADSB, ACLY, ACSS2, ALB, APOB, APOC3, APOH, CD4, CXCL1, CYP2E1, EPHX1, FASN, GC, GM2A, KIT, ME1, NFATC2, PHGDH, PLP1, SCD, SLC36A2, SLC38A2, SLCO1A2, SNCA, UGT8*	1.32E-05

Table 4. Top five canonical pathways significantly modulated in backfat adipose tissue when comparing H vs. L animals

Ingenuity canonical pathway	Genes	p-value
LXR/RXR Activation	*SCD, APOB, APOH, AHSG, PCYOX1, PON1, ALB, LYZ, APOC3, FASN, ACACA, S100A8, FGA, GC, PON3, TNFRSF11B*	6.19E-10
Ethanol Degradation II	*ALDH4A1, ADH1A, ALDH1A1, ACSS2, PECR, ADHFE1*	2.52E-05
Noradrenaline and Adrenaline Degradation	*ALDH4A1, ADH1A, ALDH1A1, COMT, PECR, ADHFE1*	3.58E-05
Acute Phase Response Signaling	*ALB, HP, RBP7, APOH, AHSG, CRABP2, FGB, FGA, RBP1, FGG, CRABP1, TNFRSF11B*	5.95E-05
Superpathway of Serine and Glycine Biosynthesis I	*PSPH, PHGDH, SHMT2*	1.14E-04

FIGURES

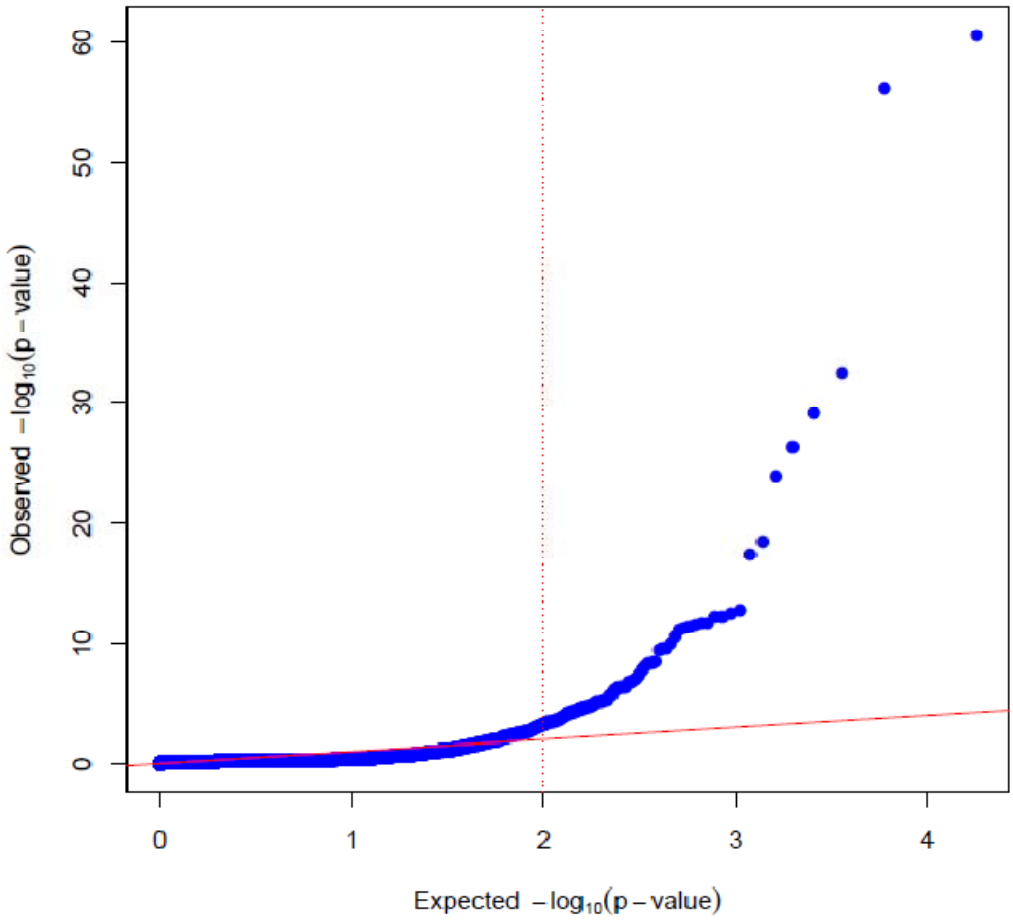

Figure 1. Q-Q plot representing the distribution of the p-value. Red line represents the expected distribution of the p-value, while the blue trend represents the observed distribution. X-axis values are Expected $-\log_{10}$ (p-value) and y-axis are the Observed $-\log_{10}$ (p-value).

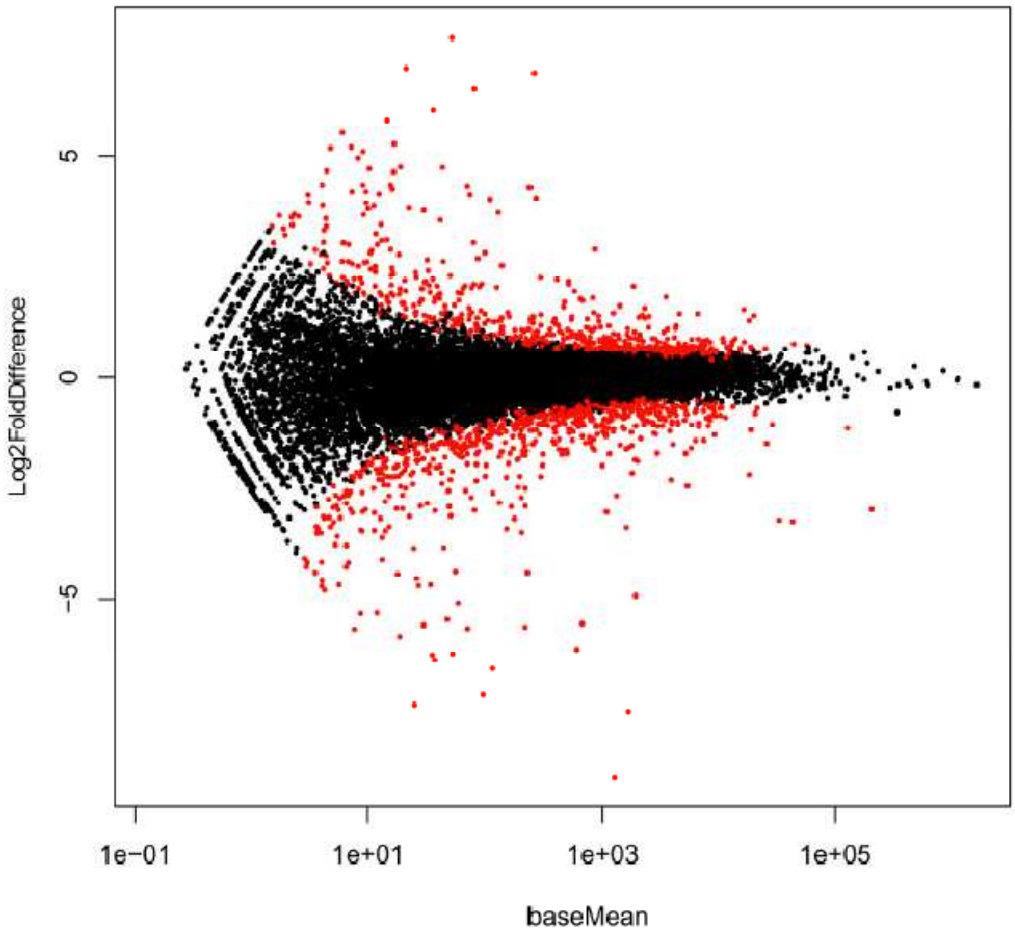

Figure 2. Representation of the 396 differentially-expressed genes (in red) with fold change ≥ 1.2 and p-value ≤ 0.01. X-axis values are base mean expression values and y-axis values are the log2 (fold change).

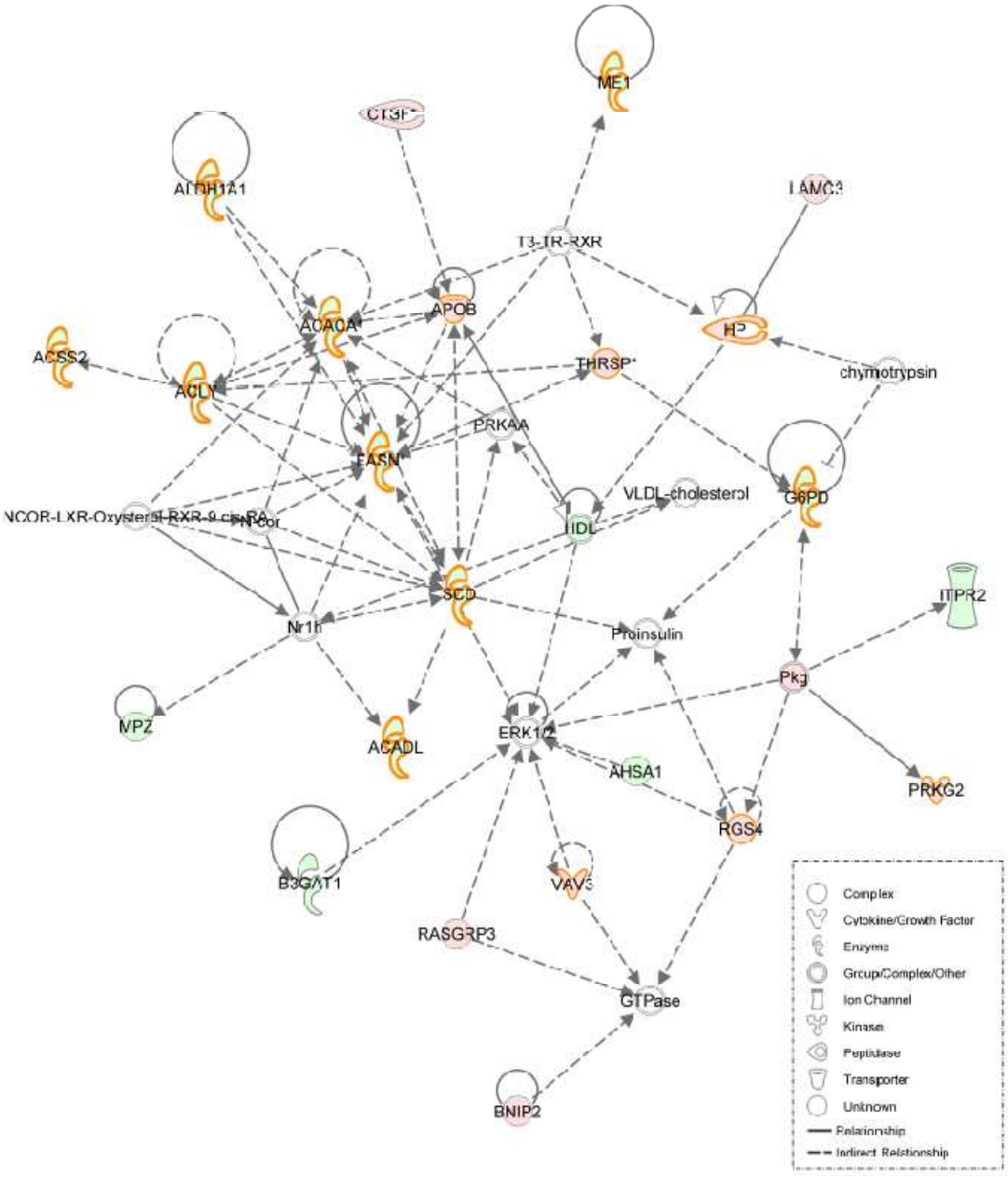

Figure 3. IPA network of genes associated with lipid metabolism, nucleic acid metabolism and small molecule biochemistry. This network diagram shows the biological association of 35 focus genes associated mainly with lipid metabolism as a graphical representation of the molecular relationship (edges) between genes/gene products (nodes). The intensity of the node color indicates the degree of expression: (red) up-regulated and (green) down-regulated in H group relative to L group. The shape of nodes indicates the functional classes of the gene products. Genes highlighted in orange are those genes related to lipid metabolism.

PAPER IV

Different patterns of methylation on ELOVL6 promoter caused by a promoter polymorphism is associated with a major QTL effect on fatty acid composition in pigs

Jordi Corominas *et al.* (Manuscript in preparation)

Different patterns of methylation on *ELOVL6* promoter caused by a promoter polymorphism is associated with a major QTL effect on fatty acid composition in pigs

Jordi Corominas[1,2] et al. (Manuscript in preparation)

[1] Centre de Recerca en Agrigenòmica (Consorci CSIC-IRTA-UAB-UB), Edifici CRAG, Campus UAB, Bellaterra, 08193, Barcelona, Spain.

[2] Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Campus UAB, Bellaterra, 08193, Barcelona, Spain.

E-mail: jordi.corominas@uab.cat

ABSTRACT

Background: Porcine *ELOVL6* gene has been proposed as the major causative gene of the QTL on SSC8, affecting palmitic and palmitoleic acid content in muscle and backfat, detected in an Iberian x Landrace cross. The *ELOVL6:c.-533C>T* SNP located in the promoter was highly associated with its own gene expression and, consequently, with the percentages of palmitic and palmitoleic acid in *longisimus dorsi* and adipose tissue. The main goal of this work was to study the major role of *ELOVL6* on these traits and to analyze *ELOVL6* gene expression regulation and the implication of *ELOVL6* polymorphisms on meat quality traits in pigs.

Results: High-throughput sequencing of BACs containing the porcine *ELOVL6* gene showed two isoforms expressed in muscle and adipose tissue that differ on their 3'UTR length. Although several polymorphisms on the 3'UTR were associated with palmitic and palmitoleic acids content, this association was lower than that observed previously with *ELOVL6:c.-533C>T* SNP located at the promoter. This SNP is in LD with the *ELOVL6:c.-394G>A* polymorphism that is affecting the ERα binding site, showing union only in animals with G allele. The ERα binding is associated with an increase of the methylation levels of *ELOVL6* promoter and, consequently with a decrease on *ELOVL6* gene expression. In addition, functional studies performed in this work validate experimentally the union of SREBF1 on *ELOVL6* promoter. Therefore, SREBF1 is a clear candidate factor, which regulatory activity may be altered by the methylation levels.

Conclusions: Our results suggest *ELOVL6:c.-394G>A* mutation as a potential causal mutation for the QTL on SSC8 affecting fatty acid composition in pigs.

Keywords: elongase, *de novo* lipogenesis, fatty acid metabolism, meat quality, pork.

BACKGROUND

Elongation of very long-chain fatty acids (*ELOVLs*) are a family of enzymes that catalyze the initial and rate-limiting condensation reaction of fatty acid elongation cycle in mammals [1-3]. To date, seven *ELOVL* proteins have been identified, with *ELOVL1*, *ELOVL3*, *ELOVL6* and *ELOVL7* preferring saturated fatty acids (SFA) and monounsaturated fatty acids (MUFA) as substrate, and *ELOVL2*, *ELOVL4* and *ELOVL5* preferring polyunsaturated fatty acids (PUFA) [4-6]. In mammals, the *ELOVL6* enzyme catalyses the elongation of long-chain saturated and monounsaturated FAs with 12-16 carbons to C18 and it is considered a key gene for controlling the overall balance of fatty acid composition [2, 7]. The expression of *ELOVL6* is highly up-regulated, both in liver and adipose tissue in the refed state, showing the major relevance of this enzyme in the synthesis of long-chain fatty acids [8].The porcine *ELOVL6* gene is located on chromosome 8 (SSC8), in a region where a quantitative trait locus (QTL) affecting palmitic and palmitoleic acid composition was previously detected [2, 9]. Moreover, this gene has been recently reported as differentially-expressed in adipose tissue of animals from an Iberian x Landrace backcross with extreme phenotypic differences in intramuscular fatty acid composition [10].

ELOVL6 gene was first identified in the liver of transgenic mice that over-express *sterol regulatory element binding transcription factors* (*SREBFs*) [1].SREBFs are transcription factors that control the expression of genes implicated in *de novo* lipogenesis [11]. In tissues that synthesize fatty acids de novo, the expression of SREBFs is highly correlated with the expression of key lipogenic genes implicated in this metabolic pathway [12].Transcriptional regulation of *ELOVL6* gene by *SREBFs* was also confirmed by both gene expression analysis using DNA-microarrays in *SREBF1* transgenic mice [13], and a detailed promoter analysis of mouse *ELOVL6* gene [14]. Kumadaki *et al.* (2008) demonstrated that in liver, the nuclear SREBF1 activate the murine *ELOVL6* promoter by interacting with two sterol response elements (SRE). In this study, no contribution of E-box on *ELOVL6* activity was observed, despite SREBF1 is capable to bind E-box motifs [14, 15]. In a previous work, we compared the murine and pig *ELOVL6* promoters, observing a conservation of SRE and E-box motifs across both species [2]. In particular, the porcine promoter presents SRE at positions -18, -450 and

-524 and E-box at position -331. Close to the most distal SRE element a polymorphism, the *ELOVL6:c.-533C>T*, was identified. This SNP was highly associated with the percentages of palmitic and palmitoleic acids in muscle and backfat, but also with the *ELOVL6* gene expression levels in backfat [2]. Other transcription factors binding sites were described in the porcine *ELOVL6* promoter, such as *SP1 transcription factor* (*SP1*) at position -470 and containing the *ELOVL6:c.-480C>T* polymorphism or the *MLX interacting protein-like* (*MLXIPL*), also called *carbohydrate response element binding protein* (*ChREBP*), at position -322 of *ELOVL6* promoter [2]. Furthermore, the five additional SNPs (*ELOVL6:c.-574C>T, ELOVL6:c.-534C>T*, *ELOVL6:c.-492G>A,ELOVL6:c.-394G>A* and *ELOVL6:c.-313C>T*) identified in *ELOVL6* promoter [2] may be also relevant for *ELOVL6* gene expression regulation.

Differences on *ELOVL6* gene expression among lipogenic tissues (liver, adipose tissues and muscle) were also observed, suggesting different regulatory mechanisms for each tissue [2]. For instance, a whole genome association study of *ELOVL6* gene expression levels (eGWAS) in the liver, adipose tissue and muscle showed different genomic regions that may be affecting the tissue-specific *ELOVL6* gene expression [2]. However, these regulatory differences may also be produced by epigenetic modifications, adding an additional level of gene expression regulation [16]. One of the major epigenetic mechanism that regulates gene transcription is DNA methylation, which has been related with the regulation of lipid metabolism related genes, such as *fatty acid desaturase 2* (*FADS2*) [17] or the *peroxisomal proliferator-activated receptor alpha* (*PPARα*) [18]. Finally, a hypothetical role of microRNAs on *ELOVL6* gene expression should not be ruled out, due to the 3'UTR of porcine *ELOVL6* gene has not been characterized.

The overall objective of the present study was to determine the different mechanisms that could contribute to the control and regulation of *ELOVL6* gene expression and its implications on meat quality traits. In this sense, we characterized the unknown 3'UTR of porcine *ELOVL6* gene where several polymorphisms were identified. In addition, a methylation study of gene promoter was performed on liver, adipose tissue and muscle, in order to evaluate the implication of epigenetic modification on *ELOVL6* gene expression regulation across tissues. Furthermore, chromatin immunoprecipitation

(ChIP) were carried out to determine the effects of both promoter polymorphisms and methylation to transcription factor binding. In this article, we present new knowledge for a better understanding of *ELOVL6* gene expression regulation and its implication on fatty acid composition and meat quality traits in pigs.

MATERIALS AND METHODS

Animal material

The population analyzed was generated by crossing three Iberian (Guadyerbas line) boars with 31 Landrace sows (IBMAP cross) [19], and contained several generation and backcrosses. In particular animals used in this study belong to a backcross (BC1_LD) generated by crossing five F1 (Iberian x Landrace) boars with 26 Landrace sows and producing 144 backcrossed animals. All pigs were raised and fed under the standard intensive system in Europe and feeding was *ad libitum* with a cereal-based commercial diet. Pigs were slaughtered at an average age of 179.8 ± 2.6 days following national and institutional guidelines for the Good Experimental Practices and approved by the Ethical Committee of the Institution (IRTA- Institut de Recerca i Tecnologia Agroalimentàries). Samples of liver, muscle (*Longissimus dorsi*) and adipose tissue (backfat) were collected, snap-frozen in liquid nitrogen and stored at -80ºC. Genomic DNA was obtained from blood samples of all animals by the phenol-chloroform method, as described elsewhere. Backfat [20] and intramuscular fatty acid composition [9] was measured with a protocol based on gas chromatography of methyl esters [21].

BAC screening and sequencing

The porcine bacterial artificial chromosome (BAC) library of the *Centre de Ressources Biologiques Génomique des Animaux Domestiques et d'Intérèt Economique* (CRB-GADIE) was used to select those BACs containing the SSC8 region where *ELOVL6* is located. This BAC library was constructed using the pBeloBAC11 vector and comprised 107,520 clones with an average insert size of 135 kb, representing a five-fold coverage of the pig haploid genome [22]. Screening of the library was performed using three pairs of primers located in the promoter region, second exon (intermediate gene region) and at the end of the coding region in exon 4, in order to encompass the major part of *ELOVL6* gene (Table S1). Primers were designed using the software PRIMER3 [23] and were validated using the software PRIMER EXPRESS^{TM} (*Applied Biosystems*). BAC screening was performed by two-step PCR (superpools and pools PCR) and

positive BAC clones were confirmed with a PCR on isolated clones, checking the expected size of amplified fragments.

BAC clones were isolated by growing up BACs on Luria-Bertani (LB) agar containing 12.5 µg/ml of chloramphenicol, and incubating the culture overnight in a 37ºC incubator. Isolated clones were added to 4 ml LB broth with 12.5 µg/ml of chloramphenicol and culture was grown overnight at 37ºC in a shaking incubator. Finally, 4ml of the overnight starter culture was inoculated in 500 ml LB-medium supplemented with chloramphenicol (12.5 µg/ml) until an optical density (OD) of 2 was obtained. BAC isolation was carried out using the plasmid DNA purification Nucleobond BAC100 kit (*Macherey-Nagel*), following the manufacturer's recommendations of the Low-copy plasmid purification (Maxi BAC100) section. DNA was quantified using the *Nano-Drop ND-1000 spectrophotometer* (*NanoDrop products*) and checked for purity and integrity with agarose gels.

For each BAC, bar-coded libraries were generated using the Ion Xpress Plus fragment library kit (*Life Technologies*) with an insert size of approximately 250 bp. Libraries were sequenced on a Personal Genome Machine (PGM) Ion Torrent instrument (*Life Technologies*) using an Ion 314R chip. More than 220,000 single-end reads were generated with an average length of 152 bp for each BAC.

De novo assembly and porcine *ELOVL6* gene characterization

Quality control and reads statistics were determined with FASTQC [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]. All reads were mapped against the *Escherichia coli* genome using the Burrows-Wheeler Alignment tool (bwa v.0.6.2) [24], in order to discard all reads corresponding to the bacterial genome. The removal of sequence adapters, read trimming and *de novo* assembly was performed using *de novo* assembler tool of CLC Genomics Workbench v.6.0.1 [http://www.clcbio.com]. *ELOVL6* 3'UTR was identified from *de novo* assembled reads by sequence similarity with the same regions in the human (GenBank:NM_001130721) and bovine (GenBank:NM_001102155) genes, using the Basic Local Alignment Search

Tool (BLAST v2.2.28) [25]. The resulting pig *ELOVL6* 3'UTR sequence was used as reference for mapping reads from the liver (12 animals BC1_LD) and adipose tissue (6 animals BC1_LD) transcriptomes [10, 26],using the software TopHat v2.0.1 [27, 28] and the pig genome (Sscrofa10.2) [http://www.ensembl.org/info/data/ftp/index.html] as reference. SNPs were identified by comparing the reference sequence with mapped reads using the Integrative Genomic Viewer(IGV v.2.1) [29, 30].

Genotyping

The SNPs *ELOVL6:c.1408C>T* and *ELOVL6:c.1922C>T* were genotyped using the platform KASP SNP genotyping system [http://www.kbioscience.co.uk/reagents/KASP/KASP.html]. A total of 176 animals were genotyped, 141 of those belong to BC1_LD and 35 parental animals of the IBMAP cross (F0 and F1).

For Genome-Wide Association Studies (GWAS), a total of 144 animals of the BC1_LD backcross were genotyped with the Porcine SNP60 BeadChip (*Illumina*) using the Infinium HD Assay Ultra protocol (*Illumina*). Raw data had a high genotyping quality (call rate > 0.99) and was visualized and analyzed with the GenomeStudio software (*Illumina*). For subsequent data analysis, a subset of 54,998 SNPs was selected by removing the SNPs with a minor allele frequency <5%, those with missing genotypes >5% and the duplicated SNPs in the Sscrofa 10.2 assembly.

Chromosome wide association analyses

Association analyses of whole-genome SNP genotypes, together with the previously identified SNPs *ELOVL6:c.-533C>T*, *ELOVL6:c.-480C>T*, *ELOVL6:c.416C>T* [2] and the new *ELOVL6:c.1408C>T* and *ELOVL6:c.1922C>T* polymorphisms were performed with the following phenotypes: RT-qPCR expression data of *ELOVL6* mRNA in backfat and C16:0 and C16:1(n-7) fatty acid composition in backfat and intramuscular fat. The position of the SNPs was based on the *Susscrofa* 10.2 genome assembly

[http://www.animalgenome.org/repository/pig/]. GWAS were performed with a mixed model [31, 32] accounting for additive effects associated with each marker (see below) by using Qxpak 5.0 software[33]:

$y_{ijlkm} = Sex_i + Batch_j + \lambda_l a_k + u_l + e_{ijlkm}$

in which y_{ijlkm} is the l-th individual record, sex (two levels) and batch (five levels) are fixed effects, λ_l is a -1, 0, +1 indicator variable depending on the l-th individual genotype for the k-th SNP, a_k represents the additive effect associated with SNP, u_l represents the infinitesimal genetic effect treated as random and distributed as N(0, $\mathbf{A}\sigma_u$) where **A** is a numerator of the kinship matrix and e_{ijlkm} is the residual. The infinitesimal effect allows us to adjust the data for family information and, thus, to correct the inter-chromosomal linkage disequilibrium effect. In this analysis, each SNP was tested individually to check the association. The R package q-value [34] was used to calculate the FDR-based q-value to measure the statistical significance at the genome-wide level for association studies. The cut-off for a significant association at the chromosome level was set at q-value ≤ 0.05. This significance threshold is likely too stringent due to the linkage association among SNP genotypes. The same model was applied in order to determine the effect of haplotypes to the traits of interest. The only difference is that haplotypes were treated as random additive effects, in contrast with the individual SNPs, which were considered as fixed additive effects.

DNA methylation analyses

DNA methylation analyses of the main lipogenic tissues (liver and backfat) were performed in 31 animals, while two additional control tissues that express low levels of *ELOVL6* gene (muscle and spleen) were analyzed in 11 animals. DNA was obtained using the phenol-chloroform method, as described elsewhere. Methylation studies were performed using the bisulfite methodology [35] and the pyrosequencing technique [36, 37]. The bisulfite gDNA conversion was performed with the EZ DNA Methylation kit (*Zymo Research*) and 500 ng of genomic DNA from each sample. The regions of interest were amplified using primers designed (Table S1) over the resulting

methylated sequence using the allele quantification assay type of the PSQ assay design software (*Biotage*). PCRs were performed in a total volume of 25 µl containing 0.6 units of AmpliTaq Gold (*Applied Biosystems*), 1.5-2.5 mM $MgCl_2$ (depending on the primers; Table S1), 0.2 mM of each dNTP, 0.5 µM of each primer and 25 ng of treated genomic DNA. Thermocycling was carried out under the following conditions: 94 ºC for 10 min, 40 cycles of 94 ºC for 1 min, 60 ºC for 1 min and 72 ºC for 1 min, with a final extension of 72 ºC for 7 min. Pyrosequencing analysis was carried out on a PSQ HS 96A system with the Pyro Gold sequence analysis (SQA) reagent (*Biotage*) and using specific pyrosequencing primers for each region (Table S1). Statistical comparison of mean methylation values between tissues and *ELOVL6* gene expression were made using a linear procedure of R software considering sex and batch. Spleen gene expression quantification was performed by real time quantitative PCR (RT-qPCR), following the procedure described in [2].

Chromatin immunoprecipitation (ChIP)

ChIP was carried out using liver samples from four homozygous animals differing on *ELOVL6:c.-533C>T* and *ELOVL6:c.-394A>G* genotypes (two per genotype). A quantity of 0.05 g of frozen tissue was chopped into small pieces and thawed in freshly prepared PBS containing 1% formaldehyde for crosslinking. The tissue was homogenized to obtain a cell pellet. Nuclei were isolated by resuspending the cell pellet in 500 µl of celllysis buffer (50 mMTris-HCl pH 8, 150 mMNaCl, 2 mM EDTA, 1% Triton X-100 and protease inhibitors). Chromatin was fragmented by subjecting the nuclei to restriction enzyme digestion with micrococcal nuclease, in order to eliminate variable results caused by the traditional method of sonication. The digestion was processed following a treatment with 2 U of micrococcal nuclease for 15 min at 37ºC, mixing by inversion every 5 min. Chromatin was isolated using nuclei lysis buffer (50 mMTris-HCl pH=8, 10 mM EDTA and 1% SDS). The lysate was used to perform the chromatin immunoprecipitation with the commercial kit *Pierce ® Agarose ChIP* (*Thermo scientific*), following the manufacturer's recommendations. Based on the known binding of RNA polymerase II on *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*)

promoter, this binding was used as a positive control for our experiments. A total of 4 µg of antibodies against RNA polymerase II (positive control), ERα and SREBP1 (*Santa Cruz Biotechnology, Inc.*) were used. The negative control was performed with 1 µg of normal rabbit IgG. Immunoprecipitated DNA was analyzed by PCR using specific primers for *ELOVL6* promoter (Table S1).

RESULTS AND DISCUSSION

Pig *ELOVL6* gene structure and identification of a new isoform

Despite the importance of *ELOVL6* gene in lipid metabolism [2, 3, 7], a comparison of the human, bovine and porcine mRNA sequences revealed that the current sequence of porcine *ELOVL6* gene is incomplete (data not shown). In order to characterize the complete sequence of the *ELOVL6* gene, a BAC screening for the promoter region, the second exon (intermediate region) and fourth exon (terminal region) was performed. A total of 7 positive BACs containing at least one of the three regions of *ELOVL6* gene were identified by PCR. BACs 651E12, 650D01 and 385A04 were positive for the promoter region; BACs 201D05, 95C02, 754E02 and 385A04 were positive for the amplicon located at the second exon and BAC 754E02 was positive for the fourth exon containing the terminal part of the coding region. From the isolated BACs, only BACs 385A04 and 754E02 were positive for two different amplicons. BAC 385A04 contained the promoter and second exon region, whereas BAC754E02 had exon two and exon four. These results suggested an overlapping of the sequences of the two BACs which all together should contain the major part of the porcine *ELOVL6* gene. Both BACs were sequenced with the PGM of Ion Torrent and around 265,000 single-end reads were generated for each BAC with an average length of 152 bp. Data generated was used to perform a *de novo* assembly of porcine *ELOVL6* gene with the CLC Genomics Workbench v.6.0.1. A total of 129,672 bp sequence of the porcine *ELOVL6* gene was obtained, containing 1,942 bp of the upstream region, three introns and four exons (Figure 1).The protein coding region starts in the first exon (at position 2,201bp of the sequence) and ends in the fourth exon (at position 123,132bp) (Figure 1). To validate the new gene annotation, RNA-Seq data from adipose tissue and liver transcriptomes [10, 26] were employed to map the reads against the new *ELOVL6* sequence. The alignments obtained were concordant with the proposed *ELOVL6* gene structure. In addition, a clear decrease in the number of mapped reads was observed in the middle of the fourth exon. Interestingly, a poly-A region was identified in this region (at position 124,578 bp), suggesting the end of one alternative isoform (Figure 1). Therefore, as occurs in human *ELOVL6* gene, two different isoforms are expressed from the porcine *ELOVL6* gene in both liver and adipose tissue. The difference between

the two isoforms is the length of the 3'UTR, being shorter in variant 1 (1,455 bp) in comparison with variant 2 (5,117 bp) (Figure 1).

Identification of polymorphisms in the 3'UTR of porcine *ELOVL6* gene

The alignment and analysis of all mapped reads from RNA-Seq data [10, 26] allowed for the identification of eleven polymorphisms (Table 1) in the 3'UTR of porcine *ELOVL6* gene: five in both variants and six present only in variant 2. All SNPs were arranged in three haplotypes, which can be distinguished by genotyping the *ELOVL6:c.1408A>G* and *ELOVL6:c.1922A>G* polymorphisms (relative to the transcription start site, TSS, of the GenBank:NW_003610943). Hence, these two tag polymorphisms were genotyped in IBMAP founders, parental BC1_LD animals and the BC1_LD population. Regarding the IBMAP founders, the *ELOVL6:c.1408G* and *ELOVL6:c.1922G* alleles were fixed in Iberian boars. The *ELOVL6:c.1408A* allele was fixed in the founder Landrace sows, whereas the allele *ELOVL6:c.1922A* had a frequency of 0.7 in these sows. In BC1_LD Landrace sows, the allelic frequencies for *ELOVL6:c.1408A* and *ELOVL6:c.1922A* were 0.94 and 0.38, respectively. Both *ELOVL6:c.1408A>G* and *ELOVL6:c.1922A>G* polymorphisms segregated in the BC1_LD animals with frequencies for allele A of 0.72 and 0.46, respectively.

It is well known that polymorphisms in the 3'UTR affect the binding of microRNAs. This binding is important for gene expression regulation, due to microRNAs cause translational repression and/or mRNA destabiliazation [38]. For instance, the microRNA miR-33a/b has been described as a potential regulator of lipid metabolism by the repression of key enzymes involved in cholesterol efflux (*ABCA1* and *NPC1*), fatty acid metabolism (*CROT* and *CPT1a*) and insulin signaling (*IRS2*) [39, 40]. In this sense, a computer-assisted identification of potential microRNA binding elements was performed with the patrocles finder tool of patrocles programme [http://www.patrocles.org/], in order to asses if polymorphisms are affecting the disruption or creation of microRNAs binding sites. A total of twelve microRNA binding sites were modified by the eleven detected polymorphisms. The microRNAs miR-524-3p, miR-525-3p, miR-18a/b, miR-204 and miR-211 were predicted to bind to both

mRNA isoforms, whereas miR-584, miR-452, miR-603, miR-1262, miR-490-5p, miR-30a/d/e and miR-335 bind only to variant 2. Therefore, these microRNAs may be implicated in the regulation of porcine *ELOVL6* gene. Thus, further studies are needed to elucidate the implication of these microRNAs in *ELOVL6* gene expression in adipose tissue, liver and muscle.

Association studies reinforce the major role of *ELOVL6:c.-533C>T* polymorphism

The two newly genotyped 3'UTR SNPs and the three mutations described in Corominas *et al.* (2013) were added to the 2,565 SNPs of SSC8 included in the Porcine SNP60 BeadChip (*Illumina*), in order to perform association analyses with 136 BC1_LD animals for FA composition in muscle and backfat. In this analysis, the *ELOVL6:c.1922A>G* polymorphism showed a significant association with the percentages of palmitic acid in muscle (p-value= $3.38x10^{-04}$) and backfat (p-value= $1.23x10^{-11}$) (Figure 2A-C). In contrast, *ELOVL6:c.1408A>G* polymorphism was only significant with the percentages of palmitic acid in backfat (p-value= $1.73x10^{-06}$) (Figure 2C). On the other hand, the percentage of palmitoleic acid was significantly associated with *ELOVL6:c.1922A>G* polymorphism in both tissues, muscle (p-value= $1.51x10^{-07}$) and backfat (p-value= $1.22x10^{-06}$) (Figure 2B-D). Lower signification was also obtained for palmitoleic acid and *ELOVL6:c.1408A>G* polymorphism in muscle (p-value= $4.86x10^{-05}$) and backfat (p-value= $4.24x10^{-04}$) (Figure 2B-D). Nevertheless, for both fatty acids, the *ELOVL6:c.-533C>T* polymorphism showed always a higher association in comparison with the 3'UTR SNPs (Figure 2), reinforcing its role in the determination of the analyzed QTL. In addition, no significant associations were observed between the 3'UTR polymorphisms and *ELOVL6* expression levels in backfat, liver and muscle.

All together these results point to the promoter *ELOVL6:c.-533C>T* polymorphism as the most promising among the genotyped SNPs of SSC8. However, to ensure the main role of this SNP an association study was performed comparing the individual effect of each polymorphism (*ELOVL6:c.-533C>T* and *ELOVL6:c.1922A>G*) against the effect of the haplotypes formed by the combination of the two SNPs. These analyses were performed with a reduced number of animals (n=88), in which the allele origin (Iberian

or Landrace) was unambiguously determined from the pedigree information. As expected, *ELOVL6:c.-533C>T* polymorphism was more associated than the haplotype in all analyzed traits (p-value $_{\text{Backfat gene expression}}$= $3.68\text{x}10^{-03}$, p-value $_{\text{IMF\_C16:0}}$= $1.33\text{x}10^{-03}$, p-value $_{\text{IMF\_C16:1(n-7)}}$= $3.72\text{x}10^{-04}$, p-value $_{\text{BF\_C16:0}}$= $6.15\text{x}10^{-10}$ and p-value $_{\text{BF\_C16:1(n-7)}}$= $9.15\text{x}10^{-04}$) (Table 2S). In conclusion, the results obtained in this work reinforce the key role of *ELOVL6:c.-533C>T* polymorphism in explaining the phenotypic variation of the QTL affecting palmitic and palmitoleic acids content in the porcine SSC8. Despite the results obtained, we cannot discard a secondary role of the 3'UTR on the regulation of *ELOVL6* gene.

Promoter methylation is an additional level of regulation of porcine *ELOVL6* gene expression

DNA bisulfite conversion were used to compare the methylation patterns in the pig *ELOVL6* promoter between liver, adipose tissue (backfat), muscle and spleen (a tissue without *ELOVL6* gene expression). The methylation study was focused on those CpG motifs whose methylated states may affect the binding of SREBF1, the most relevant *ELOVL6* transcription factor. Previously, a ChIP assay on formaldehyde-cross-linked liver samples was performed to validate the occupancy of SREBF1 on *ELOVL6* promoter *in vivo*. The chromatin immunoprecipitation with anti-SREBF1 showed an enrichment of the *ELOVL6* promoter region, validating the role of this transcription factor on *ELOVL6* gene expression regulation (Figure 3). All individual CpG-sites identified in the SRE and E-box motifs of *ELOVL6* promoter [2] were included in the study in order to determine the effect of methylation on SREBF1 binding. In addition, it was observed in several lipogenic genes that SP1 is required as an additional regulator for SREBF1 activity [1, 41]. Interestingly, a CpG-site was identified in the SP1 binding element, in which the described *ELOVL6:c-416C>T* polymorphism [2] was located (Figure 4A), making this CpG a clear candidate to be analyzed. Finally, a total of 6 CpG-sites, covering the major part of the described promoter, were analyzed (Figure 4A). The analyses with muscle showed, in all CpG analyzed, higher levels of methylation than liver and adipose tissue and similar levels to spleen (Table 2). This data suggests that

the higher methylation causes a lower *ELOVL6* gene expression in muscle and spleen, as was observed with quantitative PCR (Figure 1S).

Methylation levels of the six selected CpG-sites showed two clear regions with opposite levels of methylation: i) lower methylation levels in the proximal region (-349 pb to -1 bp) and ii) higher methylation levels in the distal region (-529 bp to -350bp) (Figure 4B). The low methylation levels in the proximal region (CpG1, CpG2 and CpG3), including tissues with low *ELOVL6* gene expression (muscle and spleen), suggest that this region is important for maintaining a basal gene expression (Table 2). Nevertheless, CpG2 showed significant differences among tissues, so further studies are needed to determine the implication of MLX interacting protein-like (MLXIPL) protein in *ELOVL6* gene expression. On the other hand, the distal promoter region showed a higher degree of methylation (CpG4, CpG5 and CpG6), suggesting that the methylation of these motifs may be relevant for the regulation of *ELOVL6* gene expression across tissues (Table 2). Two continuous CpG-sites (CpG4 and CpG5) were found in the binding core of a SRE, thus, changes in the methylation pattern of these motifs may be critical for the SREBF1 binding (Figure 3A). Statistical analyses showed significant lower methylation levels of these sites in liver in comparison with backfat (p-valueCpG4 = $7.27x10^{-09}$, p-value CpG5 = $1.18x10^{-05}$ and p-value CpG6 = $1.04x10^{-07}$), muscle (p-value CpG4 = $2.13x10^{-14}$, p-value CpG5 = $1.52x10^{-10}$ and p-value CpG6 = $1.82x10^{-14}$) and spleen (p-value CpG4 = $4.42x10^{-11}$, p-value CpG5 = $5.24x10^{-07}$ and p-value CpG6 = $1.47x10^{-10}$) (Figure 3B). Additionally, high correlations have been observed between the methylation levels of CpG4, CpG5 and CpG6 in liver and adipose tissue, being the mean correlations of 0.84 and 0.73, respectively. Despite the low number of animals used in this study, suggestive significant effects have been observed between gene expression and the methylation levels of CpG4 (p-value= $9x10^{-02}$) and CpG6 (p-value= $8x10^{-02}$) in liver, and CpG5 (p-value=$6x10^{-02}$) in adipose tissue. The methylation levels observed in adipose tissue were clearly higher than those observed in liver, despite the higher expression of *ELOVL6* gene in adipose tissue in comparison to liver [2]. The higher expression levels of *ELOVL6* gene in adipose tissue are caused by the major role of this tissue in lipogenic pathways [42, 43] and, consequently, a higher regulation by SREBF1 in this tissue is produced [42]. In

agreement, gene expression correlation analysis performed in our animal material, showed high correlation between *SREBF1* and the lipogenic genes *ELOVL6* (r = 0.77) and *SCD* (r = 0.64) in adipose tissue, but no correlation were obtained in liver (Ballester *et al.*, personal communication). In this sense, the higher methylation observed in the distal promoter region on adipose tissue may be explained by the major necessity of inhibition lipogenic genes under fasting conditions. It is well known that fasting animals inhibit the lipogenic pathway, having a greater effect on adipose tissue in comparison with liver [44]. Under fasting conditions, liver receive a large amount of fatty acids from adipose tissue and these are either reesterified to triacylglycerol and secreted as VLDL or used for β-oxidation [45]. Hence, despite the inhibition of fatty acid synthesis, a low rate of triglyceride formation happens in liver, causing a lower inhibition in this tissue. Therefore, these results allowed us to hypothesize that pigs under fasting condition inhibit *ELOVL6* gene and one mechanism for this inhibition is the methylation of the -529 bp to -350 bp region of the promoter. This region contains one methylated SRE element and one SP1 binding site. The well-known importance of these transcription factors in the regulation of lipogenic genes such as *ELOVL6* gene [1, 2, 14, 41] suggests that this region may be crucial for understanding *ELOVL6* gene expression regulation. The occupancy of SREBF1 on *ELOVL6* promoter suggests that this transcription factor may be relevant for understanding the variation in the *ELOVL6* gene expression. Nevertheless, several SRE elements have been identified in the *ELOVL6* promoter with very short distance between them [2]. Therefore, further studies are needed to determine the capacity of SREBF1 to bind the specific methylated SRE, following a site-specific ChIP approach [46, 47], and also the effects of methylation levels on SREBF1 binding, comparing the binding levels using a qPCR approach. Concerning the SP1 element, containing the *ELOVL6:c-416C>T* polymorphism detected in [2], several studies have demonstrated a protective role of SP1 protein on CpG methylation [48, 49]. However, no association was observed between *ELOVL6:c-416C>T* genotype and the levels of methylation. This lack of association may be explained by both, the lower number of animals used in the analysis and the fact that SP1 binding on *ELOVL6* promoter was not validated experimentally.

Finally, it was described that some transcription factors, such as ERα are capable to regulate dynamic methylation cycles, producing rapid changes in the methylation levels of the regulated gene promoter [50]. These additional levels of regulation may affect the promoter regulation of *ELOVL6* gene expression, as a response to physiological changes such as fasting.

ELOVL6:c.-394G>A genotype determines the ERα occupancy to *ELOVL6* gene promoter

A computer-assisted identification of putative ERα binding sites was performed in *ELOVL6* promoter using the LASAGNA-Search software [51]. Interestingly, an estrogen response element (ERE) was predicted at position -397 to -382, between the two regions with different methylation patterns. Hence, this result suggests that ERα binding may be related with the different methylation patterns detected in *ELOVL6* promoter. In addition, within the ERE is located the *ELOVL6:c.-394G>A* polymorphism [2], which is in LD with *ELOVL6:c.-533C>T* polymorphism. All together these data allowed us to hypothesize that the *ELOVL6:c.-394G>A* polymorphism is affecting ERα binding and, consequently, alterations of this binding may affect *ELOVL6* gene expression via promoter methylation. The occupancy of the *ELOVL6* promoter by ERα *in vivo* was analyzed using a ChIP assay on formaldehyde-cross-linked liver samples. Four animals differing on *ELOVL6:c.-394G>A* genotype (GG vs. AA) were used to determine the effect of *ELOVL6:c.-394G>A* polymorphism on the binding of ERα in *ELOVL6* promoter. Cross-linked sheared chromatin from hepatocytes was immunoprecipitated with anti-ERα antibody and the DNA recovered was subjected to PCR. A PCR product specific to the *ELOVL6* promoter was amplified from anti-ERα-immunoprecipitated DNA samples carrying the GG genotype, but no amplification was observed in AA animals (Figure 5).This study strongly demonstrated the differential association of ERα to the proximal region of the *ELOVL6* promoter depending on *ELOVL6:c.-394G>A* genotype. The causative role of *ELOVL6:c.-394G>A* mutation on methylation level is supported by the statistical analyses performed with *ELOVL6:c.-533C>T* polymorphism, assuming that both mutations are in LD. In this analyses,

animals homozygous for *ELOVL6:c.-533C* showed a higher methylation rate in CpG4 (p-value= $1.2x10^{-02}$) and CpG5 (p-value= $2.5x10^{-02}$) and also a suggestive higher methylation percentage in CpG6 (p-value= $1x10^{-01}$). Hence, the genotyping of *ELOVL6:c.-394G>A* polymorphism is essential to validate the role of this polymorphism as the causal mutation of the QTL on SSC8 affecting the palmitic and palmitoleic acid content.

All together, the results provided in this study suggest a new mechanism of *ELOVL6* gene expression regulation. In pigs under fasting condition, a non-lipogenic state, there is a repression of *ELOVL6* gene expression and one plausible mechanism is the methylation of specific CpG-sites in the promoter region. Previous studies observed that rapid epigenetic modifications are controlled by ERα, providing an additional level of gene expression regulation [50, 52]. Activation of ERα, by estradiol binding or the phosphorylation of Serine 118, modulates its three-dimensional surface causing a recruitment of coactivators complexes, including DNA methyltransferases (Dnmts) [50]. The alteration of ERα binding by the *ELOVL6:c.-394G>A* polymorphism may be the main factor for explaining the differential expression observed in *ELOVL6* gene. In the methylated region two SRE elements were found, being the SRE element at -460 to -449 a clear candidate to be affected by methylation. This SRE element is carrying on its core binding site two CpG-sites (CpG4 and CpG5) (Figure 3A), which methylation levels were associated with *ELOVL6:c.-394G>A* genotype. In addition, this SRE element was previously studied in mouse, showing a relevant effect on *ELOVL6* gene expression [14]. Hence, the inhibition of SREBF1 binding caused by the methylation of SRE elements in *ELOVL6* promoter, specially the SRE element located at -460 to -449, seems to be the better explanation for the reduction of *ELOVL6* gene expression. Previous works performed in our group associated the lower *ELOVL6* gene expression with an accumulation of palmitic and palmitoleic acids in muscle and adipose tissue [2]. In this sense, the variation in *ELOVL6* gene expression can modulate fatty acid composition in muscle and backfat, affecting important sensorial and technological aspects of meat quality [53] and insulin sensitivity [7]. Hence, the results obtained in this study provide enough genetic evidences to propose *ELOVL6:c.-394G>A* polymorphism as a putative causal mutation affecting the QTL on SSC8. Furthermore,

from our knowledge this is the first study that suggests a mechanism of regulation of *ELOVL6* gene expression. Therefore, considering the metabolic similarities between pigs and humans [42], the regulatory mechanism described here may be applied to improve the knowledge of human lipid-related diseases, such as obesity, diabetes or metabolic syndrome.

CONCLUSIONS

In this work, we have described the complete genetic structure of porcine *ELOVL6* gene, showing two different isoforms expressed both in liver and adipose tissue. The *ELOVL6:c.-533C>T* mutation showed a stronger association than the two genotyped SNPs of the 3'UTR region, suggesting that the main regulatory region for *ELOVL6* gene expression in pigs is the promoter. Functional analyses of *ELOVL6* promoter showed the occupancy of SREBF1 and ERα, suggesting an important role of both transcription factors on *ELOVL6*gene expression regulation. Interestingly, the *ELOVL6:c.-394G>A* polymorphism, which is in LD with *ELOVL6:c.-533C>T* SNP, determine the ERα binding on *ELOVL6* promoter. In addition, ERα binding was associated with an increase on the percentages of methylation in the region comprised between -529 bp to -350 bp. Hence, *ELOVL6:c.-394G>A* SNP is the major candidate position to explain the differential expression of *ELOVL6* gene and, consequently, the palmitic and palmitoleic acids content in muscle and backfat.

AKNOWLEDGMENTS

We wish to thank the BAC-YAC Resource Center of the Animal Genetics Department of the INRA, especially Déborah Jardet, for the BAC screening. In addition, we are indebted to the *Centre de Recerca En Sanitat Animal* (CRESA) for its help in the BAC DNA isolation, especially to Marta Pérez. Finally, we also wish to thank to Dra. Sophie Pollet for her advices in ChIP performance. This work was funded by the *Ministerio de Economía y Competitividad* project AGL2011-29821-C02, and by the Innovation Consolider-Ingenio 2010 Program (CSD2007-00036, Centre for Research in Agrigenomics). J. Corominas was funded by a FPI PhD grant from the Spanish *Ministerio de Educación* (BES-2009-081223), Y. Ramayo-Caldas was funded by a FPU PhD grant from the Spanish *Ministerio de Educación* (AP2008-01450) and A. Puig-Oliveras was funded by a PIF PhD grant from the Universitat Autónoma de Barcelona (458-01-1/2011).

REFERENCES

1. Moon Y-A, Shah NA, Mohapatra S, Warrington JA, Horton JD: Identification of a Mammalian Long Chain Fatty Acyl Elongase Regulated by Sterol Regulatory Element-binding Proteins. *Journal of Biological Chemistry* 2001, 276(48):45358-45366.
2. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, Ballester M: Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE* 2013, 8(1):e53687.
3. Jakobsson A, Westerberg R, Jacobsson A: Fatty acid elongases in mammals: Their regulation and roles in metabolism. *Progress in Lipid Research* 2006, 45(3):237-249.
4. Guillou H, Zadravec D, Martin PGP, Jacobsson A: The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. *Progress in Lipid Research* 2010, 49(2):186-199.
5. Leonard AE, Bobik EG, Dorado J, Kroeger PE, Chuang LT, Thurmond JM, Parker-Barnes JM, Das T, Huang YS, Mukerji P: Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. *Biochemical Journal* 2000, 350(3):765-770.
6. Kitazawa H, Miyamoto Y, Shimamura K, Nagumo A, Tokita S: Development of a High-Density Assay for Long-Chain Fatty Acyl-CoA Elongases. *Lipids* 2009, 44(8):765-773.
7. Matsuzaka T, Shimano H, Yahagi N, Kato T, Atsumi A, Yamamoto T, Inoue N, Ishikawa M, Okada S, Ishigaki N *et al*: Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. *Nature Medicine* 2007, 13(10):1193-1202.
8. Turyn J, Stojek M, Swierczynski J: Up-regulation of stearoyl-CoA desaturase 1 and elongase 6 genes expression in rat lipogenic tissues by chronic food restriction and chronic food restriction/refeeding. *Molecular and Cellular Biochemistry*, 345(1):181-188.

9. Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, Díaz I, Ibánez-Escriche N, Noguera JL, Pérez-Enciso M *et al*: Genome-Wide Association Study for Intramuscular Fatty Acid Composition in an Iberian x Landrace Cross. *Journal of Animal Science* 2012, 90:1-11.
10. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Estellé J, Castelló A, Alves E, Pena RN, Ballester M, Folch JM: Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics* 2013. In revision
11. Knebel B, Haas J, Hartwig S, Jacob S, Köllmer C, Nitzgen U, Muller-Wieland D, Kotzka J: Liver-Specific Expression of Transcriptionally Active SREBP-1c Is Associated with Fatty Liver and Increased Visceral Fat Mass. *PLoS ONE* 2012, 7(2):e31812.
12. Gondret F, Ferré P, Dugail I: ADD-1/SREBP-1 is a major determinant of tissue differential lipogenic capacity in mammalian and avian species. *Journal of Lipid Research* 2001, 42(1):106-113.
13. Matsuzaka T, Shimano H, Yahagi N, Yoshikawa T, Amemiya-Kudo M, Hasty AH, Okazaki H, Tamura Y, Iizuka Y, Ohashi K *et al*: Cloning and characterization of a mammalian fatty acyl-CoA elongase as a lipogenic enzyme regulated by SREBPs. *Journal of Lipid Research* 2002, 43(6):911-920.
14. Kumadaki S, Matsuzaka T, Kato T, Yahagi N, Yamamoto T, Okada S, Kobayashi K, Takahashi A, Yatoh S, Suzuki H *et al*: Mouse Elovl-6 promoter is an SREBP target. *Biochemical and Biophysical Research Communications* 2008, 368(2):261-266.
15. Amemiya-Kudo M, Shimano H, Hasty AH, Yahagi N, Yoshikawa T, Matsuzaka T, Okazaki H, Tamura Y, Iizuka Y, Ohashi K *et al*: Transcriptional activities of nuclear SREBP-1a, -1c, and -2 to different target promoters of lipogenic and cholesterogenic genes. *Journal of Lipid Research* 2002, 43(8):1220-1235.
16. Ferrari A, Fiorino E, Giudici M, Gilardi F, Galmozzi A, Mitro N, Cermenati G, Godio C, Caruso D, De Fabiani E *et al*: Linking epigenetics to lipid metabolism: focus on histone deacetylases. *Molecular Membrane Biology* 2012, 29(7):257-266.

17. Devlin AM, Singh R, Wade RE, Innis SM, Bottiglieri T, Lentz SR: Hypermethylation of Fads2 and Altered Hepatic Fatty Acid and Phospholipid Metabolism in Mice with Hyperhomocysteinemia. *Journal of Biological Chemistry* 2007, 282(51):37082-37090.

18. Wang L, Chen L, Tan Y, Wei J, Chang Y, Jin T, Zhu H: Betaine supplement alleviates hepatic triglyceride accumulation of apolipoprotein E deficient mice via reducing methylation of peroxisomal proliferator-activated receptor alpha promoter. *Lipids in Health and Disease* 2013, 12(1):34.

19. Perez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, Folch JM, Babot D, Estany J, Oliver MA, Díaz I *et al*: A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *Journal of Animal Science* 2000, 78(10):2525-2531.

20. Muñoz M, Rodríguez MC, Alves E, Folch JM, Ibañez-Escriche N, Silió L, Fernández A: Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. *Submitted* 2013.

21. Mach N, Devant M, Díaz I, Font-Furnols M, Oliver MA, García JA, Bach A: Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. *Journal of Animal Science* 2006, 84(11):3039-3048.

22. Rogel-Gaillard C, Bourgeaux N, Billault A, Vaiman M, Chardon P: Construction of a swine BAC library: application to the characterization and mapping of porcine type C endoviral elements. *Cytogenetics and cell genetics* 1999, 85(3-4):205-211.

23. Rozen S, Skaletsky H, Misener S, Krawetz SA: Primer3 on the WWW for General Users and for Biologist Programmers. In: *Bioinformatics Methods and Protocols* Edited by Walker JM, vol. 132: Humana Press; 1999: 365-386.

24. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009, 25(14):1754-1760.

25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *Journal of Molecular Biology* 1990, 215(3):403-410.

26. Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castello A, Ballester M, Estelle J, Ibanez-Escriche N, Fernandez A, Perez-Enciso M *et al*: Liver

transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics* 2012, 13(1):547.

27. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009, 25(9):1105-1111.
28. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L: Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 2012, 7(3):562-578.
29. Thorvaldsdóttir H, Robinson JT, Mesirov JP: Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 2012, 14(2):178-192.
30. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP: Integrative genomics viewer. *Nature Biotech* 2011, 29(1):24-26.
31. Henderson CR: Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics* 1975, 31(2):423-447.
32. Henderson CR: Applications of Linear Models in Animal Breeding: University of Guelph; 1984.
33. Perez-Enciso M, Misztal I: Qxpak.5: Old mixed model solutions for new genomics problems. *BMC Bioinformatics* 2011, 12(1):202.
34. Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* 2003, 100(16):9440-9445.
35. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL: A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences* 1992, 89(5):1827-1831.
36. Colella S, Shen L, Baggerly K, Issa J, Krahe R: Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 2003, 35(1):146-150.
37. Tost J, Dunker J, Gut I: Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. *Biotechniques* 2003, 35(1):152-156.

38. Bartel DP: MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 2009, 136(2):215-233.

39. Dávalos A, Goedeke L, Smibert P, Ramírez CM, Warrier NP, Andreo U, Cirera-Salinas D, Rayner K, Suresh U, Pastor-Pareja JC *et al*: miR-33a/b contribute to the regulation of fatty acid metabolism and insulin signaling. *Proceedings of the National Academy of Sciences* 2011, 108(22):9232-9237.

40. Goedeke L, Vales-Lara FM, Fenstermaker M, Cirera-Salinas D, Chamorro-Jorganes A, Ramírez CM, Mattison JA, de Cabo R, Suárez Y, Fernández-Hernando C: A Regulatory Role for MicroRNA 33* in Controlling Lipid Metabolism Gene Expression. *Molecular and Cellular Biology* 2013, 33(11):2339-2352.

41. Reed BD, Charos AE, Szekely AM, Weissman SM, Snyder M: Genome-Wide Occupancy of SREBP1 and Its Partners NFY and SP1 Reveals Novel Functional Roles and Combinatorial Regulation of Distinct Classes of Genes. *PLoS Genet* 2008, 4(7):e1000133.

42. Bergen WG, Mersmann HJ: Comparative Aspects of Lipid Metabolism: Impact on Contemporary Research and Use of Animal Models. *The Journal of Nutrition* 2005, 135(11):2499-2502.

43. O'Hea EK, Leveille GA: 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. *The Journal of Nutrition* 1969, 99(3):338-344.

44. Kersten S: Mechanisms of nutritional and hormonal regulation of lipogenesis. *Embo Journal* 2001, 2(4):282-286.

45. Kersten S, Seydoux J, Peters JM, Gonzalez FJ, Desvergne B, xE, atrice, Wahli W: Peroxisome proliferator-activated receptor alpha mediates the adaptive response to fasting. *The Journal of Clinical Investigation* 1999, 103(11):1489-1498.

46. Schuch R, Agelopoulos K, Neumann A, Brandt B, Burger H, Korsching E: Site-specific chromatin immunoprecipitation: a selective method to individually analyze neighboring transcription factor binding sites in vivo. *BMC Research Notes* 2012, 5(1):109.

47. Kang SH, Vieira K, Bungert J: Combining chromatin immunoprecipitation and DNA footprinting: a novel method to analyze protein-DNA interactions in vivo. *Nucleic Acids Research* 2002, 30(10):e44.

48. Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Names A, Temper V, Razin A, Cedar H: Spl elements protect a CpG island from de novo methylation. *Nature* 1994, 371(6496):435-438.

49. Boumber YA, Kondo Y, Chen X, Shen L, Guo Y, Tellez C, Estécio MRH, Ahmed S, Issa J-PJ: An Sp1/Sp3 Binding Polymorphism Confers Methylation Protection. *PLoS Genetics* 2008, 4(8):e1000162.

50. Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G *et al*: Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 2008, 452(7183):45-50.

51. Lee C, Huang C-H: LASAGNA: A novel algorithm for transcription factor binding site alignment. *BMC Bioinformatics* 2013, 14(1):108.

52. Fritah A, Redeuilh Gr, Sabbah Ml: Molecular cloning and characterization of the human WISP-2/CCN5 gene promoter reveal its upregulation by oestrogens. *Journal of Endocrinology* 2006, 191(3):613-624.

53. Wood JD, Richardson RI, Nute GR, Fisher AV, Campo MM, Kasapidou E, Sheard PR, Enser M: Effects of fatty acids on meat quality: a review. *Meat Science* 2004, 66(1):21-32.

FIGURES AND TABLES

Table 1. Polymorphisms identified in the 3'UTR of the *ELOVL6* gene

Position (bp)	Polymorphism	Isoforms
1408	A/G	Variant 1 and 2
1817	C/T	
1922	C/T	
2070	C/T	
2532	C/T	
3599	G/T	Variant 2
3834	A/G	
4750	A/G	
4765	G/T	
4967	A/C	
5233	A/C	

[1]Positions relative to the coding region (GenBank: NW_003610943)

[2]SNPs genotyped in the BC1_LD population

Table2. Mean values of the methylation percentages of the six CpG-sites analyzed in theliver, adipose tissue, muscle and spleen.

TISSUE	CpG6	CpG5	CpG4	CpG3	CpG2	CpG1
Liver	35.3	31.5	21.5	3.3	7.2	1.7
Adipose tissue	51.9	44.6	33.9	3.9	12.4	2.6
Muscle	74.7	58.7	47.8	7	30.3	1.7
Spleen	62.5	50.1	41.6	7.7	21.5	2.2

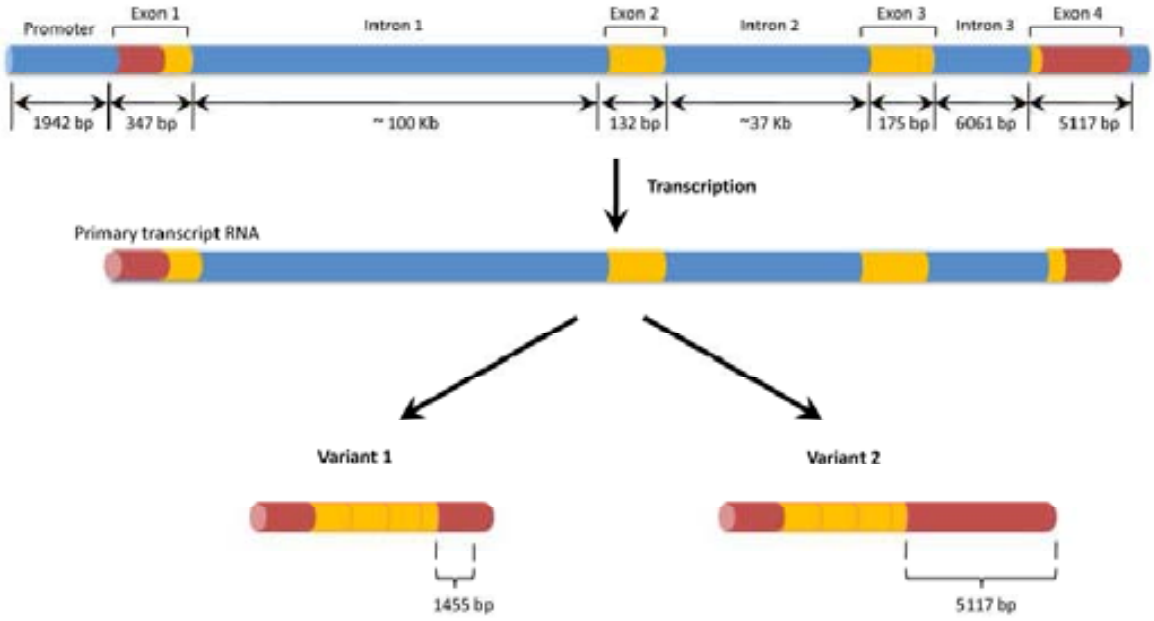

Figure 1. Genetic architecture of porcine *ELOVL6* gene, with the two transcribed variants identified by BAC screening and sequencing.

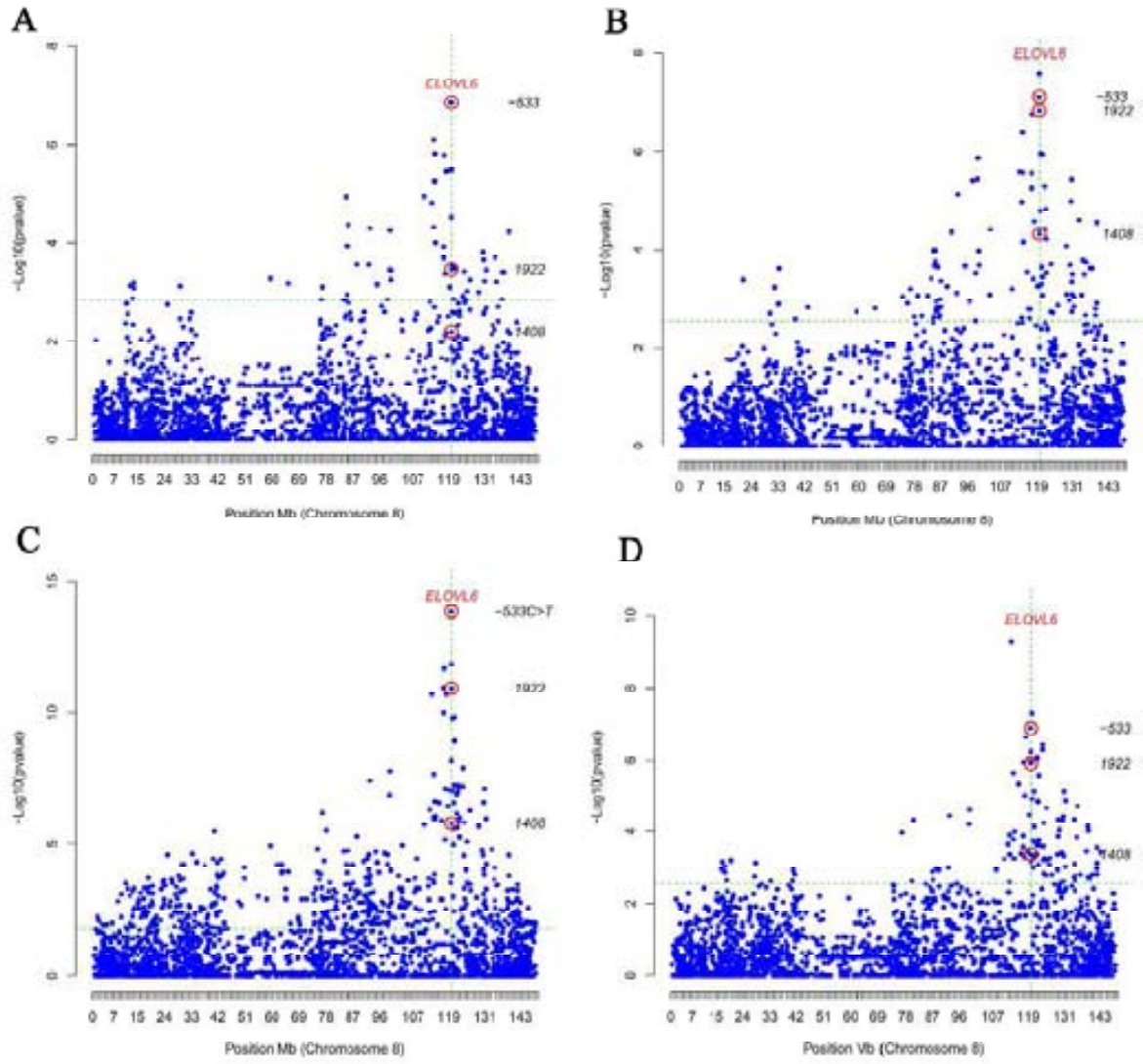

Figure2. Association analysis between SNP genotypes for SSC8 and the percentages of: palmitic (A) and palmitoleic (B) acids in muscle and palmitic (C) and palmitoleic (D) acids in backfat. *ELOVL6* polymorphisms are included and labeled with a red circle. Positions in Mb are relative to *Sscrofa 10.2 assembly* of the pig genome. Vertical dashed line indicates the position of *ELOVL6* gene and horitzontal dashed line marks the chromosome-wide significance level (FDR-based q-value ≤ 0.01).

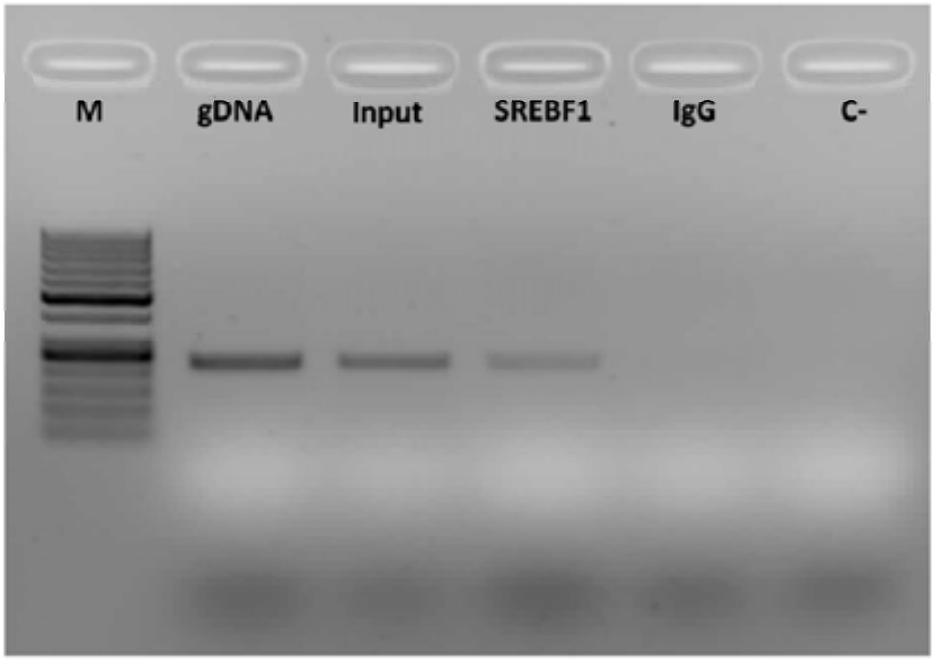

Figure 3. Ocuppancy of SREBF1 on *ELOVL6* promoter. Cross-linked chromatin from liver samples were immunoprecipitated with control IgG and anti-SREBF1 antibody. Recovered chromatin was subjected to PCR analysis using specific *ELOVL6* promoter primers.

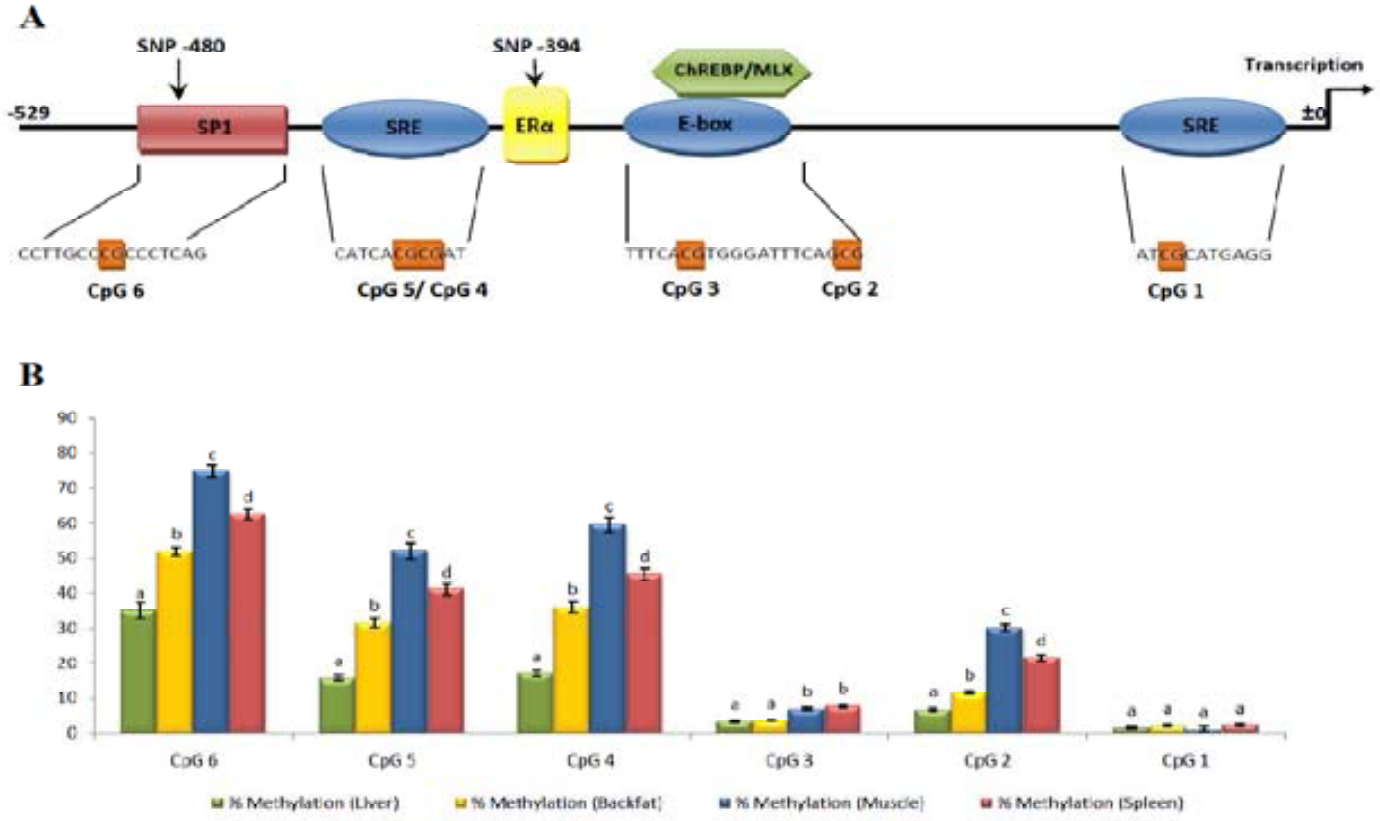

Figure 4. Characterization of the methylation patterns on *ELOVL6* gene promoter. (A) Schematic representation of the transcription factor binding elements studied, together with the CpG-sites analyzed in the methylation study. (B) Plot showing the percentages of methylation observed for each CpG-site in four porcine tissues: liver, backfat, muscle and spleen. Data represents means ± SEM. Values with different superscript letter (a, b, c and d) indicate significant differences between groups (p-value < 0.05) as determined by a linear model in R.

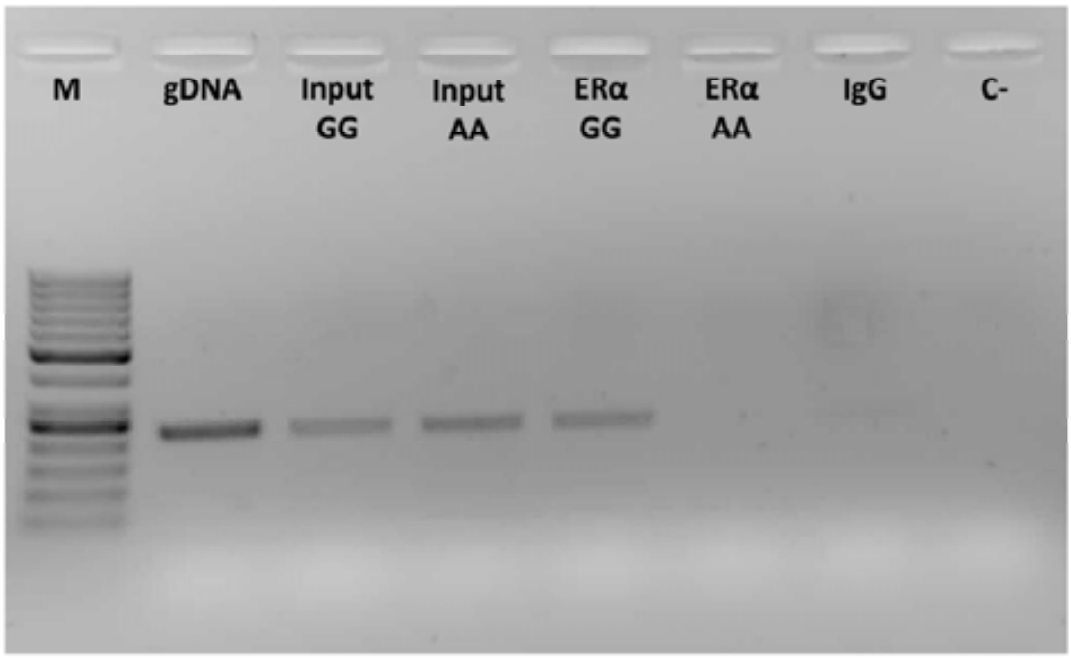

Figure 5. Differential occupancy of ERα on *ELOVL6* promoter depending on *ELOVL6:c.-394G>A* genotype. Cross-linked chromatin of hepatocytes from animals differing on *ELOVL6:c.-533C>T* genotype were immunoprecipitated with control IgG, anti-ERα antibodies. Recovered chromatin was subjected to PCR analysis using specific *ELOVL6* promoter primers.

4. GENERAL DISCUSSION

GENERAL DISCUSSION

Historically, animal breeding has been performed assuming the infinitesimal genetic model, in which traits are determined by a large number of unlinked and additive loci, each one with an small effect (Fisher 1919). Nevertheless, the lack of knowledge of the genetic architecture of the selected traits in breeding programs may produce a reduced response or the selection of undesired traits (Wood & Whittemore 2007). Hence, the understanding of the genetic basis of economically important traits related with growth, fatness, prolificacy, body composition and meat quality among others is essential to design better genetic selection strategies in livestock animals. In the last decades, the increasing interest in the study of genetic factors affecting meat quality traits led to important advances in the characterization of QTLs affecting these traits. Furthermore, new molecular genetics technologies including microarrays, high throughput SNP chips and next generation sequencing (NGS) technologies have allowed the use of new genomic approaches representing a significant improvement towards the identification of causal genes and mutations.

Despite FA composition is critical in determining pork meat quality, the molecular processes controlling FA composition are not fully understood. This PhD thesis illustrates the use of these new molecular genetic technologies, from the QTL fine-mapping to functional validation, going through the identification and characterization of candidate genes and mutations, affecting meat quality QTLs.

Additionally, the use of NGS technologies has generated an excellent opportunity to follow a more massive strategy: from individual gene analyses to a whole transcriptome profiling. The NGS approach allowed a more complete characterization of the transcriptome than microarrays does. RNA-Seq is capable to detect novel transcripts or isoforms, alternative splice sites, allele-specific expression and other advantages described in Table 1.5. Hence, we have characterized the porcine adipose tissue (backfat) transcriptome of pigs with divergent phenotypes for IMF FA composition, obtaining a more global view of the role of this tissue in lipid metabolism.

4.1.- QTL mapping: from QTL scan to GWAS and LDLA analyses

QTL mapping is focused on the identification of chromosome regions associated with a variation in phenotypic quantitative traits. This approach compares the phenotypic information of the trait of interest with the genetic information of molecular markers distributed in the whole genome. In 2003, Clop and collaborators performed the first QTL genome scan for backfat FA composition in pigs. In this study, several significant QTLs were found in SSC4, SSC6, SSC8, SSC10 and SSC12. A QTL on SSC8 (86 cM) affecting average chain length, palmitic and palmitoleic fatty acids content (Table 4.1) was confirmed when correcting either for carcass weight or backfat thickness (Clop *et al.* 2003). Later, Estellé *et al.* (2009b) evaluated the porcine *fatty acid binding protein 2* (*FABP2*) gene as a candidate gene for the SSC8 QTL. Despite no association was observed between the *FABP2:g.412T>C* polymorphism and FA composition traits, this SNP together with two additional microsatellites were used to better define the QTL position. In the same year, the *microsomal triglyceride transfer protein* (*MTTP*) was also evaluated as candidate gene for these QTLs (Estellé *et al.* 2009a) (Table 4.1).

In general, this family-based linkage analyses have been carried out with a limited number of microsatellite markers causing huge confidence intervals, which complicate the selection of candidate genes and the identification of causal mutations. Recently, the development of high-throughput technologies has facilitated the genotyping of thousands of SNPs covering the whole genome in an important number of individuals at reasonable costs. The replacement of microsatellites with high-density SNP arrays has been followed by the development of statistical approaches used for GWAS. In GWAS, the statistical association between a trait of interest and the genetic markers is analyzed, assuming that significant association can be detected because the SNPs are in LD with the causative mutation (Klein *et al.* 2005; Foulkes 2009; Mackay *et al.* 2009). In comparison with the classical QTL mapping approach, GWAS with high-density SNP arrays allows the discovery of a higher number of TASs and also facilitates the selection of candidate genes and the identification of candidate causal mutations due to the reduction of the confidence interval (Meuwissen & Goddard 2000; Goddard & Hayes 2009). In addition, GWAS has the advantage of using all recombination events after the mutations occurred, increasing the precision of the QTL localization (Meuwissen &

Goddard 2000). Moreover, GWAS considers the contribution of the variability within breeds or lines, whereas linkage analysis usually ignores it.

Table 4.1- Different QTL mapping approaches used for the SSC8 in the IBMAP cross.

Approach	Tissue	QTL trait	QTL position	Candidate gene	Reference
QTL scan	Backfat	C16:0 C16:1n-7 ACL	86 cM	----	(Clop *et al.* 2003)
QTL scan	Backfat	C16:0	88 cM	*FABP2* *MTTP*	(Estellé *et al.* 2009b) (Estellé *et al.* 2005) (Estellé *et al.* 2009a)
		C16:1n-7 ACL	95-96 cM		
GWAS (Sscrofa10)	Muscle*	C16:0 C16:1n-7 SFA C16:1/C16:0 C18:1/C16:1	92.1-96.7 Mb	----	(Ramayo-Caldas *et al.* 2012b)
GWAS & LDLA (Sscrofa10.2)	Backfat Muscle*	C16:0 C16:1n-7	117.8-119.8 Mb	*ELOVL6*	(Corominas *et al.* 2013b)
QTL scan	Backfat Muscle*	C16:0	86 -97 cM	----	(Muñoz *et al.* 2013)
		C16:1n-7			
GWAS (Sscrofa10.2)	Backfat Muscle*	C16:0	83.8-130.6 Mb	----	
		C16:1n-7	99.3-99.5 Mb 110.9-126.1 Mb		
GWAS & LDLA (Sscrofa10.2)	Backfat	C16:0 C16:1n-7 ACL	93 Mb	*MAML3* *SETD7*	(Revilla *et al.* 2013)
			119 Mb	*ELOVL6*	

* The muscle analyzed was *Longuissimus dorsi.*

Previous studies performed in our group identified QTLs related with intramuscular FA composition using genotypic information from Porcine SNP60K BeadChip (Illumina) (Ramayo-Caldas *et al.* 2012b). In this study, a strong association signal was observed in the 92.1-96.7 Mb genomic interval on SSC8 with the following traits: C16:0, C16:1n-7,

SFA, C16:1n-7/C16:0 and C18:1n-7/C16:1n-7 (Table 4.1). The positional concordance observed in QTLs affecting palmitic and palmitoleic acid content with the previously reported by Clop *et al.* (2003) suggested a pleiotropic effect of these QTLs in determining the FA composition in both, backfat and muscle. This hypothesis has recently been confirmed in a genome-wide analysis combining QTL scan and GWAS performed in the same animal material (Muñoz *et al.* 2013) (Table 4.1).

The use of a dense set of markers provides also the opportunity to simultaneously exploit the linkage and the linkage disequilibrium (LDLA) for QTL fine-mapping (Meuwissen & Goddard 2004; Druet & Georges 2009). This method was applied to reduce the confidence interval of SSC8 QTLs, obtaining the interval 117,824,822-119,887,525 bp for palmitic acid and the interval 117,824,360-119,727,822 bp for palmitoleic acid (Corominas *et al.* 2013b) (Table 4.1). The confidence interval reduction allowed us to focus all efforts on the minimal region, facilitating the causal gene detection. The QTL peak obtained from the LDLA analysis was at approximately 10 Mb from the *MTTP* gene in a region where the *ELOVL6* gene was located. Thus, we focused our studies in analyzing *ELOVL6* as candidate gene. Nevertheless, the genetic effect of *MTTP* polymorphism on the QTL could not be ruled out. In fact, the combination of GWAS and LDLA analysis showed a significant effect of *MTTP* p.Phe840Leu polymorphism in both, palmitic (p-value=6.45×10^{-05}) and palmitoleic (p-value=2.97×10^{-04}) acids. Taking into account these results, we could hypothesize that a new QTL affecting palmitic and palmitoleic acids content was located at 10 Mb from the original one, ranging from the 129.8 Mb to 131 Mb. This hypothetical second QTL was tested using models fitting one QTL against a model considering two different QTLs. Results obtained confirmed that a model with two QTLs were more likely for both palmitic (p-value= 2.17×10^{-06}) and palmitoleic acids (p-value = 3.67×10^{-07}). Hence, *MTTP* gene is still a promising candidate gene for this QTL and more functional studies of the non-synonymous mutation may be performed in a near future to elucidate its implication as a causal mutation for this QTL. Additionally, a recent association weight matrix (AWM) gene-network analysis showed a co-association between *MTTP* and *ELOVL6* genes, that was validated with a moderate significant co-expression between both genes in a liver network (r=0.53, p-value < 0.01) (Ramayo-Caldas 2013). The co-

expression observed may be caused by a co-regulation of both genes by the same regulatory factors in liver. In this sense, eQTL analyses with *MTTP* and *ELOVL6* gene expression in liver showed a common regulatory region at approximately 120 Mb of SSC8, proximal to *ELOVL6* gene (Ballester *et al.*, personal communication). This is a candidate region to contain transcription factors or microRNAs responsible of the co-expression observed between both genes. However, an alternative hypothesis is considering that a cis-acting genetic variant is modulating the *ELOVL6* gene expression and the consequently variations in FA composition controlled *MTTP* gene expression, such other lipid-related genes. In contrast, no significant co-expression was observed between these genes in a adipose tissue network (Ramayo-Caldas *et al.* 2013). The lack of co-expression observed in adipose tissue may be explained by the poor *MTTP* gene expression observed in the transcriptome of this tissue (Corominas *et al.* 2013a), suggesting a minor role of this gene in adipose tissue metabolism.

Recently, backfat fatty acid composition of 470 IBMAP F_2 animals was analyzed by combining single-marker association and haplotype-based approach (Revilla *et al.* 2013). This study was performed using a panel of 144 informative SNPs distributed along SSC8, mostly derived from the *Porcine SNP60K BeadChip*. A total of eleven FA composition traits showed significant association with the SNPs on SSC8: C14:0, C16:0, C18:0, C16:1n-7, C18:1n-9, C20:1n-6, ACL, C16:1n-7/C16:0, C18:0/C16:0, C18:1n-7/C16:1n-7 and C20:2n-6/C18:2n-6. Interestingly, Revilla and collaborators (2013) demonstrated that all traits (except the C20:2n-6/C18:2n-6 ratio) present two trait-associated SNP regions, the already known region at 119 Mb and a second one at 93 Mb of SSC8. Models fitting one QTL against a model considering two QTLs confirmed the presence of an additional QTL for the ten analyzed traits (Table 4.1). In addition, porcine *mastermind-like 3* (*MAML3*) and *SET domain containing 7* (*SETD7*) genes were analyzed as positional candidate genes of the 93-Mb region. Nevertheless, the QTL containing *ELOVL6* gene had a stronger association signal in comparison with the 93-Mb region for all the significant traits. In conclusion, SSC8 is clearly affecting FA composition, with at least three QTLs related with palmitic and palmitoleic acid content. Taking into account the three associated regions, QTL located at position 120 Mb has a higher effect compared with the other two regions. Therefore, the porcine

ELOVL6 gene is a strong positional candidate gene for explaining the phenotypic variation of FA composition in both, muscle and backfat tissues.

4.2.- Candidate genes affecting FA metabolism

The positional candidate gene analyses performed in this PhD thesis were focused on the QTLs affecting FA composition located in SSC8 (*ELOVL6*) and SSCX (*ACSL4*). Comprehensive studies of these genes are needed to determine both their effect on FA composition and also how these genes can improve the porcine selection programmes. ELOVL6 and ACSL4 proteins play crucial roles in FA metabolism, acting in the conversion of free FA to acyl-CoA and the elongation of this acyl-CoA.

4.2.2- ACSL4

In the IBMAP cross, the porcine *ACSL4* gene has been shown to be highly associated with the QTL on SSCX affecting live weight at slaughter and the percentages of oleic acid and MUFA (Mercadé *et al.* 2006a). The polymorphism identified and genotyped (*DQ144454:c.2645G>A*) was associated with the content of oleic acid and MUFA and, in consequence, this genetic variant is candidate for explaining the SSCX QTL. In this sense, we performed an expression analysis in 120 individuals with different genotypes for the *DQ144454:c.2645G>A* polymorphism. It is noteworthy the higher expression level observed in liver (mean ΔCT = 3.6) in comparison with adipose tissue (mean ΔCT = -1.9). This result may be explained by the different role of each tissue in FA β-oxidation, being pig liver the main tissue related with the degradation of FA. In agreement with (Mashek *et al.* 2006), different mechanisms controlling *ACSL4* gene expression were observed between tissues. In fact, significant differences were observed between *DQ144454:c.2645G>A* genotypes in liver, whereas in backfat no differential expression was observed. Coleman and collaborators (2002) demonstrated the major role of *ACSL4* in activating FA required for β-oxidation in rat liver. In contrast, studies performed with the pancreatic cell line INS 832/13 showed that *ACSL4* gene is important for maintaining the glucose-stimulated insulin secretion, via

the FA metabolites with signaling properties formed from arachidonic acid (C20:4n-6) (Klett *et al.* 2013). The positive effect of *ACSL4* gene on glucose metabolism observed in pancreatic cells is beneficial for the synthesis of triglycerides in lipogenic tissues, due to the production of pyruvate and consequently the formation of citrate in the Krebs cycle, which is used to initialize lipogenesis. This implication of *ACSL4* gene on glucose metabolism and lipogenesis explains the inhibitory effect of PUFA (mainly arachidonic acid) on *ACSL4* gene expression in pancreatic cells (Klett *et al.* 2013). This inhibitory effect may be extrapolated to lipogenic tissues such as adipose tissue, explaining the lower mRNA levels and the lack of differences on backfat gene expression.

Focusing on the expression analysis of *DQ144454:c.2645G>A* in liver, we observed that the G allele was associated with higher expression levels in comparison with the A allele (Corominas *et al.* 2012). In addition, the heterozygous sows showed an intermediate expression between the two homozygotes, suggesting a differential allelic expression with a higher expression of G allele in comparison to A allele. The allelic differential expression is common in mammals and it is an important genetic factor affecting heritable differences in phenotypic traits. The detection of differential allelic expression can be carried out using a Pyrosequencing method of RT-PCR products, which allowed the detection of differences as small as 4% (Neve *et al.* 2002; Wasson *et al.* 2002). This approach has been used in humans, domestic animals and plants (Sun *et al.* 2004; Schaart *et al.* 2005; Schimpf *et al.* 2008; Aminoff *et al.* 2009; Wang *et al.* 2012), but as far as we know this Pyrosequencing approach has not been used in pigs. The analysis was performed using the cDNA treated with DNases of 32 BC1_LD heterozygous sows. Promising results were obtained when the mean values were compared, observing significant higher percentages of G alleles (64%) in comparison with the percentages of the A alleles (36%) (p-value= $2.51x10^{-04}$) (Figure 4.1). In addition, significant correlations were obtained between the percentages of G allele and *ACSL4* gene expression values (r_G=0.44, p-value=0.01).

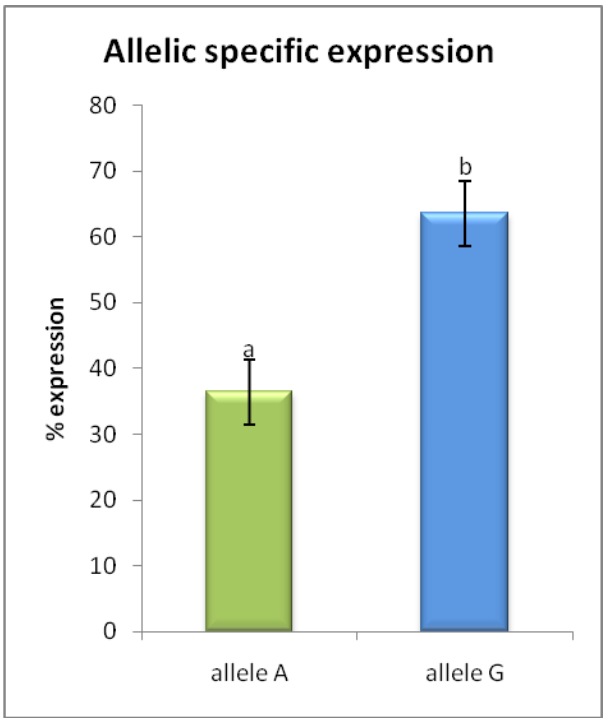

Figure 4.1- Plot representing the differences in pyrosequencing expression values between the two alleles of *DQ144454:c.2645G>A* in liver. Data represent means ± SEM. Values with different superscript letter (a and b) indicate significant differences between groups (p-value ≤ 0.05), as determined by a linear model in R.

Nevertheless, the high variability observed in the corresponding percentages of each allele (mean G allele = 63.59 ± 27.99 and mean A allele = 36.41 ± 27.99), did not allow us to conclude that the *DQ144454:c.2645G>A* polymorphism presents differential expression between both alleles. Probably, *DQ144454:c.2645G>A* polymorphism located at the 3'UTR of the gene is retaining a partial allelic differential expression (observed in the mean values) due to the fact that this SNP is in LD with the polymorphism that causes the allelic differential expression.

Differences in the expression levels between *DQ144454:c.2645G>A* genotypes observed in liver should be interpreted with caution because animals belong to an experimental cross in which linkage has an important effect. Therefore, it is possible that the significant effect observed for the *DQ144454:c.2645G>A* SNP was caused by LD between this polymorphism and the real causal mutation (Zhao *et al.* 2003; Varona

et al. 2005). Additionally, the high variability observed within groups of animals with the same genotype suggests that other polymorphisms may explain these expression differences. In order to find this hypothetical SNP, the proximal promoter was sequenced but no polymorphisms were identified (Corominas *et al.* 2012). However, a 250-bp-long CpG island was predicted in the proximal promoter, suggesting that this island may be relevant for the different regulation of *ACSL4* gene among tissues. Furthermore, the characterization of regulatory elements in *ACSL4* promoter showed similar binding elements as the ones described in the human *ACSL4* gene (Minekura *et al.* 2001). More recently, the murine *ACSL4* promoter has been characterized and transcription binding sites for *Sp1 transcription factor* (*Sp1*) and *CREB* were identified and validated experimentally (Orlando *et al.* 2013). Both Sp1 and CREB binding sites were previously identified in our porcine *ACSL4* promoter characterization, although the Sp1 binding site was not described in the manuscript (Corominas *et al.* 2012). The CREB family members are involved in the transactivation of target genes by the cAMP-dependent protein kinase (PKA) pathway, where PKA phosphorylates CREB in Ser133 leading to increased transcription (De Cesare & Sassone-Corsi 2000). In addition, the association of CREB with the *ACSL4* proximal promoter was confirmed by ChIP, validating the regulation of *ACSL4* gene expression by cAMP (Cano *et al.* 2006; Cooke *et al.* 2011; Orlando *et al.* 2013). On the other hand, Orlando and collaborators (2013) highlighted the implication of Sp1 in *ACSL4* regulation. EMSA and ChIP assays validated that this transcription factor was able to bind the *ACSL4* promoter, supporting the involvement of Sp1 in *ACSL4* gene regulation. The importance of both transcription factors in *ACSL4* gene expression is clear, and for this reason is relevant to know if these genes are located within the eQTL regions obtained from GWAS with liver *ACSL4* gene expression data. Interestingly, the porcine *CREB* gene is located at 122 Mb of SSC15, proximal to the marker MARC0039645 (119,375,975 bp of Sscrofa 10.2). This genomic region showed one of the strongest association signals with liver *ACSL4* gene expression (p-value= 1.46×10^{-04}) (Corominas *et al.* 2012). Gene annotation of the region containing MARC0039645 marker was not performed in our study, due to the lack of chromosome-wide significance obtained in SSC15. However, we suggested that associated peaks without chromosome-wide significance (i.e. *CREB* gene region) may

contain genes with biological relevance that should be considered. The case of Sp1 is clearly different because any associated peak was obtained around the localization of this gene (19 Mb of SSC5), despite the biological importance of this transcription factor.

The eQTL analysis performed in SSCX did not allowed the identification of significant associated regions with *ACSL4* gene expression, despite that the *DQ144454:c.2645G>A* polymorphism was the most associated one in the BC1_LD (Corominas *et al.* 2012). In addition, other factors located at autosomal chromosomes should not be discarded. Hence, the main objective of performing eQTL analysis was the identification of autosome regions associated with *ACSL4* gene expression. This study showed significant positions on SSC6 and SSC12, where candidate genes associated with *ACSL4* regulation were found: *acyl-CoA dehydrogenase, C-4 to C-12 straight chain* (*ACADM*) in SSC6 and *SREBP1* in SSC12 (Corominas *et al.* 2012). On the other hand, other factors such as the microRNA miR-224-5p have been described as regulators of fatty acid metabolism in adipocytes, by directly binding to *ACSL4* 3'UTR (Peng *et al.* 2013). This microRNA is located in SSCX at 35 Mb from *ACSL4* gene, very close to the second most significant marker in SSCX (MARC0052702; p-value= $1.57x10^{-03}$).

All together the results from this study were obtained using BC1_LD animals, which only carried Landrace SSCX, due to the experimental design. Therefore, a putative effect of Iberian breed cannot be observed by using these data. Hence, future studies performing GWAS with IBMAP F_2 population are needed, having special attention to the *DQ144454:c.2645G>A* polymorphism, MARC0039645 and MARC0052702 markers. These results are crucial for determining not only the relevance of Iberian breed on this QTL, but also the future experiments for validating this gene as the causal gene. Finally, taking into account the ten polymorphisms identified within the *ACSL4*-3'UTR region in Mercadé *et al.* (2006), we have identified four human microRNAs (miR-944, miR-675, miR-30c-1 and miR-30c-2) that could bind differentially in this region. The miR-30c are expressed in liver and adipose tissue and have recently been associated with a decrease of lipid synthesis and lipoprotein secretion (Soh *et al.* 2013). The miR-30c-1 and miR-30c-2 were located in the 157 Mb of SSC6 and 57 Mb of SSC1, respectively, in regions not associated with *ACSL4* gene expression. However, these

microRNAs should not be discarded, because the lack of association can be explained by absences of SNPs in this region that are segregating in our animal material. On the other hand, the genomic localization of the other two microRNAs is unknown, giving evidences of the need to improve the microRNA annotation for achieving a complete understanding of eQTL analyses.

4.2.1- ELOVL6

The content of palmitic (C16:0) and palmitoleic (C16:1n-7) acids is clearly depending on the elongation of saturated and monounsaturated FAs with 12-16 carbons to C18. This function is carried out by *ELOVL6* gene (Figure 4.2), fact that converts this gene in a promising candidate gene for explaining the phenotypic variation of the QTL on SSC8. The human *ELOVL6* gene spans more than 150 Kb which makes the sequencing of the complete gene a demanding task. Hence, we initially applied an alternative approach by sequencing only the coding region and the proximal promoter of the gene. The characterization of *ELOVL6* gene allowed the identification of eight polymorphisms: one synonymous polymorphism in exon 4 and seven in the promoter (Corominas *et al.* 2013b). In this study, we highlighted the *ELOVL6:c.-533C>T* polymorphism, which was highly associated with the palmitic and palmitoleic acid content in muscle and backfat. The analyses were carried out using BC1_LD animals of the IBMAP cross, but a recent study performed in our group has validated the same effects of *ELOVL6:c.-533C>T* polymorphism in the backfat of 168 animals from one selected family of the F_2 generation (Revilla *et al.* 2013). In this study, the *ELOVL6:c.-533C>T* polymorphism was the most associated SNP for both palmitic and palmitoleic acid content. Nevertheless, in our study the *ELOVL6:c.-533C>T* polymorphism did not show the highest association with palmitoleic acid content in backfat. These different results may be explained by both, the different number of animals (125 vs. 168) and the different genetic background (BC1_LD vs. F_2) used in each study.

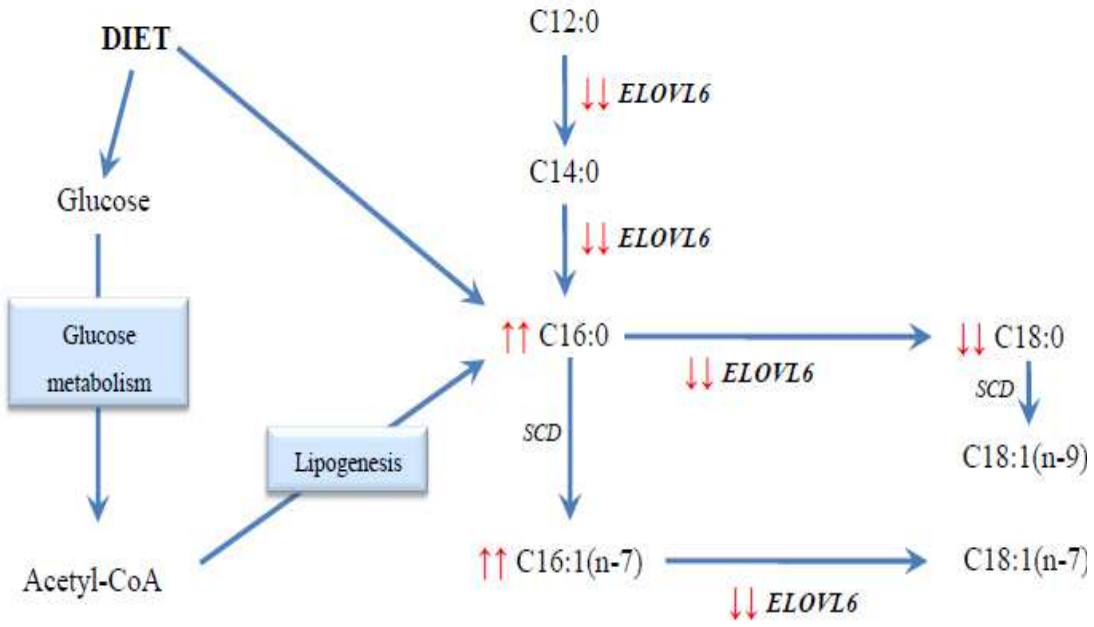

Figure 4.2- Schematic representation of the pathways controlling the composition of FA from C12 to C18. The percentages of C16:0 is affected by diet directly or through the glucose metabolism and lipogenesis. Then C16:0 is further elongated and desaturated in the endoplasmic reticulum by ELOVL6 and SCD. In red are represented the phenotypic variations observed in BC1_LD animals containing the Iberian allele. A decrease of *ELOVL6* gene expression is associated with an accumulation of C16:0 and C16:1(n-7) and a decrease on C18:0 content.

Additionally, Revilla and collaborators detected that *ELOVL6:c.-533C>T* was also the most associated SNP for the elongation ratios of C18:0/C16:0 and C18:1n-7/C16:1n-7, in concordance with the results obtained by Ramayo-Caldas and collaborators (2012b). Therefore, these results showed that the QTL on SSC8 is affecting the percentage of palmitic and palmitoleic acids content and the elongation ratios of C18:0/C16:0 and C18:1n-7/C16:1n-7 in both, IMF and backfat. However, it should be noted that the IBMAP studies that performed the mapping of SSC8 QTLs in the backfat of the F_2 generation identified significant effect on the ACL (Clop *et al.* 2003; Estellé *et al.* 2009b; Revilla *et al.* 2013). In contrast, this effect was not reported in the GWAS analysis for IMF composition performed in BC1-LD animals (Ramayo-Caldas *et al.* 2012b). This difference may be explained by both, the different genetic background of

each population and also the different metabolism of each tissue. To discard the genetic background hypothesis we performed a GWAS for ACL index on the backfat of BC1_LD animals, observing a clear significant association in this tissue (Figure 4.3A).

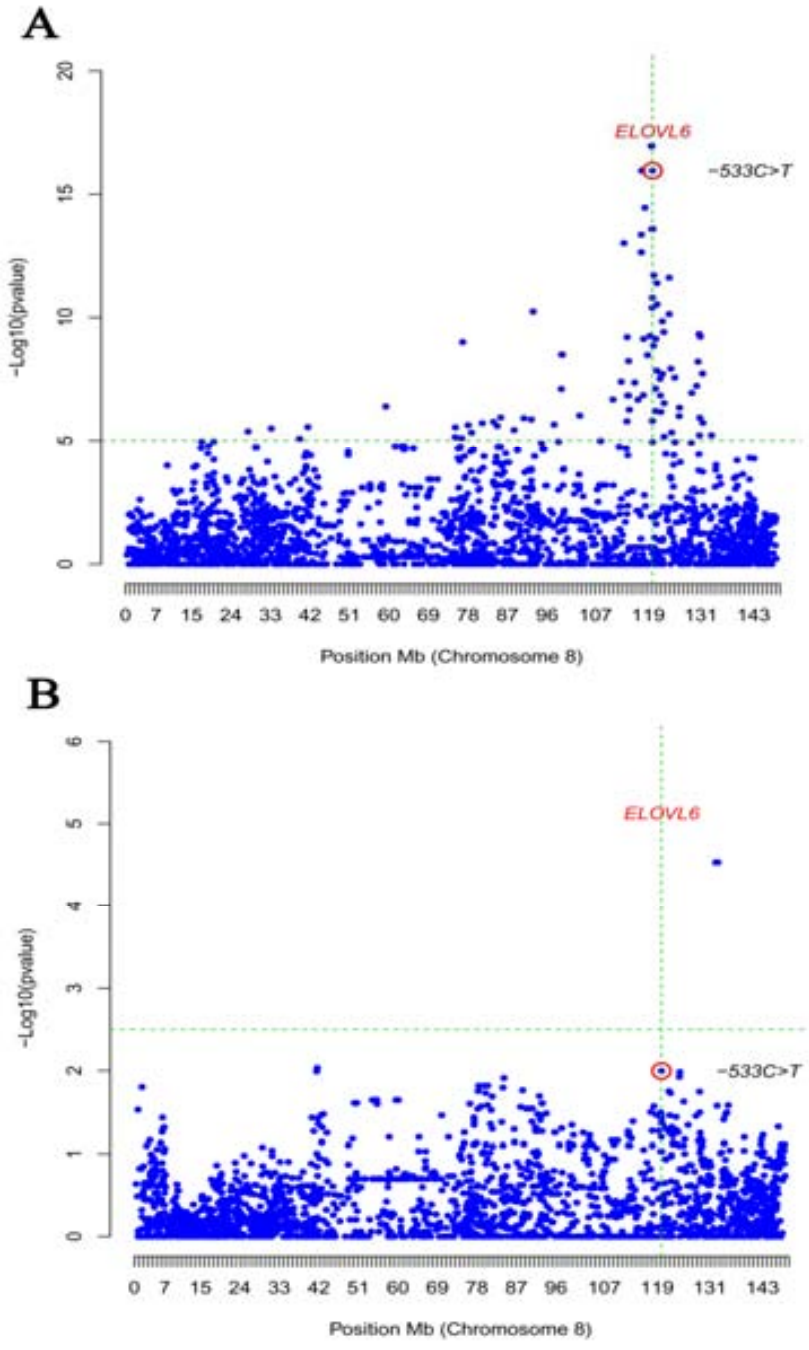

Figure 4.3- Association analysis between SNP genotypes for SSC8 and ACL index in backfat (A) and IMF (B). *ELOVL6:c.-533C>T* polymorphism is included and labeled with a red circle. Vertical, dashed lines indicate the location of *ELOVL6* gene. Horizontal, dashed lines mark the chromosome-wide significance level (FDR-based q-value ≤ 0.1). Positions in Mb are relative to *Sscrofa10.2 assembly* of the pig genome.

The *ELOVL6:c.-533C>T* polymorphism showed a high association with ACL index (p-value = $1.11x10^{-16}$), but similarly as occurs with the palmitoleic acid content in backfat this SNP is not the most significant. The clear difference observed in the ACL between tissues (Figure 4.3) is in concordance with the major lipogenic role of backfat in pigs. It is well known that the highest expression levels of *ELOVL* family genes are observed in the major lipogenic tissues: adipose tissue in pigs (Corominas *et al.* 2013b). In this sense, the higher *ELOVL6* gene expression in adipose tissue in comparison with muscle, reinforce the key role of this gene in determining the ACL of FA in backfat.

All together these results suggest a role of *ELOVL6:c.-533C>T* SNP in regulating the *ELOVL6* gene expression and, subsequently, affecting FA composition. With the purpose of validating this hypothesis, we analyzed the expression profile of *ELOVL6* gene in liver, backfat and muscle. Differences between *ELOVL6:c.-533C>T* genotypes were obtained only in backfat samples and, as was expected, the allele associated with a decrease in *ELOVL6* gene expression (C allele) showed an increase in palmitic and palmitoleic acid contents and a decrease of elongation activity ratios in muscle and backfat (Corominas *et al.* 2013b) (Figure 4.2). Results obtained showed that different levels of *ELOVL6* gene expression in adipose tissue seem to be affecting not only backfat FA composition but also that observed in muscle. This distal effect of *ELOVL6* gene expression may be caused by the fact that FA composition in IMF closely resembles that in adipose tissue (Ruiz *et al.* 1998; Yang *et al.* 2010). In this sense, we could hypothesize that variations on *ELOVL6* gene expression modify FA composition in backfat, but these changes have an effect on muscle FA composition because adipose tissue is the major source of free FA supplied to muscle (Cao *et al.* 2008). On the other hand, GWAS performed with *ELOVL6* gene expression showed that the most associated markers (ALGA0049135 and ALGA0049139) were located at 3 Mb downstream of the *ELOVL6* gene. These markers were also highly associated with the content of palmitic and palmitoleic acids in muscle (p-value $_{C16:0}$ = $1.71x10^{-06}$ and p-value $_{C16:1n-7}$ = $2.06x10^{-07}$) and backfat (p-value $_{C16:0}$ = $1.30x10^{-12}$ and p-value $_{C16:1n-7}$ = $2.12x10^{-06}$), therefore a polymorphism in LD with ALGA0049135 and ALGA0049139 markers cannot be discarded. Gene annotation of ALGA0049135 and ALGA0049139 genomic surroundings identified the *calcium/calmodulin-dependent protein kinase II*

beta (*CAMK2B*), which is a calmodulin kinase responsible for the activation of *cAMP responsive element binding protein* (*CREB*) by the phosphorylation of its Ser133. Previous studies have described the important role of CREB in controlling hepatic lipid metabolism by stimulating FA oxidation and inhibiting lipogenic pathways (Herzig *et al.* 2003). However, low levels of *CREB* gene expression (mean of 100 reads) were observed in adipose tissue. Despite that we cannot rule out a putative effect of this transcription factor in inhibiting lipogenic genes such as *ELOVL6*.

Associations observed between *ELOVL6* gene expression and the percentages of palmitic and palmitoleic acids content and the elongation ratios, convert *ELOVL6* in the main positional candidate gene for explaining the QTL in SSC8. Additionally, the *ELOVL6:c.-533C>T* SNP has been suggested as a candidate causal mutation for *ELOVL6* gene expression and FA composition traits. However, the 3'UTR of *ELOVL6* gene was not characterized in Corominas *et al.* (2013b), which is a candidate region to contain potential mutations that may alter gene expression. Unfortunately, the comparison of mRNA sequences between human, bovine and porcine confirmed that the current data of porcine *ELOVL6* gene was incomplete, lacking the *ELOVL6* 3'UTR. Thus, complicating the design of PCR primers for the analysis of this region. Nowadays, DNA sequencing of specific large DNA sequences can be achieved by several high-throughput approaches, such as sequencing of microdissected chromosomal regions (Weise *et al.* 2010), *de novo* assembly of RNA-Seq data (Zhao *et al.* 2011) or BAC screening and sequencing (Jeukens *et al.* 2011). Initially, we considered *de novo* assembly approach because we already had the RNA-Seq data of backfat transcriptome. However, *de novo* transcriptome assembly to obtain the DNA sequence of *ELOVL6* gene coding region was a demanding computational task. Finally, considering the high cost and the difficulties in equipment and management to perform the microdissection of chromosomal regions, we decided to apply the BAC screening and sequencing approach. The next generation sequencing of BACs containing the *ELOVL6* gene together with RNA-Seq data allowed us to identify two different isoforms of this gene expressed in both liver and backfat tissues (Corominas *et al.*, manuscript in preparation). Additionally, genomic data from the genome high-throughput sequencing of seven IBMAP parental animals (two Iberian boars and five Landrace

sows) and from the RNA-Seq study of twelve BC1_LD animals was used to identify polymorphisms in the 3'UTR (Corominas *et al.*, manuscript in preparation). The alignment of all mapped reads allowed for the identification of eleven polymorphisms in this region arranged into three haplotypes. These SNPs are *in silico* predicted as affecting the binding sites of several human microRNAs: miR-525-3p, miR-524-3p, miR-18a/b, miR-204 and miR-211. Only the microRNAs miR-18a and miR-204 are annotated in pigs, but located in regions where no associations with *ELOVL6* gene expression have been reported: 66 Mb of SSC11 and 250 Mb of SSC1, respectively. In addition, the miR-204 has been recently associated with a higher expression in the backfat of fatty pigs, suggesting that this microRNA may be relevant for adipogenesis and fat deposition (Chen *et al.* 2012). Other microRNAs related with the regulation of *ELOVL6* gene expression should be considered such as the miR-302 which has been associated with a decrease of hepatic *ELOVL6* gene expression in *low-density lipoprotein* (LDL) receptor knockout mice (Hoekstra *et al.* 2012). Unfortunately, the localization of this microRNA in pigs is also unknown, showing the importance of improving the microRNA annotation of the pig genome.

Association studies performed with *ELOVL6*-3'UTR polymorphisms showed significant association of *ELOVL6:c.1922C>T* SNP with the ACL of FAs in backfat and the palmitic and palmitoleic acid content in both IMF and backfat. Nevertheless, the *ELOVL6:c.-533C>T* polymorphism presents a higher association in comparison with *ELOVL6:c.1922C>T* polymorphism and also with the haplotype generated from both SNPs (Corominas *et al.*, manuscript in preparation). These results reinforce the key role of the promoter region in explaining the phenotypic variation of the QTL on SSC8. In this sense, a functional characterization of *ELOVL6* promoter was performed to determine the implication of epigenetic modifications and promoter polymorphisms on the transcription factor binding. In particular, the transcription factors analyzed were the SREBF1, for being the major known regulator of lipogenic genes (Moon *et al.* 2001; Matsuzaka *et al.* 2002; Horton *et al.* 2003; Kumadaki *et al.* 2008), and the ERα, due to its role in controlling promoter methylation levels (Metivier *et al.* 2008). Previously to the study of the occupancy of these transcription factors on *ELOVL6* promoter by ChIP approach, we checked the proper specificity of both antibodies (anti-

SREBF1 and anti-ERα) by western blotting and immunofluorescence (data not shown). Concerning the SREBF1, we showed for the first time experimental evidences regarding the union of this transcription factor on SRE elements located on porcine *ELOVL6* promoter: SRE-1 (-27 to -17 bp), SRE-2 (-460 to -449 bp) or SRE-3 (-532 to -524 bp) (Corominas *et al.*, manuscript in preparation). The regulation of *ELOVL6* gene by SREBF1 is well described in model species and for this reason we expected the union of this transcription factor. Nevertheless, the experimental design allowed only the validation of SREBF1 binding in *ELOVL6* promoter, but not at which element is bind. In mouse, Kumandaki and collaborators studies the relevance of each SREBF1 binding elements on gene expression. In this study, a minimal regulatory activity was observed on the SRE-2 upstream region (including SRE-3). In contrast, a major role of SRE-2 and SRE-1 elements on *ELOVL6* gene expression regulation was observed. Therefore, we suggest that SREBF1 may bind to at least one of these two SRE (SRE-1 and SRE-2) in animals on a fed state, enhancing *ELOVL6* gene expression. In fact, the methylation study performed gave new valuable evidences for understanding the real implication of SRE-1 and SRE-2 to enhance *ELOVL6* gene expression. SRE-1 was located at the in the region that showed low methylation pattern in all tissues analyzed, including those tissues without *ELOVL6* gene expression. On the other hand, SRE-2 includes two high methylated CpG sites in the core binding site. Consequently, results obtained in pigs suggest that SRE-2 element may have a role on *ELOVL6* gene expression, as a major regulatory mechanism or as an additional regulatory process to determine tissue-specific gene expression. However, site specific ChIP assays (Schuch *et al.* 2012) may be performed in order to validate this hypothesis.

Despite *ELOVL6:c.-533C>T* polymorphism was the tag SNP selected for genotyping, in the *ELOVL6* promoter there are four polymorphisms in LD with *ELOVL6:c.-533C>T* SNP (*ELOVL6:c.-534C>T, ELOVL6:c.-492G>A, ELOVL6:c.-394G>A* and *ELOVL6:c.-313C>T*) that may have a greater effect on fatty acid composition. A clear candidate SNP is the *ELOVL6:c.-394G>A* polymorphism, which is located within the binding site of ERα. In addition, results obtained in the ChIP analysis showed a differential binding of ERα depending on *ELOVL6:c.-394G>A* polymorphism (Corominas *et al.*, manuscript in preparation). Therefore, we suggest that this SNP may be implicated in ERα binding

and, consequently, it may be relevant on determining the mechanisms of *ELOVL6* gene expression regulation. Activation of ERα is mediated by the phosporylation of Serine 118 in the activation function 1 region (AF1) and can be potentiated by oestradiol (E_2) binding on activation function 2 (AF2) (Chen *et al.* 2000; Fritah *et al.* 2006). In fact, E_2 binding into the carboxy-terminal hydrophobic pocket of ERα produce three-dimensional changes of surfaces that allowed the recruitment of coactivators complexes (Brzozowski *et al.* 1997). At this point, two different mechanisms of transcriptional regulation by ERα were described. One mechanism consists on the recruitment of bridging proteins that help the binding of other transcription factor (Fritah *et al.* 2006). For instance, the CREB-binding protein (CREBBP) is recruited, allowing the CREB binding and enhancing the inhibitory effects of this transcription on *ELOVL6* gene expression. This data is in agreement with results obtained in the *ELOVL6* eQTL analysis with adipose tissue data (Corominas *et al.* 2013b), where significant associations were observed in the region proximal to *CAMK2B* gene, which is an activator of CREB. The alternative mechanism is based on Métivier and collaborators work (2008), in which authors demonstrated that ERα produced dynamic demethylation/methylation cycles by the recruitment of Dnmts. Interestingly, this mechanism may explain the higher methylation levels observed in the more distal region of *ELOVL6* promoter, which was suggested as the main regulatory region for *ELOVL6* gene expression.

In conclusion, our hypothesis suggest that the differential occupancy of *ELOVL6* promoter by ERα determined by *ELOVL6:c.-394G>A* SNP affects the methylation levels of SRE-2 (CpG4 and CpG5) and SP1 binding site (CpG6). Hence, animals carrying the *ELOVL6:c.-394G* allele allowed ERα binding, fact that is traduced on a higher methylation of SRE and Sp1 binding site and a decrease of *ELOVL6* gene expression. The opposite occurred in animals with *ELOVL6:c.-394A* allele, which showed lower methylation and higher *ELOVL6* gene expression. At this point, the alteration of *ELOVL6* gene expression may be the major causal factor on determining the different palmitic and palmitoleic acid composition in muscle and backfat. Therefore, the results obtained suggest the *ELOVL6:c.-394G>A* polymorphismas a likely causal factor of the QTL affecting fatty acid composition in SSC8. However, the genotyping of *ELOVL6:c.-*

394G>A SNP and the performance of association studies with the traits of interest are essential to validate this hypothesis.

4.3.- Regulation of Gene expression in FA metabolism

Dietary fat is essential for the development of all organisms because of its role as energy source, its effects on membrane lipid composition and on gene expression. Some FA or their metabolites act like hormones to control the activity or abundance of several lipid-related transcription factors, including *peroxisome proliferator-activated receptors family* (*PPAR*), *hepatocyte nuclear factor alpha* (*HNFAα*) and *SREBP1*. FA can regulate these transcription factors by a direct binding to its promoter, by activating signalling cascades that target to the nucleus via G-protein-linked cell surface receptors or by regulating the intracellular calcium levels (Jump & Clarke 1999). It has been suggested that the effect of dietary fat on gene expression is the consequence of an adaptive response to changes in the quantity and type of fat ingested (Jump 2002). For example, ingestion of high saturated fat diets increase triglycerides and promotes insulin resistance, hypertension and obesity (Kraegen *et al.* 1991; Katan *et al.* 1994; Chizzolini *et al.* 1999). Nevertheless, changing the diet to one with a major content of PUFA ameliorates the insulin resistance and abnormal lipid profile (Storlien *et al.* 1987; de Lorgeril & Salen 2012). In addition, high PUFA diets cause a rapid activation of genes related with lipid oxidation and a decrease in genes encoding enzymes of lipid synthesis (Clarke *et al.* 1997; Bergen & Mersmann 2005; Schmitz & Ecker 2008). These PUFA effects were previously observed in the liver transcriptome of two groups of pigs differing in the PUFA content. Animals with high percentage of PUFA showed and overexpression of genes regulated by *PPARα*, suggesting that PUFA promote FA oxidation in these animals (Ramayo-Caldas *et al.* 2012a). Following a similar strategy, in this thesis we have analyzed the adipose tissue transcriptome of two groups: i) animals with higher content of PUFA and lower content of SFA and MUFA (group H) and ii) animals with a lower PUFA percentage and higher SFA and MUFA percentages (group L) (Corominas *et al.* 2013a). Although all animals were fed with the same diets, clear differences were observed in intramuscular FA composition. These phenotypic

differences are likely determined by genetic variability in the absorption of essential FA or in the following steps of transport, deposition, storage or degradation of these FA.

To gain insight into the FA regulatory processes in adipose tissue transcriptome that differed between two groups of animals phenotypically extreme for intramuscular FA composition, the list of 396 differentially-expressed genes was explored using the core analysis function included in Ingenuity Pathway Analysis (IPA). Interestingly, animals with lower content of PUFA (L group) showed an overexpression of the most important genes related with *de novo* FA synthesis (*ACLY*, *ACSS2*, *ACACA*, *FASN*, *SCD* and *ELOVL6*). This result is in agreement with differences observed in the transcriptome of lean pigs (Landrace) and obese pigs (Rongchang) (Li *et al.* 2012). In addtion, the most representative canonical pathway significantly modulated between groups was the LXR/RXR activation pathway. This pathway is important for the regulation of *de novo* FA synthesis and its down-representation in animals of H group may be explained by the negative effect of PUFAs on LXR (Duplus & Forest 2002; Sampath & Ntambi 2005; Schmitz & Ecker 2008). In contrast, no effects were observed in the FA oxidation pathway, giving new evidences of the poor role of porcine adipose tissue in FA oxidation. Therefore, these results suggest that porcine adipose tissue is important for *de novo* FA synthesis (O'Hea & Leveille 1969), but less relevant for FA oxidation, which is mainly produced in porcine liver.

The functional and anatomical separation of *de novo* lipid synthesis and FA oxidation in the porcine adipose tissue and liver, respectively, suggests a tightly coordinated process between transcription factors, hormones and nuclear receptors to avoid the simultaneously activation of these antagonistic pathways. Hence, for improving our knowledge of the interconnectivity and regulation of these processes is essential to improve the porcine genome annotation to create a global view of the whole FA metabolism. For example, in the current annotation of the porcine genome the *adiponectin* (*ADIPOQ*) gene is not present. This gene encodes an adipokine secreted by adipose tissue that modulates several biological processes such as lipogenesis, gluconeogenesis, FA oxidation, glucose uptake and insulin sensitivity (Dall'Olio *et al.* 2009). Adiponectin receptors are present in numerous organs and tissues; thus, *ADIPOQ* is a clear candidate gene for explaining how adipose tissue is affecting the FA

composition in muscle. Another example is the *fatty acid binding protein 4* (*FABP4*), which is highly expressed in adipocytes and the encoding protein is responsible for the transport of FA inside the cells. It is well known that FABP4 transports FA from the cellular membrane to different cell location depending on the biological process: to mitochondria and peroxisome for FA oxidation, to endoplasmic reticulum for triglyceride and phospholipids synthesis or to the nucleus for gene expression regulation (Zimmerman & Veerkamp 2002). Currently, the detection of these and other genes with a known or unknown function in FA metabolism is difficult following a RNA-Seq approach. Nevertheless, the information regarding these genes is present in the RNA-Seq data obtained, in the group of unmapped reads. Therefore, these unmapped reads can be used to perform a single gene study, with the condition of knowing the gene sequence. *FABP4* gene was selected to perform a specific analysis with RNA-Seq data, because of its role in transporting FA in the different lipidic processes in adipose tissue, its hypothetical role as *ELOVL6* regulator and also for being a candidate gene for explaining the QTL on SSC4 related to fat deposition (Mercadé *et al.* 2006b). The pipeline followed in this study was the same as in the RNA-Seq analysis (Corominas *et al.* 2013a), but changing the genome reference and genome annotation files for modified ones with the only information of the *FABP4* sequence and annotation. Results obtained showed that mean gene expression not differed between groups (L=96,995 reads and H =92,076 reads), but the high expression of this gene in adipose tissue is in concordance with its importance in adipose tissue FA metabolism. In conclusion, the RNA-Seq approach is an excellent opportunity to analyze whole transcriptomes in a massive way and the lack of data for certain genes in the genome annotation could be solved by mapping the discarded reads to the specific gene sequence. However, this study is not possible if we are interested in genes with unknown coding sequence, due to a reference sequence is needed to map the reads generated from the high-throughput sequencing of specific transcriptomes. In our case, we provide valuable data for improving the genome annotation. In fact, we have detected 4,130 putative unannotated protein-coding sequences in intergenic regions, from which 1,596 showed similarities with predicted or known proteins in pigs. Furthermore, we have observed that an important percentage of unmapped reads

(36%) are classified as transposable elements. Nevertheless, to achieve a complete genome sequence and annotation is necessary the continuity of coordinated projects between expert groups in the field to obtain a similar annotation to that achieved in human genome.

4.4.- Future directions

In the last decade, the important reduction of the sequencing costs was crucial for the genomic studies of non-model species, such as domestic animals. Porcine genome publication has been a turning point for the genetic studies of this specie, by providing new data for candidate gene analyses and the identification of causal factors of QTLs. However, major improvements are needed for the near future at both, sequence level and genome annotation of genes, microRNAs and other non-coding RNAs. Results obtained in association analyses and RNA-Seq studies performed in this thesis are limited to the current genome annotation knowledge. Some of the unannotated genes or microRNAs that have been discussed may have relevant roles in metabolic pathways affecting lipid metabolism. Therefore, future improvements of porcine genome annotation may provide new valuable data for interpreting the biological processes under consideration.

The reductionist trend of the sequencing costs has allowed many groups to focus their genomic studies on the complete characterization of tissue transcriptomes. However, a demanding task on filtering, selecting and validating those SNPs are needed in the near future to determine the implication of these SNPs on the lipid-related QTLs and its value to perform gene-assisted selection on fatty acid composition traits without influencing other pig production traits. Additionally, in the next years the domestic species genomics will suffer a considerably increase on the number of studies focused on the complete genome sequencing of selected animals. This fact will provide more knowledge of different sources of genetic variation such as structural variants (CNVs, inversions, translocations and indels) and SNPs, including those in non-transcribed regions. In fact, the rapid evolution of the non-transcribed regulatory sequences observed in the ENCODE project has suggested that this regions may contain most of the QTL causative SNPs (Maher 2012). The continuous achievements on sequencing

techniques need to be followed by new bioinformatic tools capable to analyze this huge amount of data. Thus, bioinformatic tools must improve to facilitate the interpretation of data obtained not only from RNA-Seq analyses, but also from the current and future studies focused on the whole genome sequencing and epigenomics. Hence, the development of high-throughput sequencing techniques allows us to obtain the sequence of a whole genome, but the storage, processing, analysis and the interpretation of this huge amount of data is a significant task that holds many obstacles and challenges (Koboldt *et al.* 2010). Nowadays, emerging integrative approaches that combine multiple sources of information, such as system biology approaches, are being applied to address these challenges. Finally, advances achieved in the last decades in data obtaining and analysis gives us the necessary tools for elucidating the high degree of complexity observed in the genome functionality. Henceforth, the research communities started a long path to interpret the differences sources of data, a fact that will require the contribution of experts in different fields in a global and multidisciplinary effort.

5. CONCLUSIONS

CONCLUSIONS

1. Different regulatory mechanisms control *ACSL4* gene expression in porcine liver and adipose tissue, being most expressed in liver. In this tissue, *ACSL4* mRNA levels are significantly affected by *DQ144454:c.2645G>A* polymorphism, being higher in animals with the G allele in comparison with animals carrying the A allele. This differential expression between genotypes in liver may be explained by the significant regions observed in SSC6 and SSC12 in a GWAS analysis of *ACSL4* gene expression.

2. Fine mapping of QTL on SSC8 allowed the identification of *ELOVL6* as a promising candidate gene for explaining the phenotypic variation on palmitic and palmitoleic FA content in muscle and adipose tissue. The complete genetic architecture of *ELOVL6* gene was described for the first time in pigs, showing two different isoforms expressed in both, liver and adipose tissue.

3. The characterization of *ELOVL6* gene allowed the identification of nineteen polymorphisms distributed along the gene: seven in the promoter region, one synonymous polymorphism in exon four and the remaining eleven SNPs on the 3'UTR. Association analyses performed with five tag genotyped polymorphisms showed that the promoter *ELOVL6:c.-533C>T* polymorphism had the strongest statistical significance for *ELOVL6* gene expression and the percentages of palmitic and palmitoleic FA acids content. Therefore, this polymorphism together with the other four promoter SNPs in linkage disequilibrium (*ELOVL6:c.-534C>T*, *ELOVL6:c.-492G>A*, *ELOVL6:c.-394G>A* and *ELOVL6:c.-313C>T*) are potential candidate polymorphisms to explain the SSC8 QTL.

4. Functional analyses of *ELOVL6* promoter demonstrated the occupancy of SREBF1 and ERα, confirming the important role of both transcription factors on *ELOVL6* gene expression regulation. In fact, a differential binding of ERα caused by *ELOVL6:c.-394G>A* polymorphism was observed, showing union only in

animals carrying the G allele. Furthermore, the *ELOVL6:c.-394G* allele was associated with higher methylation levels in the distal region of *ELOVL6* promoter. Therefore, results suggested that differences on ERα binding and the methylation levels are the major factor producing the *ELOVL6* gene expression variation Assuming linkage disequilibrium between *ELOVL6:c.-394G>A* and *ELOVL6:c.-533C>T* polymorphisms we propose the *ELOVL6:c.-394G>A* SNP as a putative causal mutation for explaining the phenotypic variation of the QTL analyzed.

5. A global view of pig adipose tissue transcriptome has been obtained by RNA-Seq, allowing the identification of 4,130 putative unannotated protein-coding sequences and new transposable elements.

6. The transcriptome comparison between two groups (H and L) of pigs with extreme phenotypes for intramuscular fatty acid composition allowed the identification of 396 differentially-expressed genes. The 62% of these genes had higher expression in animals with higher content of intramuscular SFA and MUFA, while the remaining 149 genes showed higher expression in the group with higher content of PUFA.

7. Pathway analysis of differential-expressed genes revealed that the major metabolic pathway differentially modulated between H and L groups was *de novo* lipogenesis, being the group with more PUFA the one that showed lower expression of lipogenic genes. Changes in lipogenesis is crucial for determining intramuscular FA composition, therefore, the differentially-expressed genes are clear candidate genes to affect FA composition traits in pigs.

6. REFERENCES

REFERENCES

Aminoff A., Ledmyr H., Thulin P., Lundell K., Nunez L., Strandhagen E., Murphy C., Lidberg U., Westerbacka J., Franco-Cereceda A., Liska J., Nielsen L.B., Gafvels M., Mannila M.N., Hamsten A., Yki-Järvinen H., Thelle D., Eriksson P., Borén J. & Ehrenborg E. (2009) Allele-specific regulation of MTTP expression influences the risk of ischemic heart disease. *Journal of Lipid Research* **51**, 103-11.

Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nature Reviews Genetics* **2**, 130-8.

Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M., Hansson I. & al. e. (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771-4.

Archibald A., Bolund L., Churcher C., Fredholm M., Groenen M., Harlizius B., Lee K.-T., Milan D., Rogers J., Rothschild M., Uenishi H., Wang J., Schook L. & the Swine Genome Sequencing C. (2010) Pig genome sequence - analysis and publication strategy. *BMC Genomics* **11**, 438.

Archibald A.L., Haley C.S., Brown J.F., Couperwhite S., McQueen H.A., Nicholson D., Coppieters W., Weghe A., Stratil A., Winterø A.K., Fredholm M., Larsen N.J., Nielsen V.H., Milan D., Woloszyn N., Robic A., Dalens M., Riquet J., Gellin J., Caritez J.C., Burgaud G., Ollivier L., Bidanel J.P., Vaiman M., Renard C., Geldermann H., Davoli R., Ruyter D., Verstege E.J.M., Groenen M.A.M., Davies W., Høyheim B., Keiserud A., Andersson L., Ellegren H., Johansson M., Marklund L., Miller J.R., Anderson Dear D.V., Signer E., Jeffreys A.J., Moran C., Tissier P., Muladno, Rothschild M.F., Tuggle C.K., Vaske D., Helm J., Liu H.C., Rahman A., Yu T.P., Larson R.G. & Schmitz C.B. (1995) The PiGMaP consortium linkage map of the pig (Sus scrofa). *Mammalian Genome* **6**, 157-75.

Bartel D.P. (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136**, 215-33.

Béjar F., Dobao M.T., García J., Rodrigáñez J., Rodríguez M.C., Silió L., Toro M.A., Campo J.M., Rodríguez R. & Sánchez-Palomo J. (1992) Investigaciones en la Piara de Cerdo Ibérico "El Dehesón del Encinar". *Junta de Castilla-La Mancha, Toledo, Spain*.

Bergen W.G. & Mersmann H.J. (2005) Comparative Aspects of Lipid Metabolism: Impact on Contemporary Research and Use of Animal Models. *The Journal of Nutrition* **135**, 2499-502.

Boumber Y.A., Kondo Y., Chen X., Shen L., Guo Y., Tellez C., Estécio M.R.H., Ahmed S. & Issa J.-P.J. (2008) An Sp1/Sp3 Binding Polymorphism Confers Methylation Protection. *PLoS Genetics* **4**, e1000162.

Bourneuf E., Du Z.-Q., Estellé J., Gilbert H.l.n., Créchet F.o., Piton G., Milan D., Geffrotin C., Lathrop M., Demenais F., Rogel-Gaillard C. & Vincent-Naulleau S. (2011) Genetic and functional evaluation of MITF as a candidate gene for cutaneous melanoma predisposition in pigs. *Mammalian Genome* **22**, 602-12.

Brzozowski A.M., Pike A.C.W., Dauter Z., Hubbard R.E., Bonn T., Engstrom O., Ohman L., Greene G.L., Gustafsson J.-A. & Carlquist M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753-8.

Burgos-Paz W., Souza C.A., Megens H.J., Ramayo-Caldas Y., Melo M., Lemus-Flores C., Caal E., Soto H.W., Martinez R., Alvarez L.A., Aguirre L., Iniguez V., Revidatti M.A., Martinez-Lopez O.R., Llambi S., Esteve-Codina A., Rodriguez M.C., Crooijmans R.P.M.A., Paiva S.R., Schook L.B., Groenen M.A.M. & Perez-Enciso M. (2013) Porcine colonization of the Americas: a 60k SNP story. *Heredity* **110**, 321-30.

Cameron N.D. & Enser M.B. (1991) Fatty acid composition of lipid in Longissimus dorsi muscle of Duroc and British Landrace pigs and its relationship with eating quality. *Meat Science* **29**, 295-307.

Cano F., Poderoso C., Cornejo Maciel F., Castilla R., Maloberti P., Castillo F., Neuman I., Paz C. & Podestá E.J. (2006) Protein tyrosine phosphatases regulate arachidonic acid release, StAR induction and steroidogenesis acting on a hormone-dependent arachidonic acid-preferring acyl-CoA synthetase. *The Journal of Steroid Biochemistry and Molecular Biology* **99**, 197-202.

Cao H., Gerhold K., Mayers J.R., Wiest M.M., Watkins S.M. & Hotamisligil G.k.S. (2008) Identification of a Lipokine, a Lipid Hormone Linking Adipose Tissue to Systemic Metabolism. *Cell* **134**, 933-44.

Cao Y., Dave K.B., Doan T.P. & Prescott S.M. (2001) Fatty Acid CoA Ligase 4 Is Up-Regulated in Colon Adenocarcinoma. *Cancer Research* **61**, 8429-34.

Cao Y., Traer E., Zimmerman G.A., McIntyre T.M. & Prescott S.M. (1998) Cloning, Expression, and Chromosomal Localization of Human Long-Chain Fatty Acid-CoA Ligase 4 (FACL4). *Genomics* **49**, 327-30.

Carey M. & Smale S.T. (2000) Transcriptional regulation in eukaryotes: concepts, strategies, and techniques. *Cold Spring Harbor Laboratory Press New York.*

Carrapiso A.I., Bonilla F. & García C. (2003) Effect of crossbreeding and rearing system on sensory characteristics of Iberian ham. *Meat Science* **65**, 623-9.

Casellas J., Noguera J.L., Reixach J., Díaz I., Amills M. & Quintanilla R. (2010) Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs. *Journal of Animal Science* **88**, 2246-54.

Čepica S., Bartenschlager H. & Geldermann H. (2007) Mapping of QTL on chromosome X for fat deposition, muscling and growth traits in a wild boar × Meishan F2 family using a high-density gene map. *Animal Genetics* **38**, 634-8.

Čepica S., Masopust M., Knoll A., Bartenschlager H., Yerle M., Rohrer G.A. & Geldermann H. (2006) Linkage and RH mapping of 10 genes to a QTL region for fatness and muscling traits on pig chromosome X. *Animal Genetics* **37**, 603-4.

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

Clarke S.D., Turini M. & Jump D. (1997) Polyunsaturated fatty acids regulate lipogenic and peroxisomal gene expression by independent mechanisms. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **57**, 65-9.

Clop A., Cercós A., Tomàs A., Pérez-Enciso M., Varona L., Noguera J.L., Sànchez A. & Amills M. (2002) Assignment of the 2,4-dienoyl-CoA reductase (DECR) gene to porcine chromosome 4. *Animal Genetics* **33**, 164-5.

Clop A., Marcq F., Takeda H., Pirottin D., Tordoir X., Bibe B., Bouix J., Caiment F., Elsen J.-M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C. & Georges M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics* **38**, 813-8.

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

Coleman R.A., Lewin T.M., Van Horn C.G. & Gonzalez-Baró M.R. (2002) Do Long-Chain Acyl-CoA Synthetases Regulate Fatty Acid Entry into Synthetic Versus Degradative Pathways? *The Journal of Nutrition* **132**, 2123-6.

Cooke M., Orlando U., Maloberti P., Podestá E.J. & Maciel F.C. (2011) Tyrosine phosphatase SHP2 regulates the expression of acyl-CoA synthetase ACSL4. *Journal of Lipid Research* **52**, 1936-48.

Cooper D.K. (2012) A brief history of cross-species organ transplantation. *Baylor University Medical Center Proceedings* **25**, 49-57.

Coppieters W., Van de Weghe A., Depicker A., Coppieters J., Peelman L., Van Zeveren A. & Bouquet Y. (1995) Polymorphic CAC/T repetitive sequences in the pig genome 1. *Animal Genetics* **26**, 327-30.

Corominas J., Ramayo-Caldas Y., Castelló A., Muñoz M., Ibáñez-Escriche N., Folch J.M. & Ballester M. (2012) Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs. *Animal Genetics* **43**, 714-20.

Corominas J., Ramayo-Caldas Y., Puig-Oliveras A., Estellé J., Castelló A., Alves E., Pena R.N., Ballester M. & Folch J.M. (2013a) Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics. In revision.*

Corominas J., Ramayo-Caldas Y., Puig-Oliveras A., Pérez-Montarelo D., Noguera J.L., Folch J.M. & Ballester M. (2013b) Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE* **8**, e53687.

Covault J., Pettinati H., Moak D., Mueller T. & Kranzler H.R. (2004) Association of a long-chain fatty acid-CoA ligase 4 gene polymorphism with depression and with enhanced niacin-induced dermal erythema. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **127B**, 42-7.

Chen C., Ai H., Ren J., Li W., Li P., Qiao R., Ouyang J., Yang M., Ma J. & Huang L. (2011) A global view of porcine transcriptome in three tissues from a full-sib pair with extreme phenotypes in growth and fat deposition by paired-end RNA sequencing. *BMC Genomics* **12**, 448.

Chen C., Deng B., Qiao M., Zheng R., Chai J., Ding Y., Peng J. & Jiang S. (2012a) Solexa Sequencing Identification of Conserved and Novel microRNAs in Backfat of Large White and Chinese Meishan Pigs. *PLoS ONE* **7**, e31426.

Chen C., Qiao R., Wei R., Guo Y., Ai H., Ma J., Ren J. & Huang L. (2012b) A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC Genomics* **13**, 733.

Chen D., Riedl T., Washbrook E., Pace P.E., Coombes R.C., Egly J.-M. & Ali S. (2000) Activation of Estrogen Receptor alpha by S118 Phosphorylation Involves a Ligand-Dependent Interaction with TFIIH and Participation of CDK7. *Molecular Cell* **6**, 127-37.

Chizzolini R., Zanardi E., Dorigoni V. & Ghidini S. (1999) Calorific value and cholesterol content of normal and low-fat meat and meat products. *Trends in Food Science & Technology* **10**, 119-28.

Dall'Olio S., Davoli R., Buttazzoni L., Zambonelli P. & Russo V. (2009) Study of porcine adiponectin (ADIPOQ) gene and association of a missense mutation with EBVs for production and carcass traits in Italian Duroc heavy pigs. *Livestock Science* **125**, 101-4.

Dávalos A., Goedeke L., Smibert P., Ramírez C.M., Warrier N.P., Andreo U., Cirera-Salinas D., Rayner K., Suresh U., Pastor-Pareja J.C., Esplugues E., Fisher E.A., Penalva L.O.F., Moore K.J., Suárez Y., Lai E.C. & Fernández-Hernando C. (2011) miR-33a/b contribute to the regulation of fatty acid metabolism and insulin signaling. *Proceedings of the National Academy of Sciences* **108**, 9232-7.

Davies W., Høheim B., Chaput B., Archibald A.L. & Frelat G. (1994) Characterization of microsatellites from flow-sorted porcine Chromosome 13. *Mammalian Genome* **5**, 707-11.

Davoli R. & Braglia S. (2008) Molecular approaches in pig breeding to improve meat quality. *Briefings in Functional Genomics & Proteomics* **6**, 313-21.

De Cesare D. & Sassone-Corsi P. (2000) Transcriptional regulation by cyclic AMP-responsive factors. In: *Progress in Nucleic Acid Research and Molecular Biology* (pp. 343-69). Academic Press.

de Lorgeril M. & Salen P. (2012) New insights into the health effects of dietary saturated and omega-6 and omega-3 polyunsaturated fatty acids. *BMC Medicine* **10**, 50.

Denic V. & Weissman J.S. (2007) A Molecular Caliper Mechanism for Determining Very Long-Chain Fatty Acid Length. *Cell* **130**, 663-77.

Druet T. & Georges M. (2009) A Hidden Markov Model Combining Linkage and Linkage Disequilibrium Information for Haplotype Reconstruction and Quantitative Trait Locus Fine Mapping. *Genetics* **184**, 789-98.

Duijvesteijn N., Knol E., Merks J., Crooijmans R., Groenen M., Bovenhuis H. & Harlizius B. (2010) A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genetics* **11**, 42.

Duplus E. & Forest C. (2002) Is there a single mechanism for fatty acid regulation of gene transcription? *Biochemical Pharmacology* **64**, 893-901.

Duthie C.-A., Simm G., Perez-Enciso M., Doeschl-Wilson A., Kalm E., Knap P. & Roehe R. (2009) Genomic scan for quantitative trait loci of chemical and physical body composition and deposition on pig chromosome X including the pseudoautosomal region of males. *Genetics Selection Evolution* **41**, 27.

Eaton S., Bartlett K. & Pourfarzam M. (1996) Mammalian mitochondrial beta-oxidation. *Biochemical Journal* **320**, 345-57.

Echard G., Milan D., Yerle M., Lahbib-Mansais Y. & Gellin J. (1992) The gene map of the pig (Sus scrofa domestica L.): a review. *Cytogenetic and Genome Research* **61**, 146-51.

Ernst C.W. & Steibel J.P. (2013) Molecular advances in QTL discovery and application in pig breeding. *Trends in Genetics* **29**, 215-24.

Estellé J., Fernández A.I., Pérez-Enciso M., Fernández A., Rodríguez C., Sánchez A., Noguera J.L. & Folch J.M. (2009a) A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. *Animal Genetics* **40**, 813-20.

Estellé J., Mercadé A., Noguera J.L., Pérez-Enciso M., Óvilo C., Sánchez A. & Folch J.M. (2005a) Effect of the porcine IGF2-intron3-G3072A substitution in an outbred Large White population and in an Iberian x Landrace cross. *Journal of Animal Science* **83**, 2723-8.

Estellé J., Mercadé A., Pérez-Enciso M., Pena R.N., Silió L., Sánchez A. & Folch J.M. (2009b) Evaluation of FABP2 as candidate gene for a fatty acid composition QTL in porcine chromosome 8. *Journal of Animal Breeding and Genetics* **126**, 52-8.

Estellé J., Pérez-Enciso M., Mercadé A., Varona L., Alves E., Sánchez A. & Folch J.M. (2006) Characterization of the porcine FABP5 gene and its association with the FAT1 QTL in an Iberian by Landrace cross. *Animal Genetics* **37**, 589-91.

Estellé J., Sánchez A. & Folch J.M. (2005b) Assignment of the microsomal triglyceride transfer protein large subunit (MTP) gene to porcine chromosome 8. *Animal Genetics* **36**, 354-5.

Esteve-Codina A., Kofler R., Palmieri N., Bussotti G., Notredame C. & Perez-Enciso M. (2011) Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC Genomics* **12**, 552.

Felig P. & Wahren J. (1975) Fuel Homeostasis in Exercise. *New England Journal of Medicine* **293**, 1078-84.

Fernández A., Pérez-Montarelo D., Barragán C., Ramayo-Caldas Y., Ibánez-Escriche N., Castelló A., Noguera J., Silió L., Folch J. & Rodríguez M. (2012) Genome-wide linkage analysis of QTL for growth and body composition employing the PorcineSNP60 BeadChip. *BMC Genetics* **13**, 41.

Ferrari A., Fiorino E., Giudici M., Gilardi F., Galmozzi A., Mitro N., Cermenati G., Godio C., Caruso D., De Fabiani E. & Crestani M. (2012) Linking epigenetics to lipid metabolism: focus on histone deacetylases. *Molecular Membrane Biology* **29**, 257-66.

Fisher R.A. (1919) The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* **52**, 399-433.

Foulkes A. (2009) Applied Statistical Genetics with R. For Population-based Association Studies. *New York*.

Freeman T., Ivens A., Baillie J.K., Beraldi D., Barnett M., Dorward D., Downing A., Fairbairn L., Kapetanovic R., Raza S., Tomoiu A., Alberio R., Wu C., Su A., Summers K., Tuggle C., Archibald A. & Hume D. (2012) A gene expression atlas of the domestic pig. *BMC Biology* **10**, 90.

Fritah A., Redeuilh G.r. & Sabbah M.l. (2006) Molecular cloning and characterization of the human WISP-2/CCN5 gene promoter reveal its upregulation by oestrogens. *Journal of Endocrinology* **191**, 613-24.

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.

Galas D.J. & Schmitz A. (1978) DNAase footprinting a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Research* **5**, 3157-70.

Gaspa D.B., Novell C.S., Fabregat A.C., Rodríguez J.A., Hernández E.G., Romero K.L., Solans L.N., Tinoco D.C., Abill G.B., Fosas F.R. & Ribalta A.C. (2011) Informe del sector porcino. *In: A. i. M. N. Departament d'Agricultura Ramaderia Pesca. Mayo.*

Gazou A., Riess A., Grasshoff U., Schäferhoff K., Bonin M., Jauch A., Riess O. & Tzschach A. (2012) Xq22.3–q23 deletion including ACSL4 in a patient with intellectual disability. *American Journal of Medical Genetics Part A* **161**, 860-4.

Georges M. (2007) Mapping, Fine Mapping, and Molecular Dissection of Quantitative Trait Loci in Domestic Animals. *Annual Review of Genomics and Human Genetics* **8**, 131-62.

Goddard M.E. & Hayes B.J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics* **10**, 381-91.

Gregersen V., Conley L., Sorensen K., Guldbrandtsen B., Velander I. & Bendixen C. (2012) Genome-wide association scan and phased haplotype construction for quantitative trait loci affecting boar taint in three pig breeds. *BMC Genomics* **13**, 22.

Groenen M.A.M., Archibald A.L., Uenishi H., Tuggle C.K., Takeuchi Y., Rothschild M.F., Rogel-Gaillard C., Park C., Milan D., Megens H.-J., Li S., Larkin D.M., Kim H., Frantz L.A.F., Caccamo M., Ahn H., Aken B.L., Anselmo A., Anthon C., Auvil L., Badaoui B., Beattie C.W., Bendixen C., Berman D., Blecha F., Blomberg J., Bolund L., Bosse M., Botti S., Bujie Z., Bystrom M., Capitanu B., Carvalho-Silva D., Chardon P., Chen C., Cheng R., Choi S.-H., Chow W., Clark R.C., Clee C., Crooijmans R.P.M.A., Dawson H.D., Dehais P., De Sapio F., Dibbits B., Drou N., Du Z.-Q., Eversole K., Fadista J., Fairley S., Faraut T., Faulkner G.J., Fowler K.E., Fredholm M., Fritz E., Gilbert J.G.R., Giuffra E., Gorodkin J., Griffin D.K., Harrow J.L., Hayward A., Howe K., Hu Z.-L., Humphray S.J., Hunt T., Hornshoj H., Jeon J.-T., Jern P., Jones M., Jurka J., Kanamori H., Kapetanovic R., Kim J., Kim J.-H., Kim K.-W., Kim T.-H., Larson G., Lee K., Lee K.-T., Leggett R., Lewin H.A., Li Y., Liu W., Loveland J.E., Lu Y., Lunney J.K., Ma J., Madsen O., Mann K., Matthews L., McLaren S., Morozumi T., Murtaugh M.P., Narayan J., Truong Nguyen D., Ni P., Oh S.-J., Onteru S., Panitz F., Park E.-W., Park H.-S., Pascal G., Paudel Y., Perez-Enciso M., Ramirez-Gonzalez R., Reecy J.M., Rodriguez-Zas S., Rohrer G.A., Rund L., Sang Y., Schachtschneider K., Schraiber J.G., Schwartz J., Scobie L., Scott C., Searle S., Servin B., Southey B.R., Sperber G., Stadler P., Sweedler J.V., Tafer H., Thomsen B., Wali R., Wang J., Wang J., White S., Xu X., Yerle M., Zhang G., Zhang J., Zhang J., Zhao S., Rogers J., Churcher C. & Schook L.B. (2012) Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393-8.

Groenen M.A.M., Ruyter D., Verstege E.J.M., Vries M.d. & Poel J.J.v.d. (1995) Development and mapping of ten porcine microsatellite markers. *Animal Genetics* **26**, 115-8.

Guillaumond F., Gréchez-Cassiau A., Subramaniam M., Brangolo S., Peteri-Brünback B., Staels B., Fiévet C., Spelsberg T.C., Delaunay F. & Teboul M. (2010) Kruppel-Like Factor KLF10 Is a Link between the Circadian Clock and Metabolism in Liver. *Molecular and Cellular Biology* **30**, 3059-70.

Guillevic M., Kouba M. & Mourot J. (2009) Effect of a linseed diet or a sunflower diet on performances, fatty acid composition, lipogenic enzyme activities and stearoyl-CoA-desaturase activity in the pig. *Livestock Science* **124**, 288-94.

Guillou H., Zadravec D., Martin P.G.P. & Jacobsson A. (2010) The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. *Progress in Lipid Research* **49**, 186-99.

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 × Erhualian intercross F2 population. *Animal Genetics* **40**, 185-91.

Haley C.S., Archibald A.L., Andersson L., Bosma A.A., Davies W., Fredholm M., Geldermann H., Groenen M., Gustavsson I., Ollivier L., Tucker E.M., World A.V.d.W.T.p.g.m.p.p.I.t. & Congress on Genetics Applied to Livestock Production E.p. (1990) The pig gene mapping project - pigmap. *In: 4th World Congress on Genetics Applied to Livestock Production, Edinburgh*, p 67-70.

Harlizius B., Rattink A.P., de Koning D.J., Faivre M., Joosten R.G., van Arendonk J.A.M. & Groenen M.A.M. (2000) The X Chromosome harbors quantitative trait loci for backfat thickness and intramuscular fat content in pigs. *Mammalian Genome* **11**, 800-2.

Herzig S., Hedrick S., Morantte I., Koo S.-H., Galimi F. & Montminy M. (2003) CREB controls hepatic lipid metabolism through nuclear hormone receptor PPAR-[gamma]. *Nature* **426**, 190-3.

Hoekstra M., van der Sluis R.J., Kuiper J. & Van Berkel T.J.C. (2012) Nonalcoholic fatty liver disease is associated with an altered hepatocyte microRNA profile in LDL receptor knockout mice. *The Journal of Nutritional Biochemistry* **23**, 622-8.

Hoffman K. (1994) What is quality. *Meat Focus*, 73-82.

Horton J.D., Shah N.A., Warrington J.A., Anderson N.N., Park S.W., Brown M.S. & Goldstein J.L. (2003) Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proceedings of the National Academy of Sciences* **100**, 12027-32.

Hu Z.-L., Park C.A., Wu X.-L. & Reecy J.M. (2013) Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* **41**, D871-D9.

Huff-Lonergan E., Baas T.J., Malek M., Dekkers J.C.M., Prusa K. & Rothschild M.F. (2002) Correlations among selected pork quality traits. *Journal of Animal Science* **80**, 617-27.

Jakobsson A., Westerberg R. & Jacobsson A. (2006) Fatty acid elongases in mammals: Their regulation and roles in metabolism. *Progress in Lipid Research* **45**, 237-49.

Jayakumar A., Tai M.H., Huang W.Y., al-Feel W., Hsu M., Abu-Elheiga L., Chirala S.S. & Wakil S.J. (1995) Human fatty acid synthase: properties and molecular cloning. *Proceedings of the National Academy of Sciences* **92**, 8695-9.

Jeon J.-T., Carlborg O., Tornsten A., Giuffra E., Amarger V., Chardon P., Andersson-Eklund L., Andersson K., Hansson I., Lundstrom K. & Andersson L. (1999) A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. *Nature Genetics* **21**, 157-8.

Jeukens J., Boyle B., Kukavica-Ibrulj I., St-Cyr J., Lévesque R.C. & Bernatchez L. (2011) BAC library construction, screening and clone sequencing of lake whitefish (Coregonus clupeaformis, Salmonidae) towards the elucidation of adaptive species divergence. *Molecular Ecology Resources* **11**, 541-9.

Jump D.B. (2002) Dietary polyunsaturated fatty acids and regulation of gene transcription. *Current Opinion in Lipidology* **13**.

Jump D.B. & Clarke S.D. (1999) Regulation of gene expression by dietary fat. *Annual Review of Nutrition* **19**, 63-90.

Jung W.Y., Kwon S.G., Son M., Cho E.S., Lee Y., Kim J.H., Kim B.-W., Park D.H., Hwang J.H., Kim T.W., Park H.C., Park B.Y., Choi J.-S., Cho K.K., Chung K.H., Song Y.M., Kim I.S., Jin S.K., Kim D.H., Lee S.-W., Lee K.W., Bang W.Y. & Kim C.W. (2012) RNA-Seq Approach for Genetic Improvement of Meat Quality in Pig and Evolutionary Insight into the Substrate Specificity of Animal Carbonyl Reductases. *PLoS ONE* **7**, e42198.

Kantojärvi K., Kotala I., Rehnström K., Ylisaukko-oja T., Vanhala R., von Wendt T.N., von Wendt L. & Järvelä I. (2011) Fine mapping of Xq11.1-q21.33 and

mutation screening of RPS6KA6, ZNF711, ACSL4, DLG3, and IL1RAPL2 for autism spectrum disorders (ASD). *Autism Research* **4**, 228-33.

Katan M., Zock P. & Mensink R. (1994) Effects of fats and fatty acids on blood lipids in humans: an overview. *The American Journal of Clinical Nutrition* **60**, 1017S-22S.

Kim J.H., Lim H.T., Park E.W., Rodríguez C., Silio L., Varona L., Mercade A., Jeon J.T. & Ovilo C. (2006) Polymorphisms in the promoter region of the porcine acyl-coA dehydrogenase, medium-chain (ACADM) gene have no effect on fat deposition traits in a pig Iberian × Landrace cross. *Animal Genetics* **37**, 430-1.

Kim K.S., Larsen N.J. & Rothschild M.F. (2000) Rapid communication: linkage and physical mapping of the porcine melanocortin-4 receptor (MC4R) gene. *Journal of Animal Science* **78**, 791-2.

Kitazawa H., Miyamoto Y., Shimamura K., Nagumo A. & Tokita S. (2009) Development of a High-Density Assay for Long-Chain Fatty Acyl-CoA Elongases. *Lipids* **44**, 765-73.

Klein R.J., Zeiss C., Chew E.Y., Tsai J.-Y., Sackler R.S., Haynes C., Henning A.K., SanGiovanni J.P., Mane S.M., Mayne S.T., Bracken M.B., Ferris F.L., Ott J., Barnstable C. & Hoh J. (2005) Complement Factor H Polymorphism in Age-Related Macular Degeneration. *Science* **308**, 385-9.

Klett E.L., Chen S., Edin M.L., Li L.O., Ilkayeva O., Zeldin D.C., Newgard C.B. & Coleman R.A. (2013) Diminished acyl-CoA synthetase isoform 4 activity in INS 832/13 cells reduces cellular epoxyeicosatrienoic acid levels and results in impaired glucose-stimulated insulin secretion. *Journal of Biological Chemistry* **288(30)**, 21618-29.

Knight J.C. (2003) Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clinical Science* **104**, 493-501.

Koboldt D.C., Ding L., Mardis E.R. & Wilson R.K. (2010) Challenges of sequencing human genomes. *Briefings in Bioinformatics* **11**, 484-98.

Kornberg A. & Pricer W.E. (1953) Enzymatic synthesis of the coenzyme A derivatives of long chain fatty acids. *Journal of Biological Chemistry* **204**, 329-43.

Kraegen E.W., Clark P.W., Jenkins A.B., Daley E.A., Chisholm D.J. & Storlien L.H. (1991) Development of Muscle Insulin Resistance After Liver Insulin Resistance in High-Fat-Fed Rats. *Diabetes* **40**, 1397-403.

Kumadaki S., Matsuzaka T., Kato T., Yahagi N., Yamamoto T., Okada S., Kobayashi K., Takahashi A., Yatoh S., Suzuki H., Yamada N. & Shimano H. (2008) Mouse Elovl-6 promoter is an SREBP target. *Biochemical and Biophysical Research Communications* **368**, 261-6.

Lee J.-S., Kim J.-M., Hong J.-S., Lim K.-S., Hong K.-C. & Lee Y. (2012) Effects of polymorphisms in the 3' untranslated region of the porcine PPARGC1A gene on muscle fiber characteristics and meat quality traits. *Molecular Biology Reports* **39**, 3943-50.

Leonard A.E., Bobik E.G., Dorado J., Kroeger P.E., Chuang L.T., Thurmond J.M., Parker-Barnes J.M., Das T., Huang Y.S. & Mukerji P. (2000) Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. *Biochemical Journal* **350**, 765-70.

Leonard A.E., Pereira S.L., Sprecher H. & Huang Y.-S. (2004) Elongation of long-chain fatty acids. *Progress in Lipid Research* **43**, 36-54.

Lewin T.M., Van Horn C.G., Krisans S.K. & Coleman R.A. (2002) Rat liver acyl-CoA synthetase 4 is a peripheral-membrane protein located in two distinct subcellular organelles, peroxisomes, and mitochondrial-associated membrane. *Archives of Biochemistry and Biophysics* **404**, 263-70.

Li G., Li Y., Li X., Ning X., Li M. & Yang G. (2011) MicroRNA identity and abundance in developing swine adipose tissue as determined by solexa sequencing. *Journal of Cellular Biochemistry* **112**, 1318-28.

Li H.-Y., Xi Q.-Y., Xiong Y.-Y., Liu X.-L., Cheng X., Shu G., Wang S.-B., Wang L.-N., Gao P., Zhu X.-T., Jiang Q.-Y., Yuan L. & Zhang Y.-L. (2012a) Identification and comparison of microRNAs from skeletal muscle and adipose tissues from two porcine breeds. *Animal Genetics* **43**, 704-13.

Li X.J., Yang H., Li G.X., Zhang G.H., Cheng J., Guan H. & Yang G.S. (2012b) Transcriptome profile analysis of porcine adipose tissue by high-throughput sequencing. *Animal Genetics* **43**, 144-52.

Liang Y., Wu C., Chu J., Wang C., Hung L., Wang Y., Ho Y., Chang J. & Lin S. (2005) Involvement of fatty acid-CoA ligase 4 in hepatocellular carcinoma growth:

Roles of cyclic AMP and p38 mitogen-activated protein kinase. *World Journal of Gastroenterology* **11**, 2557-63.

Lopez-Bote C.J., Rey A.I., Sanz M., Gray J.I. & Buckley D.J. (1997) Dietary Vegetable Oils and alpha-Tocopherol Reduce Lipid Oxidation in Rabbit Muscle. *The Journal of Nutrition* **127**, 1176-82.

Lunney J. (2007) Advances in Swine Biomedical Model Genomics. *International Journal of Biological Sciences* **3**, 179-84.

Ma J., Gilbert H., Iannuccelli N., Duan Y., Guo B., Huang W., Ma H., Riquet J., Bidanel J.-P., Huang L. & Milan D. (2013) Fine mapping of fatness QTL on porcine chromosome X and analyses of three positional candidate genes. *BMC Genetics* **14**, 46.

Ma J., Ren J., Guo Y., Duan Y., Ding N., Zhou L., Li L., Yan X., Yang K., Huang L., Song Y., Xie J. & Milan D. (2009) Genome-wide identification of quantitative trait loci for carcass composition and meat quality in a large-scale White Duroc × Chinese Erhualian resource population. *Animal Genetics* **40**, 637-47.

Mackay T.F.C., Stone E.A. & Ayroles J.F. (2009) The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* **10**, 565-77.

Maher B. (2012) Encode: The human encyclopaedia. *Nature* **489**, 46-8.

Mashek D.G., Bornfeldt K.E., Coleman R.A., Berger J., Bernlohr D.A., Black P., DiRusso C.C., Farber S.A., Guo W., Hashimoto N., Khodiyar V., Kuypers F.A., Maltais L.J., Nebert D.W., Renieri A., Schaffer J.E., Stahl A., Watkins P.A., Vasiliou V. & Yamamoto T.T. (2004) Revised nomenclature for the mammalian long-chain acyl-CoA synthetase gene family. *Journal of Lipid Research* **45**, 1958-61.

Mashek D.G., Li L.O. & Coleman R.A. (2006) Rat long-chain acyl-CoA synthetase mRNA, protein, and activity vary in tissue distribution and in response to diet. *Journal of Lipid Research* **47**, 2004-10.

Maston G.A., Evans S.K. & Green M.R. (2006) Transcriptional Regulatory Elements in the Human Genome. *Annual Review of Genomics and Human Genetics* **7**, 29-59.

Matsuzaka T., Shimano H., Yahagi N., Kato T., Atsumi A., Yamamoto T., Inoue N., Ishikawa M., Okada S., Ishigaki N., Iwasaki H., Iwasaki Y., Karasawa T.,

Kumadaki S., Matsui T., Sekiya M., Ohashi K., Hasty A.H., Nakagawa Y., Takahashi A., Suzuki H., Yatoh S., Sone H., Toyoshima H., Osuga J.-i. & Yamada N. (2007) Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. *Nature Medicine* **13**, 1193-202.

Matsuzaka T., Shimano H., Yahagi N., Yoshikawa T., Amemiya-Kudo M., Hasty A.H., Okazaki H., Tamura Y., Iizuka Y., Ohashi K., Osuga J.-i., Takahashi A., Yato S., Sone H., Ishibashi S. & Yamada N. (2002) Cloning and characterization of a mammalian fatty acyl-CoA elongase as a lipogenic enzyme regulated by SREBPs. *Journal of Lipid Research* **43**, 911-20.

McGarry J.D., Mannaerts G.P. & Foster D.W. (1977) A possible role for malonyl-CoA in the regulation of hepatic fatty acid oxidation and ketogenesis. *The Journal of Clinical Investigation* **60**, 265-70.

Meloni I., Muscettola M., Raynaud M., Longo I., Bruttini M., Moizard M.-P., Gomot M., Chelly J., des Portes V., Fryns J.-P., Ropers H.-H., Magi B., Bellan C., Volpi N., Yntema H.G., Lewis S.E., Schaffer J.E. & Renieri A. (2002) FACL4, encoding fatty acid-CoA ligase 4, is mutated in nonspecific X-linked mental retardation. *Nature Genetics* **30**, 436-40.

Mercadé A., Estellé J., Noguera J., Folch J., Varona L., Silió L., Sánchez A. & Pérez-Enciso M. (2005a) On growth, fatness, and form: A further look at porcine Chromosome 4 in an Iberian x Landrace cross. *Mammalian Genome* **16**, 374-82.

Mercadé A., Estellé J., Pérez-Enciso M., Varona L., Silió L., Noguera J.L., Sánchez A. & Folch J.M. (2006a) Characterization of the porcine acyl-CoA synthetase long-chain 4 gene and its association with growth and meat quality traits. *Animal Genetics* **37**, 219-24.

Mercadé A., Pérez-Enciso M., Varona L., Alves E., Noguera J.L., Sánchez A. & Folch J.M. (2006b) Adipocyte fatty-acid binding protein is closely associated to the porcine FAT1 locus on chromosome 4. *Journal of Animal Science* **84**, 2907-13.

Mercadé A., Sánchez A. & Folch J.M. (2005b) Exclusion of the acyl CoA:diacylglycerol acyltransferase 1 gene (DGAT1) as a candidate for a fatty acid composition QTL on porcine chromosome 4. *Journal of Animal Breeding and Genetics* **122**, 161-4.

Mercadé A., Sánchez A. & Folch J.M. (2007) Characterization and Physical Mapping of the Porcine CDS1 and CDS2 Genes. *Animal Biotechnology* **18**, 23-35.

Metivier R., Gallais R., Tiffoche C., Le Peron C., Jurkowska R.Z., Carmouche R.P., Ibberson D., Barath P., Demay F., Reid G., Benes V., Jeltsch A., Gannon F. & Salbert G. (2008) Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452**, 45-50.

Meuwissen T. & Goddard M. (2004) Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genetics Selection Evolution* **36**, 261-79.

Meuwissen T.H.E. & Goddard M.E. (2000) Fine Mapping of Quantitative Trait Loci Using Linkage Disequilibria With Closely Linked Marker Loci. *Genetics* **155**, 421-30.

Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundström 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.

Minekura H., Kang M.-J., Inagaki Y., Cho Y.-Y., Suzuki H., Fujino T. & Yamamoto T.T. (2001) Exon/Intron Organization and Transcription Units of the Human Acyl-CoA Synthetase 4 Gene. *Biochemical and Biophysical Research Communications* **286**, 80-6.

Modi H.R., Basselin M., Taha A.Y., Li L.O., Coleman R.A., Bialer M. & Rapoport S.I. (2013) Propylisopropylacetic acid (PIA), a constitutional isomer of valproic acid, uncompetitively inhibits arachidonic acid acylation by rat acyl-CoA synthetase 4: A potential drug for bipolar disorder. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1831**, 880-6.

Moon Y.-A., Shah N.A., Mohapatra S., Warrington J.A. & Horton J.D. (2001) Identification of a Mammalian Long Chain Fatty Acyl Elongase Regulated by Sterol Regulatory Element-binding Proteins. *Journal of Biological Chemistry* **276**, 45358-66.

Morcillo S., Martín-Núñez G.M., Rojo-Martínez G., Almaraz M.C., García-Escobar E., Mansego M.L., de Marco G., Chaves F.J. & Soriguer F. (2011) ELOVL6 Genetic

Variation Is Related to Insulin Sensitivity: A New Candidate Gene in Energy Metabolism. *PLoS ONE* **6**, e21198.

Muñoz G., Alves E., Fernández A., Óvilo C., Barragán C., Estellé J., Quintanilla R., Folch J.M., Silió L., Rodríguez M.C. & Fernández 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. *Animal Genetics* **38**, 639-46.

Muñoz M., Rodríguez M.C., Alves E., Folch J.M., Ibañez-Escriche N., Silió L. & Fernández A. (2013) Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. *Submitted*.

Neely J.R. & Morgan H.E. (1974) Relationship Between Carbohydrate and Lipid Metabolism and the Energy Balance of Heart Muscle. *Annual Review of Physiology* **36**, 413-59.

Neve B., Froguel P., Corset L., Vaillant E., Vatin V. & Boutin P. (2002) Rapid SNP allele frequency determination in genomic DNA pools by pyrosequencing™. *Biotechniques* **32**, 1138-42.

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 × Large White intercross. *Animal Genetics* **37**, 342-7.

Nobrega M.A., Ovcharenko I., Afzal V. & Rubin E.M. (2003) Scanning Human Gene Deserts for Long-Range Enhancers. *Science* **302**, 413.

Ntawubizi M., Colman E., Janssens S., Raes K., Buys N. & De Smet S. (2010) Genetic parameters for intramuscular fatty acid composition and metabolism in pigs. *Journal of Animal Science* **88**, 1286-94.

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. *The Journal of Nutrition* **99**, 338-44.

Odriozola M. (1976) Investigación sobre los datos acumulados en dos piaras experimentales. *Ministerio de Agricultura, Madrid.*

Orlando U., Cooke M., Cornejo Maciel F., Papadopoulos V., Podestá E.J. & Maloberti P. (2013) Characterization of the mouse promoter region of the acyl-CoA

synthetase 4 gene: Role of Sp1 and CREB. *Molecular and Cellular Endocrinology* **369**, 15-26.

Óvilo C., Clop A., Noguera J.L., Oliver M.A., Barragán C., Rodríguez C., Silió L., Toro M.A., Coll A., Folch J.M., Sánchez A., Babot D., Varona L. & Pérez-Enciso M. (2002) Quantitative trait locus mapping for meat quality traits in an Iberian x Landrace F2 pig population. *Journal of Animal Science* **80**, 2801-8.

Óvilo C., Fernández A., Noguera J.L., Barragán C., Letón R., Rodríguez C., Mercadé A., Alves E., Folch J.M., Varona L. & Toro M. (2005) Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross. *Genetics Research* **85**, 57-67.

Ovilo C., Oliver A., Noguera J., Clop A., Barragan C., Varona L., Rodriguez C., Toro M., Sanchez A., Perez-Enciso M. & Silio L. (2002) Test for positional candidate genes for body composition on pig chromosome 6. *Genetics Selection Evolution* **34**, 465-79.

Óvilo C., Pérez-Enciso M., Barragán C., Clop A., Rodríguez C., Oliver M.A., Toro M.A. & Noguera J.L. (2000) A QTL for intramuscular fat and backfat thickness is located on porcine Chromosome 6. *Mammalian Genome* **11**, 344-6.

Peng Y., Xiang H., Chen C., Zheng R., Chai J., Peng J. & Jiang S. (2013) MiR-224 impairs adipocyte early differentiation and regulates fatty acid metabolism. *The International Journal of Biochemistry & Cell Biology* **45**, 1585-93.

Pérez-Enciso M., Clop A., Folch J.M., Sánchez A., Oliver M.A., Óvilo C., Barragán C., Varona L. & Noguera J.L. (2002) Exploring Alternative Models for Sex-Linked Quantitative Trait Loci in Outbred Populations: Application to an Iberian x Landrace Pig Intercross. *Genetics* **161**, 1625-32.

Pérez-Enciso M., Clop A., Noguera J.L., Ovilo C., Coll A., Folch J.M., Babot D., Estany J., Oliver M.A., Díaz I. & Sánchez A. (2000) A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *Journal of Animal Science* **78**, 2525-31.

Pérez-Enciso M., Ferraz A., Ojeda A. & López-Béjar M. (2009) Impact of breed and sex on porcine endocrine transcriptome: a bayesian biometrical analysis. *BMC Genomics* **10**, 89.

Pérez-Enciso M., Mercadé A., Bidanel J.P., Geldermann H., Cepica S., Bartenschlager H., Varona L., Milan D. & Folch J.M. (2005) Large-scale, multibreed, multitrait analyses of quantitative trait loci experiments: The case of porcine X chromosome. *Journal of Animal Science* **83**, 2289-96.

Pérez-Montarelo D., Hudson N.J., Fernández A.I., Ramayo-Caldas Y., Dalrymple B.P. & Reverter A. (2012) Porcine Tissue-Specific Regulatory Networks Derived from Meta-Analysis of the Transcriptome. *PLoS ONE* **7**, e46159.

Piccini M., Vitelli F., Bruttini M., Pober B.R., Jonsson J.J., Villanova M., Zollo M., Borsani G., Ballabio A. & Renieri A. (1998) FACL4, a New Gene Encoding Long-Chain Acyl-CoA Synthetase 4, Is Deleted in a Family with Alport Syndrome, Elliptocytosis, and Mental Retardation. *Genomics* **47**, 350-8.

Ponsuksili S., Murani E., Phatsara C., Jonas E., Walz C., Schwerin M., Schellander K. & Wimmers K. (2008) Expression Profiling of Muscle Reveals Transcripts Differentially Expressed in Muscle That Affect Water-Holding Capacity of Pork. *Journal of Agricultural and Food Chemistry* **56**, 10311-7.

Postic C. & Girard J. (2008) Contribution of de novo fatty acid synthesis to hepatic steatosis and insulin resistance: lessons from genetically engineered mice. *The Journal of Clinical Investigation* **118**, 829-38.

Poudyal H., Panchal S.K., Diwan V. & Brown L. (2011) Omega-3 fatty acids and metabolic syndrome: Effects and emerging mechanisms of action. *Progress in Lipid Research* **50**, 372-87.

Puig-Oliveras A., Ramayo-Caldas Y., Corominas J., Estellé J., Fernández A., Casellas J., Ballester M. & Folch J.M. Differences in swine muscle transcriptome among animals phenotipically extremes for fatty acid composition. Manuscript in preparation.

Ramayo-Caldas Y. (2013) *Dissecting the Genetic Basis of Pig Intramuscular Fatty Acid Composition.* PhD thesis. Universitat Autònoma de Barcelona.

Ramayo-Caldas Y., Ballester M., Fortes M., Esteve-Codina A., Castelló A., Noguera J.L., Fernandez A.I., Pérez-Enciso M., Reverter A. & Folch J.M. (2013) From SNP co-association to gene co-expression: Evidence of pleiotropic expression-QTL in lipid metabolism in pigs. *BMC Genomics. Submitted*.

Ramayo-Caldas Y., Castello A., Pena R., Alves E., Mercade A., Souza C., Fernandez A., Perez-Enciso M. & Folch J. (2010) Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* **11**, 593.

Ramayo-Caldas Y., Mach N., Esteve-Codina A., Corominas J., Castello A., Ballester M., Estelle J., Ibanez-Escriche N., Fernandez A., Perez-Enciso M. & Folch J. (2012a) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics* **13**, 547.

Ramayo-Caldas Y., Mercadé A., Castelló A., Yang B., Rodríguez C., Alves E., Díaz I., Ibánez-Escriche N., Noguera J.L., Pérez-Enciso M., Fernández A.I. & Folch J.M. (2012b) Genome-Wide Association Study for Intramuscular Fatty Acid Composition in an Iberian x Landrace Cross. *Journal of Animal Science* **90**, 1-11.

Ramos A.M., Crooijmans R.P.M.A., Affara N.A., Amaral A.J., Archibald A.L., Beever J.E., Bendixen C., Churcher C., Clark R., Dehais P., Hansen M.S., Hedegaard J., Hu Z.-L., Kerstens H.H., Law A.S., Megens H.-J., Milan D., Nonneman D.J., Rohrer G.A., Rothschild M.F., Smith T.P.L., Schnabel R.D., Van Tassell C.P., Taylor J.F., Wiedmann R.T., Schook L.B. & Groenen M.A.M. (2009) Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. *PLoS ONE* **4**, e6524.

Revilla M., Ramayo-Caldas Y., Castelló A., Corominas J., Puig-Oliveras A., Ibáñez-Escriche N., Muñoz M., Ballester M. & Folch J.M. (2013) Association study for backfat fatty acid composition in an Iberian x Landrace cross with a 144 snp chip of SSC8. *Genetics Selection Evolution. Submitted.*

Rothschild M.F. & Ruvinsky A. (2011) *The Genetics of the Pig*. CABI.

Rudel L.L., Parks J.S. & Sawyer J.K. (1995) Compared With Dietary Monounsaturated and Saturated Fat, Polyunsaturated Fat Protects African Green Monkeys From Coronary Artery Atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **15**, 2101-10.

Ruiz J., Cava R., Antequera T., Martín L., Ventanas J. & López-Bote C.J. (1998) Prediction of the feeding background of Iberian pigs using the fatty acid profile of subcutaneous, muscle and hepatic fat. *Meat Science* **49**, 155-63.

Sampath H. & Ntambi J.M. (2005) Polyunsaturated fatty acid regulation of genes of lipid metabolism. *Annual Review of Nutrition* **25**, 317-40.

Schaart J.G., Mehli L. & Schouten H.J. (2005) Quantification of allele-specific expression of a gene encoding strawberry polygalacturonase-inhibiting protein (PGIP) using Pyrosequencing™. *The Plant Journal* **41**, 493-500.

Schimpf S., Fuhrmann N., Schaich S. & Wissinger B. (2008) Comprehensive cDNA study and quantitative transcript analysis of mutant OPA1 transcripts containing premature termination codons. *Human Mutation* **29**, 106-12.

Schmitz G. & Ecker J. (2008) The opposing effects of n-3 and n-6 fatty acids. *Progress in Lipid Research* **47**, 147-55.

Schook B.L., Beever J.E., Rogers J. & et al. (2005a) Swine Genome Sequencing Consortium (SGSC): A Strategic Roadmap for Sequencing The Pig Genome. *Comparative and Functional Genomics* **6**, 251-5.

Schook L., Beattie C., Beever J., Donovan S., Jamison R., Zuckermann F., Niemi S., Rothschild M., Rutherford M. & Smith D. (2005b) Swine in Biomedical Research: Creating the Building Blocks of Animal Models. *Animal Biotechnology* **16**, 183-90.

Schuch R., Agelopoulos K., Neumann A., Brandt B., Burger H. & Korsching E. (2012) Site-specific chromatin immunoprecipitation: a selective method to individually analyze neighboring transcription factor binding sites in vivo. *BMC Research Notes* **5**, 109.

Sellier P. (1998) Genetic improvement of the pig. *In: C. International (ed.) The genetics of the pig*, 511-40., Wallingford, Oxon.

Serra X., Gil F., Pérez-Enciso M., Oliver M.A., Vázquez J.M., Gispert M., Díaz I., Moreno F., Latorre R. & Noguera J.L. (1998) A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livestock Production Science* **56**, 215-23.

Simopoulos A. (2001) n-3 fatty acids and human health: Defining strategies for public policy. *Lipids* **36**, S83-S9.

Soh J., Iqbal J., Queiroz J., Fernandez-Hernando C. & Hussain M.M. (2013) MicroRNA-30c reduces hyperlipidemia and atherosclerosis in mice by decreasing lipid synthesis and lipoprotein secretion. *Nature Medicine* **19**, 892-900.

Storlien L.H., Kraegen E.W., Chisholm D.J., Ford G.L., Bruce D.G. & Pascoe W.S. (1987) Fish oil prevents insulin resistance induced by high-fat feeding in rats. *Science* **237**, 885-8.

Sun A., Ge J., Siffert W. & Frey U.H. (2004) Quantification of allele-specific G-protein [beta]3 subunit mRNA transcripts in different human cells and tissues by Pyrosequencing. *European Journal of Human Genetics* **13**, 361-9.

Sung Y.K., Park M.K., Hong S.H., Hwang S.Y., Kwack M.H., Kim J.C. & Kim M.K. (2007) Regulation of cell growth by fatty acid-CoA ligase 4 in human hepatocellular carcinoma cells. *Experimental & Molecular Medicine* **39**, 477-82.

Suzuki H., Kawarabayasi Y., Kondo J., Abe T., Nishikawa K., Kimura S., Hashimoto T. & Yamamoto T. (1990) Structure and regulation of rat long-chain acyl-CoA synthetase. *Journal of Biological Chemistry* **265**, 8681-5.

Thomsen K., Laposata M., Njoroge S., Umunakwe O., Katrangi W. & Seegmiller A. (2011) Increased Elongase 6 and Δ9-Desaturase Activity are Associated with n-7 and n-9 Fatty Acid Changes in Cystic Fibrosis. *Lipids* **46**, 669-77.

Timoneda O., Balcells I., Núñez J.I., Egea R., Vera G., Castelló A., Tomás A. & Sánchez A. (2012) miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds. *PLoS ONE* **8**, e55402.

Torres-Rovira L., Astiz S., Caro A., Lopez-Bote C., Ovilo C., Pallares P., Perez-Solana M.L., Sanchez-Sanchez R. & Gonzalez-Bulnes A. (2011) Diet-Induced Swine Model with Obesity/Leptin Resistance for the Study of Metabolic Syndrome and Type 2 Diabetes. *The Scientific World Journal* **2012**, 8.

van der Steen H.A.M., Prall G.F.W. & Plastow G.S. (2005) Application of genomics to the pork industry. *Journal of Animal Science* **83**, E1-E8.

Van Horn C.G., Caviglia J.M., Li L.O., Wang S., Granger D.A. & Coleman R.A. (2005) Characterization of Recombinant Long-Chain Rat Acyl-CoA Synthetase Isoforms 3 and 6: Identification of a Novel Variant of Isoform 6. *Biochemistry* **44**, 1635-42.

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.

Varona L., Gómez-Raya L., Rauw W.M. & Noguera J.L. (2005) A simulation study on the detection of causal mutations from F2 experiments. *Journal of Animal Breeding and Genetics* **122**, 30-6.

Varona L., Ovilo C., Clop A., Noguera J., Pérez-Enciso M., Coll A., Folch J.M., Barragán C., Toro M.A., Babot D. & Sánchez A. (2002) QTL mapping for growth and carcass traits in an Iberian by Landrace pig intercross: additive, dominant and epistatic effects. *Genetics Research* **80**, 145-54.

Verot L., Alloisio N., Morlé L., Bozon M., Touraine R., Plauchu H. & Edery P. (2003) Localization of a non-syndromic X-linked mental retardation gene (MRX80) to Xq22-q24. *American Journal of Medical Genetics Part A* **122A**, 37-41.

Vincent-Naulleau S., Le Chalony C., Leplat J.-J., Bouet S., Bailly C., Spatz A., Vielh P., Avril M.-F., Tricaud Y., Gruand J., Horak V., Frelat G. & Geffrotin C. (2004) Clinical and Histopathological Characterization of Cutaneous Melanomas in the Melanoblastoma-Bearing Libechov Minipig Model. *Pigment Cell Research* **17**, 24-35.

Wakil S.J., Stoops J.K. & Joshi V.C. (1983) Fatty Acid Synthesis and its Regulation. *Annual Review of Biochemistry* **52**, 537-79.

Wang J., Jiang J., Fu W., Jiang L., Ding X., Liu J.-F. & Zhang Q. (2012a) A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* **13**, 273.

Wang J., Pitarque M. & Ingelman-Sundberg M. (2006) 3'-UTR polymorphism in the human CYP2A6 gene affects mRNA stability and enzyme expression. *Biochemical and Biophysical Research Communications* **340**, 491-7.

Wang X., Miller D.C., Clark A.G. & Antczak D.F. (2012b) Random X inactivation in the mule and horse placenta. *Genome Research* **22**, 1855-63.

Wang Y., Botolin D., Christian B., Busik J., Xu J. & Jump D.B. (2005) Tissue-specific, nutritional, and developmental regulation of rat fatty acid elongases. *Journal of Lipid Research* **46**, 706-15.

Wasson J., Skolnick G., Love-Gregory L. & Permutt A. (2002) Assessing allele frequencies of single nucleotide polymorphisms in DNA pools by pyrosequencing technology. *Biotechniques* **32**, 1144-50.

Weise A., Timmermann B., Grabherr M., Werber M., Heyn P., Kosyakova N., Liehr T., Neitzel H., Konrat K., Bommer C., Dietrich C., Rajab A., Reinhardt R., Mundlos

S., Lindner T.H. & Hoffmann K. (2010) High-throughput sequencing of microdissected chromosomal regions. *European Journal of Human Genetics* **18**, 457-62.

Wood J. & Whittemore C. (2007) Pig Meat and Carcass Quality. In: *Whittemore's Science and Practice of Pig Production* (pp. 4-64). Blackwell Publishing Ltd.

Wood J.D. & Enser M. (1997) Factors influencing fatty acids in meat and the role of antioxidants in improving meat quality. *British Journal of Nutrition* **78**, S49-S60.

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**, 343-58.

Wood J.D., Richardson R.I., Nute G.R., Fisher A.V., Campo M.M., Kasapidou E., Sheard P.R. & Enser M. (2004) Effects of fatty acids on meat quality: a review. *Meat Science* **66**, 21-32.

Xi S., Yin W., Wang Z., Kusunoki M., Lian X., Koike T., Fan J. & Zhang Q. (2004) A minipig model of high-fat/high-sucrose diet-induced diabetes and atherosclerosis. *International Journal of Experimental Pathology* **85**, 223-31.

Yang B., Navarro N., Noguera J.L., Muñoz M., Guo T.F., Yang K.X., Ma J.W., Folch J.M., Huang L.S. & Pérez-Enciso M. (2011) Building phenotype networks to improve QTL detection: a comparative analysis of fatty acid and fat traits in pigs. *Journal of Animal Breeding and Genetics* **128**, 329-43.

Yang B., Zhang W., Zhang Z., Fan Y., Xie X., Ai H., Ma J., Xiao S., Huang L. & Ren J. (2013) Genome-Wide Association Analyses for Fatty Acid Composition in Porcine Muscle and Abdominal Fat Tissues. *PLoS ONE* **8**, e65554.

Yang K.X., Ma J.W., Guo Y.M., Guo T.F., Zhao Y.G., Ding N.S., Betti M., Plastow G.S. & Huang L.S. (2010) Correlations between fat depot traits and fatty acid composition in abdominal subcutaneous adipose tissue and longissimus muscle: Results from a White Duroc x Erhualian intercross F2 population. *Journal of Animal Science* **88**, 3538-45.

Zhao H., Rothschild M., Fernando R. & Dekkers J.M. (2003) Tests of candidate genes in breed cross populations for QTL mapping in livestock. *Mammalian Genome* **14**, 472-82.

Zhao Q.-Y., Wang Y., Kong Y.-M., Luo D., Li X. & Hao P. (2011) Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC Bioinformatics* **12**, S2.

Zimmerman A.W. & Veerkamp J.H. (2002) New insights into the structure and function of fatty acid-binding proteins. *Cellular and Molecular Life Sciences CMLS* **59**, 1096-116.

7. ANNEXES

ANNEXES

7.1.- Supplementary material of paper I "Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs"

Supplementary table

Table S1. Primers for the ACSL4 promoter sequencing (R) and RT-qPCR (RT) study.

Name	Sequence	Amplicon lenght (bp)	Tm	[$MgCl_2$]
ACSL4-P1-Fw (R)	5'-CGTTTTCTTTCTTGCGGTGATA-3'	600	62ºC	2mM
ACSL4-P1-Rv (R)	5'-AACGTGATGGAGCGTGTTTTC-3'			
ACSL4-P2-Fw (R)	5'-CATTACAGGGTGGAGCCCTAAG-3'	584	63ºC	1.5mM
ACSL4-P2-Rv2 (R)	5'-GGGCTGCAGTTACTCACCAGA -3'			
ACSL4-RT-Fw2 (RT)	5'-GGGACCAAAGGACACATATATCG -3'	125	60ºC	2.5mM
ACSL4-RT-Rv2 (RT)	5'-GGTCAGAGAGTGTCAGTGGAGAAG -3'			
HPRT1-RT-Fw (RT)	5'-TCATTATGCCGAGGATTTGGA-3'	91	60ºC	5mM
HPRT1-RT-Rv (RT)	5'-CTCTTTCATCACATCTCGAGCAA-3'			

Supplementary figures

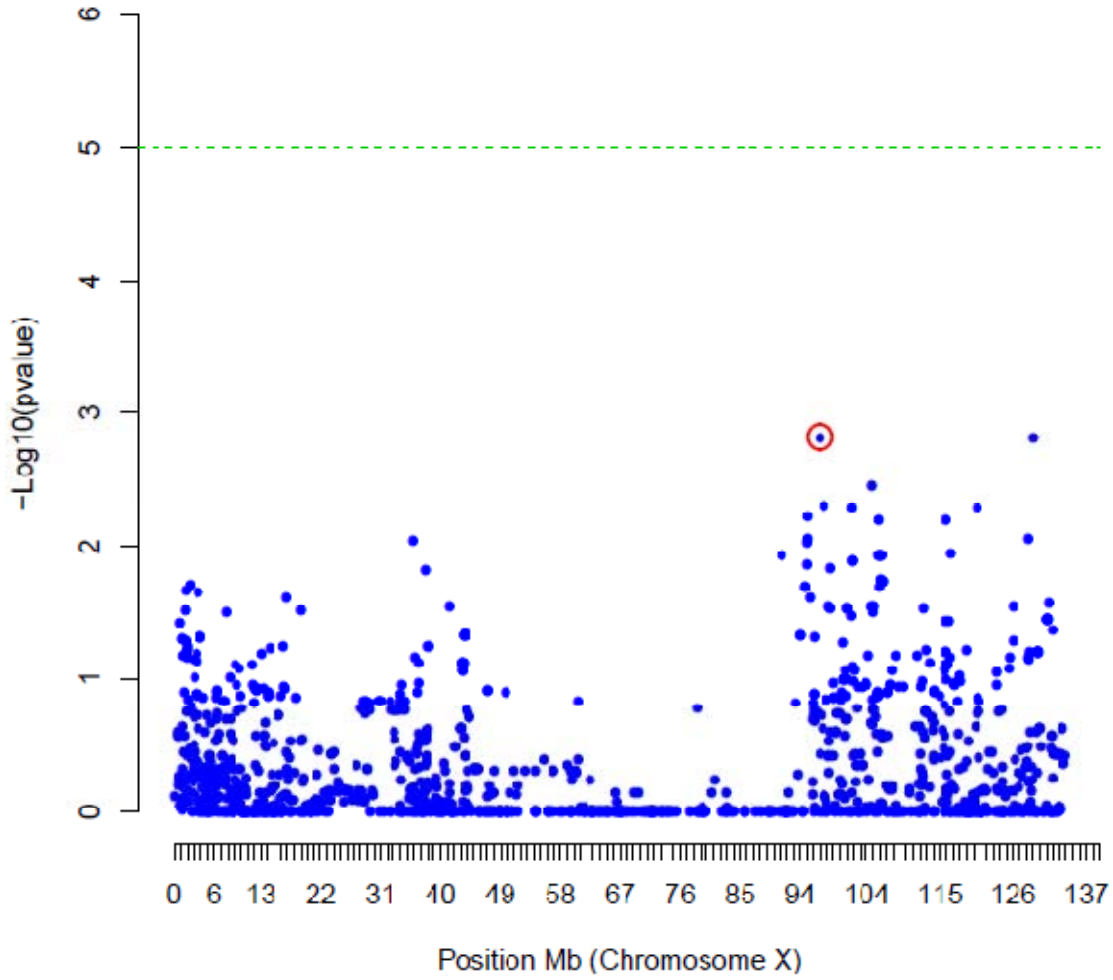

Figure 1S. Association analysis of ACSL4 expression level in liver with SNP genotypes for chromosome X including the c.2645G>A ACSL4 polymorphism (labeled with a red circle). Positions in Mb are relative to the Sus scrofa Assembly 10.

Figure 2S. Nucleotide sequences of the 5'-flanking region of porcine ACSL4 gene. Pig promoter displayed 74.5% similarity with the human promoter. Potential binding sites for transcription factors (>85 cut-off score) are underlined. Positions in: green, indicates transcription factors expressed in liver and conserved in human promoter; blue, indicates transcription factors expressed in liver and red, transcription factors conserved in human promoter.

7.2.- Supplementary material of paper II "Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs"

Supplementary tables

Table S1. List of SNPs for SSC8 linkage map and haplotype reconstruction

See table at:

http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0053687#s6

Table S2. Primers for *ELOVL6* mRNA sequencing (R), promoter sequencing (P) and RT-qPCR (RT) study.

Name	Sequence	Amplicon lenght (bp)	Tm	[$MgCl_2$]
ELOVL6-Fw1 (R)	5'-GGAAGCAGACAGGAGAACACTC-3'	688	58ºC	2mM
ELOVL6-Rv1 (R)	5'-TGATGTGGTGATACCAGTGCAG-3'			
ELOVL6-Fw2 (R)	5'-TCACTGTGCTCCTGTACTCTTGG-3'	499	62ºC	2.5mM
ELOVL6-Rv2 (R)	5'-TAAGCTGCCTTGGGTTTTGTG-3'			
ELOVL6-P-Fw1 (P)	5'-GAGAGCAGGGGTTCAGTAGAGG-3'	604	62ºC	2mM
ELOVL6-P-Rv1 (P)	5'-AGGAAGTGGTGTCGAGGTCATC-3'			
ELOVL6-P-Fw2 (P)	5'-CCAGAGCTGGCAGGTTTTACTA-3'	605	62ºC	2mM
ELOVL6-P-Rv2 (P)	5'-CGGAGTCGCTACGTGTTCTCTA-3'			
ELOVL6-RT-Fw (RT)	5'- AGCAGTTCAACGAGAACGAAGCC -3'	103	60ºC	5mM
ELOVL6-RT-Rv (RT)	5'- TGCCGACCGCCAAAGATAAAG -3'			
HPRT1-RT-Fw (RT)	5'-TCATTATGCCGAGGATTTGGA-3'	91	60ºC	5mM
HPRT1-RT-Rv (RT)	5'-CTCTTTCATCACATCTCGAGCAA-3'			
β2M-RT-Fw (RT)	5'-ACCTTCTGGTCCACACTGAGTTC-3'	100	60ºC	5mM
β2M -RT-Rv (RT)	5'-GGTCTCGATCCCACTTAACTATCTTG-3'			

Supplementary figures

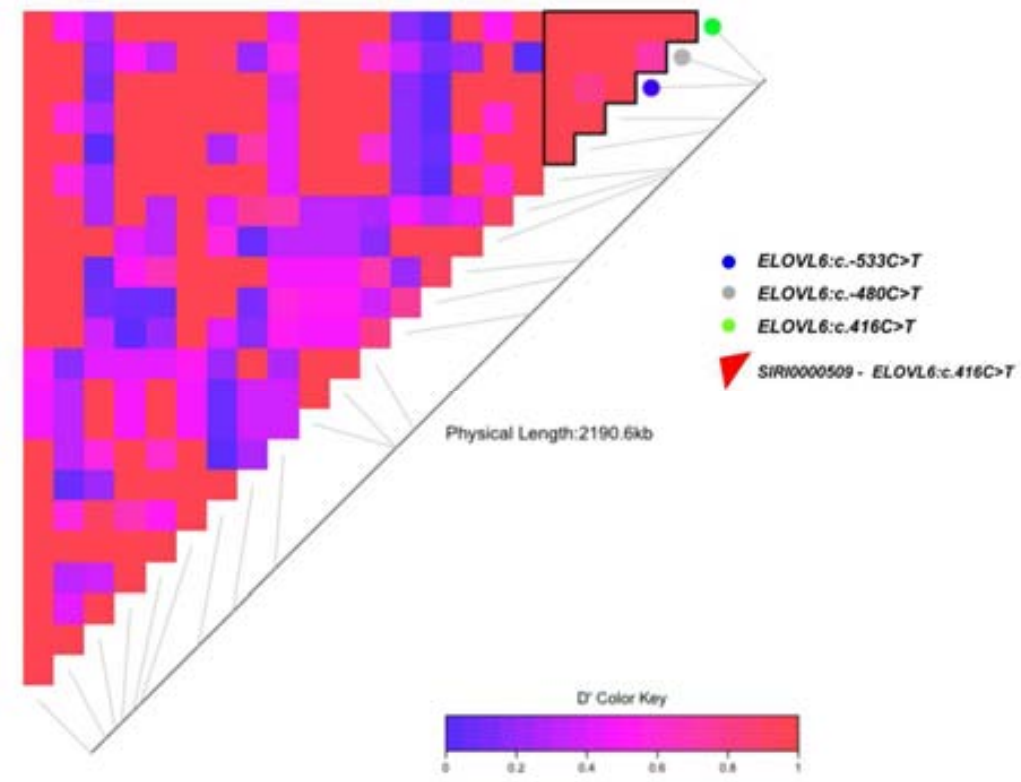

Figure S1. Linkage disequilibrium among *ELOVL6* polymorphisms. Pattern of linkage disequilibrium analysis between the three identified polymorphisms on the *ELOVL6* gene and the most significant SNP detected in both GWAS and fine mapping. Figure colored from blue to red according to LD strength between consecutive markers.

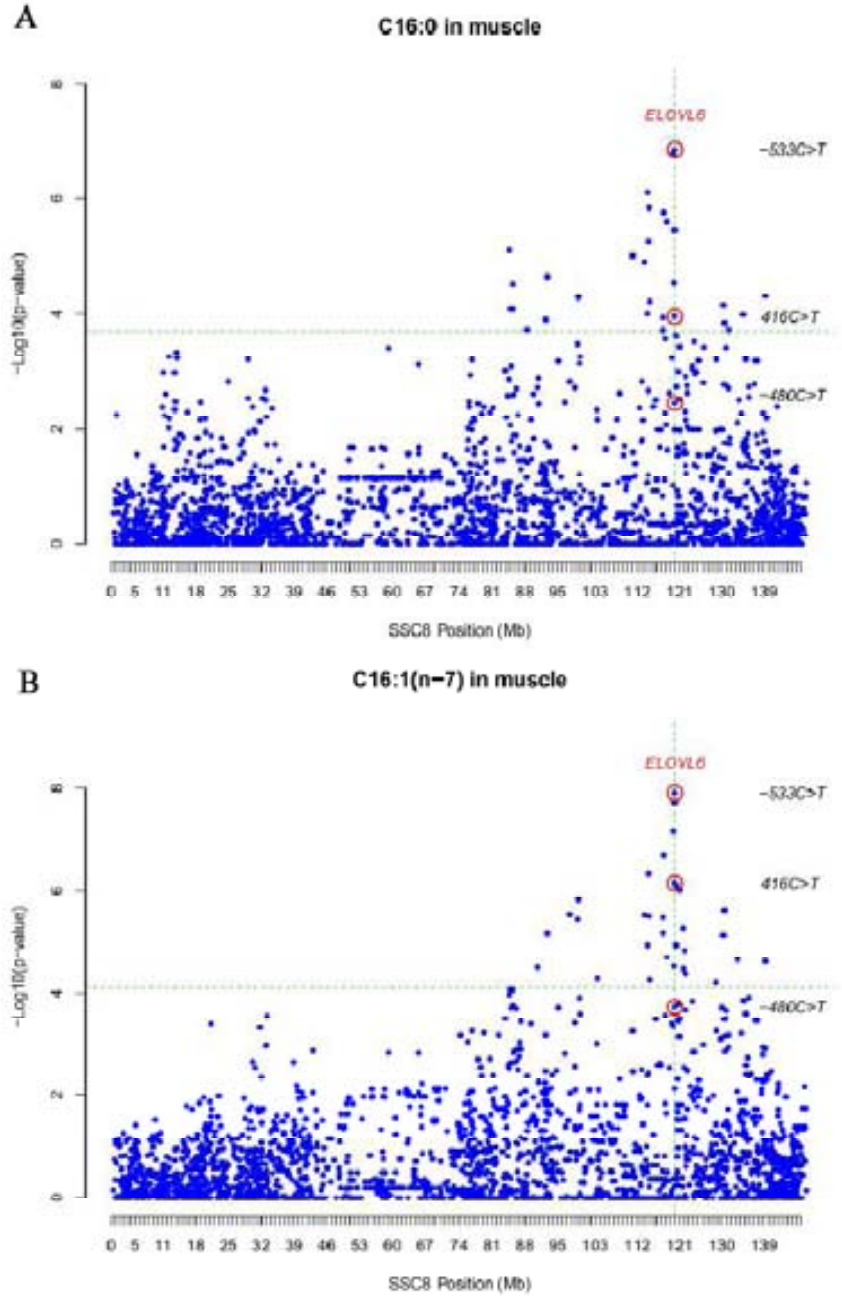

Figure S2. Association of SNPs from SSC8 and *ELOVL6* polymorphims with palmitic and palmitoleic acid content. Association analyses of C16:0 (A) and C16:1(n-7) (B) with genotypes of markers included in the Porcine SNP60 Bead-Chip (Illumina). *ELOVL6* polymorphisms are included and labeled with a red circle. Positions in Mb are relative to the *Sscrofa10.2 assembly* of the pig genome. The horizontal, dashed line indicates the genome-wide significance level (FDR-based q-value ≤ 0.05).

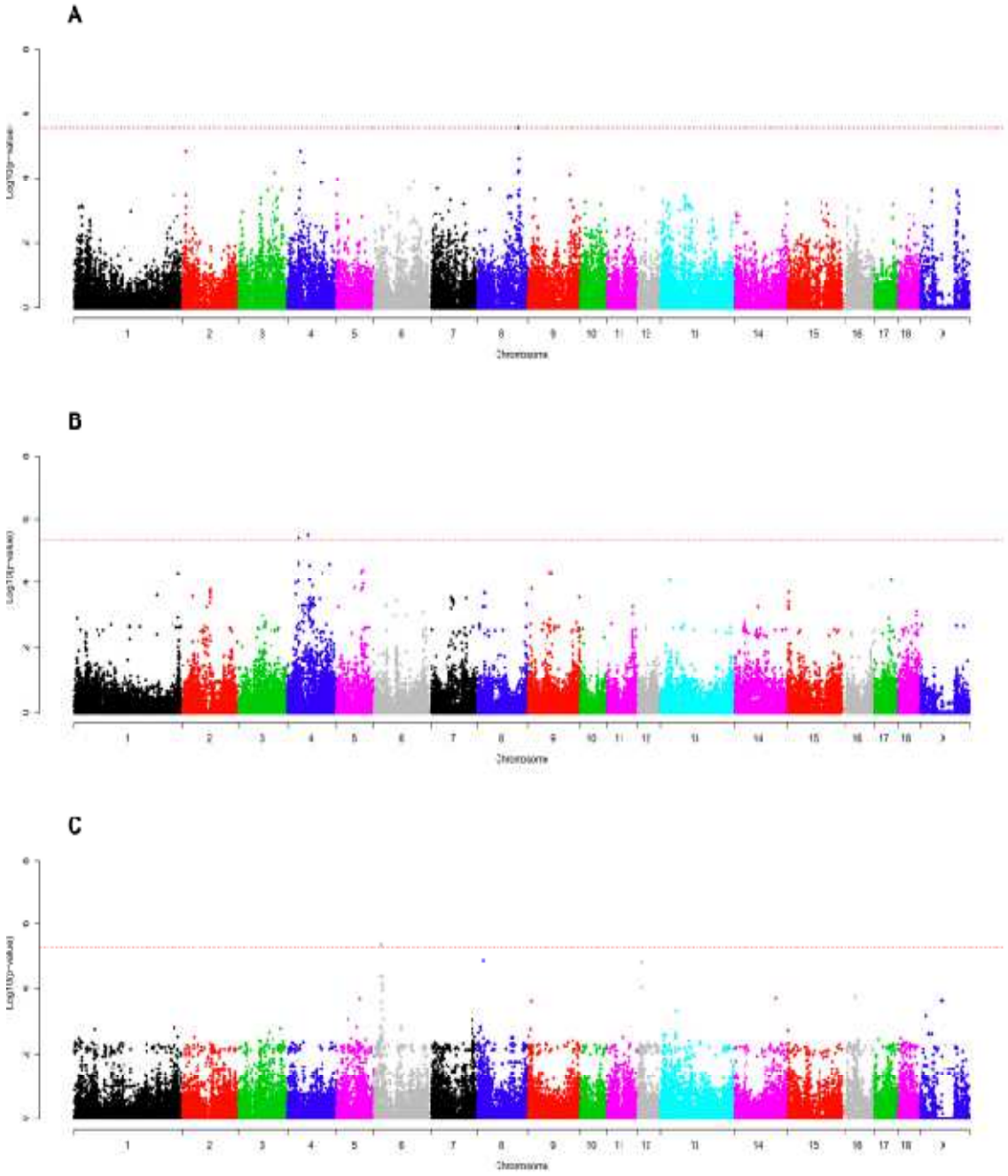

Figure S3. GWAS for *ELOVL6* gene expression in backfat, liver and muscle. Association analyses of *ELOVL6* expression levels in backfat (A), liver (B) and muscle (C) with genotypes of markers included in the Porcine SNP60 Bead-Chip (Illumina). Positions in Mb are relative to the *Sscrofa10.2 assembly* of the pig genome. The horizontal, dashed line indicates the genome-wide significance level (FDR-based q-value ≤ 0.1).

7.3.- Supplementary material of paper IV "Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition"

Supplementary tables

Table S1. Mean (±SEM) comparison between H and L groups of the intramuscular fatty acid composition traits.

Carcass quality	Group H	Group L	Significance	p-value
Intramuscular fat (IMF)	2.27 ± 0.63	1.76 ± 0.75	NS	0.41796
Carcass weight (CW)	64.10 ± 16.13	67.50 ± 8.55	NS	0.76321
Ham weight (HW)	17.72 ± 3.15	17.18 ± 0.87	NS	0.78914
Shoulder weight (SW)	6.04 ± 1.04	5.67 ± 0.34	NS	0.59235
Backfat thickness (BFT)	13.33 ± 2.52	18 ± 5.20	NS	0.23410
Fatty acids				
Saturated FA				
Myristic acid (C14:0)	1.09 ± 0.18	1.26 ± 0.08	NS	0.21009
Palmitic acid (C16:0)	21.45 ± 0.70	24.01 ± 0.65	**	0.0095
Heptadecenoic acid (C17:0)	0.38 ± 0.05	0.21 ± 0.04	*	0.0101
Stearic acid (C18:0)	14.00 ± 1.11	13.90 ± 1.02	NS	0.91878
Arachidic acid (C20:0)	0.33 ± 0.03	0.26 ± 0.05	NS	0.12730
Monounsaturated FA				
Palmitoleic acid (C16:1 n-7)	2.25 ± 0.43	2.95 ± 0.20	·	0.06403
Heptadecenoic acid (C17:1)	0.36 ± 0.10	0.26 ± 0.03	NS	0.17130
Oleic acid (C18:1 n-9)	35.16 ± 3.94	43.17 ± 1.26	*	0.02840
Octadecenoic acid (C18:1 n-7)	3.86 ± 0.20	4.10 ± 0.19	NS	0.21116
Eicosenoic acid (C20:1 n-9)	0.88 ± 0.18	0.80 ± 0.02	NS	0.47216
Polyunsaturated FA				
Linoleic acid (C18:2 n-6)	13.64 ± 1.99	6.84 ± 0.63	**	0.00485
α-Linolenic acid (C18:3 n-3)	1.43 ± 0.44	0.52 ± 0.04	*	0.02448
Eicosadienoic acid (C20:2 n-6)	0.51 ± 0.12	0.37 ± 0.06	NS	0.13036
Eicosatrienoic acid (C20:3 n-6)	0.50 ± 0.21	0.21 ± 0.01	*	0.04215
Arachidonic acid (C20:4 n-6)	3.46 ± 1.54	0.74 ± 0.23	*	0.03899
Metabolic ratios				
Average Chain Length (ACL)	17.46 ± 0.01	17.37 ± 0.01	***	0.000678
Saturated FA (SFA)	37.23 ± 0.38	39.64 ± 1.17	*	0.02783
Monounsaturated FA (MUFA)	42.89 ± 3.64	51.63 ± 1.46	*	0.01810
Polyunsaturated FA (PUFA)	19.54 ± 3.95	8.61 ± 0.80	**	0.00929
PUFA(n-3)/PUFA(n-6)	0.08 ± 0.02	0.07 ± 0.01	NS	0.32977
Peroxidability index (PI)	33.92 ± 8.78	13.51 ± 1.40	*	0.01644
Double-bond index (DBI)	0.48 ± 0.11	0.19 ± 0.02	*	0.01324
Unsaturated index (UI)	0.90 ± 0.08	0.71 ± 0.02	*	0.01319

NS: p-value > 0.1, · p-value > 0.05, * p-value < 0.05, **p-value < 0.01, ***p-value <0.001

Table S2. Cufflinks transcript assembly (TA) statistics for each sample

Animals	BC1		BC2		BC3		BC4		BC5		BC6	
	TA	%	TA	%	TA	%	TA	%	TA	%	TA	%
=	25,919	47	25,858	45	25,991	42	25,919	45	25,958	43	25,905	45
c	4	0	4	0	5	0	3	0	3	0	3	0
e	1,722	3	1,651	3	2,063	3	1,742	3	2,191	4	2,018	4
i	3,865	7	4,514	8	7,185	12	4,717	8	6,773	11	5,091	9
j	11,404	21	11,436	20	12,072	20	11,526	20	11,557	19	11,800	21
o	763	1	865	2	810	1	808	1	762	1	831	1
p	1,274	2	1,359	2	1,682	3	1,439	3	1,502	2	1,339	2
s	14	0	13	0	13	0	13	0	7	0	10	0
u	9,544	17	11,196	20	11,580	19	10,896	19	11,800	19	10,156	18
x	336	1	390	1	307	0	348	1	328	1	340	1
Total	54,845	100	57,286	100	61,708	100	57,411	100	60,881	100	57,493	100

BC1 to BC3 correspond to animals of the L group, while BC4 to BC6 correspond to animals of the H group. Class codes described by Cuffcompare: "=" Exactly equal to the reference annotation, "c " Contained in the reference annotation, "e" Possible pre-mRNA molecule, "i " An exon falling into an intron of the reference, "j " New isoforms, "o" Unknown, generic overlap with reference, "p" Possible polymerase run-on fragment, “s” An intron of the transfrag overlaps a reference intron on the opposite strand, "u" Unknown, intergenic transcript, “x” Exonic overlap with reference on the opposite strand.

Table S3. Gene ontology (GO) of the novel predicted proteins in adipose tissue transcriptome.

Pathway	Sequences in Pathway	Enzyme	Number of sequence
Purine metabolism	49	adenylpyrophosphatase	30
Purine metabolism	49	DNA polymerase	4
Purine metabolism	49	phosphatase	3
Purine metabolism	49	RNA polymerase	3
Purine metabolism	49	cyclase	3
Purine metabolism	49	cyclase	2
Purine metabolism	49	kinase	2
Purine metabolism	49	diphosphatase	1
Purine metabolism	49	adenylyltransferase	1
Purine metabolism	49	kinase	1
Phosphatidylinositol signaling system	11	3-kinase	9
Pyrimidine metabolism	11	DNA polymerase	4
Pyrimidine metabolism	11	synthase	4
Pyrimidine metabolism	11	RNA polymerase	3
Phosphatidylinositol signaling system	11	kinase C	1
Phosphatidylinositol signaling system	11	phospholipase C	1
Inositol phosphate metabolism	10	3-kinase	9
Inositol phosphate metabolism	10	phospholipase C	1
T cell receptor signaling pathway	5	phosphatase	3
T cell receptor signaling pathway	5	protein-tyrosine kinase	2
One carbon pool by folate	4	cyclo-ligase	4
Arachidonic acid metabolism	4	epoxide hydrolase	3
Arachidonic acid metabolism	4	A2	1
Chloroalkane and chloroalkene degradation	3	epoxide hydrolase	3
Thiamine metabolism	3	phosphatase	3
Nitrogen metabolism	3	reductase (H+-translocating)	2
Porphyrin and chlorophyll metabolism	3	ceruloplasmin	2
Nitrogen metabolism	3	reductase	1
Porphyrin and chlorophyll metabolism	3	synthase	1
Glycosaminoglycan degradation	3	hexosaminidase	1
Glycosaminoglycan degradation	3	chondroitin sulfatase	1
Glycosaminoglycan degradation	3	N-acetyltransferase	1
Glycosaminoglycan degradation	3	chondroitinsulfatase	1
Glycosphingolipid biosynthesis - ganglio series	3	alpha-2,8-sialyltransferase	1
Glycosphingolipid biosynthesis - ganglio series	3	hexosaminidase	1
Glycosphingolipid biosynthesis - ganglio series	3	alpha-2,3-sialyltransferase	1
mTOR signaling pathway	2	protein kinase	2

Lysine degradation	2	N-methyltransferase	2
Oxidative phosphorylation	2	reductase (H+-translocating)	2
Glycerophospholipid metabolism	2	N-methyltransferase	1
Glycerophospholipid metabolism	2	N-methyltransferase	1
Glycerophospholipid metabolism	2	lecithinase B	1
Glycerophospholipid metabolism	2	A2	1
Amino sugar and nucleotide sugar metabolism	2	hexosaminidase	1
Amino sugar and nucleotide sugar metabolism	2	lyase	1
Glycosphingolipid biosynthesis - globo series	2	alpha-2,8-sialyltransferase	1
Glycosphingolipid biosynthesis - globo series	2	hexosaminidase	1
Steroid hormone biosynthesis	1	17-dehydrogenase	1
Aminobenzoate degradation	1	organophosphate hydrolase	1
Ether lipid metabolism	1	A2	1
Glycine, serine and threonine metabolism	1	synthase	1
Polycyclic aromatic hydrocarbon degradation	1	hydrolase	1
Starch and sucrose metabolism	1	amylo-1,6-glucosidase	1
Starch and sucrose metabolism	1	disproportionating enzyme	1
Pentose phosphate pathway	1	phosphohexokinase	1
Various types of N-glycan biosynthesis	1	hexosaminidase	1
Aminoacyl-tRNA biosynthesis	1	ligase	1
Arginine and proline metabolism	1	reductase	1
Other glycan degradation	1	hexosaminidase	1
Steroid biosynthesis	1	3-dehydrogenase	1
Methane metabolism	1	phosphohexokinase	1
Sulfur metabolism	1	adenylyltransferase	1
Sulfur metabolism	1	kinase	1
Selenocompound metabolism	1	adenylyltransferase	1
Histidine metabolism	1	hydratase	1
Galactose metabolism	1	phosphohexokinase	1
Fructose and mannose metabolism	1	phosphohexokinase	1
Nicotinate and nicotinamide metabolism	1	adenylyltransferase	1
Nicotinate and nicotinamide metabolism	1	adenylyltransferase	1
alpha-Linolenic acid metabolism	1	A2	1
Bisphenol degradation	1	A-esterase	1
Linoleic acid metabolism	1	A2	1
Glycolysis / Gluconeogenesis	1	phosphohexokinase	1
Glycosphingolipid biosynthesis - lacto and neolacto series	1	alpha-2,8-sialyltransferase	1
Retinol metabolism	1	dehydrogenase	1
Sphingolipid metabolism	1	N-acyltransferase	1

Table S4. Description of the repetitive elements identified in the pig adipose tissue transcriptome.

	Number of elements	Length occupied (bp)	Percentage of sequence
SINEs:	92.007	18.038.833	13,96
Alu/B1	0	0	0
MIRs	18.722	2.649.677	2,05
LINEs:	43.428	20.510.590	15,88
LINE1	31.214	17.423.284	13,49
LINE2	10.583	2.716.949	2,1
L3/CR1	1.355	298.567	0,23
RTE	260	69.824	0,05
LTR elements:	14.137	5.213.341	4,04
ERVL	3.408	1.372.601	1,06
ERVL-MaLRs	6.428	2.130.280	1,65
ERV_classI	3.558	1.479.543	1,15
ERV_classII	203	104.274	0,08
DNA elements:	12.695	2.658.910	2,06
hAT-Charlie	7.871	1.501.938	1,16
TcMar-Tigger	2.129	607.881	0,47
Unclassified:	118	24.837	0,02
Total interspersed		46.446.511	35,95
Small RNA:	245	22.571	0,02
Satellites:	25	3.993	0
Simple repeats:	22.123	917.390	0,71
Low complexity:	17.158	688.060	0,53

Total length:	129.196.334
GC level:	43,95%
Bases masked:	48,071,297 (37.21%)

Table S5. Differential-expressed genes between H and L groups with a fold difference ≥ 1.2 and a p-value ≤ 0.01.

Gene name	p-value	Fold difference_IPA
ENSSSCG00000012178	5.66E-13	-532.4303313
SGSH	2.52E-07	-532.4303313
ENSSSCG00000024291	4.51E-05	-532.4303313
CILP2	3.17E-65	-532.4303313
COMP	1.96E-60	-188.2422282
VCP	3.48E-07	-170.1213445
PHACTR2	1.84E-16	-145.5586597
GABARAP	6.94E-16	-97.67583212
PA2G4	1.63E-09	-84.42223011
BANF1	2.94E-08	-77.69393434
PIN1	4.94E-11	-76.01928082
COL11A2	3.46E-33	-71.20911731
CRABP1	5.92E-06	-57.44388231
DDX3Y	8.77E-12	-50.7863987
MAP6	3.91E-21	-49.685235
CLK1	1.41E-06	-48.1437998
TNMD	2.83E-30	-46.91141157
ATP6V1B2	8.02E-09	-44.11473892
ENSSSCG00000030665	9.52E-05	-39.70114676
GPAA1	3.06E-10	-34.14858475
MYL6	1.22E-36	-30.54078385
NAA38	2.48E-06	-26.10340507
AHSA1	4.73E-06	-25.72409838
C22orf32	7.51E-06	-23.19250678
NDUFA4L2	4.32E-05	-21.61318447
ACAN	1.11E-14	-21.13155497
TNFRSF11B	1.09E-07	-20.70999325
ENSSSCG00000029302	0.001551544	-17.12268369
WNT11	0.000244186	-14.33521226
NEFL	1.39E-05	-14.19048443
ENSSSCG00000001230	0.002460497	-12.21820734
PGLYRP2	2.72E-11	-11.26863596
FHL3	0.003101127	-11.01703936
ENSSSCG00000030475	0.007127068	-10.9361492
ANGPTL5	3.46E-08	-10.64063269
RDH5	2.59E-22	-10.40408649
CD1E	0.001156221	-10.29159064
FASN	5.68E-15	-9.498110275
FASN	3.28E-16	-9.304104803
MOGAT2	1.62E-09	-9.098109838
FBXW4	0.002184785	-9.062396579
ENSSSCG00000026762	0.000796892	-8.725646682

AQP4	0.000212097	-8.660855304
DPYSL5	0.007401313	-8.29857662
COL11A1	7.71E-15	-8.201936331
ENSSSCG00000025853	1.65E-14	-8.15881097
DUS3L	0.000213703	-7.998458928
THBS4	7.99E-07	-7.954382035
SCD	1.95E-06	-7.872690839
TNR	0.002631676	-7.757619353
MFRP	0.004080445	-7.747098757
FOLR1	0.000111656	-7.496266711
CXCL1	4.38E-08	-7.49359219
ARMC10	3.60E-10	-7.41722734
ENSSSCG00000001689	0.001264956	-7.40768803
PDE4A	0.000232243	-7.107310606
SHISA2	3.15E-06	-7.055422261
ANKRD34B	0.004560926	-6.914085152
OTOR	8.02E-07	-6.480862888
PDE6C	5.83E-05	-6.47543775
AKR1CL1	2.72E-15	-6.272353017
ENSSSCG00000021783	0.005084902	-6.186075194
CAPZA2	3.08E-05	-6.038944725
TMEM18	7.28E-05	-5.831016588
GTPBP2	6.99E-05	-5.820081475
CD4	4.99E-06	-5.614199944
O3FAR1	2.31E-06	-5.525986606
THBS4	1.05E-12	-5.4176729
TMEM176A	0.007485571	-5.252166651
NAT8L	0.001712556	-5.211647667
MEPCE	2.91E-05	-4.998469885
THBS4	1.32E-11	-4.92573364
KLHL31	0.000702205	-4.911885184
CTDNEP1	5.79E-06	-4.728696114
C2orf40	6.15E-06	-4.67697509
CD164L2	1.15E-05	-4.656232019
C19orf53	0.00027439	-4.646468254
FMOD	3.99E-09	-4.531021667
PRG4	1.64E-09	-4.43458057
METTL11B	0.005259007	-4.432321045
CTCFL	0.007262627	-4.420704688
CLCA2	0.009619478	-4.354736955
ATP6V1B2	0.003691395	-4.276549538
USP9Y	0.005109625	-4.034431542
SLC26A11	0.001580537	-3.98081099
TNC	1.38E-07	-3.940411033
RP11-48B14.2	1.13E-05	-3.922378132
COL8A2	0.003280829	-3.919121361
PODNL1	0.004887071	-3.852016136

B3GNT4	0.002316484	-3.822399162
UGT8	0.001804491	-3.799129272
CPXM2	4.23E-08	-3.685087943
GLT25D2	0.000575717	-3.611885311
AL354898.1	1.53E-07	-3.599630457
GALNTL5	0.000112366	-3.540783606
UNC119	0.002388667	-3.457697487
B3GAT1	0.006045165	-3.434238092
BBS1	0.00824057	-3.430423133
ATXN10	0.004488698	-3.405358077
IGFN1	0.00672629	-3.381084411
PON1	0.000407806	-3.360630898
PON1	0.000488007	-3.348568863
CHST13	0.000119938	-3.314096596
FAM180A	0.000319077	-3.302039513
S100A8	0.005863355	-3.295446049
PIK3C2G	3.45E-07	-3.280209359
SAL1	6.11E-05	-3.266767852
RET	0.00244771	-3.262207452
RDH16	9.19E-08	-3.261029687
MPZ	3.91E-06	-3.25349973
NTM	0.002496208	-3.167854142
KEAP1	0.009662243	-3.148418187
SNCA	0.00493107	-3.140127294
PSPH	0.00036368	-3.140043572
PLEKHB1	0.000817333	-3.100861098
FAM180B	0.000282746	-3.090020731
HIST1H1D	1.08E-05	-3.08951783
DRP2	0.001285392	-3.068136549
C1QTNF3	4.99E-06	-3.050359771
GRIA2	0.001649437	-3.018440147
ACACA	0.009761151	-3.018002793
CYP2A6	4.70E-07	-3.006134895
PLP1	0.000270828	-2.983372642
SNCG	1.89E-06	-2.98225412
LRRC66	0.005352342	-2.976243593
CENPF	0.008953206	-2.944930742
AIFM3	0.000577269	-2.90438047
UBE2QL1	0.002214124	-2.893739993
CNN1	4.04E-07	-2.881626352
DES	5.15E-07	-2.86221602
ECHDC1	2.82E-07	-2.833927966
IQGAP3	0.002031281	-2.821995041
CA3	1.49E-05	-2.794437198
G0S2	0.000935079	-2.794250472
CTSF	6.30E-06	-2.768286468
KERA	0.000156295	-2.761036529

ENSSSCG00000004911	0.003040118	-2.698456778
CAPN6	0.003248354	-2.683033425
ATHL1	0.002158125	-2.675888687
ACACA	9.20E-07	-2.670036874
IVD	3.21E-06	-2.660981179
CHAD	0.00102904	-2.657836303
ESRRG	0.005394296	-2.64487737
CRABP2	0.000185736	-2.622730965
ANKRD35	0.001014586	-2.618595365
TNC	4.61E-05	-2.579105101
CCL21	0.001494198	-2.553323186
SLC2A12	0.00518463	-2.539135509
MSH3	0.009518608	-2.501565825
DGKB	0.001021612	-2.495436636
AP3M2	0.00629032	-2.492494338
RDH16	0.000223418	-2.482126114
ACTG2	1.81E-05	-2.480054497
P2RX6	0.006415296	-2.445482594
ACE2	1.31E-05	-2.4378672
NRSN1	0.004162272	-2.42403552
LOXL3	0.000820672	-2.419256601
F5	1.68E-05	-2.413775111
CTNNAL1	8.76E-05	-2.383544297
S100A12	0.000375281	-2.377232456
PRIM2	0.003037368	-2.361948403
CES1	0.000561666	-2.349428426
ALDH4A1	0.000110231	-2.347847948
OSR2	0.008784584	-2.327816297
PAK1	0.000450063	-2.326536162
ALDH1A1	0.000183279	-2.32109559
AACS	2.66E-05	-2.294951256
RP11-152F13.10	0.008224686	-2.280284182
CIDEC	0.000463677	-2.253783048
SHMT2	0.005719892	-2.250858177
ACVR1C	0.009612864	-2.235748995
LSS	0.000227378	-2.226387586
ENSSSCG00000021041	0.005169772	-2.219707252
BCHE	0.000657346	-2.219078372
FMO3	0.007993607	-2.195166902
ENSSSCG00000022258	0.007443947	-2.194932507
GDF11	0.000560949	-2.185476117
FKBP14	0.001532501	-2.153632168
EPHX1	0.000414271	-2.149367748
TMEM25	0.001227589	-2.142664786
ROBO1	0.000662689	-2.141240334
ENSSSCG00000006793	0.007382972	-2.139745981
HSD17B12	0.000254482	-2.133730353

RBP1	0.004118959	-2.13192444
IER3	0.004996634	-2.123414376
ACSS2	9.58E-05	-2.122301899
CES1	0.00138795	-2.120488448
SLC36A2	0.00298249	-2.119100557
SLC25A34	0.000704192	-2.118725828
ME1	0.000214541	-2.118509859
MFNG	0.000161349	-2.115076689
HVCN1	0.002233433	-2.098657839
DTX1	0.002281264	-2.090007353
AK4	0.000264048	-2.079638538
SEMA3B	0.003064484	-2.077206648
ALDOC	0.002881688	-2.072032177
ACBD4	0.009412571	-2.07114357
ACLY	0.000247635	-2.069772756
LSS	0.000356208	-2.056023893
PHGDH	0.001313512	-2.049103598
DHCR24	0.00174974	-2.045100756
CXCL1	0.000969432	-2.0187661
PON3	0.000219844	-1.992815652
RCN3	0.004186217	-1.98080087
IGFBP6	0.004819041	-1.978434283
ADIG	0.0038187	-1.974381039
ACADSB	0.000974698	-1.969862221
KIAA0408	0.008634817	-1.960840217
NIPSNAP1	0.007718584	-1.941844802
C10orf116	0.003238919	-1.940418229
SGCB	0.001488587	-1.938229041
ELOVL6	0.000718364	-1.936584081
ISLR	0.001451758	-1.933242428
LIMCH1	0.001142306	-1.926439273
ACADL	0.001607819	-1.915630421
TTC39A	0.009138062	-1.910663156
COMT	0.006756288	-1.907578713
ZFP106	0.001383212	-1.884362914
HOGA1	0.007789795	-1.879012459
DHRS7	0.002760099	-1.868053067
PON2	0.004853953	-1.865705121
CA5B	0.002666396	-1.858996089
GM2A	0.0023546	-1.856534491
PMVK	0.009010919	-1.853034819
PECR	0.008620335	-1.852303985
CDO1	0.004772841	-1.841121655
BMPER	0.003983451	-1.826647744
PPIF	0.00741548	-1.818486116
PAQR6	0.004869107	-1.814782948
GBE1	0.005776636	-1.803044441

TUBB	0.004509324	-1.798463902
RASAL1	0.008421448	-1.79626184
CYP4B1	0.009575914	-1.786960222
COL15A1	0.004162214	-1.785410289
AZGP1	0.005233488	-1.772977502
SLC25A1	0.004395382	-1.76484891
P4HA1	0.008323737	-1.762772656
PCYOX1	0.005843978	-1.755930087
ENSSSCG00000029421	0.007909072	-1.738354792
ENSSSCG00000027204	0.006738768	-1.714461811
ENSSSCG00000005381	0.008734334	-1.709890326
ADHFE1	0.007609502	-1.708709955
DDO	0.009340931	-1.708124809
G6PD	0.008248002	-1.704138304
MPC2	0.009114085	-1.694488906
ITPR2	0.007596296	-1.660324395
ANXA4	0.009824087	-1.639247459
CCBP2	0.009011993	1.652955705
PTPRU	0.006809935	1.702210207
ELN	0.005593613	1.731347077
RHBDL3	0.003393649	1.767121622
CAPZA2	0.002574277	1.788791015
ELTD1	0.009860879	1.797312784
RASGRP3	0.007214174	1.805942197
KLF3	0.008422732	1.814216922
SIPA1L2	0.006399625	1.814435669
PSME4	0.006134508	1.824842589
ETS1	0.0063827	1.831933574
RSAD2	0.003152959	1.837470625
PEAR1	0.005128563	1.838074305
CCRL1	0.004637115	1.842322391
DAPK1	0.004689277	1.856734506
KALRN	0.003642149	1.857596683
CMPK2	0.004521331	1.859830201
CHAC1	0.001960082	1.865484786
UNC13D	0.009493029	1.866005574
NFATC2	0.000673339	1.866047435
S100A4	0.005581437	1.881649135
GNAT1	0.00496559	1.897395244
ZBP1	0.00977599	1.930598225
SMARCC2	0.001633284	1.951604367
LYZ	0.002596715	1.966453421
TTYH2	0.002202806	1.96808591
CLIC5	0.006796291	1.980813865
CCT6A	0.006983581	1.992926578
C10orf116	0.000263764	2.026819006
AQPEP	0.001872454	2.030253032

ENSSSCG00000021357	0.004669966	2.056351559
FERMT1	0.006953469	2.069316362
CLEC2D	0.006781065	2.111000545
COL5A1	0.005119283	2.125339246
ENSSSCG00000022741	0.000370451	2.133092843
SMOC1	0.000369899	2.139651893
CDHR4	0.002057382	2.157993728
MT1A	0.00063135	2.159456729
OAS1	6.71E-05	2.16051868
SH2D6	0.001012049	2.181374544
CD300LG	0.00083435	2.185899807
PDE2A	0.000942739	2.197241707
AFAP1L1	0.000697612	2.234334025
NNAT	4.90E-05	2.240026829
HP	0.001850731	2.241151035
ATP5G2	0.000999494	2.294039213
TMCC3	0.004793843	2.345739998
COL13A1	0.004723534	2.380797682
STEAP4	0.00113289	2.398676717
TUBB2B	0.006251621	2.416286875
KDR	0.000455549	2.425503606
RGS4	0.001965773	2.463719379
DUOX1	0.002353967	2.470685589
ITIH1	0.002883322	2.483536042
LAMC3	0.000186593	2.487623311
SLC6A6	6.70E-05	2.524577909
RBP7	5.29E-05	2.554065875
IFITM2	8.04E-05	2.584818659
MT2A	2.57E-06	2.59469943
RN7SK	0.007581826	2.595136858
REC8	2.92E-05	2.610555877
DNAJB1	3.95E-05	2.618831128
APAF1	2.90E-05	2.647643366
ZSWIM2	0.001793095	2.666793992
FCN2	2.31E-05	2.699212614
OAS2	1.39E-06	2.701150706
PREX2	3.46E-05	2.732639521
ENSSSCG00000016483	2.65E-05	2.754778228
FGG	0.002265151	2.762308509
FGA	0.008455825	2.765313401
REV3L	0.006433996	2.841078624
SLPI	8.44E-06	2.88522132
ENSSSCG00000015140	2.72E-06	2.940537307
VAV3	0.000162089	3.007738406
ALB	1.23E-07	3.053678724
VWCE	2.30E-05	3.237435343
USP43	9.15E-07	3.308164939

AHSG	0.006927697	3.327600993
MMRN2	3.40E-06	3.406328398
THRSP	3.36E-06	3.461208999
KIT	8.37E-08	3.494825324
SH3BGRL	0.004154428	3.607157611
CYP2E1	0.002400592	3.713102454
EPS8	0.008571313	3.752503416
TGFBR2	4.87E-07	3.823016858
MUT	0.004126102	3.976233239
APOB	0.001290194	4.019266209
RNF220	0.00028436	4.058064333
ENSSSCG00000017285	0.003539543	4.07114075
REXO1	0.000736512	4.1134457
ENSSSCG00000015142	1.55E-11	4.140189439
ADH1A	0.005622083	4.152581298
APOH	0.001003643	4.168903272
CXCL1	0.000556466	4.264085863
THRSP	1.53E-07	4.280545416
APOC3	0.000344775	4.281593868
ITGAD	0.007172446	4.401572865
GC	0.005329491	4.478682817
CTSF	6.48E-10	4.492466575
FAM5B	4.44E-07	4.59642738
ESPN	0.008887793	4.622666632
TRIP10	3.51E-08	4.775985798
WDR37	0.008883159	4.781546553
AWN	0.003111639	4.917121345
CH242-486P11.2	0.002536155	5.16785482
ENSSSCG00000024402	0.001732742	5.347486711
CTD-2287O16.3	0.008251975	5.555462277
RXFP2	3.39E-05	5.785426338
CTAGE5	0.000287564	6.152925403
CYC1	0.007102594	6.301342144
THEM5	8.41E-05	6.448370625
HMGXB3	0.003137691	6.82507769
ATP5J2	6.50E-05	7.038248956
MRPS16	0.00291673	7.389435942
PXDN	7.55E-16	7.407529482
C8A	0.004397145	7.980161437
SEMA4C	6.55E-05	8.23074755
PHF10	0.004019416	8.536063834
BCL10	0.002934164	8.613620115
BNIP2	0.003417659	9.465571411
ENSSSCG00000017756	0.002287998	10.93461552
GNB2	9.24E-05	11.66748447
SLC38A2	5.56E-10	13.18686394
LENG8	0.000115972	13.71072504

CTHRC1	0.000105201	14.16819579
TCN1	0.001394294	14.68527123
BANF1	9.14E-09	16.14164231
SLCO1A2	1.98E-13	16.4625269
PRKG2	2.27E-08	17.39766169
ENSSSCG00000010549	0.000443573	17.59430053
PSMB8	0.000791869	19.04164236
ENSSSCG00000024911	3.30E-15	19.62304182
SLC26A11	2.71E-09	20.00884964
PQLC1	0.000114999	20.27142706
ENSSSCG00000030198	0.000487812	24.95967879
TCP11L2	5.31E-08	26.64082525
PIGV	3.79E-05	26.86451444
HIF1AN	2.81E-05	38.53786801
SPG21	4.77E-05	56.05307363
ZNF251	1.35E-09	65.80344146
ENSSSCG00000007025	5.79E-14	90.99083406
ATP6V0A4	8.25E-28	116.8528926
TRIO	3.19E-07	125.8886615
KHDRBS1	7.51E-13	203.7357741
DIABLO	2.62E-08	203.7357741
FGB	1.62E-10	203.7357741
ENSSSCG00000024171	9.82E-08	203.7357741
APRT	2.96E-06	203.7357741
MED23	1.73E-07	203.7357741

Table S6. Genetic networks generated from the differential expressed genes between H and L animals.

ID	Score	Focus Molecules	Top Functions
1	36	23	Lipid Metabolism. Nucleic Acid Metabolism. Small Molecule Biochemistry
2	36	23	Connective Tissue Disorders. Developmental Disorder. Hereditary Disorder
3	36	23	Digestive System Development and Function. Neurological Disease. Visual System Development and Function
4	36	23	Neurological Disease. Cell Morphology. Cellular Assembly and Organization
5	31	23	Developmental Disorder. Hematological Disease. Hereditary Disorder
6	30	20	Molecular Transport. Small Molecule Biochemistry. Cell Cycle
7	24	17	Molecular Transport. Cellular Movement. Embryonic Development
8	23	17	Cellular Function and Maintenance. Cellular Growth and Proliferation. Hematological System Development and Function
9	23	17	Drug Metabolism. Small Molecule Biochemistry. Lipid Metabolism
10	22	16	Molecular Transport. Energy Production. Amino Acid Metabolism
11	20	15	Cancer. Connective Tissue Disorders. Developmental Disorder
12	20	15	Developmental Disorder. Endocrine System Disorders. Hereditary Disorder
13	19	16	Cancer. Hematological Disease. Immunological Disease
14	18	14	Cell Cycle. Developmental Disorder. Cellular Development
15	16	13	Cell Death and Survival. Tumor Morphology. Cellular Function and Maintenance
16	15	13	Cellular Development. Connective Tissue Development and Function. Molecular Transport
17	15	12	Organ Morphology. Skeletal and Muscular System Development and Function. Amino Acid Metabolism
18	14	12	Cancer. Reproductive System Disease. Carbohydrate Metabolism
19	14	12	Cancer. Behavior. Amino Acid Metabolism
20	2	1	Developmental Disorder. Hereditary Disorder. Metabolic Disease

Supplementary figures

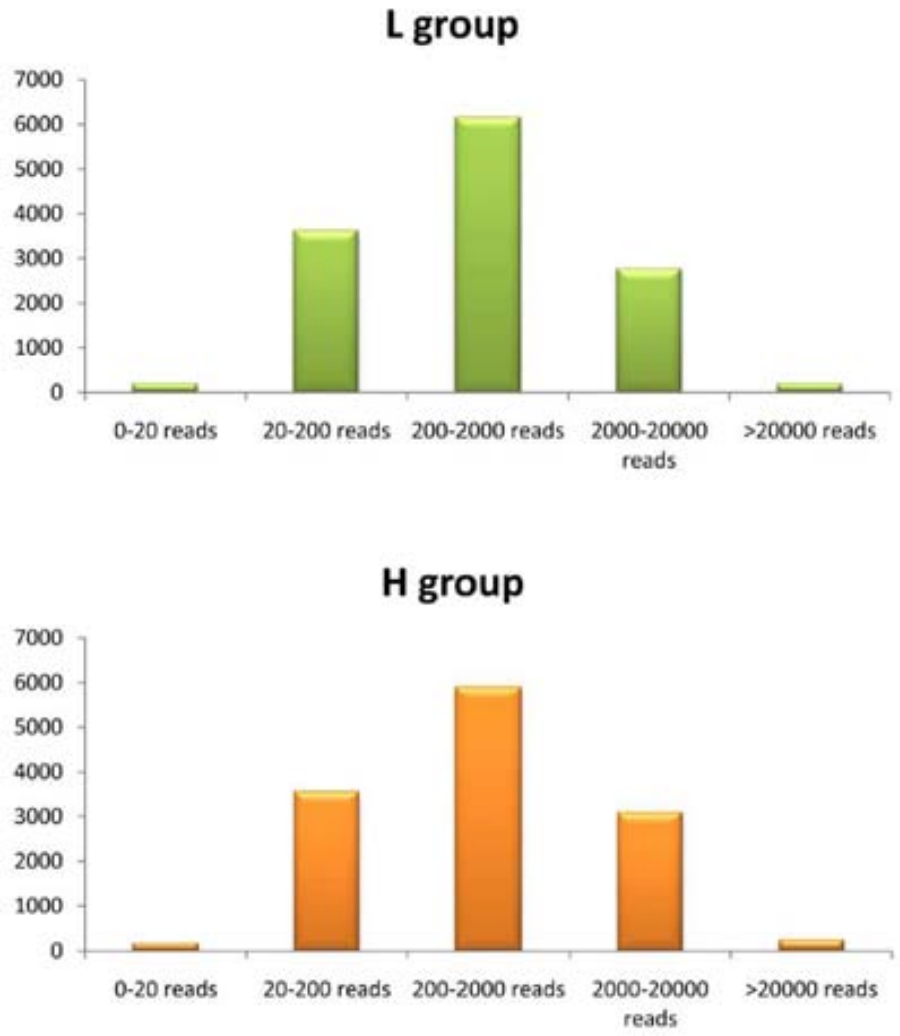

Figure S1. Distribution of gene expression levels in both H (High) and L (Low) groups.

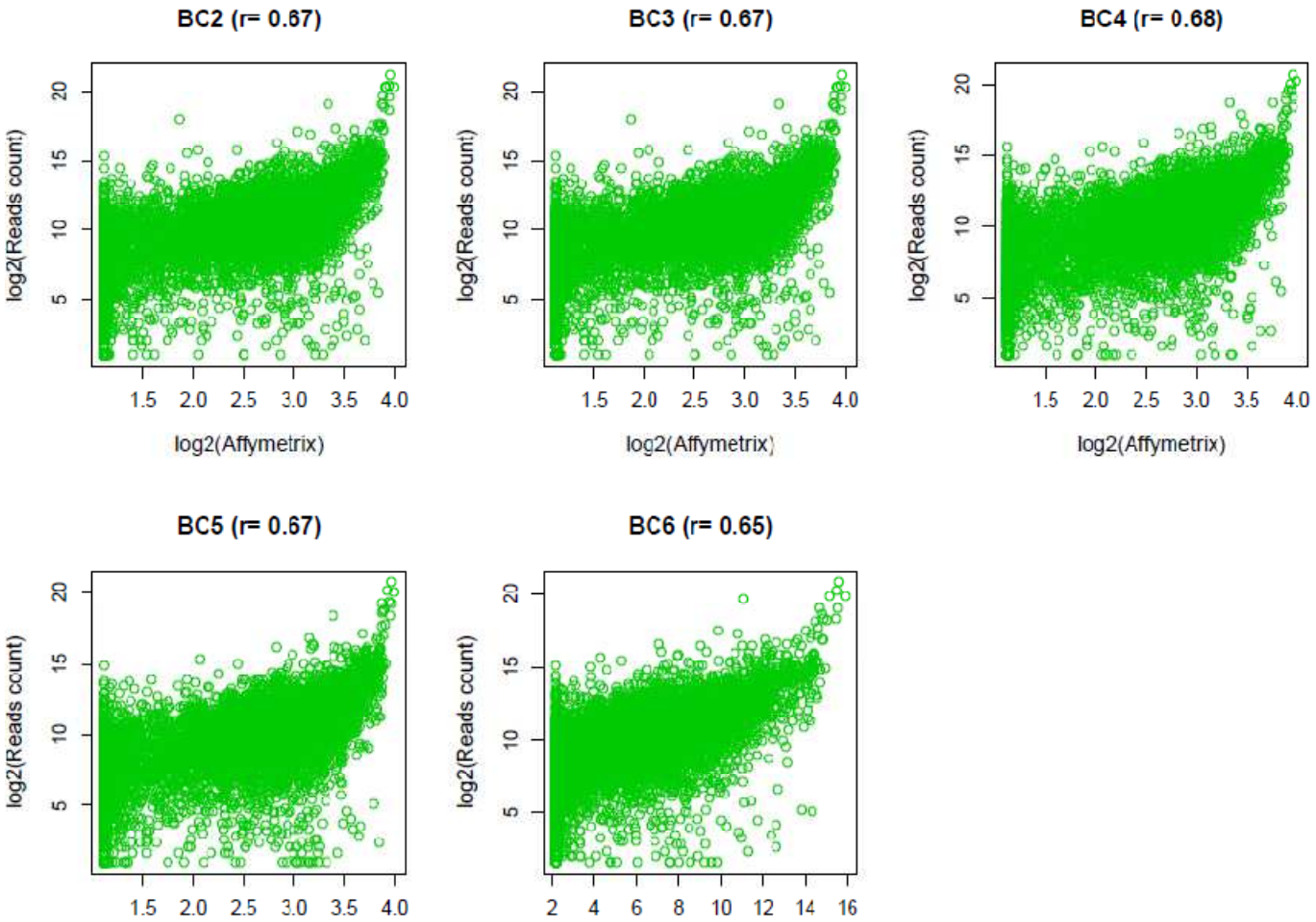

Figure S2. Correlation between expression values of RNA-Seq and Affymetrix microarray. X-axis values are the log2 of expression quantified with Affymetrix microarray technology and y-axis are values of log2(counts).

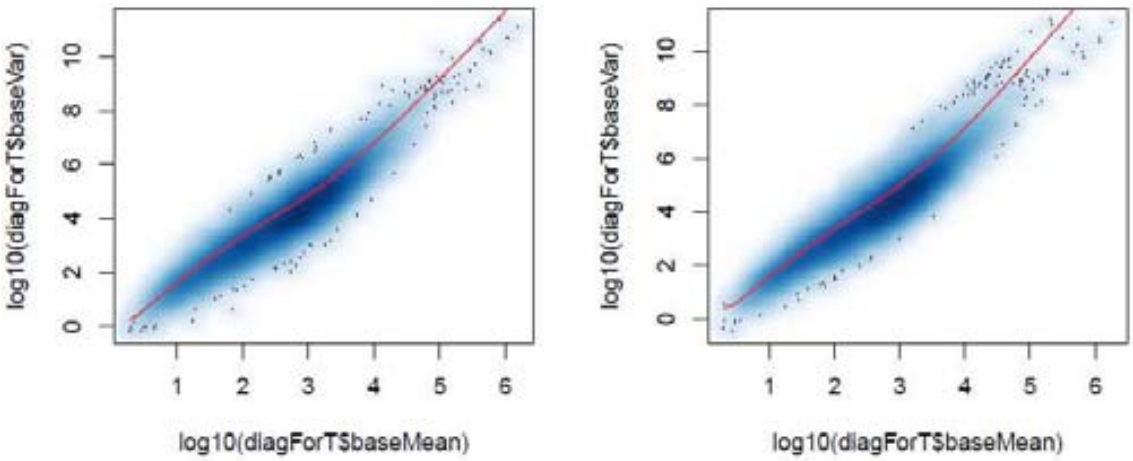

Figure S3. Per-gene estimates of the base variance against the base level. The red line represents the fit variance. X-axis is the log10 of the base mean and y-axis values are the log10 of the base variance.

7.3.- Supplementary material of paper III "Different patterns of methylation on ELOVL6 promoter caused by a promoter polymorphism explains the major QTL effect on fatty acid composition in pigs"

Supplementary tables

Table 1S. Primers for the BAC screening (S). methylation study (M) and site-specific ChIP (C)

Name	Sequence (5'- 3')	Amplicon lenght (bp)	Tm	[$MgCl_2$]
ELOVL6_P_Fw (S)	CCAGAGCTGGCAGGTTTTACTA	605	60ºC	2mM
ELOVL6_P_Rv (S)	CGGAGTCGCTACGTGTTCTCTA			
ELOVL6_e2_Fw (S)	CCTGGTTTCTGCTCTGTATGCT	94	60ºC	2mM
ELOVL6_e2_Rv (S)	CAGCACTAATGGCTTCCTCAGTT			
ELOVL6_e4_Fw (S)	TCACTGTGCTCCTGTACTCTTGG	499	60ºC	2mM
ELOVL6_e4_Rv (S)	TAAGCTGCCTTGGGTTTTGTG			
Met_F1_Fw (M)	TGTGTTTTGTATTGGATTAGTTGG	309	60ºC	1.5mM
Met_F1_Rv (M)	TCCCACRTAAAAAAATCAAACTTC			
Met_Seq1 (M)	YGTTTTTAGTATTTTTAGATAT	----	----	----
Met_F2_Fw (M)	GATTTGGAGGGTGTGGTAAGAGTA	232	60ºC	1.5mM
Met_F2_Rv (M)	TCATCCACAACCTCAATCCT			
Met_Seq2 (M)	GAGGTGGGAAGTTTGA	----	----	----
Met_F3_Fw (M)	TAGGATTGAGGTTGTGGATGATT	200	60ºC	1.5mM
Met_F3_Rv (M)	TCCATCACCCTTTTTACTTATCTACA			
Met_Seq3 (M)	TTTTTTTYGTAAAGGGTTAAT	----	----	----
Met_F4_Fw (M)	TGATTTTTTTTTTTTGGTTATTAG	183	58ºC	2.5mM
Met_F4_Rv (M)	TGGTTATTAGTTATTTTTTATTTA			
Met_Seq4 (M)	AATCAAACTTCCCACCTCCTTAC	----	----	----
ChIP_Fw (C)	CCAAACCTACCACCTGACCTCT	216	60ºC	1.5mM
ChIP_Rv (C)	CTGAAATCCCACGTGAAAGGAT			

Table 2S. Association between phenotypic data and the selected polymorphisms (*ELOVL6:c.-533C>T* and *ELOVL6:c.1922A>G*) and the haplotype formed with both SNPs.

Genetic variants	Backfat gene expression	IMF		Backfat	
		C16:0	C16:1(n-7)	C16:0	C16:1(n-7)
ELOVL6:c.-533C>T	$3.68x10^{-03}$	$1.33x10^{-03}$	$3.72x10^{-04}$	$6.15x10^{-10}$	$9.15x10^{-04}$
ELOVL6:c.1922A>G	$4.46x10^{-02}$	$6.96x10^{-02}$	$6.98x10^{-04}$	$5.44x10^{-08}$	$3.57x10^{-03}$
Haplotype	$7.22x10^{-03}$	$1.81x10^{-03}$	$6.68x10^{-03}$	$1.42x10^{-09}$	$9.66x10^{-04}$

Supplementary figures

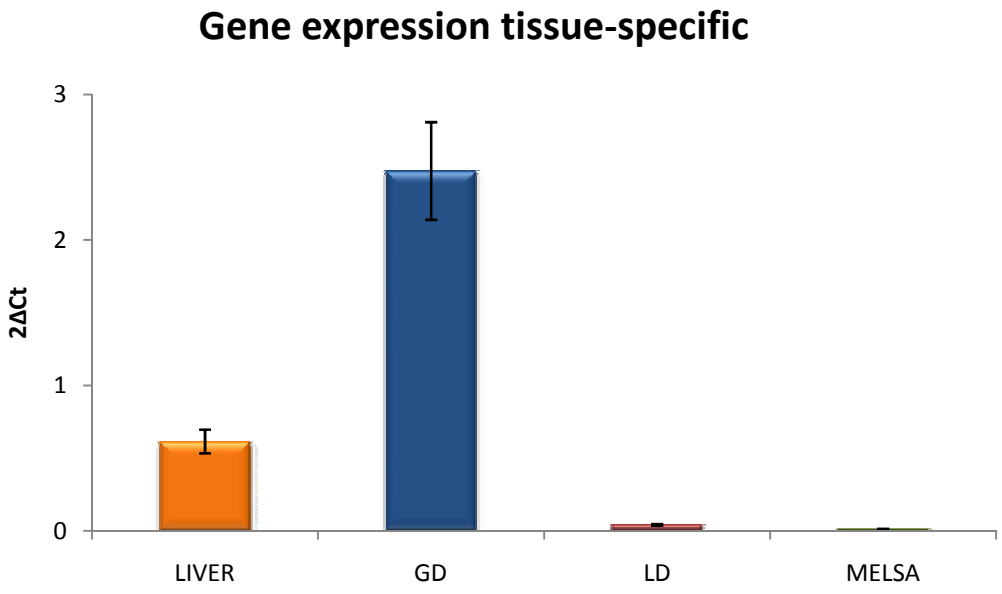

Figure 1S. Tissue-specific differences on *ELOVL6* gene expression among liver, adipose tissue, muscle and spleen. Gene expression was compared using the 2ΔCt data obtained from quantitative PCR analyses. Data represent mean ± SEM.