Universitat Autònoma de Barcelona

Genome-wide association analysis of meat quality and gene expression phenotypes in Duroc pigs

Rayner González Prendes

Universitat Autònoma de Barcelona

Departament de Ciència Animal i dels Aliments
Facultat de Veterinària

Centre de Recerca en Agrigenòmica

Departament de Genètica Animal

Genome-wide association analysis of meat quality and gene expression phenotypes in Duroc pigs

PhD, Thesis in Animal Production

Rayner González Prendes

Supervisor

Dr. Marcel Amills Eras

Bellaterra, March 2017

El Dr. Marcel Amills Eras, professor agregat del Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de Barcelona,

CERTIFICA:

Que Rayner González Prendes ha realitzat sota la seva direcció el treball de tesis doctoral "***Genome-wide association analysis of meat quality and gene expression phenotypes in Duroc pigs*** " per obtenir el grau de Doctor per la Universitat Autònoma de Barcelona.

Que aquest investigació 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 al Departament de Genètica Animal del Centre de Recerca en Agrigenòmica.

Bellaterra, a 23 de Març del 2017.

Dr. Marcel Amills Eras

Summary

The main objective of this thesis was to identify genomic regions associated with technological and lipid composition meat quality traits. In this way, we carried out a GWAS for 57 phenotypes measured in the *gluteus medius* (GM) and *longissimus dorsi* (LD) muscles of pigs from a commercial Duroc line. In general, SNPs included in the PorcineSNP60 BeadChip only explained a limited amount of the phenotypic variance of the meat quality traits recorded in our population (0-51%). Moreover, we detected 40 and 101 genome- and chromosome-wide significant associations respectively. The majority of these associations were muscle-specific, maybe because the GM and LD muscles have different profiles of mRNA expression. Several of these regions were associated with more than one trait, suggesting the existence of pleiotropic effects. Specifically relevant was the genomic region located on SSC14 (120-124 Mb) which was associated with stearic, linoleic, unsaturated, and saturated fatty acids in both LD and GM muscles. We also investigated if QTL regions contain expression QTL (eQTL) influencing the mRNA levels of loci transcribed in the GM muscle. The number of eQTL co-localizing with QTL for meat technological traits (5 *cis*-eQTLs) was lower than that of QTL for intramuscular fat (IMF) composition traits (20 *cis*-eQTL). Besides, we detected SNPs mapping to IMF QTL with *trans*-regulatory effects on gene expression (116 *trans*-eQTL).

We were also interested in analysing the genetic regulation of lipid genes. With this goal, we have performed an eQTL scan for 63 loci that are known to have a key role in lipid metabolism. Our results revealed 13 *cis*- and 18 *trans*-eQTL modulating the expression of 19 loci with a broad variety of biochemical functions. Moreover, we did not detect a clear predominance of either *cis*- or *trans*- effects on gene expression and none of the 31 eQTLs mapped to QTLs for lipid traits. This finding suggests that detected eQTLs have effects on gene expression but not on fatness phenotypes.

We have also investigated the existence of eQTL regulating gene expression in the GM muscle and liver, two tissues with a key role in the regulation of energy homeostasis. In this way, we have mapped 436 *cis*- and 450 *trans*-eQTLs in the GM muscle, while for hepatic genes the number of *cis*- and *trans*-eQTLs was more unbalanced *i.e.* 504 *cis*- vs 3,228 *trans*-eQTLs. The positional concordance between eQTLs maps generated in both tissues was weak, suggesting that the determinism of gene expression is mostly tissue-specific. In addition, we have used SNP data to identify 104 copy number variant regions, 47% of which co-localize with structural variants reported in previous studies. Approximately 39% of these CNVR co-localized with *cis*-eQTL signals, whilst the co-localization of CNVR and *trans*-eQTL was somewhat higher (≈60%). The consistency of these co-localizations in the liver and muscle was weak.

Resumen

El principal objetivo de la presente tesis fue identificar regiones genómicas asociadas con la variación fenotípica de caracteres relacionados con la calidad de la carne. Para ello se realizaron estudios de asociación genómica (GWAS) para el contenido y la composición de la grasa intramuscular así como para la conductividad eléctrica, el pH, y el color de la carne. Dichos fenotipos se midieron en los músculos *gluteus medius* (GM) y *longissimus dorsi* (LD) de cerdos pertenecientes a una línea comercial Duroc. En general, los SNPs incluidos en el PorcineSNP60 BeadChip de Illumina explicaron un porcentaje bajo o moderado (0-51%) de la varianza fenotípica de los caracteres evaluados en nuestra población. Mediante la aproximación GWAS, se identificaron un total de 40 QTLs significativos a nivel genómico y 101 QTLs significativos a nivel cromosómico. Se observó que la mayoría de los QTLs detectados fueron específicos de cada músculo y esto probablemente se deba a que los perfiles de expresión génica de ambos músculos son diferentes. Se detectaron además, varios QTLs con efecto en más de un fenotipo, lo que sugiere la existencia de regiones con efectos pleiotrópicos. Es importante destacar el QTL localizado en el cromosoma porcino SSC14, que presentó asociaciones significativas con el ácido esteárico, linoleico y los porcentajes de ácidos grasos saturados e insaturados en los dos músculos LD y GM. Además, se investigó si las regiones QTL para la calidad de la carne contienen QTL con efectos sobre la expresión génica (eQTL). Con este enfoque se detectaron cinco eQTLs cuyas posiciones coincidieron con QTLs para el pH, la conductividad eléctrica y el color de la carne. Por otra parte, también se identificaron 20 *cis*-eQTL y 116 *trans*-eQTL que mostraban concordancia posicional con QTLs para el contenido y la composición de la grasa intramuscular.

De forma independiente, se analizó la regulación de la expresión génica de 63 loci cuyas funciones están relacionadas con el metabolismo de los lípidos. Nuestros resultados revelaron 13 *cis*-y 18 *trans*-eQTLs asociados a la expresión de 19 loci. No se observó un claro predominio de los *cis*- o *trans*-eQTLs y además ninguno de los 31 eQTLs co-localizó con las regiones QTL para caracteres de engrasamiento. Este hallazgo sugiere que los eQTLs detectados tienen efectos sobre la expresión génica, pero no sobre la variabilidad fenotípica.

El estudio de la regulación de la expresión génica en el músculo GM y en el hígado nos ha permitido detectar 436 *cis*- y 450 *trans*-eQTLs en el músculo GM, mientras que para los genes expresados en el tejido hepático el número de *cis*- y *trans*-eQTLs fue más desequilibrado (504 *cis*- vs 3.228 *trans*-eQTLs). La concordancia posicional de los mapas de QTLs en ambos tejidos fue baja, sugiriendo que existe un determinismo genético que es específico de cada órgano. Por otra parte, se han empleado los datos del PorcineSNP60 BeadChip para identificar 104 regiones genómicas con variaciones en el número de copias (CNVR). El 47 % del total de los CNVR detectados fueron previamente reportados en otras poblaciones. Además, aproximadamente un 39 % de los CNVR co-localizaron con *cis*-eQTLs mientras que las co-localizaciones con los *trans*-eQTLs fueron mayores (≈60%). En general se obtuvo un número bajo de CNVR que co-localizaron simultáneamente con *cis*- o *trans*-eQTLs en el músculo GM y en el tejido hepático.

Index

Chapter I. Introduction

1. Basics of the pig production system.

Pigs are one of the most important sources of meat for humans. Indeed, pork production represents more than 40% of the meat consumed worldwide, and in 2025 an increment of the global demand in 248,795 thousands of tons is expected (OECD-FAO , 2016). Around 90% of the total pork production comes from China, Europe and the United States of America (USDA, 2016). Genetic improvement is one of the main factors that have contributed to the progress and success of pig production, which is often organized in three strata: the nuclei of selection, where purebred F_0 individuals with a high genetic value are raised; the multiplier stratum, where F_1 hybrid individuals are generated by crossing F_0 hogs and sows from two maternal breeds; and the commercial stratum where F_1 individuals are crossed with a third paternal breed, characterized by its excellence in carcass or meat quality traits, to produce F_2 pigs to be slaughtered. The nuclei of selection are generally managed by large integrator companies that own such genetic material and assume several financial risks, such as those associated with the variation of pig feed (grain) and meat prices. These companies sign contracts with farmers which contribute facilities and the labour force necessary to generate the final commercial product. In Spain, approximately 70% of pig production is integrated. Selection of purebred parental individuals relies on the estimation of their genetic value for several traits of economic interest. Such estimates are obtained with the best linear unbiased predictor (BLUP) method, and they are usually combined in an index where each trait is balanced in accordance with its economic importance.

2. The nutritional properties and quality of pig meat

The meat of pigs is mostly made up of proteins (26%), having a low level of carbohydrate and a proportion of fat that ranges from 10 to 16% (Ciobanu et al., 2011). Pig meat is a source of high quality proteins, vitamin B_{12} and iron, that are important for many

physiological processes such as muscle growth and the formation of red blood cells (Paddon-Jones et al., 2008). Besides, the low level of carbohydrates is positively associated with a reduced glycemic index, which is assumed to have beneficial effects on susceptibility to obesity, diabetes and some types of cancer (Biesalski, 2005). On the other hand, an excessive consumption of pig meat can be detrimental because of its cholesterol and saturated fatty acid (FA) contents, particularly if accompanied by a low consumption of fruit and vegetables, reduced physical activity, smoking, and overeating (Eckel et al., 2014).

The definition of pig meat quality is complex and encompasses multiple technological and organoleptic attributes. Besides, it also depends on the preferences of producers and consumers. For example, meat pH, water-holding capacity, cooking loss and oxidative stability are important for the processing industry, because they determine if meat can be safely commercialized (Maltin and Balcerzak, 2003). For the consumer, color, marbling, tenderness, juiciness and flavour are key factors that determine meat acceptance (Maltin and Balcerzak, 2003). Besides genetics, meat quality traits are influenced by a wide array of factors such as muscle type (fiber class and size), management conditions (nutrition, growth rate and age of slaughter) and the very same process of slaughtering (Rosenvold and Andersen, 2003). The transformation of muscle into meat involves a progressive depletion of energy produced by the exhaustion of ATP, a pH decrease from neutrality to acidity (pH $\approx$ 5-5.8), an increase in ionic strength and also the degradation of myofibrillar, cytoskeletal and intermediate filament proteins (Ciobanu et al., 2011). As previously said, important processing meat quality traits are pH, water-holding capacity (ability of the post-mortem muscle to retain water thus avoiding moisture loss), color, cooking loss (change of meat weight due to moisture loss associated with cooking) and firmness (Ciobanu et al., 2011). The rate of post-mortem pH decline is particularly critical: if it is too accelerated and the pH is low (pH < 6, 45 minutes after slaughter), the water-holding capacity of meat will be also low resulting in a high drip loss (pale, soft and exudative meat). On the other hand, a high ultimate pH (pH > 6, 12-48 h after slaughter) will result in an undesirable dark, firm, dry (DFD) meat (Adzitey and Nurul, 2011). Meat quality can be predicted on the basis of muscle electrical conductivity. In general PSE, reddish-pink, soft and exudative (RSE), and reddish-pink, firm and non-exudative (RFN) meats show an increased conductivity, while

DFD meat would display the opposite trend (Lee et al., 2000). With regard to meat color, it is mostly explained by myoglobin content, a parameter that depends on multiple factors such as nutrition, pH dynamics, rate of chilling and storage conditions. The oxidation of myoglobin and oxymyoglobin results in the generation of metmyoglobin, that confers a brown coloration to meat (Mancini and Hunt, 2005). Meat color can be measured by colorimetry, being the Minolta chroma-meter one of the most widely used instruments to achieve such goal. In general, meat quality traits are highly interrelated and they often depend on common factors such as glycogen content (Ciobanu et al., 2011).

Meat quality and nutritional properties are also affected by the content and composition of intramuscular fat (IMF). Brewer et al. (2001) found that chops with a low-middle fat infiltration were more acceptable than chops with a high fat content; and they were also juicier, more tender, oily and flavorful than leaner chops. However, Rincker et al. (2008) found a limited effect of IMF on the juiciness and flavour of meat. Indeed, the effects of IMF on the organoleptic characteristics of meat may vary amongst breeds *i.e* Fernandez et al. (1999) detected significant associations between IMF and meat juiciness and flavour in Meslan × Landrace pigs, but significance of such associations was lower in a Duroc × Landrace cross. Blanchard et al. (2000) detected correlations close to zero between IMF and juiciness, tenderness and flavor, in Large White, British Landrace and Duroc pigs. Differences amongst studies show that the relationship between IMF and meat attributes is complex and depends on many genetic and environmental factors.

With regard to human health, elevated ratios ofpolyunsaturated/monounsaturated fatty acids (PUFA/MUFA) and of omega-3 to omega-6 FA (n-3/n-6) are desirable because PUFA are associated with reduced concentrations of low density lipoproteins and total cholesterol, thus helping to decrease the incidence of coronary heart diseases (Astrup et al., 2011). In contrast, saturated FA (SFA) are associated with obesity, high plasma cholesterol concentration, and an increased risk to suffer coronary heart diseases (Keys et al., 1986) and cancer (Wynder et al., 1997). From a technological point of view, PUFAs have a negative effect on the oxidative stability of meat, causing a rancid flavour, an increment of drip loss, a darker color, a low firmness and a decrease of shelf life (Wood et al., 2008). In

stark contrast, SFA have been associated to desirable sensorial characteristics of meat (Chizzolini et al., 1998; Wood et al., 2008).

3. Heritability of meat quality traits.

Narrow-sense heritability (h^2) refers to the proportion of phenotypic variance explained by additive genetic effects (Falconer and MacKay, 1996). Additive genetic variance can be estimated from the similarity between related individuals by relating the phenotypic covariance of a quantitative trait with the proportion of the genome for which two relatives share genes. Heritabilities are calculated using pedigree information to infer the relationships between individuals in a population (Falconer and MacKay, 1996). Usually, heritabilities of meat quality traits range between 0.10 and 0.30 (Ciobanu et al., 2011, **Table 1**). As reported by Ciobanu et al., (2011), tenderness appears to be more heritable (h^2 = 0.25-0.30) than flavour and juiciness ($h^2 < 0.10$), while in the group of technological traits, meat color displays higher heritabilities (h^2 = 0.15-0.57) than drip loss (h^2 = 0.16) or ultimate pH (h^2 = 0.21). In general, heritabilities for IMF content and composition traits are moderate to high, as shown in **Table 1.**

Table 1. Range of the heritabilities for meat quality traits. (Adapted from Ciobanu et al., 2011)

Traits	Range of heritability
Technological traits	
One hour post-mortem pH	0.04–0.41
Ultimate post-mortem pH	0.07–0.39
Colour (light reflectance, CIE L* value)	0.15–0.57
Water-holding capacity	0.01–0.43
Drip loss	0.01–0.31
Cooking loss	0.00–0.51
Technological yield (cooked ham processing)	0.09–0.40
Napole yield	0.26–0.78
Meat quality index	0.11–0.33
Visual score of meat quality	0.10–0.37
Eating quality traits	
Tenderness (instrumental determination)	0.17–0.46
Tenderness (sensory panel score)	0.18–0.70
Flavour (sensory panel score)	0.01–0.16
Juiciness (sensory panel score)	0.00–0.28
Overall acceptability (sensory panel score)	0.16–0.34
Muscle composition traits	
% water	0.14–0.52
% lipid	0.26–0.86
% glycogen (glycolytic potential)	0.25–0.90
Fat composition traits (backfat)	
% water	0.27–0.42
% stearic acid (C18:0)	0.30–0.57
% linoleic acid (C18:2)	0.59–0.67
Androstenone level (entire males)	0.25–0.88

Heritability for a given trait varies depending on the genetic background of each population under analysis. For instance, Suzuki et al. (2006) reported a heritability for meat lightness (L*) of 0.16 in Duroc pigs, while Gjerlaug-Enger et al. (2010) provided a higher estimate ($h^2 = 0.41$) in Landrace swine. Similarly, the heritability of palmitic acid content in the *longissimus dorsi* (LD) muscle ranged from 0.07 in Sutai pigs (Yang et al., 2013) to 0.47 in Duroc pigs (Casellas et al., 2010); and the heritability of muscle stearic acid content was 0.24 in Iberian × Landrace swine (Ramayo-Caldas et al., 2012) and 0.54 in Duroc pigs

(Gjerlaug-Enger et al., 2010). Such heterogeneity was also found when measuring h^2 of traits recorded in different tissues or body locations *i.e* Larzul et al. (1999) found h^2 of 0.03 and 0.23 for L* measured in the *gluteus profundus* and *longissimus* muscles, respectively. The same authors reported that h^2 for pH_{24} measured in 4 different muscles oscillated between 0.17 (*longissimus*) and 0.39 (*biceps femoris*).

4. The search for causal mutations in pigs

4.1. The microsatellite era

A quantitative trait locus (QTL) is a genomic region that contains a polymorphism with significant effects on a quantitative trait. Until 2010, the majority of studies identified QTL by exploiting the existence of linkage disequilibrium between microsatellite markers and causal mutations with effects on traits of economic interest (Dekkers, 2004). If a QTL is linked to a marker locus, then individuals with different marker locus genotypes will have different mean values of the quantitative trait (Mackay et al., 2009). The two main strategies to detect QTL are shown in **Figure 1.**

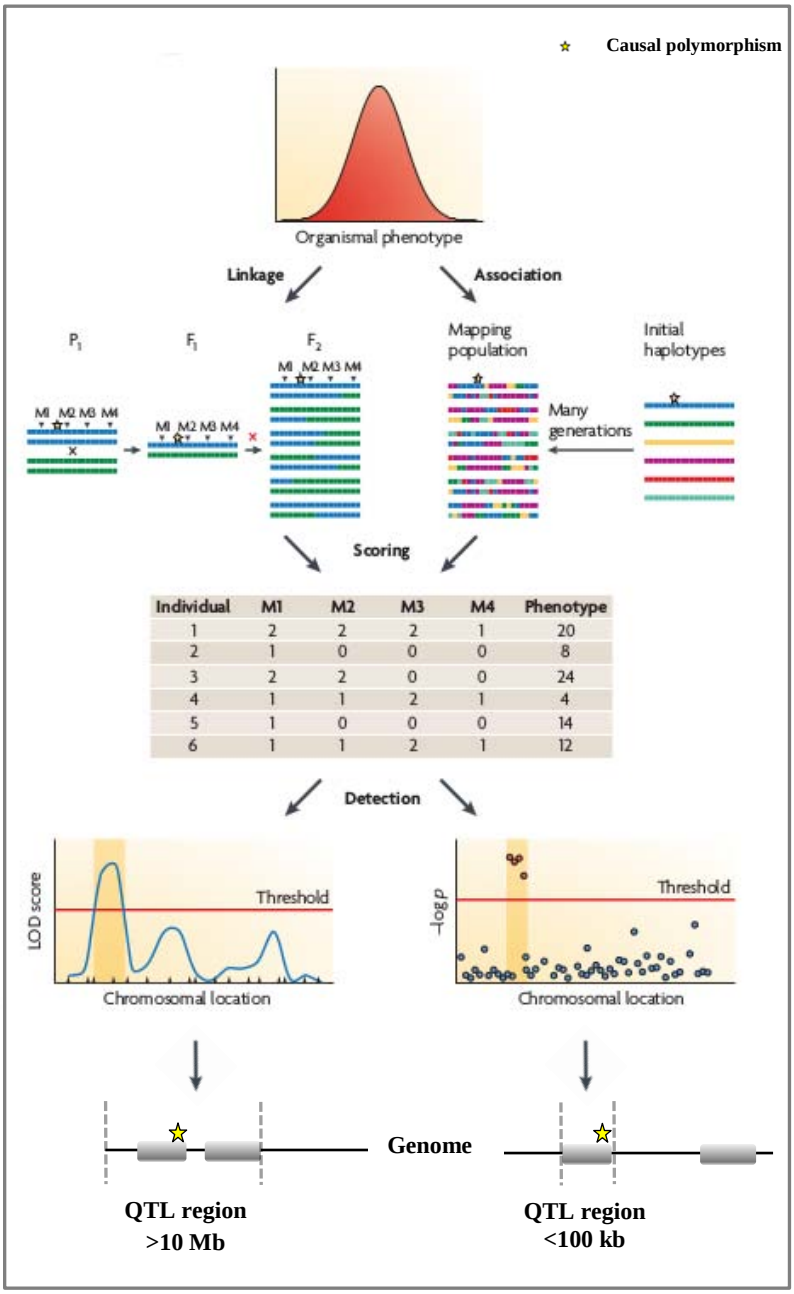


Figure 1. Representation of Linkage (left) and association (right) mapping (Mackay et al., 2009). The yellow star indicates the position of a causal mutation. Both approaches require phenotypic and genotypic information. Linkage mapping: the parental generation (P_1) consists of two genetically divergent inbred lines that are crossed to create the F_1 generation. Crossing individuals from the F_1 generation yields the F_2 mapping population. Recombination in the F_2 population creates new haplotypes and can uncouple marker genotypes from the causal locus. In contrast, association mapping uses historical recombination. In this case, recombination shuffles the initial haplotypes uncoupling all but the most tightly linked markers from the causal locus, thus allowing to map it with precision (reprinted from Nature Reviews Genetics 10, 565-577).

In livestock species, most of QTL studies were performed in either F_2 divergent crosses or half-sib families. In pigs, the first QTL mapping study was performed by Andersson et al. (1994), who reported QTL for backfat thickness and fat deposition in a Wild Boar × Large White cross. This seminal study proved the existence of genetic variation influencing fatness phenotypes in pigs. During the last couple of decades, a total of 16,560 QTLs were reported in the Pig QTL database (Hu et al., 2013) as shown in the **Figure 2**.

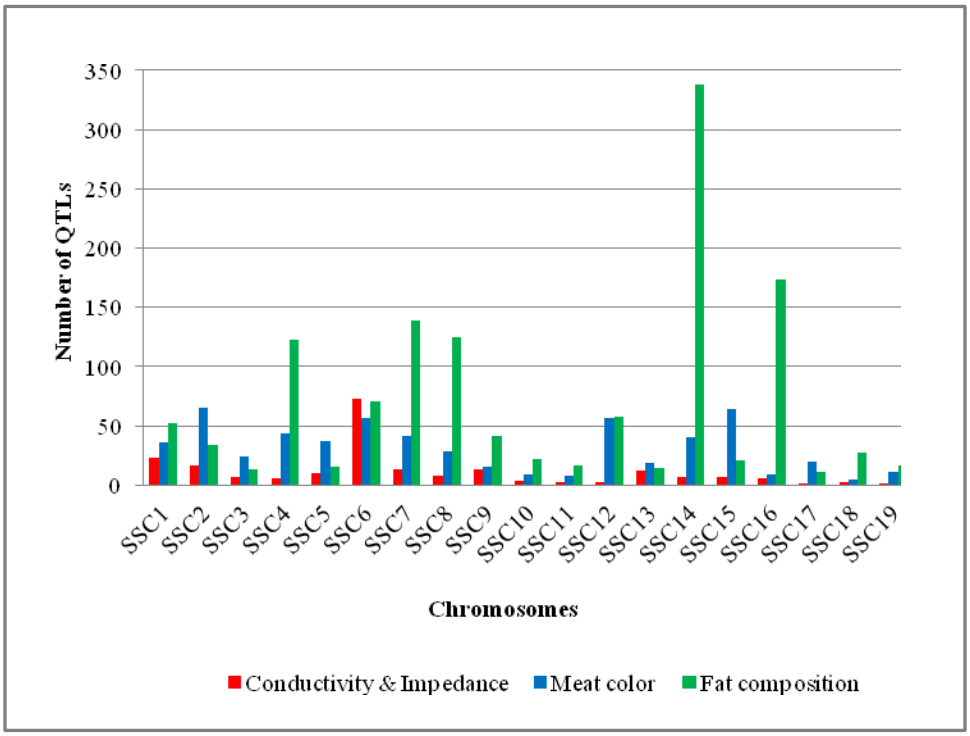


Figure 2. Distribution of reported QTLs for electric conductivity and impedance, fat composition and colour. The X-axis represents pig chromosomes and the Y-axis shows the number of QTLs detected for each trait. All data were retrieved from the Pig QTL database (Hu et al., 2013).

Although these studies resulted in the identification of a large number of QTL, they were flawed by a number of important methodological limitations: 1) the size of the populations was generally small, usually in the range of a few hundred individuals, making difficult the identification of mutations with modest effects, and 2) the number of microsatellites (100-200) was also small, and by this reason the confidence intervals of the QTL were quite large (Nagamine et al., 2003). However, several causal mutations were identified with this

approach *e.g.* a nucleotide substitution at intron 3 of the *IGF2* gene with regulatory effects on its expression (by abrogating the binding of a repressor factor) and with strong impact on muscle growth (Van Laere et al., 2003); and a missense substitution at the *PRKAG3* gene that influences muscle glycogen content, yielding meat with a low ultimate pH and a reduced water-holding capacity (Milan et al., 2000). In other instances, causal mutations were identified without performing previous QTL studies. A well-known example is the halothane gene, whose variation is associated with the porcine stress syndrome and the production of pale, soft and exudative meat (MacLennan et al., 1990). Fujii et al. (1991) sequenced a candidate gene, the ryanodine receptor 1 locus, and found one missense polymorphism that causes the malfunctioning of this molecule and the deterioration of meat quality due to the sharp decline of the ultimate pH, protein denaturing and increased drip loss.

4.2. The sequencing of the pig genome

The sequencing of the pig genome has been one of the main hallmarks of porcine genomics (Groenen, 2016). A large international consortium generated a 2.6 Gigabase draft sequence that was thoroughly annotated, evidencing the existence of 21,640 protein-coding genes, 380 pseudogenes and 2,965 non-coding RNAs (ncRNAs), which is probably an underestimate of the true number of ncRNAs. The evolution of *Sus scrofa* in Eurasia was investigated by sequencing ten wild boars, an experiment that yielded approximately 17 million single nucleotide polymorphisms (SNPs, *i.e.* single-nucleotide substitutions of one base for another that occur in more than one percent of the general population). Analysis of these SNPs demonstrated that Asian wild boars are much more diverse than their European counterparts, a feature explained by the fact that wild boars are originary from Asia as well as because of the occurrence of a strong founder effect in Europe. The split between European and Asian wild boars probably took place 1.6–0.8 Myr ago, but gene flow between the wild and domesticate forms, and between the Asian and European gene pools was quite frequent. Genome sequencing of 16 wild boars and pigs from Europe and Asia also revealed the existence of copy number variation (CNV, chromosomal duplications or

deletions with sizes between 50 bp and several megabases) in the pig genome (Paudel et al., 2013). The average size of the 3,118 CNV found by Paudel et al. (2013) was 13 kb and they comprised about 1% of the pig genome and 545 genes. Copy number variant regions (CNVRs) in pigs were enriched for genes related to sensory perception, neurological process, and response to stimulus, and the majority of them were shared between pigs and wild boars (Paudel et al., 2013). One of the most characterized CNV in pigs is the one including the *KIT* gene, that is closely associated with coat color *i.e.* white individuals have duplicated or triplicated copies of this locus while individuals with a solid coat have just a single copy (Pielberg et al., 2002; Johansson et al., 2005).

4.3. The SNP era

The implementation of next generation sequencing methods making possible to generate millions of SNP markers in a single experiment enhanced the development of porcine high density SNP panels (Ramos et al., 2009) *i.e*. the Illumina Porcine SNP60 BeadChip, with 64,000 SNP markers, and the Axiom chip, that appeared some time later, with 700,000 SNPs (Samorè and Fontanesi, 2016). The SNPs contained within the Porcine SNP60 BeadChip were detected by sequencing pigs belonging to diverse domestic swine breeds (Duroc, Pietrain, Large White and Landrace), and European and Far Eastern wild boars. The average distance between SNPs was 30-40 kb, with gap sizes larger than 250 kb on SSC14 and SSCX (Ramos et al., 2009).

Chips have been widely used not only to genotype SNPs but to detect CNV in pigs. Such studies are based on two metrics: the logged ratio of observed probe intensity to expected intensity (LRR, deviations from zero indicate the existence of a copy number change) and B-allele frequency (BAF) *i.e.* the proportion of hybridized sample that carries the B allele, which should take values of 0.0 (AA), 0.5 (AB), and 1.0 (BB) depending on the genotype of the individual. For instance, a BAF of 0.66 for a locus may provide evidence of the existence of two copies of the B allele and one copy of the A allele. The first study in which SNP arrays were used to detect pig CNVs was carried out by Ramayo-

Caldas et al. (2010), who discovered 49 CNVRs, ranging from 44.7 kb to 10.7 Mb (mean size: 754.6 kb), in a sample of 55 Iberian × Landrace pigs. Many other studies have focused on the description of CNVs in the genome of pigs (Chen et al., 2012; Wang et al., 2012) and even on the association of CNVs with meat quality traits (Wang et al., 2015).

The development of SNP chips was also fundamental to carry out genome-wide association studies (GWAS). The conceptual premises of GWAS are similar to those employed in QTL mapping with microsatellites, but the higher marker density makes possible to narrow QTL regions as well as to discover QTL in genomic regions not well covered with microsatellite markers. Besides, meta-analysis of data generated in different laboratories is easier because the same panel of SNP markers is used everywhere. Depending on the type of trait (categorical or continuous) different statistical methods are used to analyze the data. In this regard, mixed linear models have been increasingly used to carry out GWAS because sample structure (*i.e.* relatedness, geographic structure, etc., can be easily modeled by building a genetic relationship matrix). The contribution of SNPs to phenotypic variance of traits under analysis is estimated using a random-effects model (with or without fixed effects) and its significance is inferred on the basis of association statistics (Yang et al., 2014). Finally, a correction for multiple testing needs to be implemented in order to control the rate of false positives.

In pigs, the genetic basis of multiple traits has been explored by performing GWAS *e.g.* fat deposition (Duijvesteijn et al., 2010), IMF content and composition (Ramayo-Caldas et al., 2012; Hernández-Sánchez et al., 2013; Yang et al., 2013), lipid serum concentrations (Chen et al., 2013; Manunza et al., 2014; Ding et al., 2015; Yang et al., 2015), meat quality traits (Luo et al., 2012; Ma et al., 2013, Sanchez et al., 2014), reproduction (Onteru et al., 2011, 2012; Schneider et al., 2012), and clinical parameters (Boddicker et al., 2012; Fu et al., 2012). By using a GWAS approach, Ma et al. (2014) demonstrated that a splice mutation in the *PHKG1* gene causes a 32 bp deletion and a premature stop codon, events that decrease the catalytic activity of this enzyme thus preventing glycogen breakdown. The muscle accumulation of glycogen decreases the ultimate pH of meat and increases drip loss (Ma et al., 2014).

4.4. Genome-wide association studies for meat quality traits in pigs

The implementation of GWAS studies in pigs made possible to reduce the confidence intervals of QTL previously detected with microsatellite markers as well as to identify thousands of new associations between genomic regions and relevant productive traits. One of the first GWAS published was performed by Duijvesteijn et al., (2010), who reported 37 SNPs, on SSC1 and SSC6, associated with androstenone levels in fat tissue. Many GWAS studies have targeted porcine meat quality traits. For instance, Ma et al. (2013) carried out a GWAS, in Sutai and Duroc × Erhualian F_2 pigs, for meat quality traits recorded in the LD and *semimembranosus* muscles. In this study, the main genome-wide associations mapped to SSC3 (pH_{24}) and SSC15 (drip loss). Similarly, Liu et al. (2015) evaluated multiple meat quality traits in Western Duroc × (Landrace × Yorkshire) and Erhualian pigs. Only a few QTL were shared between these populations, probably reflecting differences in the genetic architecture of meat quality traits between Asian and European breeds. Moreover, QTL on SSC12 and SSC15 showed pleiotropic effects on lightness (L*) of meat, color score, firmness and marbling. Zhang et al. (2015) detected six genomic regions associated with pH, redness (a*) and yellowness (b*) by carrying out a GWAS based on phenotypes and genotypes recorded in 1,943 crossbred commercial pigs. Five genomic regions, located on SSC1, SSC5, SSC9, SSC16 and SSCX, were associated with meat color traits. Interestingly, the SSC15 (133–134 Mb) region has been associated with a*, b*, pH_{24}, shear force and cook loss in many independent studies (Ponsuksili et al., 2014; Bernal Rubio et al., 2015; Liu et al., 2015; Zhang et al., 2015). Interestingly, this SSC15 region contains the protein kinase AMP-activated non-catalytic subunit gamma 3 (*PRKAG3*) gene, whose polymorphism has causal effects on muscle glycogen depletion, a parameter that can have a strong influence on meat quality traits (Milan et al., 2000). Although electrical conductivity has been widely studied by performing QTL scans based on microsatellites (Cepica et al., 2003; Evans et al., 2003; Gallardo et al., 2012), this phenotype has been only included in a few GWAS studies despite being an important predictor of meat quality (Ponsuksili et al., 2014; González-Prendes et al., 2017). Ponsuksili et al. (2014) reported 15 SNPs distributed on SSC1, SSC6, SSC8 and SSC13 which happened to be associated with electrical conductivity measured 24 hours after slaughter in the LD muscle. On SSC6 (54.4 Mb),

there was one SNP with pleiotropic effects on electrical conductivity as well as on pH, impedance and percentage of weight loss in muscle at 24 hours post-mortem.

Several GWAS for IMF content and composition traits have been performed so far. Ramayo-Caldas et al. (2012) genotyped a backcross population (25% Iberian × 75% Landrace) with the Porcine SNP60 BeadChip and detected 813 SNPs, distributed in 43 chromosome regions, displaying significant associations with IMF phenotypes. Particularly interesting was the region detected on SSC8 (**Figure 3**), with pleiotropic effects on palmitic, palmitoleic, palmitoleic/palmitic ratio, oleic/palmitoleic ratio and percentage of saturated fatty acids. In a subsequent study in the same population, Corominas et al. (2013) reported that the *ELOVL6*:c.-533C>T polymorphism may be the causal mutation explaining the associations found. The mechanism of action of this mutation might be related with the methylation status of the *ELOVL6* promoter. More recently, Zhang et al. (2016[a]) reported 26 genome-wide significant QTLs, distributed on eight chromosomes, for eight fatty acids. In this study, one QTL for stearic acid on SSC14 was detected in a crossbred Duroc × (Landrace × Yorkshire) population as well as in a meta-analysis with five different populations. This SSC14 region has been also identified in other studies as significantly associated with muscle FA composition traits (Yang et al., 2013; Ros-Freixedes et al., 2016; Zhang et al., 2016[a], 2016[b]). One SNP (g.2228T>C) at the promoter of the stearoyl-CoA desaturase (*SCD*) gene, which catalyses the D9-*cis* desaturation of a range of fatty acyl-CoA substrates (Paton and Ntambi, 2009) and maps to SSC14 (120.9 Mb), has been shown to have causal effects on muscle stearic and oleic contents in Duroc pigs (Estany et al., 2014).

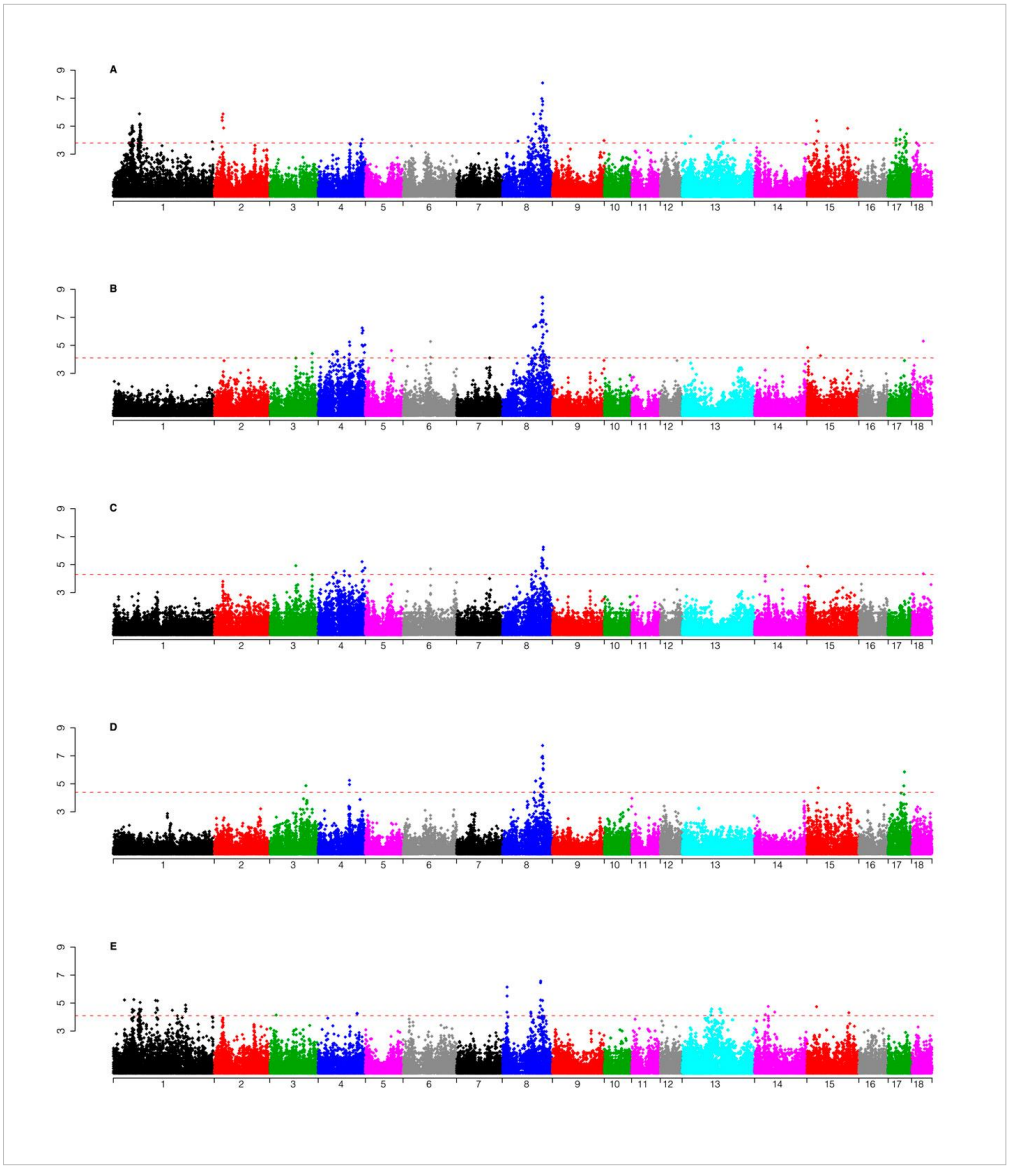


Figure 3. Manhattan plot of a GWAS (Ramayo-Caldas et al., 2012) for A) palmitic acid (C16:0), B) palmitoleic acid (C16:1 n-7), C) ratios of C16:1(n-7)/C16:0, D) ratios of C18:1(n-7)/C16:1(n-7), and E) saturated fatty acids. The x-axis represents the chromosomes containing the QTLs and the y-axis shows the $-\log_{10}$ (*P*-value) of the reported associations. The horizontal line indicates the threshold of significance (q-value ≤ 0.05).

4.5. Dissecting the regulatory basis of gene expression and complex phenotypes

The development of the microarray technology represented an important advance in animal genomics, providing a high throughput tool to characterize the transcriptome of multiples tissues (Pena et al., 2014). Several studies have focused on the evaluation of transcriptome variability in pigs with extreme phenotypes (Ponsuksili et al., 2008; Liu et al., 2009; Cánovas et al., 2010; Hamill et al., 2012; Pena et al., 2016). In this way, the analysis of gene expression in the *gluteus medius* and/or *longissimus dorsis* muscles from pigs with divergent lipid phenotypes revealed that fatter animals had higher mRNA levels of both lipogenic and lipolytic enzymes (Liu et al., 2009; Cánovas et al., 2010; Hamill et al., 2012; Pena et al., 2016). Similarly, Ponsuksili et al. (2008) detected that pigs with lower water-holding capacity have reduced expression of lipid metabolism genes.

The combination of high throughput genotyping and microarray data has made possible to map the genetic determinants of gene expression, the so-called expression QTL (eQTL). An eQTL is a genomic region whose variation partly influences the genetic variance of a gene expression phenotype (Nica and Dermitzakis, 2013). In general, eQTL can be classified as *cis*-acting (or *cis*-eQTL), when they map close (± 1 Mb, though this distance is arbitrary) to the regulated gene, or *trans*-acting (or *trans*-eQTL) when they do not. The two types of eQTLs are shown in **Figure 4.**

Since eQTLs may contain variants with regulatory effects on both gene expression and phenotypic variation, they can provide valuable information about candidate genes to be further investigated (Nicolae et al., 2010; Nica and Dermitzakis, 2013; Torres et al., 2014). One of the first porcine eQTL maps was reported by Ponsuksili et al. (2010). These authors carried out an eQTL analysis in pigs with divergent phenotypes for a combination of technological traits such as pH, conductivity and color. A total of 9,180 eQTLs were reported, of which 653 were classified as putative *cis*-acting. In another study, Ponsuksili et al. (2014) mapped meat quality QTL to SSC4 and SSC6 and identified SNPs associated with the expression of genes mapping to QTL regions *i.e. ZNF704* (SSC4) and *PIH1D1*, *SIGLEC10, TBCB*, *LOC100518735*, *KIF1B*, *LOC100514845* (SSC6). The expression

values of these genes were significantly correlated with meat quality traits as pH, colour and electric conductivity.

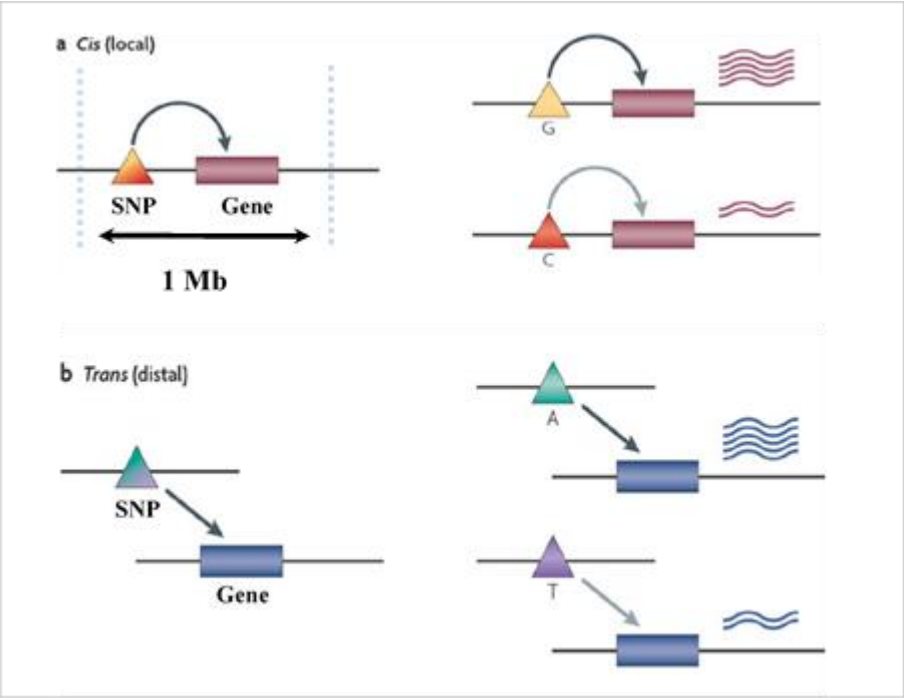


Figure 4. Representation of a) *cis*- and b) *trans*-acting regulating the gene expression (Cheung and Spielman, 2009). In the a) *cis*-eQTL the expression value of a gene close to the casual variant is incremented with the presence of G allele. For the b) *trans*-eQTLs the causal polymorphism is far away from the target gene that is influenced by the effect of the A allele (Reprinted from: Nature Reviews Genetics 10, 595–604).

In another study, Steibel et al. (2011) detected two eQTLs for the *AKR7A2* and *TXNDC12* genes, that are involved in lipid metabolism and map to marbling, intramuscular fat content and loin muscle area QTL. Muñoz et al. (2013) also combined QTL and eQTL mapping to identify candidate genes with potential effects on backfat thickness and intramuscular fat composition. By doing so, they proposed as candidate gene the *ELOVL6* locus that plays a key role in fatty acid metabolism catalyzing the elongation of fatty acid

with 12–16 carbons to C18 (Jakobsson et al., 2006). Subsequently, Corominas et al. (2013) demonstrated that a mutation at the promoter of this gene may have causal effects on the variation of C16:0 and C16:1 (n-7) fatty acids.

5- Using genomic information for the genetic improvement of pigs

One of the potential practical applications of porcine GWAS studies would be to identify causal mutations that can be used as selection criteria in breeding schemes. The success of such strategy has been limited because resource populations are usually small, the density of the Porcine SNP60 BeadChip is low and also because the intrinsic difficulty of distinguishing a causal mutation from a nearby marker. However, there have been stories of success that support the validity of the GWAS approach to detect causal mutations *e.g.* a major QTL for glycolytic potential was detected on SSC3 by Ma et al. (2014), and the subsequent refinement of the QTL and the sequencing of the *PHKG1* gene made possible to demonstrate that a mutation at a splice site has causal effects on glycolytic potential and water-holding capacity.

Another consequence of the invention of the Porcine SNP60 BeadChip has been the implementation of genomic selection in porcine production. Genomic selection consists in estimating genomic breeding values using recorded phenotypic information and a large number of markers spread across the whole genome in a reference population (Meuwissen and Goddard, 2000; Ibañez-Escriche and Gonzalez-Recio, 2011). In genomic selection, markers can be included in the genomic evaluation without pre-selection based on QTL significance and location. Genomic selection allows the reduction of the generation interval because breeding values can be estimated very early in the productive life of animals (Hayes et al., 2009). Although the generation interval is short in pigs, the implementation of genomic selection could be useful to improve economically important traits with low heritabilities or that are difficult or expensive to measure. A thorough review about the current status and perspectives of genomic selection in pigs has been elaborated by Samorè and Fontanesi, (2016). According to these authors, the progress in sequencing methods will

make possible to decrease genotyping costs and, in parallel, to increase the amount of genomic information produced, thus generating prediction equations that are more accurate and stable across time (Samorè and Fontanesi, 2016).

Chapter II. Objectives

General goal:

The main objective of this thesis was to identify genomic regions and candidate genes associated with meat quality and gene expression traits in a commercial line of Duroc pigs. This research forms part of the project “Study of traits related with lipid metabolism and pork quality using high throughput sequences and gene expression" (AGL2010-22208-C02-02).

Specific goals:

1- To identify regions of the pig genome associated with the phenotypic variation of ultimate pH, electric conductivity, color (a*, b*, and L*) and intramuscular fat content and composition traits recorded in two different muscles (*longissimus dorsi* and *gluteus medius*) from Duroc pigs

2- To investigate if genomic regions associated with such traits contain expression quantitative trait loci (eQTL) regulating the expression of genes with a potential impact on muscle physiology and metabolism.

3- To compare the eQTL landscape of two tissues, skeletal muscle and liver, with central roles on energy homeostasis and to ascertain if eQTL co-localize with copy number variants.

References

Adzitey, F., Nurul, H., 2011. Pale soft exudative (PSE) and dark firm dry (DFD) meats: causes and measures to reduce these incidences-a mini review. Int. Food Res. J. 18, 11–20.

Andersson, L., Haley, C.S., Ellegren, H., Knott, S.A., Johansson, M., Andersson, K., Andersson-Eklund, L., et al, 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263, 1771–1774.

Astrup, A., Dyerberg, J., Elwood, P., Hermansen, K., Hu, F.B., Jakobsen, M.U., Kok, F.J., et al., 2011. The role of reducing intakes of saturated fat in the prevention of cardiovascular disease: where does the evidence stand in 2010? Am. J. Clin. Nutr. 93, 684–688.

Bernal Rubio, Y.L., Gualdrón Duarte, J.L., Bates, R.O., Ernst, C.W., Nonneman, D., Rohrer, G.A., King, D.A., et al., 2015. Implementing meta-analysis from genome-wide association studies for pork quality traits. J. Anim. Sci. 93, 5607–5617.

Biesalski, H.K., 2005. Meat as a component of a healthy diet - are there any risks or benefits if meat is avoided in the diet? Meat Sci. 70, 509–524.

Blanchard, P.J., Willis, M.B., Warkup, C.C., Ellis, M., 2000. The influence of carcass backfat and intramuscular fat level on pork eating quality. J Sci Food Agric. 80, 145–151.

Boddicker, N., Waide, E.H., Rowland, R.R., Lunney, J.K., Garrick, D.J., Reecy, J.M., Dekkers, J.C., 2012. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. Anim. Sci. 90, 1733–1746.

Brewer, M.S., Zhu, L.G., McKeith, F.K., 2001. Marbling effects on quality characteristics of pork loin chops: consumer purchase intent, visual and sensory characteristics. Meat Sci. 59, 153–163.

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

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. J. Anim. Sci. 88, 2246–2254.

Cepica, A., Stratil, M., Kopecny, P., Blazkova, J., Schröffel, R., Davoli, L., Fontanesi, G., et al., 2003. Linkage and QTL mapping for Sus scrofa chromosome 5. J. Anim. Breed. Genet. 120, 38–44.

Chen, C., Qiao, R., Wei, R., Guo, Y., Ai, H., Ma, J., Ren, J., et al., 2012. 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, C., Yang, B., Zeng, Z., Yang, H., Liu, C., Ren, J., Huang, L., 2013. Genetic dissection of blood lipid traits by integrating genome-wide association study and gene expression profiling in a porcine model. BMC Genomics 14, 848.

Cheung, V.G., Spielman, R.S., 2009. Genetics of human gene expression: mapping DNA variants that influence gene expression. Nat. Rev. Genet. 10, 595–604.

Chizzolini, R., Novelli, E., Zanardi, E., 1998. Oxidation in traditional mediterranean meat products. Meat Sci. 49, S87-99.

Ciobanu, D.C., Lonergan, S.M., Huff-Lonergan, E.J., 2011. Genetics of meat quality and carcass traits. in: Rothschild, M.F., Ruvinsky, A. (2dn.), The Genetics of the Pig. CABI, Wallingford, pp. 355–389.

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+.

Dekkers, J.C., 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. J. Anim. Sci. 82, E313-328.

Ding, Y.Y., Zhang, W., Zhang, M.Q., Fu, K., Chen, W.P., Ding, C., He, X.L., et al., 2015. Functional and association studies of the cholesteryl ester transfer protein (*CETP*) gene in a Wannan Black pig model. Anim. Genet. 46, 702–706.

Duijvesteijn, N., Knol, E.F., Merks, J.W, Crooijmans, R.P, Groenen, M.A., 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 Genet. 11, 42.

Eckel, R.H., Jakicic, J.M., Ard, J.D., Jesus, J.M. de, Miller, N.H., Hubbard, V.S., Lee, I.M., et al., 2014. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk. Circulation 129, S76–99.

Estany, J., Ros-Freixedes, R., Tor, M., Pena, R.N., 2014. A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. PloS One 9, e86177.

Evans, G.J., Giuffra, E., Sanchez, A., Kerje, S., Davalos, G., Vidal, O., Illán, S., et al., 2003. Identification of quantitative trait loci for production traits in commercial pig populations. Genetics 164, 621–627.9

Falconer, D.S., MacKay, T.F., 1996. Introduction to quantitative genetics, 4th ed. Longmans Green, Harlow, Essex, UK

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

Fu, W.X., Liu, Y., Lu, X., Niu, X.Y., Ding, X.D., Liu, J.F., Zhang, Q., 2012. A genome-wide association study identifies two novel promising candidate genes affecting escherichia coli F4ab/F4ac susceptibility in swine. PloS One 7, e32127+.

Fujii, J., Otsu, K., Zorzato, F., de Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., et al., 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. Science 253, 448–451.

Gallardo, D., Pena, R.N., Quintanilla, R., Ramírez, O., Almuzara, D., Noguera, J.L., Amills, M., 2012. Quantitative trait loci analysis of a Duroc commercial population highlights differences in the genetic determination of meat quality traits at two different muscles. Anim. Genet. 43, 800–804.

Gjerlaug-Enger, E., Aass, L., Odegård, J., Vangen, O., 2010. Genetic parameters of meat quality traits in two pig breeds measured by rapid methods. Animal. 4, 1832–1843.

González-Prendes, R., Quintanilla, R., Cánovas, A., Manunza, A., Figueiredo-Cardoso, T., Jordana, J., Noguera, J.L., et al., 2017. Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. Sci. Rep. 7, 39830.

Groenen, M.A., 2016. A decade of pig genome sequencing: a window on pig domestication and evolution. Genet Sel Evol. 48, 23.

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 F_2 population. Anim. Genet. 40, 185–191.

Hamill, R.M., McBryan, J., McGee, C., Mullen, A.M., Sweeney, T., Talbot, A., Cairns, M.T., et al., 2012. Functional analysis of muscle gene expression profiles associated with tenderness and intramuscular fat content in pork. Meat Sci. 92, 440–450.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., Goddard, M.E., 2009. Invited review: Genomic selection in dairy cattle: progress and challenges. J. Dairy Sci. 92, 433–443.

Hernández-Sánchez, J., Amills, M., Pena, R.N., Mercadé, A., Manunza, A., Quintanilla, R., 2013. Genomic architecture of heritability and genetic correlations for intramuscular and back fat contents in Duroc pigs. J. Anim. Sci. 91, 623–632.

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 Res. 41, D871–D879.

Ibañez-Escriche, N., Gonzalez-Recio, O., 2011. Review. Promises, pitfalls and challenges of genomic selection in breeding programs. Span. J. Agric. Res. 9, 404-413.

Jakobsson, A., Westerberg, R., Jacobsson, A., 2006. Fatty acid elongases in mammals: their regulation and roles in metabolism. Prog. Lipid Res. 45, 237–249.

Johansson, A., Pielberg, G., Andersson, L., Edfors-Lilja, I., 2005. Polymorphism at the porcine Dominant white/KIT locus influence coat colour and peripheral blood cell measures. Anim. Genet. 36, 288–296.

Keys, A., Menotti, A., Karvonen, M.J., Aravanis, C., Blackburn, H., Buzina, R., Djordjevic, B.S., et al., 1986. The diet and 15-year death rate in the seven countries study. Am. J. Epidemiol. 124, 903–915.

Larzul, C., Roy, P., Gogué, J., Talmant, A., Jacquet, B., Lefaucheur, L., Ecolan, P., et al., 1999. Selection for reduced muscle glycolytic potential in Large White pigs. II. Correlated responses in meat quality and muscle compositional traits. Genet. Sel. Evol. 31, 1–16.

Lee, S., Norman, J.M., Gunasekaran, S., van Laack, R.L., Kim, B.C., Kauffman, R.G., 2000. Use of electrical conductivity to predict water-holding capacity in post-rigor pork. Meat Sci. 55, 385–389.

Liu, J., Damon, M., Guitton, N., Guisle, I., Ecolan, P., Vincent, A., Cherel, P., et al., 2009. Differentially-expressed genes in pig *longissimus* muscles with contrasting levels of fat, as identified by combined transcriptomic, reverse transcription PCR, and proteomic analyses. J. Agric. Food Chem. 57, 3808–3817.

Liu, X., Xiong, X., Yang, J., Zhou, L., Yang, B., Ai, H., Ma, H., et al., 2015. Genome-wide association analyses for meat quality traits in Chinese Erhualian pigs and a Western Duroc × (Landrace × Yorkshire) commercial population. Genet. Sel. Evol. 47, 44.

Luo, W., Cheng, D., Chen, S., Wang, L., Li, Y., Ma, X., Song, X., Liu, X., et al., 2012. Genome-wide association analysis of meat quality traits in a porcine Large White × Minzhu intercross population. Int. J. Biol. Sci. 8, 580–595.

Ma, J., Yang, J., Zhou, L., Ren, J., Liu, X., Zhang, H., Yang, B., et al., 2014. A splice mutation in the *PHKG1* gene causes high glycogen content and low meat quality in pig skeletal muscle. PLoS Genet 10, e1004710.

Ma, J., Yang, J., Zhou, L., Zhang, Z., Ma, H., Xie, X., Zhang, F., et al., 2013. Genome-wide association study of meat quality traits in a White Duroc×Erhualian F2 intercross and chinese sutai pigs. PloS One 8, e64047.

Mackay, T.F., Stone, E.A., Ayroles, J.F., 2009. The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. 10, 565–577.

MacLennan, D.H., Duff, C., Zorzato, F., Fujii, J., Phillips, M., Korneluk, R.G., Frodis, W., et al., 1990. Ryanodine receptor gene is a candidate for predisposition to malignant hyperthermia. Nature 343, 559–561.

Maltin, C., Balcerzak, D., 2003. Determinants of meat quality: tenderness. Proc. Nutr. Soc. 62, 337–347.

Mancini, R.A., Hunt, M.C., 2005. Current research in meat color. Meat Sci. 71, 100–121.

Manunza, A., Casellas, J., Quintanilla, R., González-Prendes, R., Pena, R.N., Tibau, J., Mercadé, A., et al., 2014. A genome-wide association analysis for porcine serum lipid traits reveals the existence of age-specific genetic determinants. BMC Genomics 15, 758.

Meuwissen, T.H, Goddard, M.E., 2000. Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. Genetics 155, 421–430.

Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., et al., 2000. A mutation in *PRKAG3* associated with excess glycogen content in pig skeletal muscle. Science 288, 1248–1251.

Muñoz, M., Rodríguez, M.C., Alves, E., Folch, J.M., Ibañez-Escriche, N., Silió, L., Fernández, A.I., 2013. Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. BMC Genomics 14, 845.

Nagamine, Y., Haley, C.S., Sewalem, A., Visscher, P.M., 2003. Quantitative trait loci variation for growth and obesity between and within lines of pigs (Sus scrofa). Genetics 164, 629–635.

Nica, A.C., Dermitzakis, E.T., 2013. Expression quantitative trait loci: present and future. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 20120362.

Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Dolan, M.E., Cox, N.J., 2010. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet. 6, e1000888.

OECD-FAO., 2016, OECD-FAO Agricultural Outlook 2016-2025, OECD Publishing, Paris. http://dx.doi.org/10.1787/agr_outlook-2016-en. Accessed 27 Feb. 2017.

Onteru, S.K., Fan, B., Du, Z.Q., Garrick, D.J., Stalder, K.J., Rothschild, M.F., 2012. A whole-genome association study for pig reproductive traits. Anim. Genet. 43, 18–26.

Onteru, S.K., Fan, B., Nikkilä, M.T., Garrick, D.J., Stalder, K.J., Rothschild, M.F., 2011. Whole-genome association analyses for lifetime reproductive traits in the pig. J. Anim. Sci. 89, 988–995.

Paddon-Jones, D., Westman, E., Mattes, R.D., Wolfe, R.R., Astrup, A., Westerterp-Plantenga, M., 2008. Protein, weight management, and satiety. Am. J. Clin. Nutr. 87, 1558S–1561S.

Paton, C.M., Ntambi, J.M., 2009. Biochemical and physiological function of stearoyl-CoA desaturase. Am. J. Physiol. Endocrinol. Metab. 297, E28-37.

Paudel, Y., Madsen, O., Megens, H.J., Frantz, L.A., Bosse, M., Bastiaansen, J.W., Crooijmans, R.P., et al., 2013. Evolutionary dynamics of copy number variation in pig genomes in the context of adaptation and domestication. BMC Genomics 14, 449.

Pena, R.N., Quintanilla, R., Manunza, A., Gallardo, D., Casellas, J., Amills, M., 2014. Application of the microarray technology to the transcriptional analysis of muscle phenotypes in pigs. Anim. Genet. 45, 311–321.

Pena, R.N., Ros-Freixedes, R., Tor, M., Estany, J., 2016. Genetic marker discovery in complex traits: a field example on fat content and composition in pigs. Int. J. Mol. Sci. 17.

Pielberg, G., Olsson, C., Syvänen, A.C., Andersson, L., 2002. Unexpectedly high allelic diversity at the *KIT* locus causing dominant white color in the domestic pig. Genetics 160, 305–311.

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

Ponsuksili, S., Murani, E., Schwerin, M., Schellander, K., Wimmers, K., 2010. Identification of expression QTL (eQTL) of genes expressed in porcine M. *longissimus dorsi* and associated with meat quality traits. BMC Genomics 11, 572.

Ponsuksili, S., Murani, E., Trakooljul, N., Schwerin, M., Wimmers, K., 2014. Discovery of candidate genes for muscle traits based on GWAS supported by eQTL-analysis. Int. J. Biol. Sci. 10, 327–337.

Ramayo-Caldas, Y., Castelló, A., Pena, R.N., Alves, E., Mercadé, A., Souza, C.A., Fernández, A.I., et al., 2010. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. BMC Genomics 11, 593.

Ramayo-Caldas, Y., Mercadé, A., Castelló, A., Yang, B., Rodríguez, C., Alves, E., Díaz, I., et al., 2012. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. Anim. Sci. 90, 2883–2893.

Ramos, A.M., Crooijmans, R.P., Affara, N.A., Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., 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, e6524.

Rincker, P.J., Killefer, J., Ellis, M., Brewer, M.S., McKeith, F.K., 2008. Intramuscular fat content has little influence on the eating quality of fresh pork loin chops. J. Anim. Sci. 86, 730–737.

Rosenvold, K., Andersen, H.J., 2003. Factors of significance for pork quality-a review. Meat Sci. 64, 219–237.

Ros-Freixedes, R., Gol, S., Pena, R.N., Tor, M., Ibáñez-Escriche, N., Dekkers, J.C., Estany, J., 2016. Genome-wide association study singles out *SCD* and *LEPR* as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. PloS One 11, e0152496.

Samorè, A.B., Fontanesi, L., 2016. Genomic selection in pigs: state of the art and perspectives. Ital. J. Anim. Sci. 15, 211–232.

Sanchez, M.P., Tribout, T., Iannuccelli, N., Bouffaud, M., Servin, B., Tenghe, A., Dehais, P., et al., 2014. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. Genet. Sel. Evol. 46, 12.

Schneider, J.F., Rempel, L.A., Rohrer, G.A., 2012. Genome-wide association study of swine farrowing traits. Part I: genetic and genomic parameter estimates. J. Anim. Sci. 90, 3353–3359.

Steibel, J.P., Bates, R.O., Rosa, G.J., Tempelman, R.J., Rilington, V.D., Ragavendran, A., Raney, N.E., et al., 2011. Genome-wide linkage analysis of global gene expression in loin muscle tissue identifies candidate genes in pigs. PloS One 6, e16766.

Suzuki, K., Ishida, M., Kadowaki, H., Shibata, T., Uchida, H., Nishida, A., 2006. Genetic correlations among fatty acid compositions in different sites of fat tissues, meat production, and meat quality traits in Duroc pigs. J. Anim. Sci. 84, 2026–2034.

Torres, J.M., Gamazon, E.R., Parra, E.J., Below, J.E., Valladares-Salgado, A., Wacher, N., Cruz, M., et al., 2014. Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. Am. J. Hum. Genet. 95, 521–534.

USDA, United States Department of Agriculture, 2016 .USDA Foreign Agricultural Service: Livestock and poultry: world markets and trade report. Available from:

https://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf. Accessed: 28 Mar. 2017.

Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., et al., 2003. A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. Nature 425, 832–836.

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

Wang, L., Xu, L., Liu, X., Zhang, T., Li, N., Hay, E.H., Zhang, Y., et al., 2015. Copy number variation-based genome wide association study reveals additional variants contributing to meat quality in swine. Sci. Rep. 5, 12535.

Wood, J.D., Enser, M., Fisher, A.V., Nute, G.R., Sheard, P.R., Richardson, R.I., Hughes, S.I., et al., 2008. Fat deposition, fatty acid composition and meat quality: A review. Meat Sci. 78, 343–358.

Wynder, E.L., Cohen, L.A., Winters, B.L., 1997. The challenges of assessing fat intake in cancer research investigations. J. Am. Diet. Assoc. 97 :S5-8.

Yang, B., Zhang, W., Zhang, Z., Fan, Y., Xie, X., Ai, H., Ma, J., et al., 2013. Genome-wide association analyses for fatty acid composition in porcine muscle and abdominal fat tissues. PloS One 8, e65554+.

Yang, H., Huang, X., Zeng, Z., Zhang, W., Liu, C., Fang, S., Huang, L., et al., 2015. Genome-wide association analysis for blood lipid traits measured in three pig populations reveals a substantial level of genetic heterogeneity. PloS One 10, e0131667.

Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., Price, A.L., 2014. Advantages and pitfalls in the application of mixed model association methods. Nat. Genet. 46, 100–106.

Zhang, C., Wang, Z., Bruce, H., Kemp, R.A., Charagu, P., Miar, Y., Yang, T., et al., 2015. Genome-wide association studies (GWAS) identify a QTL close to *PRKAG3* affecting meat pH and colour in crossbred commercial pigs. BMC Genet. 16, 33.

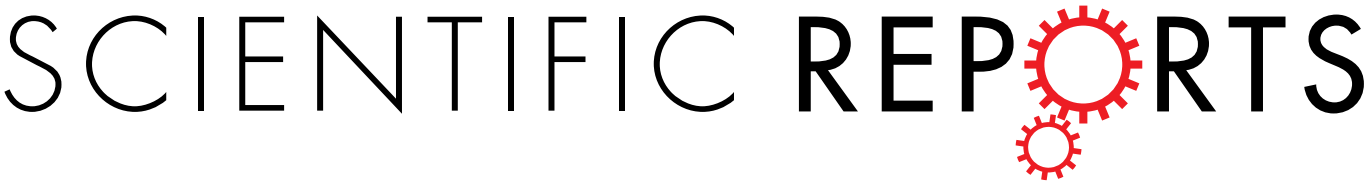
Zhang, W., Yang, B., Zhang, J., Cui, L., Ma, J., Chen, C., Ai, H., et al., 2016[a]. Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. Sci. Rep. 6, 24718.

Zhang, W., Zhang, J., Cui, L., Ma, J., Chen, C., Ai, H., Xie, X., et al., 2016[b]. Genetic architecture of fatty acid composition in the *longissimus dorsi* muscle revealed by genome-wide association studies on diverse pig populations. Genet. Sel. Evol. 48, 5.

Chapter III.

Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits

Rayner González-Prendes, Raquel Quintanilla, Angela Cánovas, Arianna Manunza,Tainã Figueiredo Cardoso, Jordi Jordana, José Luis Noguera, Ramona N. Pena &Marcel Amills (2017). *Scientific Reports* 7, 39830.



OPEN

Received: 10 June 2016
Accepted: 29 November 2016
Published: 05 January 2017

Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits

Rayner González-Prendes[1], Raquel Quintanilla[2], Angela Cánovas[1], Arianna Manunza[1], Tainã Figueiredo Cardoso[1,3], Jordi Jordana[4], José Luis Noguera[2], Ramona N. Pena[5] & Marcel Amills[1]

Meat quality traits have an increasing importance in the pig industry because of their strong impact on consumer acceptance. Herewith, we have combined phenotypic and microarray expression data to map loci with potential effects on five meat quality traits recorded in the *longissimus dorsi* (LD) and *gluteus medius* (GM) muscles of 350 Duroc pigs, *i.e.* pH at 24 hours post-mortem (pH_{24}), electric conductivity (CE) and muscle redness (a*), lightness (L*) and yellowness (b*). We have found significant genome-wide associations for CE of LD on SSC4 (~104 Mb), SSC5 (~15 Mb) and SSC13 (~137 Mb), while several additional regions were significantly associated with meat quality traits at the chromosome-wide level. There was a low positional concordance between the associations found for LD and GM traits, a feature that reflects the existence of differences in the genetic determinism of meat quality phenotypes in these two muscles. The performance of an eQTL search for SNPs mapping to the regions associated with meat quality traits demonstrated that the GM a* SSC3 and pH_{24} SSC17 QTL display positional concordance with cis-eQTL regulating the expression of several genes with a potential role on muscle metabolism.

The physicochemical properties of the porcine muscle and its post-mortem maturation determine the organoleptic properties of fresh meat and cured products and, consequently, their acceptance by consumers[1]. The genetic determinism of electrical conductivity, acidity and color, which have been often used as predictors of meat quality, has been explored by performing genome-wide association studies (GWAS) in F_2 populations[2–4] as well as in purebred pigs[5,6]. An important limitation of using F_2 intercrosses in GWAS studies is that they are not representative of the purebred populations that constitute the selection nuclei of breeding companies. On the other hand, certain breeds, such as Large White, have been strongly introgressed with Asian alleles that do not segregate in other European porcine populations[7].

In a previous study, we measured electrical conductivity at 24 hours (CE), pH at 24 hours (pH_{24}) and color (lightness or L*, redness or a*, and yellowness or b*) in *gluteus medius* (GM) and *longissimus dorsi* (LD) samples from 350 Duroc pigs (Lipgen population)[8]. Performance of a genome scan with 105 microsatellites revealed that the QTL maps for these two muscles were quite different[8]. Indeed, the only QTL that remained significant at the genome-wide level were those associated with GM a*, on *Sus scrofa* chromosome 13 (SSC13, 84 cM), and GM b* (SSC15, 108 cM). Unfortunately, the confidence intervals of these QTL were quite large due to the poor resolution of the microsatellite-based analysis. Moreover, we may have missed many QTL due to the relatively large spacing between markers. In the current work, we aimed to circumvent these limitations by employing a GWAS approach to identify meat quality QTL in the Lipgen population mentioned above. Taking advantage that microarray measurements of gene expression in the GM muscle were available for 104 Lipgen pigs, we have

[1]Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Campus Universitat Autònoma de Barcelona, Bellaterra 08193, Spain. [2]Animal Breeding and Genetics Program, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Torre Marimon, Caldes de Montbui 08140, Spain. [3]CAPES Foundation, Ministry of Education of Brazil, Brasilia D. F., Zip Code 70.040-020, Brazil. [4]Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain. [5]Department of Animal Science, University of Lleida - Agrotecnio Center, Lleida 25198, Spain. Correspondence and requests for materials should be addressed to M.A. (email: marcel.amills@uab.cat)

performed an additional analysis where we have investigated the co-localization between GM QTL and expression QTL in *cis* (cis-eQTL).

Materials and Methods

Ethics approval. The manipulation of Duroc pigs followed Spanish national guidelines and it was approved by the Ethical Committee of Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Measurement of phenotypic and expression data. Phenotypic records were collected in a commercial Duroc line of 350 barrows distributed in five half-sib families (Lipgen population). A detailed description of the management conditions of this commercial line has been previously reported[9]. Meat quality analyses were performed 24 h after slaughter at the IRTA-Centre of Food Technology by using 200 g samples of the LD and GM muscles. Electrical conductivity was estimated with a Pork Quality Meter (Intek GmbH) while pH_{24} was measured with a pH-meter equipment with a Xerolyte electrode (Crison). Meat L*, a* and b* color parameters were determined with a Minolta Chroma-Meter CR-200 (Konica Minolta) equipment (light source C and aperture 2). Microarray expression data of GM samples from 104 Duroc pigs were obtained in a previous study (data can be found in the Gene Expression Omnibus public repository, accession number: GSE19275) based on the use of GeneChip Porcine Genomic arrays (Affymetrix, Inc., Santa Clara, CA)[10]. A detailed description of the techniques and methods used to perform the RNA purification and microarray hybridization steps can be found in Canovas *et al.*[10]. Briefly, GM samples from 104 pigs were grinded in liquid nitrogen and homogenized with a mechanical rotor. Total RNA was purified with an acid phenol protocol[11] and it was subsequently used as a template to synthesize double stranded cDNA with the One Cycle cDNA Synthesis Kit (Affymetrix, Inc.). cRNAs were purified with the GeneChip Sample Cleanup Module (Affymetrix, Inc.), fragmented and added to a hybridisation cocktail[10]. The GeneChip Porcine Genome Array was equilibrated to room temperature and prehybridised with 1× hybridisation buffer at 45 °C for 10 min[10]. The hybridisation cocktail was heated to 99 °C for 5 min in a heat block and cooled to 45 °C for 5 min. Subsequently, a hybridization step was carried out at 45 °C for 16 hours. GeneChips were washed and labeled with streptavidin phycoerythrin in a Fluidics Station 450 (Affymetrix, Inc) and they were scanned in an Agilent G3000 GeneArray Scanner (Agilent Technologies, Inc.). The "Affy" and "Sympleaffy" packages from the Bioconductor project[12] were employed to establish a set of quality control metrics to assess the quality of RNA samples and the efficiencies of the labelling and hybridisation steps. Data pre-processing and normalization were carried out with the BRB-ArrayTools software version 3.7.1[13]. Genes displaying more than 20% of expression values over ±1.5 times the median expression of all arrays were retained for further analysis.

Genome-wide association analysis for meat quality and expression data. Genotyping was performed with the Porcine SNP60 BeadChip (Illumina, San Diego, CA) which contains 62,163 single nucleotide polymorphisms (SNPs). Quality genotyping analyses were carried out with the GenomeStudio software (Illumina), as previously reported[14]. We removed SNPs (a) mapping to the X chromosome, (b) with a rate of missing genotypes higher than 5%, (c) that did not conform Hardy-Weinberg expectations (threshold set at a P-value ≤ 0.001), (d) that had a minor allele frequency below 0.05, (e) that had a GenCall score < 0.15, (f) that had a call rate $< 95\%$ or (g) that could not be mapped to the pig genome (*Sus scrofa* 10.2 assembly). After filtering the raw data, a GWAS was carried out with 36,710 SNPs. Single-SNP association analyses were performed with the Genome-wide Efficient Mixed-Model Association (GEMMA) software[15] under an additive genetic model that included the genomic kinship matrix to account for relatedness. The statistical model assumed in this analysis was:

$$\boldsymbol{y}_{ijklm} = \boldsymbol{\mu} + \boldsymbol{batch}_j + \boldsymbol{\beta}\,\boldsymbol{weight}_k + \boldsymbol{\delta}\boldsymbol{g}_l + \boldsymbol{e}_{ijklm} \quad (1)$$

where $\boldsymbol{y}_{ijklm}$ is the vector of phenotypic observations *i.e.* pH_{24}, CE, L*, a* and b* measured at the GM and LD muscles of the i^{th} individual; $\boldsymbol{\mu}$ is the population mean of each trait; $\boldsymbol{batch}_j$ is a systematic effect of the j^{th} fattening batch, with 4 categories; $\boldsymbol{\beta}$ is the regression coefficient on the covariate *weight at slaughter* ($\boldsymbol{weight}_k$); $\boldsymbol{\delta}$ is the SNP allelic effect, estimated as a regression coefficient on the corresponding $\boldsymbol{g}_l$ genotype (values −1, 0, 1) of the l^{th} SNP; and $\boldsymbol{e}_{ijklm}$ is the residual effect. The statistical relevance of the systematic environmental sources of variation and the covariates included in the model were previously reported by Gallardo *et al.*[8] and Casellas *et al.*[16]. Correction for multiple testing was implemented with a false discovery rate approach[17].

Microarray data were available exclusively for GM muscle samples[10]. Following the strategy employed in the Genotype-Tissue Expression (GTEx) pilot analysis[18], we primarily searched for cis-eQTL because they are expected to have larger effects than their trans-counterparts. We used two different strategies: **Analysis 1**, we retrieved 12 genes localized within GM QTL regions and we looked for cis-eQTL that might regulate their expression and **Analysis 2**, we made a search for cis-eQTL at a whole genome scale and we analyzed if there was a positional concordance between GWAS signals and cis-eQTL identified in this way. This second strategy made possible to identify cis-eQTL that might be located in the vicinity of GWAS signals. Genes corresponding to each probe included in the GeneChip Porcine Genomic array (Affymetrix, Inc., Santa Clara, CA) were identified in the BioMart database[19]. The statistical model assumed in this analysis was:

$$\boldsymbol{y}_{ijklm} = \boldsymbol{\mu} + \boldsymbol{batch}_j + \boldsymbol{lab}_k + \boldsymbol{\delta}\boldsymbol{g}_l + \boldsymbol{e}_{ijklm} \quad (2)$$

where $\boldsymbol{y}_{ijklm}$ is the vector that defines the expression of each gene in the GM muscle of the i^{th} individual; $\boldsymbol{\mu}$ is the mean expression of each gene in the population; $\boldsymbol{batch}_j$ and $\boldsymbol{lab}_k$ are the systematic effects *i.e.* $\boldsymbol{batch}_j$ of fattening (with 4 categories) and $\boldsymbol{lab}_k$ (microarray data were generated in two different laboratories); $\boldsymbol{\delta}$ is the SNP allelic effect estimated as a regression coefficient on the corresponding $\boldsymbol{g}_l$ genotype (values −1, 0, 1) of the l^{th} SNP; and

Phenotype	$h^2_{SNP} \pm SE$	
	LD muscle	GM muscle
Electric conductivity (CE)	0.20 ± 0.07	0.11 ± 0.08
pH at 24 hours (pH_{24})	0.17 ± 0.10	0.12 ± 0.09
Minolta redness (a*)	0.41 ± 0.11	0.45 ± 0.11
Minolta yellowness (b*)	0.29 ± 0.12	0.00 ± 0.14
Minolta lightness (L*)	0.00 ± 0.25	0.00 ± 0.05

Table 1. Proportion of phenotypic variance of meat quality traits recorded in the *longissimus dorsi* (LD) and *gluteus medius* (GM) muscles of Duroc pigs explained by SNP markers (h^2_{SNP}) and its standard error (SE).

$\boldsymbol{e}_{ijklm}$ is the residual effect. Correction for multiple testing was implemented with a false discovery rate approach[17]. The threshold of significance in **Analysis 1** took into consideration the number of SNPs contained within 2 Mb windows around each one of the 12 genes under consideration, while in **Analysis 2** such threshold was established by taking into account the 36,710 SNPs typed in the Duroc population.

Results and Discussion

The SNPs arrayed in the Porcine SNP60 BeadChip explain a limited amount of the phenotypic variance of meat quality traits. By using the GEMMA software, we have estimated the proportion of phenotypic variance explained by the 36,710 SNPs (h^2_{SNP}) genotyped with the Porcine SNP60 BeadChip (Table 1). In general, estimates of h^2_{SNP} ranged from low to moderate and differed between muscles. Discrepancies in the genealogic heritability (h^2) estimates of meat quality traits recorded in different skeletal muscle samples were previously reported by Larzul *et al.*[20]. In this way, these authors found h^2 of 0.03 and 0.23 for L* measured in the *gluteus profundus* and *longissimus* muscles, respectively. Similarly, the h^2 values of pH_{24} measured in 4 different muscles oscillated between 0.17 (*longissimus*) and 0.39 (*biceps femoris*)[20]. When Gallardo *et al.*[8] performed a QTL scan for meat quality traits in the Lipgen population, they also found that QTL maps differed markedly amongst traits recorded in the GM and LD muscles. As a whole, these results suggest that there are muscle-specific factors that modulate the genetic determinism of meat quality traits. Indeed, Quintanilla *et al.*[21] identified remarkable differences in the gene expression patterns of the LD and GM muscles, a feature that was especially prominent for genes involved in muscle tissue development, cell proliferation and migration and muscle contraction.

Several h^2_{SNP} values obtained by us were comparable to genealogic heritabilities estimated for porcine meat quality traits in previous studies. For instance Gjerlaug-Enger *et al.*[22] reported heritabilities for a* of 0.43 and 0.46 in Duroc and Landrace pigs, respectively. Similarly, Van Wijk *et al.*[23] and Gjerlaug-Enger *et al.*[22] described heritabilities of 0.11 (crossbred pigs) and from 0.12 (Landrace) to 0.27 (Duroc) for pH_{24}. More unexpected were the null h^2_{SNP} values obtained in the current work for traits such as b* (in GM) and L* (in both muscles). We attribute these null heritabilities to our inability to detect genetic variants that may have small effects or that segregate at very low frequencies[24].

Environmental variables may also obscure the contribution of genetic factors. Indeed, meat quality traits can be affected by poor on-farm handling, mixing of unfamiliar animals and high pig density and long travel distance during transportation[25]. Such events may increase the stress of the swine brought to the abattoir and, consequently, they may have negative consequences on meat quality[25]. At the abattoir, extended lairage time can increase the incidence of dark, firm and dry (DFD) meat, while a short lairage time has been associated with an increased proportion of pale, soft and exudative (PSE) meat[25]. Electrical stunning induces a more rapid pH fall early post mortem and an inferior water-holding capacity than CO_2 stunning, while an accelerated chilling may have negative consequences on meat tenderness and water-holding capacity[25]. In summary, all these factors, and others that are not mentioned, can have a strong impact on the post-mortem pH, electrical conductivity and color of pig meat and "dilute" the contribution of polygenes[25].

Genome-wide and chromosome-wide associations with meat quality traits in Duroc pigs. At the genome-wide level, we found significant associations between CE of LD and three genomic regions on SSC4, SSC5 and SSC13 (Table 2). The SSC4, 104 megabase (Mb) region, lies close to a previously reported QTL for CE identified by Cepica *et al.*[26]. We also found positional concordance between the SSC13 (137.0 Mb) region associated with LD CE and a *semimembranosus* CE QTL reported by Evans *et al.*[27]. At the chromosome-wide level, a coincidence was detected between a a* QTL on SSC3 (50–57 Mb, Table 3) and a QTL for the same trait reported by Li *et al.*[28] on SSC3 (55 Mb). Overall, our results confirm the existence of differences in the genetic determinism of meat quality traits recorded in the GM and LD muscles. The only exception was a region on SSC5 that significantly affected CE in both LD and GM muscles (Table 3). When we compared these data with the set of QTL previously reported by Gallardo *et al.*[8] in the same Lipgen population we found one coincidence *i.e.* the GWAS signal identified on SSC4 (132 Mb) for CE in LD overlapped the confidence interval of a LD CE QTL (S0097 marker, ~133 Mb) detected by these authors[8].

In general the positional coincidence between GWAS signals detected by us and those reported in previous studies was weak, indicating that the majority of associations reported in the current work are new. For instance, when we compared our a*, b* and pH_{24} data with those described in six additional GWAS studies[4,6,29–32] we only found one positional coincidence between the SSC10 (70.6 Mb) genomic region associated with LD a* in the Lipgen

Trait	SSC	N	SNP	Location (Mb)	*P*-value	*q*-value	δ ± SE	A1	MAF
LD CE	4	4	H3GA0013593	104.2–104.8	6.19E-06	0.04	0.28 ± 0.06	A	0.39
	5	1	ASGA0024711	15.4	2.46E-06	0.04	−0.32 ± 0.07	G	0.18
	13	1	ALGA0027007	137.0	7.34E-06	0.04	0.27 ± 0.06	A	0.39

Table 2. Genomic regions significantly associated at the genome-wide level with meat quality traits in Duroc pigs. LD: *longissimus dorsi* muscle, **CE:** Electrical conductivity at 24 hours post-mortem, **N**: Number of SNPs significantly associated with the trait under study, **SSC**: porcine chromosome, **SNP**: SNP displaying the most significant association with the trait under study, **Location (Mb)**: region containing SNPs significantly associated with the trait under study, ***P*-value**: nominal P-value, ***q*-value**: q-value calculated with a false discovery rate approach, ***δ***: allelic effect and its standard error (**SE**), **A1**: minority allele, **MAF**: frequency of the minority allele.

Trait	SSC	N	SNP	Location (Mb)	*P*-value	*q*-value	δ ± SE	A1	MAF
LD CE	4	9	ALGA0026686	93.5–98.8	1.54E-05	0.01	−0.28 ± 0.06	G	0.50
		32	H3GA0013593	104.2–107.1	6.19E-06	0.01	0.28 ± 0.06	A	0.39
		1	ALGA0028809	131.0	2.04E-04	0.02	−0.26 ± 0.07	A	0.17
	5	11	ASGA0024711	14.4–16.1	2.46E-06	0.004	−0.32 ± 0.07	G	0.18
GM CE	5	5	ASGA0024564	13.0–14.7	3.15E-05	0.03	−0.37 ± 0.09	A	0.39
LD pH_{24}	16	3	MARC0086782	6.0–6.4	5.27E-04	0.05	0.08 ± 0.02	G	0.09
		2	ALGA0089269	17.3–18.5	5.09E-04	0.05	−0.06 ± 0.02	G	0.19
		10	ASGA0091353	20.9–29.5	4.01E-04	0.05	0.05 ± 0.02	G	0.41
GM pH_{24}	17	2	MARC0038923	14.2–16.4	9.11E-05	0.04	−0.06 ± 0.02	A	0.48
		5	MARC0101162	53.1–57.2	2.70E-04	0.04	0.07 ± 0.02	G	0.29
		3	H3GA0049744	64.5–65.3	1.81E-04	0.04	−0.06 ± 0.02	G	0.38
LD a*	10	1	ALGA0113811	70.6	2.99E-05	0.04	0.46 ± 0.11	A	0.36
GM a*	3	3	H3GA0009494	16.6–17.0	7.85E-05	0.01	0.70 ± 0.17	A	0.16
		27	H3GA0009489	50.2–57.2	1.27E-04	0.01	0.65 ± 0.17	A	0.18
		4	ALGA0021059	119.7–119.9	7.85E-04	0.04	0.48 ± 0.14	A	0.24
		4	ALGA0021078	120.0–120.4	7.85E-04	0.04	0.48 ± 0.14	A	0.24
GM L*	16	1	MARC0073433	3.5	3.45E-05	0.04	1.23 ± 0.29	C	0.24

Table 3. Genomic regions associated at the chromosome-wide level with meat quality traits in Duroc pigs. GM: *gluteus medius* muscle, **LD:** *longissimus dorsi* muscle, **CE**: Electrical conductivity at 24 hours post-mortem, **pH_{24}**: pH at 24 hours post-mortem; **a***: Minolta redness; **L***: Minolta lightness, **N**: Number of SNPs significantly associated with the trait under study, **SSC**: porcine chromosome, **SNP**: SNP displaying the most significant association with the trait under study, **Location (Mb)**: region containing SNPs significantly associated with the trait under study, ***P*-value**: nominal P-value, ***q*-value**: q-value calculated with a false discovery rate approach, ***δ***: allelic effect and its standard error (**SE**), **A1**: minority allele, **MAF**: frequency of the minority allele.

population (Table 3) and the SSC10 (72.8 Mb) region identified by Ma *et al.*[4] as associated with the same trait in the *semimembranosus* muscle of White Duroc × Erhualian F_2 pigs.

The level of coincidence of trait-associated regions between these six GWAS for a*, b* and pH_{24} traits was also quite low. Only about 20% of the regions identified as significantly associated with any of these phenotypes were shared between two studies or more, indicating that the majority of associations are population-specific. These shared regions were: (a*) SSC4 (80–85 Mb)[6,30], SSC6 (17–22 Mb)[4,30], SSC7 (31–32 Mb)[4,31], SSC12 (58–63 Mb)[30,31], SSC15 (133–136 Mb)[30–32]; (b*) SSC15 (129–133 Mb)[30,32]; and (pH_{24}), SSC3 (15–19 Mb)[30,31], SSC15 (133–136 Mb)[29,32]. This latter region on SSC15 (133–136 Mb) appeared to be associated with a*, b*, pH_{24}, shear force and cook loss in many independent studies[29–32] but not in ours. Interestingly, this SSC15 region contains the protein kinase AMP-activated non-catalytic subunit gamma 3 (*PRKAG3*) gene, whose polymorphism has causal effects on muscle glycogen depletion, a parameter that can have a strong influence on meat quality traits[33].

Besides technical and methodological reasons, a probable cause for the lack of positional concordance between GWAS studies would be genetic heterogeneity[34]. Indeed, Yang *et al.*[34] performed a GWAS for blood lipid traits in 2,400 Laiwu, Erhualian and Duroc × (Landrace × Yorkshire) pigs and they identified a total of 22 QTL. Notably, only six regions were identified in more than one population, and 16 were detected in a single population.

Positional concordance between cis-eQTL for genes expressed in the GM muscle and QTL for GM traits. In general, eQTL are highly enriched in variants with causal effects on phenotypic variation and they can provide valuable information about candidate genes to be further investigated. Integrative analyses of QTL and eQTL data have been performed in pigs, making possible to combine the power of recombination with

QTLs			Genes			Cis-eQTLs									
Trait	SSC	Location (Mb)	Names	SSC	Location (Mb)	SSC	N	SNPs	Location (Mb)	*P*-value	*q*-value	*B*	δ ± SE	A1	MAF
GM a*	3	16.6–17.0	*GUSB*	3	16.9	3	3	ALGA0104024	16.4–17.6	1.60E-03	0.02	0.04	0.28 ± 0.09	A	0.46
GM pH_{24}	17	53.1–57.2	*CTSA*	17	53.7	17	1	ALGA0095491	53.7	1.91E-05	6.11E-04	6.11E-04	−0.37 ± 0.08	G	0.25
		64.5–65.3	*FAM210B*		64.0		16	ALGA0096195	64.1–65.7	4.53E-11	1.99E-09	1.99E-09	−0.53 ± 0.07	G	0.22

Table 4. List of significant cis-eQTLs mapping within QTL regions for *gluteus medius* meat quality traits. **a*:** Minolta redness, **pH_{24}**: pH at 24 hours post-mortem**, N:** number of significant SNPs, **SNP**: marker displaying the most significant association with the trait under study, **Location (Mb)**: region containing SNPs significantly associated with the trait under study, ***P*-value:** nominal P-value, ***q*-value:** q-value calculated with a false discovery rate approach, ***B***: P-value corrected for multiple testing with the Bonferroni method, ***δ***: allelic effect and its standard error (**SE**), **A1:** minority allele, **MAF:** frequency of the minority allele.

QTLs			Genes			Cis-eQTLs									
Traits	SSC	Location (Mb)	Names	SSC	Location (Mb)	SSC	N	SNPs	Location (Mb)	*P*-value	*q*-value	*B*	δ ± SE	A1	MAF
GM a*	3	50.2–57.2	*IGKC*	3	59.8	3	20	ALGA0019294	58.0–61.9	7.54E-11	4.60E-07	2.15E-06	−1.6 ± 0.22	A	0.19
		120.0–120.4	*ADCY3*		121.1–121.2		3	ALGA0103469	120.0–121.9	2.28E-06	0.05	0.06	−0.83 ± 0.17	A	0.08
GM pH_{24}	17	53.1–57.2	*SLPI*	17	53.1	17	16	ALGA0095584	52.3–55.9	6.00E-08	3.48E-04	1.64E-03	2.30 ± 0.40	A	0.13

Table 5. List of significant cis-eQTLs mapping close to QTL regions for *gluteus medius* meat quality traits. **a*:** Minolta redness, **pH_{24}**: pH at 24 hours post-mortem, **N:** number of significant SNPs, **SNPs**: marker displaying the most significant association with the trait under study, **Location (Mb)**: region containing SNPs significantly associated with the trait under study, ***P*-value:** nominal P-value, ***q*-value:** q-value calculated with a false discovery rate approach, ***B***: P-value corrected for multiple testing with the Bonferroni method, ***δ***: allelic effect and its standard error (**SE**), **A1:** minority allele, **MAF:** frequency of the minority allele.

expression studies in order to identify promising candidate genes[35]. For instance, multiple associations between SNPs mapping to porcine chromosomes 4 and 6 and meat quality traits have been detected[30]. Through an eQTL approach, it was possible to identify several genes on SSC4 (*ZNF704, IMPA1* and *OXSR1*) and SSC6 (*IH1D1, SIGLEC10, TBCB, LOC100518735, KIF1B, LOC100514845*) whose variation is concomitantly associated with gene expression and phenotype data[30]. Similarly, Ma *et al.*[36] used a genetical genomics approach to demonstrate that a splice mutation in the *PHKG1* gene is the causal mutation for a glycolytic potential QTL mapping to SSC3.

We have used this integrative strategy to identify potential candidate genes for meat quality traits in a dataset of 12 loci that mapped to GM QTL regions (**Analysis 1**). In doing so, we have detected 3 cis-eQTLs (Table 4) that co-localize with three chromosome-wide QTLs. One of them maps to SSC3 (16.6–17.06 Mb) and displays associations with a* (Fig. 1a); while the other two are located on SSC17 (53.1–57.2; 64.5–65.3) and show significant associations with GM pH_{24} (Fig. 1b and c). Interestingly, two of the three cis-regulated genes encode lysosomal enzymes, *i.e.* cathepsin A (*CTSA*) and glucuronidase β (*GUSB*), that might be released during the post-mortem maturation of meat[37,38]. Cathepsin A is a lysosomal serine protease that can also protect galactosidase β from intralysosomal proteolysis[38], while glucuronidase β is mainly involved in the degradation of glycosaminoglycans[39]. Interestingly, there are evidences that galactosidase β and glucuronidase β might affect the degradation of the collagen mucopolysaccharide, thus having a potential impact on meat ultrastructural properties[40].

In **Analysis 2**, we have identified three additional cis-eQTL that map near to the SSC3 QTL for a* and the SSC17 QTL for pH_{24} (Table 5). The *ADCY3* locus, that co-localizes with the SSC3 QTL for GM a* (Fig. 2a), encodes an adenylate cyclase catalysing the conversion of ATP into cyclic adenosine-3′,5′-monophosphate (cAMP), a secondary messenger that can have broad effects on muscle metabolism[41]. Indeed, AMPc is an activator of the cAMP-dependent protein kinase, a molecule involved in the phosphorylation of enzymes that promote the conversion of glycogen into glucose[41]. Noteworthy, the amount of glycogen stored in the muscle determines the post-mortem production of lactic acid, a molecule that has strong effects on meat color. Another eQTL of interest is the one influencing the mRNA levels of the secretory leukocyte peptidase inhibitor (*SLPI*) gene. This cis-eQTL co-localizes with the SSC17 QTL for GM pH_{24} (Fig. 2b). The *SLP1* gene encodes a serine-protease that inhibits protein-degrading enzymes with strong effects on meat tenderization *i.e.* when the skeletal muscle is being degraded and transformed into meat, SLPI attenuates muscle proteolysis by binding to proteases and rendering them inactive[42]. Finally, the co-localization of the *IGKC* cis-eQTL and the SSC3 QTL for a* (Fig. 2c) does not have an obvious biological interpretation because this gene is mainly related with humoral immunity.

Conclusions

We have detected genome-wide and chromosome-wide significant QTL for meat quality traits recorded in a Duroc commercial line with a population size that was moderate but comparable to the ones used in other porcine GWAS[43–45]. The limited positional concordance between the set of QTL detected by us and those reported by other authors in purebred populations suggests the existence of a significant amount of genetic heterogeneity

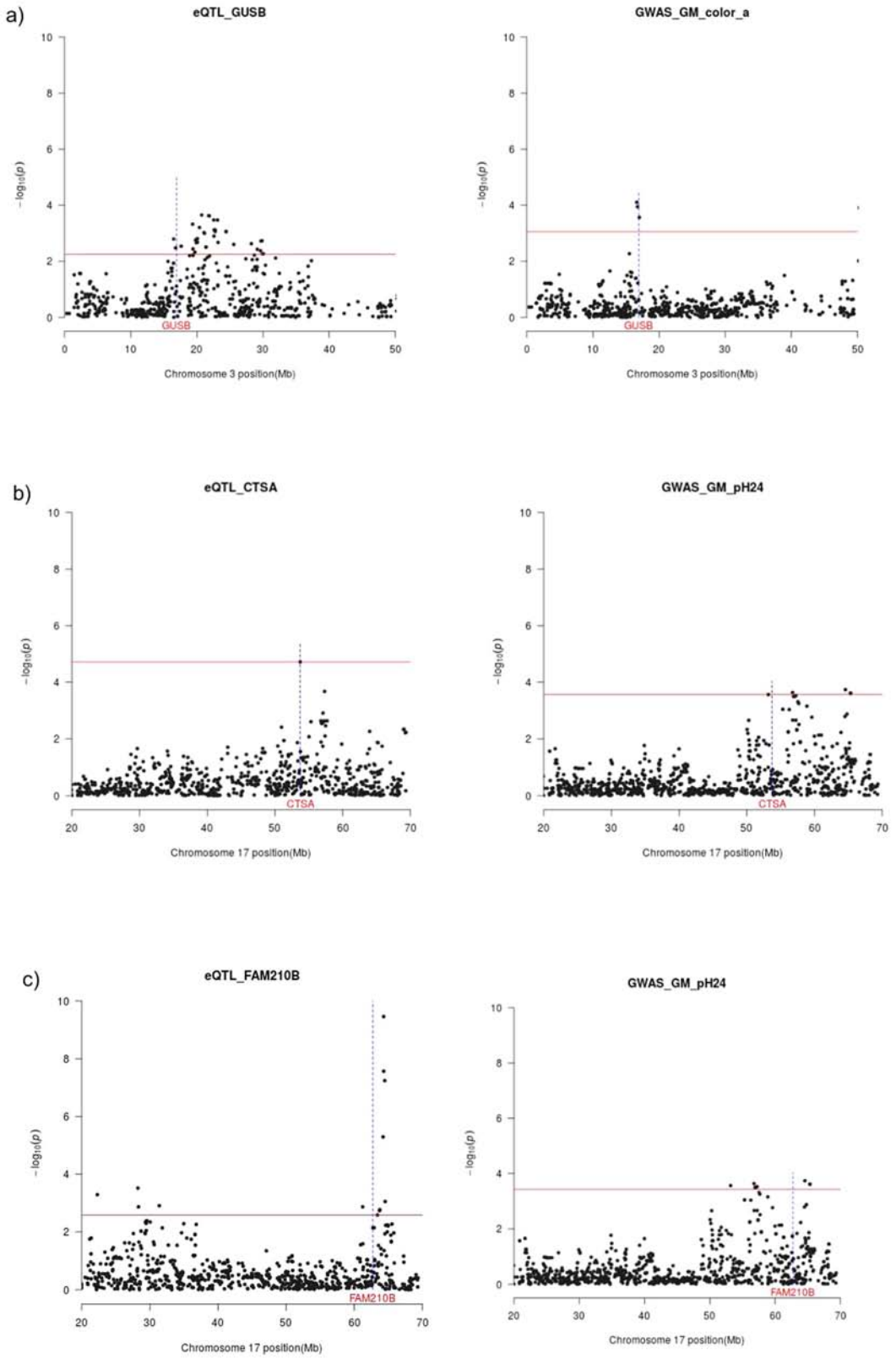


Figure 1. Cis-eQTL (left panel) for the *GUSB* (1**a**), *CTSA* (1**b**) and *FAM210B* (1**c**) genes which map to QTL regions associated with meat quality traits recorded in the *gluteus medius* muscle (right panel). The *x*-axis represents chromosome length (Mb), and the *y*-axis shows the −log10 (*P*-value) of the associations found. The horizontal line indicates the threshold of significance (q-value ≤ 0.05). The vertical line depicts the genomic location of the *GUSB, CTSA* and *FAM210B* genes.

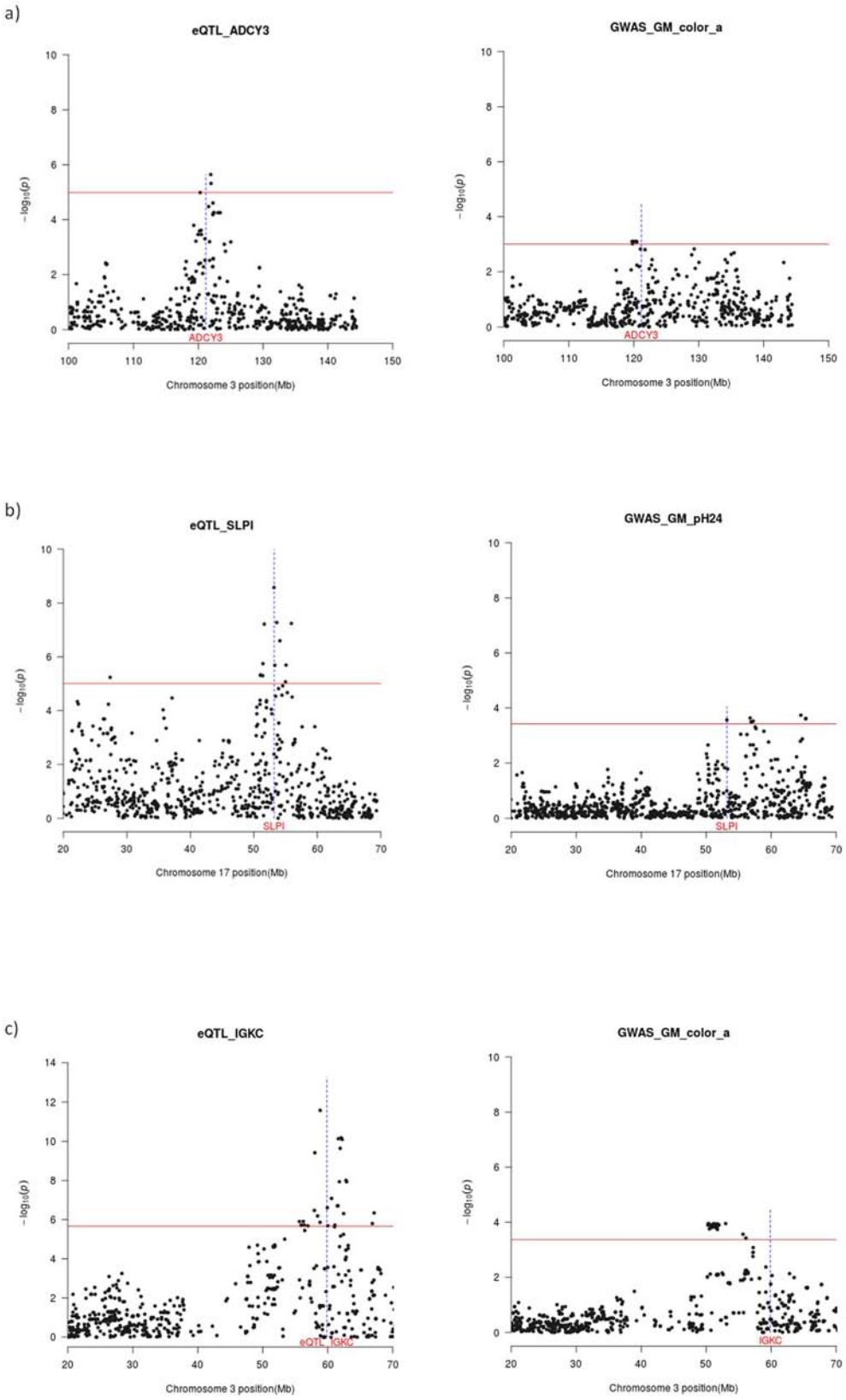


Figure 2. Co-localization of cis-eQTL (left panel) for the *ADCY3* (2**a**), *SLP1* (2**b**) and *IGKC* (2**c**) genes and QTL for meat quality traits recorded in the *gluteus medius* muscle (right panel). The *x*-axis represents chromosome length (Mb), and the *y*-axis shows the –log10 (*P*-value) of the associations found. The horizontal line indicates the threshold of significance (q-value ≤ 0.05). The vertical line depicts the genomic location of the *ADCY3, SLP1 and IGKC* genes.

for meat quality traits in porcine breeds. We have found remarkable differences between the QTL maps for the LD and GM muscles, suggesting that meat quality is determined to a great extent by genetic factors that are muscle-specific. Finally, we have observed a number of cis-eQTL that co-localize with meat quality QTL regions. Several of these cis-eQTL regulate the expression of genes which may play important roles in muscle physiology and post-mortem meat maturation. Sequencing of the regulatory regions of these loci might be useful to uncover the identity of the causal mutations explaining the existence of these QTLs.

References

1. Ciobanu, D. C., Lonergan, S. M. & Huff-Lonergan, E. J. Genetics of meat and carcass traits In *The Genetics of the Pig*. 2nd edn. (ed. Rothschild, M. F. & Ruvinsky, A.) 356 (CABI, 2011).
2. Luo, W. *et al.* Genome-wide association analysis of meat quality traits in a porcine Large White × Minzhu intercross population. *Int J Biol Sci*. **8,** 580–595 (2012).
3. Nonneman, D. *et al.* Genome-wide association of meat quality traits and tenderness in swine. *J. Anim. Sci.* **91,** 4043–4050 (2013).
4. Ma, J. *et al.* Genome-wide association study of meat quality traits in a White Duroc × Erhualian F_2 intercross and Chinese Sutai pigs. *PloS One* **8,** e64047 (2013).
5. Becker, D., Wimmers, K., Luther, H., Hofer, A. & Leeb, T. A Genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars. *PloS One* **8,** e5595 (2013).
6. Sanchez, M. P. *et al.* Genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. *Genet Sel Evol*. **46,** 12 (2014).
7. Fang, M. & Andersson, L. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. *Proc. Biol. Sci*. **273,** 1803–1810 (2006).
8. Gallardo, D. *et al.* Quantitative trait loci analysis of a Duroc commercial population highlights differences in the genetic determination of meat quality traits at two different muscles. *Anim. Genet*. **43,** 800–804 (2012).
9. Gallardo, D. *et al.* Polymorphism of the pig acetyl-coenzyme A carboxylase α gene is associated with fatty acid composition in a Duroc commercial line. *Anim. Genet*. **40,** 410–417 (2009).
10. Canovas, A., Quintanilla, R., Amills, M. & Pena, R. N. Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. *BMC Genomics* **11,** 372 (2010).
11. Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162,** 156–159 (1987).
12. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80 (2004).
13. Xu, X., Zhao, Y. & Simon, R. Gene set expression comparison kit for BRB-Array tools. *Bioinformatics* **24,** 137–139 (2008).
14. Manunza, A. *et al.* A genome-wide association analysis for porcine serum lipid traits reveals the existence of age-specific genetic determinants. *BMC Genomics* **15,** 758 (2014).
15. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44,** 821–824 (2012).
16. Casellas, J. *et al.* Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs. *J. Anim. Sci.* **88,** 2246–2254 (2010).
17. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57,** 289–300 (1995).
18. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348,** 648–660 (2015).
19. Smedley, D. *et al.* The BioMart community portal: an innovative alternative to large centralized data repositories. *Nucleic Acids Res.* **43,** W589–598 (2015).
20. Larzul, C. *et al.* Selection for reduced muscle glycolytic potential in Large White pigs. II. Correlated responses in meat quality and muscle compositional traits. *Genet. Sel. Evol.* **31,** 61–76 (1999).
21. Quintanilla, R., Pena, R. N., Cánovas, A. & Amills, M. Differential gene expression profile between two porcine skeletal muscles: *longissimus dorsi* and *gluteus medius*. In *Book of abstracts of the 32th International Conference on Animal Genetics* (2010).
22. Gjerlaug-Enger, E., Aass, L., Ødegard, J. & Vangen, O. Genetic parameters of meat quality traits in two pig breeds measured by rapid methods. *Animal* **4,** 1832–1843 (2010).
23. Van Wijk, H. J. *et al.* Genetic parameters for carcass composition and pork quality estimated in a commercial production chain. *J. Anim. Sci*. **83,** 324–333 (2005).
24. Zuk, O., Hechter, E., Sunyaev, S. & Lander, E. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. USA* **109,** 1193–1198 (2012).
25. Rosenvold, K. & Andersen, H. J. Factors of significance for pork quality-a review. *Meat Sci*. **64,** 219–237 (2003).
26. Cepica, S. *et al.* Linkage and QTL mapping for sus scrofa chromosome 4. *J. Anim. Breed. Genet.* **120,** 28–37 (2003).
27. Evans, G. J. *et al.* Identification of quantitative trait loci for production traits in commercial pig populations. *Genetics* **164,** 621–627 (2003).
28. Li, H. D. *et al.* Quantitative trait loci analysis of swine meat quality traits. *J. Anim. Sci*. **88,** 2904–2912 (2010).
29. Bernal Rubio, Y. L. *et al.* Implementing meta-analysis from genome-wide association studies for pork quality traits. *J. Anim. Sci*. **93,** 5607–5617 (2015).
30. Ponsuksili, S., Murani, E., Trakooljul, N., Schwerin, M. & Wimmers, K. Discovery of candidate genes for muscle traits based on GWAS supported by eQTL-analysis. *Int. J. Biol. Sci*. **10,** 327–337 (2014).
31. Liu, X. *et al*. Genome-wide association analyses for meat quality traits in Chinese Erhualian pigs and a Western Duroc × (Landrace × Yorkshire) commercial population. *Genet. Sel. Evol*. **47,** 44 (2015).
32. Zhang, C. *et al.* Genome-wide association studies (GWAS) identify a QTL close to *PRKAG3* affecting meat pH and colour in crossbred commercial pigs. *BMC Genetics* **16,** 33 (2015).
33. Milan, D. *et al.* A mutation in *PRKAG3* associated with excess glycogen content in pig skeletal muscle. *Science* **288,** 1248–1251 (2000).
34. Yang, H. *et al.* Genome-wide association analysis for blood lipid traits measured in three pig populations reveals a substantial level of genetic heterogeneity. *PloS One* **10,** e0131667 (2015).
35. Ernst, C. W. & Steibel, J. P. Molecular advances in QTL discovery and application in pig breeding. *Trends Genet.* **29,** 215–224 (2013).
36. Ma, J. *et al.* A Splice mutation in the *PHKG1* gene causes high glycogen content and low meat quality in pig skeletal muscle. *PLoS Genet.* **10,** e1004710 (2014).
37. Dutson, T. R. & Lawrie, R. A. Release of lysosomal enzymes during post mortem conditioning and their relationship to tenderness. *Int. J. Food Sci. Tech*. **9,** 43–50 (2007).
38. Hiraiwa, M. Cathepsin A/protective protein: an unusual lysosomal multifunctional protein. *Cell. Mol. Life Sci*. **56,** 894–907 (1999).
39. Naz, H. *et al.* Human β-glucuronidase: structure, function, and application in enzyme replacement therapy. *Rejuvenation Res*. **16,** 352–363 (2013).

40. Chang, H. J., Xu, X. L., Zhou, G. H., Li, C. B. & Huang, M. Effects of characteristics changes of collagen on meat physicochemical properties of beef semitendinosus muscle during ultrasonic processing. *Food Bioprocess. Technol.* **5,** 285–297 (2009).
41. Hong, S. *et al. et al.* Upregulation of adenylate cyclase 3 (*ADCY3*) increases the tumorigenic potential of cells by activating the CREB pathway. *Oncotarget* **4,** 1791–1803 (2013).
42. Urso, L. *et al.* Alterations in mRNA expression and protein products following spinal cord injury in humans. *J. Physiol.* **579,** 877–892 (2007).
43. Ramayo-Caldas, Y. *et al.* Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. *Anim. Sci.* **90,** 2883–2893 (2012).
44. Fontanesi, L. *et al.* A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genomics* **13,** 583 (2012).
45. Becker, D., Wimmers, K., Luther, H., Hofer, A. & Leeb, T. A genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars. *PLoS One* **8,** e55951 (2013).

Acknowledgements

The authors are indebted to Selección Batallé S.A. for providing the animal material. We gratefully acknowledge to J. Reixach (Selección Batallé), I. Díaz (IRTA) and J. Soler (IRTA) for their collaboration in the experimental protocols. Thanks to Anna Mercadé and Anna Castelló for their technical support. This research was partially funded with projects AGL2013-48742-C2-1-R and AGL2013-48742-C2-2-R, awarded by the Spanish Ministry of Economy and Competitiveness (MINECO). We also acknowledge the support of the Spanish Ministry of Economy and Competitiveness for the *Center of Excellence Severo Ochoa 2016–2019* (SEV-2015-0533) grant awarded to the Center for Research in Agricultural Genomics. Rayner Gonzalez-Prendes was funded by a FPU Ph.D. grant from the Spanish Ministerio de Educación (FPU12/00860). Thanks also to the CERCA Programme of the Generalitat de Catalunya.

Author Contributions

R.Q., M.A., J.L.N., A.M. and J.J. conceived the study and designed the experiment; R.Q. and J.L.N. produced the animal material and collected the phenotypic data; T.F. contributed to molecular tasks; R.G.-P. carried out the genome-wide association analyses for meat quality phenotypes and expression data; R.N.P. and A.C. contributed to the analysis of microarray data; R.G.-P. and M.A. wrote the manuscript. All authors helped to draft the manuscript and read and approved its final version.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: González-Prendes, R. *et al.* Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. *Sci. Rep.* **7**, 39830; doi: 10.1038/srep39830 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter IV.

Identifying genomic regions associated with muscle lipid phenotypes and gene expression traits in Duroc pigs

Rayner González-Prendes, Raquel Quintanilla, Tainã Figueiredo Cardoso, Arianna Manunza, Joaquim Casellas, Ángela Cánovas, Isabel Díaz, José Luis Noguera, Anna Castelló, Anna Mercadé, Marcel Amills. (Submitted to BMC Genomics)

Abstract

Background

Intramuscular fat (IMF) content and composition have a strong impact on the nutritional and organoleptic properties of porcine meat. Performance of genome-wide association studies (GWAS) has contributed to dissect the genomic basis of these traits and to identify mutations with causal effects. The goal of the current work was to investigate the genomic architecture of IMF traits in pigs.

Results

By performing a GWAS for 54 IMF content and composition traits, recorded in the *longissimus dorsi* (LD) and *gluteus medius* (GM) muscles of 350 Duroc pigs (Lipgen population), we have identified 37 genome-wide and 83 chromosome-wide QTL. Importantly, we have observed a low positional concordance between QTL detected for traits recorded in the GM and LD muscles ($\approx$ 85% of the QTL happened to be muscle-specific). We have also investigated if QTL regions contain expression QTL (eQTL) with either *cis*- or *trans*-regulatory effects on mRNA levels estimated with microarrays (N=104). Such analysis made evident the co-localization of QTL for IMF traits and *cis*-eQTL affecting the expression of 20 loci. By using microarray data, we have also shown that SNPs mapping to QTLs distributed on 8 chromosome regions *trans*-regulate the expression levels of 103 loci.

Conclusions

The low positional concordance of QTL detected in the GM and LD muscles may have practical implications in the framework of genomic selection schemes aimed to improve IMF content and composition traits. Indeed, our results suggest that such selection may have heterogeneous consequences depending on the muscle under consideration (unless SNPs with consistent effects across muscles are selected as markers). We have also observed the existence of a substantial number of co-localizations between IMF QTL and *cis*- and *trans*-eQTL regulating gene expression. Further research will be needed to ascertain if such co-localizations are fortuitous or if they reflect the existence of causal mutations with regulatory effects on the expression levels of lipid-related genes as well as on the phenotypic variation of intramuscular fat phenotypes.

Keywords: Intramuscular fat, fatty acid, genome-wide association analysis, GWAS, Duroc pig, gene expression.

Background

Intramuscular fat (IMF) content and composition have important effects on the oxidative stability, tenderness and juiciness of pig meat [1]. These traits are moderately heritable and, in consequence, they can be improved through artificial selection [2]. Many genome scans have been carried out in pigs to identify quantitative trait loci (QTL) with effects on IMF phenotypes [3]. These studies have revealed that all porcine chromosomes harbour at least one IMF QTL, and that there are IMF QTL hotspots on chromosomes 4, 6 and 7 [3]. Recently, several genome-wide association studies (GWAS) for IMF traits have been performed in divergent crosses and purebred pig populations [4–10], thus providing a comprehensive and high resolution picture about the genetic basis of such phenotypes.

Genome-wide association studies performed in humans evidence that most regions displaying significant associations with phenotypes do not encode proteins, suggesting that the majority of causal mutations may have regulatory effects [11]. In pigs, multiple genome scans for expression QTL (eQTL) have been carried out as a strategy to elucidate the genomic architecture of traits of economic interest [12–16]. In this way, hundreds of eQTL associated with muscle gene expression phenotypes have been identified, and several of them have been shown to co-localize with QTL for fatness traits [6,15,17,18]. In a previous study, 80 positional concordances between 14 *cis*-acting and 66 *trans*-acting eQTL and IMF QTL were detected [15]. Amongst them, five IMF QTL showed positional concordance with eQTL regulating the expression of genes involved in lipid metabolism and adipose function. The goal of the current work was to identify quantitative trait loci (QTL) for IMF content and composition traits in the *gluteus*

medius (GM) and *longissimus dorsi* (LD) muscles of Duroc pigs with the aim of ascertaining their levels of positional concordance. Moreover, we have investigated the co-localization of IMF QTL with SNPs regulating gene expression, either in *cis*- or in *trans*-. The identification of such co-localizations might be considered as a first step towards detecting regulatory mutations with causal effects on IMF traits.

Methods

Animal material and phenotype recording

Phenotypes were recorded in 350 barrows from a commercial Duroc line (Lipgen population) generated by crossing 5 boars with 400 sows. After weaning, this pig population was transferred to the experimental test station at the Centre de Control Porcí (CCP) of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). A detailed description of the experimental population and management conditions can be found in Gallardo et al. [19,20]. Pigs were slaughtered at an approximate age of 190 days and a live weight of 122 kg. A near infrared transmittance device (NIT, Infratec 1625, Tecator Hoganas, Sweden) was used to determine IMF content in the GM and LD muscles. The measurement of fatty acid (FA) composition (C:12 to C:22 range) in the GM and LD muscles was achieved with a technique based on the gas chromatography of methyl esters [21]. A complete list of the IMF content and composition traits measured in the current experiment is shown in **Supplementary Table 1**.

High throughput genotyping with the Porcine SNP60 BeadChip

Genotyping of the 350 Duroc pigs was achieved with the Porcine SNP60 BeadChip (Illumina, SanDiego, CA) which contains probes for 62,163 SNPs. Analyses related with the quality of the genotyping results were performed with the GenomeStudio software (Illumina). By using PLINK [22], we filtered SNPs with minor allele frequencies (MAF) below 5%, rates of missing genotypes above 10% or showing highly significant departures from the Hardy-Weinberg expectation (threshold set at a *P*-value of 0.001). We also excluded SNPs that did not map to the porcine reference genome (Sscrofa10.2 assembly) and those located in sexual chromosomes. After these filtering steps, we obtained a subset of 36,710 SNPs that were used as markers in the GWAS analysis.

Microarray analyses of gene expression in the *gluteus medius* muscle

GeneChip Porcine Genomic arrays (Affymetrix, Inc., Santa Clara, CA) were used to measure gene expression in GM samples from 104 Duroc pigs (data are available in the Gene Expression Omnibus public repository of the National Center for Biotechnology Information, accession number: GSE19275). A detailed description of the techniques and methods used to perform RNA purification and microarray hybridization can be found in Canovas et al. [23]. Briefly, GM muscle samples were pulverized with a mortar and a pestle in liquid nitrogen and homogenized with a polytron device. Total RNA was isolated with an acid phenol protocol [24] and the One Cycle cDNA Synthesis Kit (Affymetrix, Inc.) was employed to synthesize double

stranded cDNA. cRNAs were purified with the GeneChip Sample Cleanup Module (Affymetrix, Inc.), fragmented and added to a hybridisation mixture. GeneChip Porcine Genome Arrays were prehybridised with 1× hybridisation buffer at 45 °C for 10 min [23]. Subsequently, the arrays were hybridized with the mixture containing cRNAs at 45 °C for 16 hours. GeneChips were washed and labelled with streptavidin phycoerythrin in a Fluidics Station 450 (Affymetrix, Inc) and they were scanned in an Agilent G3000 GeneArray Scanner (Agilent Technologies, Inc.). The "Affy" and "Sympleaffy" tools of the Bioconductor project [25] were used to establish a set of quality control metrics to evaluate RNA quality and the efficiencies of the labelling and hybridisation steps. Data pre-processing and normalization were carried out with the BRB-ArrayTools software version 3.7.1 [26]. Genes displaying more than 20% of expression values over ±1.5 times the median expression of all arrays were retained for further analysis. Official gene names and positions of each probe included in the GeneChip Porcine Genomic array (Affymetrix, Inc., Santa Clara, CA) were identified in the BioMart database [27, 28]

Statistical analyses

Identification of QTL for intramuscular fat content and composition traits

Statistical methods employed in the current work have been previously reported in Gonzalez-Prendes et al. [29]. In this way, mixed-model association analyses were carried out

with the Genome-wide Efficient Mixed-Model Association (GEMMA) software, developed by Zhou and Stephens [30]. This method corrects population structure by considering the relatedness matrix, built on the basis of all genome-wide SNPs as a random effect. We used the following statistical model to analyze IMF content and composition traits:

$$\mathbf{y} = \mathbf{W}\alpha + \mathbf{x}\beta + \mathbf{Zu} + \varepsilon$$

where **y** is the vector of trait values for all individuals; **W** is a matrix of fixed effects ("batch of fattening" with 4 categories) and covariates that depend on the trait: (1) IMF content in GM (for fatty acid traits measured in the GM muscle), (2) IMF content in LD (for fatty acid traits measured in the LD muscle), (3) backfat thickness (for IMF content measured in GM and LD); α is a vector of the corresponding coefficients including the intercept; **x** is a vector of marker genotypes; β is the effect size of the marker; **u** is a vector of random individual effects with a n-dimensional multivariate normal distribution $MVN_n(0, \lambda \tau^{-1} K)$, where τ^{-1} is the variance of the residual errors; λ is the ratio between the two variance components and **K** is a known relatedness matrix derived from SNPs; and ε is a vector of errors. The statistical relevance of the systematic environmental sources of variation and the covariates were previously corroborated by Gallardo et al. [19] and Casellas et al. [31]. The proportion of phenotypic variance explained by the 36,710 SNPs (*i.e.* "chip heritability") was calculated as follows:

$$h^2_{snp} = \frac{\sigma^2_g}{\sigma^2_g + \sigma^2_e}$$

where σ_g^2 is the additive variance calculated from markers and σ_e^2 is the residual variance. Correction for multiple testing was implemented with the false discovery rate (FDR) approach of Benjamini and Hochberg [32].

Co-localization between expression QTL and QTL for intramuscular fat traits

We performed a genome scan to identify potential *cis*-eQTL regulating the expression of 66 genes mapping to QTL determining IMF traits recorded in the GM muscle by using a previously reported methodology [29]. For the QTL regions with only one significant SNP, we selected genes located ±1 Mb around it. By using microarray data, we also investigated if SNPs located within GM IMF QTL regions affect in *trans*- the expression of 2,974 genes expressed in the GM muscle.

$$\mathbf{y} = \mathbf{W}\alpha + \mathbf{x}\beta + \mathbf{Zu} + \varepsilon$$

where **y** is the vector of trait values for all individuals; **W** is a matrix of covariates *i.e.* "batch of fattening" (with 4 categories) and "laboratory" (microarray data were generated in two different laboratories); α is a vector of the corresponding coefficients including the intercept; **x** is a vector of genotypes of a marker; β is the effect size of the marker; **u** is a vector of random individual effects with a n-dimensional multivariate normal distribution MVN_n (0, $\lambda \tau^{-1}$ K), where τ^{-1} is the variance of the residual errors; λ is the ratio between the two variance components and **K** is a known relatedness matrix derived from SNPs; and ε is a vector of

errors. Correction for multiple testing was implemented with the FDR approach mentioned before [32]. As previously reported [29], the threshold of significance in the analysis of *cis*-eQTL was based on the number of SNPs contained within 2 Mb windows (*i.e.* ± 1 Mb around the analysed gene), while in the *trans*-eQTL analysis we took into account the whole set of SNPs mapping to QTL regions (459 SNPs).

Results

The description of the IMF content and composition phenotypes analyzed in the current work and the percentage of phenotypic variance explained by the 36,710 SNPs (h^2_{snp}) can be found in **Supplementary Table 1** and **Table 1**, respectively. The h^2_{snp} values ranged between 0.00-0.46 and 0.00-0.51 for GM and LD traits, respectively. Substantial discrepancies in the magnitudes of h^2_{snp} estimates corresponding to LD and GM traits were observed, a feature that suggests the existence of muscle-specific differences in the genetic determinism of IMF content and composition traits. For instance, h^2_{snp} values for omega-6 to -3 ratio (LD = 0.07, GM = 0.12), palmitic (LD = 0.13, GM = 0.26), and unsaturated (LD = 0.14, GM = 0.37) FA were clearly different in both muscles.

Table 1. Proportion of the phenotypic variance (and its standard error) of intramuscular fat content and composition traits explained by the 36,710 SNP markers analysed in the current work.

Phenotype (%)	**Symbol**	***longissimus dorsi* muscle**	***gluteus medius* muscle**
		h^2_{snp} ± SE	h^2_{snp} ± SE
Intramuscular fat	IMF	0.51 ± 0.10	0.46 ± 0.09
Saturated FA	SFA	0.14 ± 0.07	0.37 ± 0.12
Capric	C10:0	0.03 ± 0.05	0.06 ± 0.06
Lauric	C12:0	0.08 ± 0.09	0.00 ± 0.00
Myristic	C14:0	0.20 ± 0.09	0.00 ± 0.03
Palmitic	C16:0	0.13 ± 0.06	0.26 ± 0.12
Margaric	C17:0	0.14 ± 0.08	0.05 ± 0.05
Stearic	C18:0	0.26 ± 0.12	0.29 ± 0.10
Arachidic	C20:0	0.00 ± 0.06	0.02 ± 0.07
Unsaturated FA	UFA	0.14 ± 0.07	0.37 ± 0.12
Monounsaturated FA	MUFA	0.03 ± 0.05	0.00 ± 0.04
Palmitoleic	C16:1(n-7)	0.20 ± 0.08	0.25 ± 0.08
Palmitelaidic	C16:1(n-9)	0.35 ± 0.11	0.18 ± 0.10
Heptadecenoic	C17:1	0.12 ± 0.07	0.15 ± 0.07
Oleic	C18:1(n-9)	0.05 ± 0.05	0.00 ± 0.04
Gondoic	C20:1	0.02 ± 0.05	0.12 ± 0.10
Polyunsaturated FA	PUFA	0.00 ± 0.04	0.01 ± 0.03
Linoleic	C18:2	0.00 ± 0.03	0.01 ± 0.04
α-Linolenic	C18:3 (n-3)	0.10 ± 0.06	0.05 ± 0.05
Eicosadienoic	C20:2 (n-6)	0.07 ± 0.06	0.01 ± 0.05
Eicosatrienoic	C20:3 (n-3)	0.00 ± 0.04	0.02 ± 0.04
Arachidonic	C20:4	0.00 ± 0.03	0.00 ± 0.03
Eicosapentaenoic	C20:5	0.00 ± 0.00	0.04 ± 0.05
Docosahexaenoic	C22:6	0.00 ± 0.03	0.00 ± 0.03
Omega-3 FA	FA n-3	0.00 ± 0.05	0.06 ± 0.05
Omega-6 FA	FA n-6	0.00 ± 0.04	0.00 ± 0.03
Omega-6 to -3 ratio	n-6/n-3	0.07 ± 0.07	0.12 ± 0.10

Performance of a GWAS revealed the existence of 37 QTL displaying genome-wide significant associations with IMF phenotypes under study (**Table 2**). Besides, we also detected 83 chromosome-wide significant associations (**Supplementary Table 2**). Pig chromosomes SSC2 (9-11 Mb), SSC4 (63.9-64 Mb), SSC5 (71-79 Mb) and SSC14 (87-99 Mb, 120-124 Mb) harboured the majority of significant QTL for muscle FA composition traits (**Table 2, Supplementary Table 2),** and such statement was particularly true for the SSC14 (120-124 Mb) region which was associated with multiple traits both in the GM and LD muscles **(Figure 1**). In general, there was a lack of positional concordance between GWAS signals detected for traits recorded in the LD and GM muscles (**Figure 2**). Indeed, when we considered the whole set of genome-wide and chromosome-wide significant associations, around 85% IMF QTL happened to be muscle-specific. It should be taken into account, however, that there were positional coincidences for QTL regulating different but related traits in both muscles. For instance, two QTL on SSC4 (63.9 Mb) and SSC14 (87.8-87.9 Mb) were associated with stearic (C18:0) content in the GM muscle and oleic (C18:1) percentage in the LD muscle. Similarly, one QTL on SSC14 (93.7-94.9 Mb) was associated with stearic in the GM muscle and with saturated and unsaturated FA content in the LD muscle.

Table 2. Genome-wide QTL for intramuscular fat content and composition traits recorded in the *gluteus medius* (GM) and *longissimus dorsi* (LD) muscles of Duroc pigs.

Majority FA (*gluteus medius*)										
Traits	**SSC**	**N**	**SNP**	**Region(Mb)**	***P*-value**	***q*-value**	***B***	**δ ± SE**	**A_1**	**MAF**
GM C16:0	5	8	ALGA0033025	71.7-79.8	0.00	0.03	0.17	-0.75 ± 0.16	A	0.12
		4	INRA0020052	80.0-80.1	0.00	0.03	0.17	-0.75 ± 0.16	G	0.12
GM C18:0	4	3	DIAS0001351	63.9	0.00	0.00	0.04	-0.55 ± 0.11	G	0.29
	14	4	ASGA0063465	58.6-59.4	0.00	0.03	1.00	-0.44 ± 0.11	A	0.23
		9	ALGA0078300	65.5-67.9	0.00	0.01	0.41	-0.48 ± 0.11	A	0.23
		5	ALGA0079209	87.8-87.9	0.00	0.01	0.62	0.43 ± 0.09	A	0.48
		3	ASGA0064951	92.9-97.0	0.00	0.01	0.60	0.46 ± 0.09	G	0.41
		43	ALGA0081091	120.4-124.4	0.00	0.00	0.00	-0.63 ± 0.10	C	0.35
GM C18:1 (n-9)	14	19	ALGA0081091	120.9-122.4	0.00	0.02	0.26	0.13 ± 0.03	C	0.35
GM SFA	4	3	DIAS0001351	63.9	0.00	0.04	1.00	-0.81 ± 0.19	G	0.29
	14	38	ALGA0081091	120.9-123.8	0.00	0.01	0.11	-0.87 ± 0.18	C	0.35
GM UFA	4	3	DIAS0001351	63.9	0.00	0.04	1.00	0.81 ± 0.19	G	0.29
	14	38	ALGA0081091	120.9-123.8	0.00	0.01	0.11	0.87 ± 0.18	C	0.35
Majority FA (*longissimus dorsi*)										
LD C18:0	14	2	CASI0010207	93.7	0.00	0.04	1.00	0.47 ± 0.11	A	0.41
		40	ALGA0081091	120.4-123.4	0.00	0.00	0.01	-0.62 ± 0.10	C	0.35
LD C18:1 (n-9)	4	3	MARC0050687	63.9	0.00	0.00	0.05	0.17 ± 0.03	C	0.29
		2	MARC0071018	134.9	0.00	0.04	1.00	-0.12 ± 0.03	G	0.32
	10	1	ALGA0057858	27.0	0.00	0.01	0.61	0.20 ± 0.05	G	0.11
	14	4	ALGA0079221	87.8-87.9	0.00	0.03	1.00	-0.13 ± 0.03	A	0.45
		44	ALGA0081091	120.4-124.3	0.00	0.00	0.00	0.18 ± 0.03	C	0.35
		6	ALGA0082693	144.7-148.1	0.00	0.00	0.16	-0.13 ± 0.03	C	0.46
LD SFA	12	1	ALGA0110494	42.0	0.00	0.04	1.00	0.81 ± 0.18	A	0.31
	14	6	ASGA0064951	93.2-94.9	0.00	0.03	0.99	0.81 ± 0.17	G	0.41
		37	ALGA0081091	120.4-123.4	0.00	0.00	0.05	-0.95 ± 0.19	C	0.35
LD UFA	12	1	ALGA0110494	42.0	0.00	0.05	1.00	-0.80 ± 0.18	A	0.31
	14	1	ASGA0064951	94.9	0.00	0.04	1.00	-0.80 ± 0.17	G	0.41
		35	ALGA0081091	120.4-123.4	0.00	0.00	0.06	0.94 ± 0.19	C	0.35

Minority FA (g*luteus medius*)										
Traits	**SSC**	**N**	**SNP**	**Region (Mb)**	***P*-value**	***q*-value**	***B***	**δ ± SE**	**A_1**	**MAF**
GM C10:0	12	2	DRGA0011702	31.4-31.6	0.00	0.04	0.06	-0.04 ± 0.01	G	0.13
GM C20:3 (n-3)	9	1	ALGA0121521	9.3	0.00	0.05	0.41	0.11 ± 0.03	A	0.07
	18	8	ASGA0097792	43.2-46.7	0.00	0.01	0.02	0.12 ± 0.02	G	0.07
Minority FA (*longissimus dorsi*)										
LD C14:0	9	32	M1GA0026515	11.5-14.5	0.00	0.02	0.28	0.09 ± 0.02	G	0.41
LD C16:1 (n-9)	2	7	H3GA0006290	23.8-24.3	0.00	0.01	0.03	-0.02 ± 0.01	G	0.25
	6	2	ASGA0087502	80.1	0.00	0.03	0.22	-0.02 ± 0.00	G	0.33
	12	1	ASGA0053255	13.6	0.00	0.04	0.39	-0.02 ± 0.00	G	0.48
LD C17:0	2	10	MARC0050503	10.1-11.1	0.00	0.00	0.00	-0.03 ± 0.01	G	0.30
	9	1	ASGA0009038	139.3	0.00	0.02	0.06	-0.03 ± 0.01	G	0.43
	17	1	INRA0052734	14.8	0.00	0.04	0.48	-0.03 ± 0.01	G	0.20

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with the trait under study, SNP: SNP displaying the most significant association with the trait under study, Region (Mb): region containing SNPs significantly associated with the trait under study, *P*-value: nominal *P*-value, *q*-value: *q*-value calculated with a false discovery rate approach, *B*: Bonferroni corrected *P*-values, δ: allelic effect and its standard error (SE), A_1: minority allele, MAF: frequency of the minority allele.

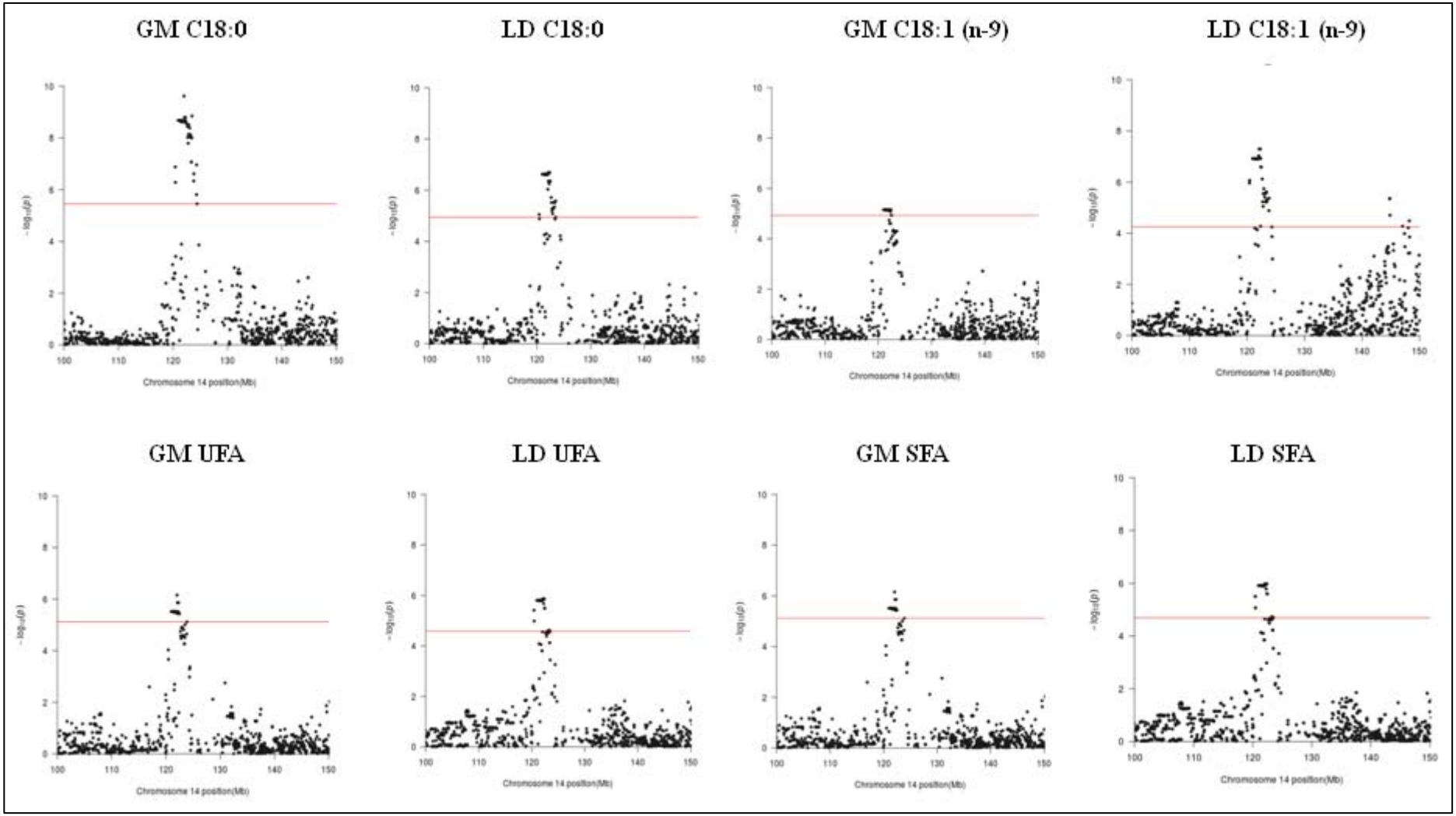


Figure 1. Manhattan plot of a genomic region on pig chromosome 14 (120-124 Mb) displaying pleiotropic genome-wide significant associations with C18:0, C18:1(n-9), unsaturated (UFA) and saturated (SFA) fatty acids contents in the *gluteus medius* (GM) and *longissimus dors*i (LD) muscles of Duroc pigs. The x-axis represents the chromosomal region (Mb) containing the QTL and the y-axis shows the $-\log_{10}$ (*P*-value) of the reported associations. The horizontal line indicates the threshold of significance (q-value ≤ 0.05).

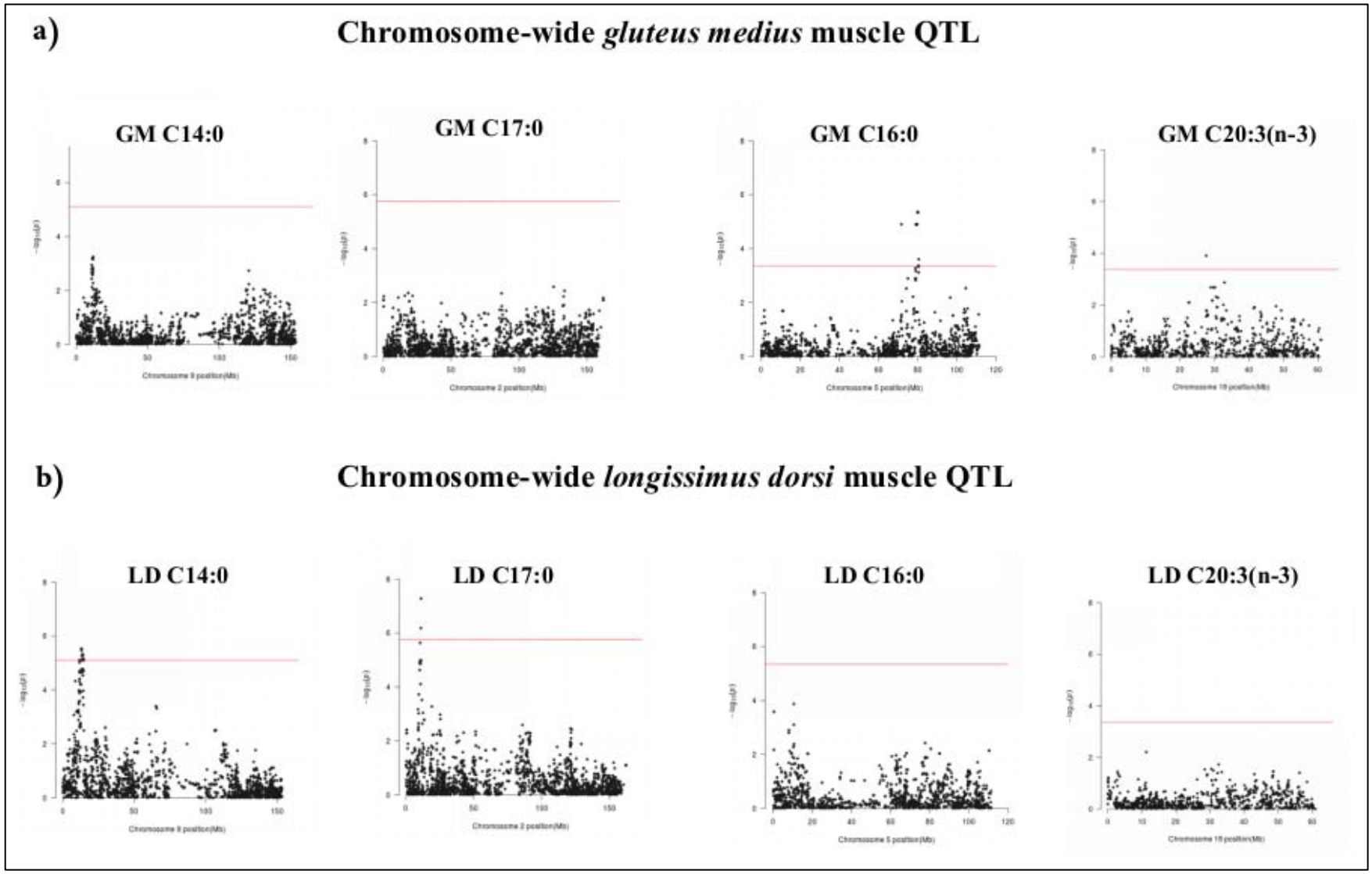


Figure 2. Comparison of the Manhattan plots of C14:0, C17:0, C16:0 and C20:3(n-3) fatty acid traits recorded in the (a) *gluteus medius* (GM) and (b) *longissimus dorsi* (LD) muscles of Duroc pigs. It can be observed that none of these 4 QTL affects IMF composition in both muscles. The x-axis represents the chromosomal region (Mb) containing the QTL and the y-axis shows the $-\log_{10}$ (*P*-value) of the reported associations. The horizontal line indicates the threshold of significance (q-value ≤ 0.05).

Performance of an eQTL scan for 66 genes, located within GM QTL regions and with available microarray measurements of gene expression, made possible to identify 20 *cis*-eQTL regulating the expression of 20 loci (**Table 3**, **Figure 3**). As shown in **Table 3**, chromosome 14 encompassed the majority of these *cis*-eQTL, which regulated 9 genes co-localizing with three GM C18:0 QTL at 6.1-7.6 Mb (*FAM160B2* and *POLR3D*), 56.2-59.8 Mb (*LGALS8* and *LYST*) and 64.8 Mb (*RAB4A*), and one GM C16:1 (n-7) QTL at 143.4 Mb (*PSTK*). The remaining three *cis*-regulated genes (*BLOC1S2*, *COX15* and *KCNIP2*) co-localized with the SSC14 (120-124 Mb) region displaying pleiotropic effects on muscle FA composition. We also identified *cis*-regulated genes mapping to QTL on SSC4, SSC5, SSC6, SSC9, SSC12, SSC13 and SSC18 (**Table 3**). Finally, an eQTL scan based on microarray data allowed us identifying ≈ 116 *trans*-eQTL regulating the expression of 103 genes located in multiple genomic regions (**Supplementary Table 3**).

Table 3. List of the co-localizations between QTL for IMF traits recorded in the *gluteus medius* (GM) muscle and *cis*-eQTL regulating the mRNA levels of genes expressed in the GM muscle and mapping to QTL regions.

***Cis*-eQTLs**									**GM QTLs**		**Genes**	
SSC	**N**	**SNPs**	**Region (Mb)**	***P*-value**	***q*-value**	**B**	**δ±SE**	**A_1**	**MAF**	**Traits**	**Region (Mb)**	**Names**	**Region (Mb)**
4	12	ALGA0026676	98.1-99.6	0.00	0.00	0.00	-0.50±0.06	G	0.48	C20:4	99.0-99.0	*PEX19*	98.0
	12	ASGA0020832	98.1-99.1	0.00	0.00	0.00	0.76 ±0.12	G	0.40			*ATP1A2*	98.2-98.3
5	1	ALGA0032768	72.9	0.00	0.05	0.05	0.29 ±0.11	A	0.22	C14:0; C16:0; C18:2; FA n-6;PUFA	71.7-80.1	*BID*	72.1
	13	MARC0048694	75.7-77.7	0.00	0.01	0.01	0.39 ±0.11	A	0.19			*ZCRB1*	76.4
6	17	ASGA0028321	62.1-63.9	0.00	0.00	0.00	0.64 ±0.12	C	0.37	C17:0	63.3-63.5	*SLC25A33*	64.1
	10	ALGA0037549	146.2-147.9	0.00	0.00	0.00	0.43 ±0.07	G	0.26		146.8-146.8	*TTC4*	145.6-145.7
9	14	ASGA0041324	8.0-9.3	0.00	0.00	0.00	0.51 ±0.08	G	0.35	C20:3 (n-3)	9.3-9.3	*PLEKHB1*	8.8
12	14	ALGA0120489	26.3-27.9	0.01	0.02	0.30	-0.31±0.13	G	0.14	C10:0	26.8-27.4	*NME1*	27.4-27.5
13	8	ALGA0067450	2.0-3.8	0.00	0.00	0.00	0.41 ±0.07	G	0.31	C20:3 (n-3)	2.2- 2.2	*SH3BP5*	2.4-2.5
	2	ASGA0095016	214.0-215.8	0.00	0.01	0.01	-0.35±0.10	A	0.16	C17:0	214.6-214.6	*SLC37A1*	216.2-216.3
14	5	ALGA0081104	121.4-121.9	0.00	0.00	0.00	0.34 ±0.08	G	0.14	C16:1 (n-7); C18:0; C18:1 (n-9); SFA; UFA	120.2-125.9	*BLOC1S2*	120.8-120.9
	17	MARC0043866	118.3-121.5	0.00	0.00	0.00	0.42 ±0.07	G	0.49			*COX15*	120.3
	20	DIAS0001040	6.1-7.5	0.00	0.00	0.00	-0.34±0.06	A	0.37	C18:0	7.5- 7.5	*FAM160B2*	6.7-6.8
	10	ASGA0066137	122.0-123.8	0.01	0.01	0.07	0.24 ±0.09	G	0.30	C16:1 (n-7);C18:0; C18:1 (n-9);SFA; UFA	120.4-124.3	*KCNIP2*	122.7
	14	ASGA0063513	56.2-59.8	0.00	0.00	0.00	0.68 ±0.08	G	0.37	C18:0	55.5-59.4	*LGALS8*	58.8-59.6
	1	M1GA0018688	59.4-59.4	0.00	0.03	0.03	0.24 ±0.08	C	0.45			*LYST*	59.5-59.6
	25	H3GA0038597	6.1-7.6	0.00	0.00	0.00	0.73 ±0.07	G	0.38		7.5- 7.5	*POLR3D*	6.9
	1	H3GA0042863	143.4	0.00	0.04	0.04	-0.29±0.09	C	0.24	C16:1 (n-7)	143.0-144.4	*PSTK*	144.1
	1	ALGA0078128	64.8	0.00	0.03	0.03	0.30 ±0.09	G	0.13	C18:0	64.7-68.2	*RAB4A*	65.3
18	18	ASGA0090003	46.1-47.7	0.00	0.01	0.02	0.29 ±0.08	A	0.35	C20:3 (n-3)	40.2-46.7	*FKBP14*	47.2-47.3

[1]SSC: porcine chromosome , N: number of SNPs significantly associated with the trait under study , SNP: SNP displaying the most significant association with the trait under study, Region (Mb): region containing SNPs significantly associated with the trait under study or gene position, *P*-value: nominal *P*-value, *q*-value: *q*-value calculated with a false discovery rate approach, B: Bonferroni-corrected *P*-value, δ:allelic effect and its standard error (SE), A_1: minority allele, MAF: frequency of the minority allele.

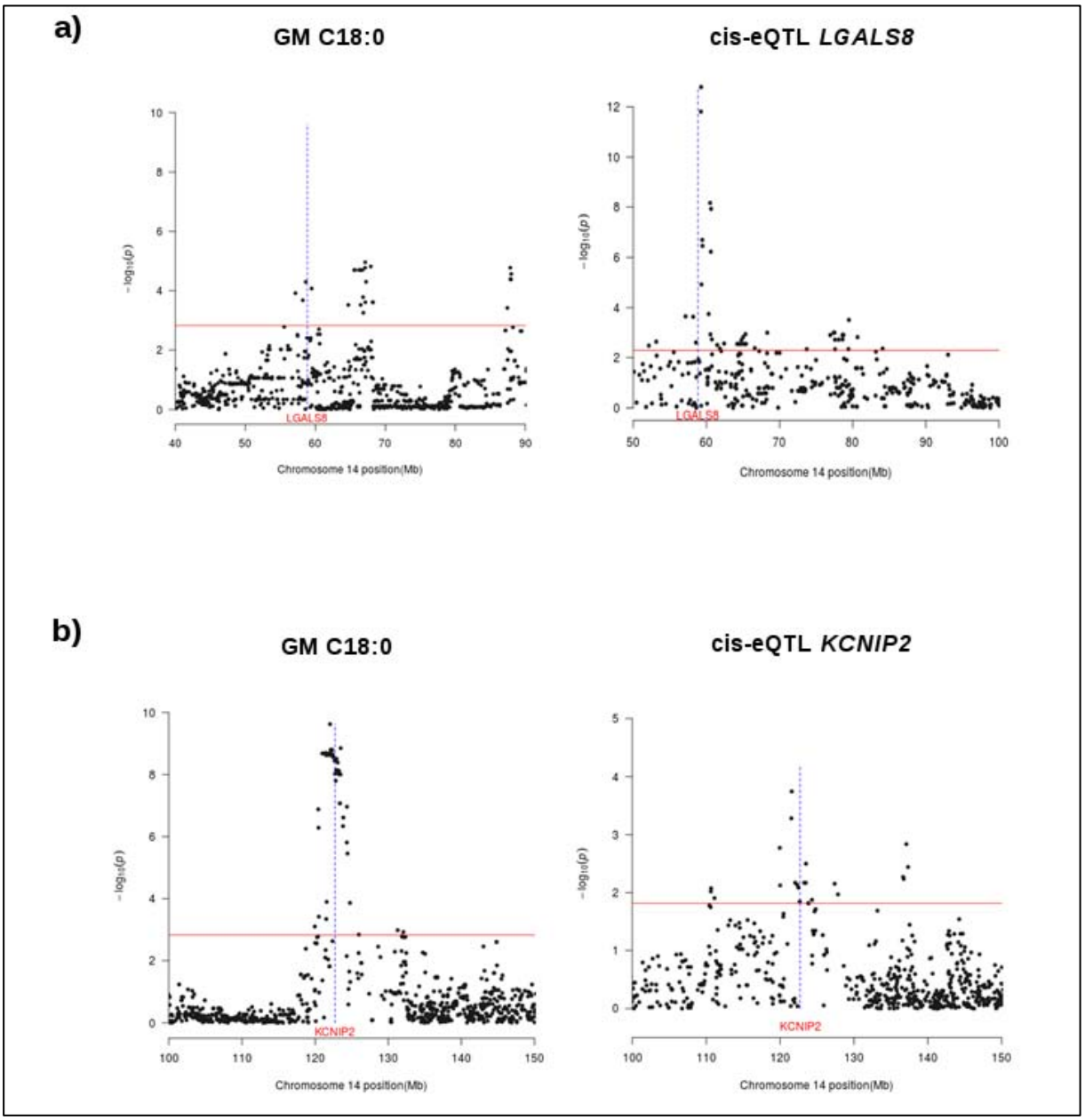


Figure 3. Co-localization of *cis*-eQTL (right panel) for the (a) *LGALS8* and (b) *KCNIP2* genes and two QTL regions (left panel) for *gluteus medius* (GM) C18:0 traits: (a) SSC14, 55-59 Mb, (b) SSC14, 120-124 Mb. The x-axis represents the chromosomal region (Mb) containing the co-localizing QTL and eQTL and the y-axis shows the $-\log_{10}$ (*P*-value) of the reported associations. The horizontal line indicates the threshold of significance (q-value ≤ 0.05). The vertical line depicts the genomic location of the *LGALS8* and *KCNIP2* genes.

Discussion

A genome-wide association analysis reveals that four chromosomes harbour the main determinants of intramuscular fat composition in a Duroc line

The amount of phenotypic variance explained by the SNPs contained in the Porcine SNP60 BeadChip was quite modest (**Table 1**). For instance, h^2_{snp} values for C16:0, C18:0, C18:1(n-9) and C18:2 at LD and GM ranged from 0.00-0.26 and 0.00-0.29, respectively. In contrast, genealogic heritabilities estimated by Casellas et al. [31] for the same set of traits happened to be considerable higher *i.e*. 0.25-0.47 for LD and 0.32-0.44 for GM. These results suggest that part of the phenotypic variation of IMF-related traits is not captured by the set of Porcine SNP60 BeadChip markers. The size of our resource population is modest, but similar to those employed in previously published studies [4,33,34]. Indeed, we were able to identify 37 QTL with genome-wide significant effects on muscle FA composition (**Table 2**), and 83 QTL displaying chromosome-wide significant associations (**Supplementary Table 2**).

Genomic regions on SSC2 (9-11 Mb), SSC4 (63.9-64 Mb), SSC5 (71-79 Mb) and SSC14 (87-99 Mb, 120-124 Mb) were the ones that harboured the main determinants of muscle FA composition in our Duroc commercial line (**Table 2, Supplementary Table 2)**. Most of these genetic determinants had pleiotropic effects on several FA traits. For instance, the SSC14 (120-124 Mb) region displayed consistent and strong associations with saturated and unsaturated FA both in the GM and the LD muscles (**Table 2, Figure 1**). This SSC14 region has been also identified in previous GWAS studies as significantly associated with muscle FA composition

traits [5,8–10]. One SNP (g.2228T>C) at the promoter of the stearoyl-CoA desaturase (*SCD*) gene, which catalyses the D^9-*cis* desaturation of a range of fatty acyl-CoA substrates [35] and maps to SSC14 (120.9 Mb), has been shown to be strongly associated with muscle stearic and oleic contents in Duroc pigs [36]. Interestingly, our data suggest the existence of additional determinants of FA composition on SSC14 *i.e.* those mapping to the 58.6-59.4 Mb, 65.5-67.9 Mb, 87.8- 87.9 Mb and 92.9-97.0 Mb intervals (**Table 2**).

Besides SSC14, there were three additional chromosomes displaying significant associations with IMF composition traits. In SSC2, two continuous regions at 10.2-10.9 Mb (LD) and 11.1 Mb (GM) were associated with C17:1 as well as with C17:0 at 9.4 Mb in the LD (**Supplementary Table 2**). In a previous report, Zhang et al. [8] described this region as associated with the LD C20:3n6/C18:2n6 and the C20:4n6/C20:3n6 ratios in Erhualian pigs. Moreover, they proposed the fatty acid desaturase 2 (*FADS2*) gene as a probable candidate locus to explain the associations found. We also detected a significant GWAS signal on SSC5 for GM polyunsaturated FA, and n-6, C14:0, C16:0 and C18:2 FA (71.7-79.8 Mb, **Table 2, Supplementary Table 2**). This region also displayed significant associations with C20:0 in the LD muscle of White Duroc × Erhualian F_2 pigs [5]. Moreover, Ros-Freixedes et al. [10] found associations with polyunsaturated FA content recorded in the GM muscle of Duroc pigs in a SSC5 region (84-85 Mb) that lies close to the one reported by us. Finally, the SSC4 (63.9-64 Mb) region identified in our commercial Duroc line as associated with GM C18:0 and LD C18:1(n-9) contents has not been reported in previous studies, including the one using a closely related Duroc population [10].

Other regions associated with majority and minority FA happened to be scattered throughout

porcine chromosomes SSC1, SSC3, SSC4, SSC5, SSC6, SSC7, SSC8, SSC9, SSC10, SSC11, SSC12, SSC13, SSC17 and SSC18 (**Table 2, Supplementary Table 2**). Most of these associations were less significant than the ones reported in the previous paragraph, affected only specific FA and showed a poor consistency across muscles (**Figure 2**). These results agree well with those of Quintanilla et al. [37], who evidenced the existence of substantial discrepancies between the QTL maps of FA traits recorded in the GM and LD muscles of Duroc pigs. They interpreted this finding as evidence of the existence of muscle-specific factors modulating the penetrance of causal polymorphisms with effects on FA composition [37]. Indeed, correlations between the FA composition of LD and GM muscles are moderate [37] and microarray analysis has demonstrated the existence of significant differences in the GM and LD expression of genes influencing cell differentiation, muscle development and function, and lipid metabolism (our unpublished results).

Co-localization of QTL for IMF related traits and genes regulated by eQTL

A proportion of the QTL detected by us could be explained by the existence of regulatory polymorphisms with causal effects on both gene expression and phenotypic variation. In consequence, we have investigated the existence of eQTL regulating gene expression in *cis*- and *trans*-. As shown in **Table 3**, we detected co-localizations between 20 *cis*-eQTL and 23 QTL for IMF traits (out of 45 GM QTL). Unexpectedly, we did not detect any *cis*-eQTL for the *SCD* gene despite the fact that a polymorphism in its promoter has been defined as the causal mutation explaining the SSC14 QTL for saturated and unsaturated FA [36]. Amongst the *cis*-regulated

genes, it is worth to highlight the peroxisomal biogenesis factor 19 (*PEX19*), which maps to a SSC4 region (98-99 Mb) with effects on C20:4 content (**Table 3**). Peroxisomes play an essential role in lipid catabolism and, more particularly, on the β-oxidation of long-chain and very-long-chain FA [38]. Moreover, intramyocellular triacylglycerol content has been associated with peroxisomal biogenesis and *PEX19* levels in skeletal muscle from lean and obese humans [39]. Another interesting candidate gene is *KCNIP2*, which encodes a potassium voltage-gated channel interacting protein 2 (**Table 3** and **Figure 3**). Interestingly, the potassium two pore domain channel subfamily K member 10 (*KCNK10*) gene has been shown to play a critical role in adipocyte differentiation and the accumulation of triacylglycerols by controlling *C/EBPβ* and *C/EBPδ* expression and insulin signaling [40]. Several of the *cis*-regulated genes detected by us have been reported in previous studies as associated with meat quality traits. For instance, Fontanesi et al. [41] demonstrated the existence of an association between the polymorphism of the ATPase Na^+/K^+ transporting subunit α_2 (*ATP1A2*) gene and backfat thickness in Italian Large White sows. Variability at the NME/NM23 nucleoside diphosphate kinase 1 (*NME1*) gene has been also associated with meat quality traits in pigs [42,43]. Moreover, the pleckstrin homology domain containing B1 (*PLEKHB1*) and zinc finger CCHC-type and RNA binding motif containing 1 (*ZCRB1*) genes are differentially expressed in the adipose tissue of pigs with distinct fatness profiles [44,45] and in pre- vs post-natal muscle samples [46], respectively.

We also explored if the SNPs contained within GM QTL regions are associated with the expression of genes mapping to other genomic regions or chromosomes. In this way, we identified ≈ 116 *trans*-eQTL regulating the expression of 103 genes (**Figure 4**, **Supplementary Table 3**). Several of these genes are *trans*-regulated by two or more genetic determinants *e.g.*

MTMR1 (SSC14, 95.6-96.2 Mb; 143.3 Mb), *PI4KA* (SSC13, 2.2 Mb; SSC14, 55.5-59.5 Mb), *RUSC2* (SSC13, 27 Mb ; SSC14, 93.4-96.2 Mb), *ADAM9* (SSC13, 207 Mb; SSC18, 37 Mb and 40.5-46.5 Mb), *KRR1* (SSC5, 74.0-78.6 Mb ; SSC9, 9.3 Mb) and *MTAP* (SSC5, 78.5 Mb ; SSC9, 9.3 Mb). In other cases, the existence of two *trans*-regulatory factors is less obvious because they map to adjacent regions and, in consequence, they could be considered as a single genetic determinant *e.g. MGAT1* (SSC5, 76.5-78.6 Mb; 80 Mb) and *DDX21* (SSC5, 77.2-77.7 Mb; 80.5 Mb). On the other hand, we also detected several genomic region *trans*-regulating two or more genes *e.g.* SSC5 (76.5-78.6 Mb) which modulates the expression of *YIPF1, MRPS33* and *MGAT1* mRNAs, SSC14 (57.7 Mb) which affects *RASIP1, PTOV1* and *CH242-204P3.4* mRNA levels, and the ALGA0121521 SNP (SSC9, 9.3 Mb), associated with the expression of 23 different loci. The high number of *trans*-eQTL co-localizing with GM QTL suggests that this could be an important source of variation of IMF traits, Indeed, several of the *trans*-regulated genes shown in **Supplementary Table 3** have an important role in lipid metabolism *e.g.* apolipoprotein E (*APOE*) which plays a key role in the transportation and storage of lipids [47] and acyl-CoA dehydrogenase, C-2 to C-3 short chain (*ACADS*), which catalyses the first step of the FA β-oxidation pathway [48].

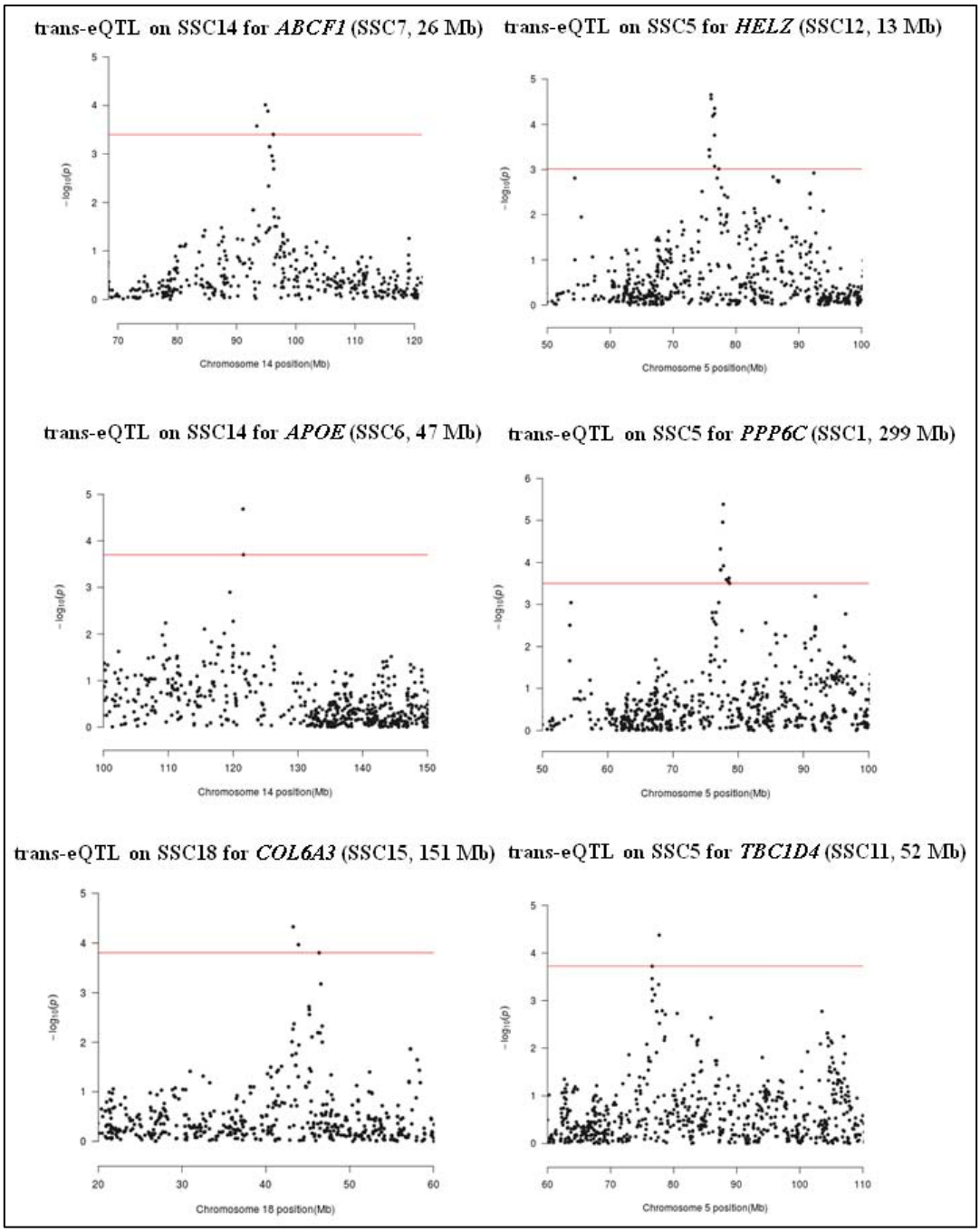


Figure 4. Manhattan plots depicting *trans*-eQTL regulating the expression of the ATP binding cassette subfamily F member 1 (*ABCF1*), apolipoprotein E (*APOE*), collagen type VI α_3 chain (*COL6A3*), helicase with zinc finger (*HELZ*), protein phosphatase 6 catalytic Subunit (*PPP6C*) and TBC1 domain family member 4 (*TBC1D4*) genes. The x-axis represents the chromosomal region (Mb) containing the eQTL and the y-axis shows the –log10 (P-value) of the reported associations. The horizontal line indicates the threshold of significance (q-value ≤ 0.05).

Conclusions

The low positional concordance of QTL detected in the GM and LD muscles may have practical implications in the framework of genomic selection schemes aimed to improve IMF content and composition traits. Indeed, our results suggest that such selection may have heterogeneous consequences depending on the muscle under consideration (unless SNPs with consistent effects across muscles are selected as markers). We have also observed the existence of a substantial amount of co-localizations between GM QTL and *cis*- and *trans*- eQTL regulating the expression of genes with potential effects on lipid metabolism. Further research will be needed to ascertain if such co-localizations are fortuitous or if they reflect the existence of causal mutations with regulatory effects on gene expression and phenotypic consequences.

Declarations

Ethics approval

Animal care, management procedures and blood sampling were performed following national guidelines for the Good Experimental Practices and they were approved by the Ethical Committee of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Consent for publication

Not applicable

Availability of data and material

Upon acceptance of the paper, data will be available at Figshare.

Competing interests

The authors declare no competing interests.

Funding

Part of the research presented in this publication was funded by grants AGL2013-48742-C2-1-R and AGL2013-48742-C2-2-R awarded by the Spanish Ministry of Economy and Competitiveness and grant 2014 SGR 1528 from the Agency for Management of University and Research Grants of the Generalitat de Catalunya. We also acknowledge the support of the Spanish Ministry of Economy and Competitiveness for the *Center of Excellence Severo Ochoa 2016-2019* (SEV-2015-0533) grant awarded to the Center for Research in Agricultural Genomics. Tainã F Cardoso was funded with a fellowship from the CAPES Foundation-Coordination of Improvement of Higher Education, Ministry of Education (MEC) of the Federal Government of Brazil. Rayner Gonzalez-Prendes was funded with a FPU Ph.D. grant from the Spanish Ministry of Education (FPU12/00860). Thanks also to the CERCA Programme of the Generalitat de Catalunya.

Authors' contributions

MA and RQ designed the experiment; RQ and JLN were responsible for the generation of animal material and phenotype recording; ID measured FA composition traits; ACas and AMe performed genotyping tasks; RGP carried out the GWAS analyses for phenotypes and expression traits in collaboration with AMa, JC and TFC; ACan participated in the bioinformatic analyses of microarray expression data; MA and RGP wrote the paper; all authors read and approved the manuscript.

Acknowledgements

The authors are indebted to *Selección Batallé S.A.* for providing the animal material. We gratefully acknowledge to J.L. Noguera (IRTA) and J. Soler (IRTA) for their collaboration in the experimental protocols.

References

1. Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, et al. Fat deposition, fatty acid composition and meat quality: A review. Meat Sci. 2008;78:343–58.

2. Sellier P, Maignel L, Bidanel JP. Genetic parameters for tissue and fatty acid composition of backfat, perirenal fat and longissimus muscle in Large White and Landrace pigs. Anim. 2010;4:497–4.

3. Casellas J, Vidal O, Pena RN, Gallardo D, Manunza A, Quintanilla R, et al. Genetics of serum and muscle lipids in pigs. Anim. Genet. 2013;44:609–19.

4. Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, et al. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. Anim. Sci. 2012;90:2883–93.

5. Yang B, Zhang W, Zhang Z, Fan Y, Xie X, Ai H, et al. Genome-wide association analyses for fatty acid composition in porcine muscle and abdominal fat tissues. PLoS One. 2013;8:e65554.

6. Muñoz M, Rodríguez MC, Alves E, Folch JM, Ibañez-Escriche N, Silió L, et al. Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. BMC Genomics 2013;14:845.

7. Davoli R, Luise D, Mingazzini V, Zambonelli P, Braglia S, Serra A, et al. Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. J. Anim. Breed. Genet. 2016;133:277–82.

8. Zhang W, Yang B, Zhang J, Cui L, Ma J, Chen C, et al. Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. Sci. Rep. 2016;6:24718.

9. Zhang W, Zhang J, Cui L, Ma J, Chen C, Ai H, et al. Genetic architecture of fatty acid composition in the *longissimus dorsi* muscle revealed by genome-wide association studies on diverse pig populations. Genet. Sel. Evol. 2016;48:5.

10. Ros-Freixedes R, Gol S, Pena RN, Tor M, Ibáñez-Escriche N, Dekkers JCM, et al. Genome-wide association study singles out *SCD* and *LEPR* as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. PloS One. 2016;11:e0152496.

11. Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics 2011;187:367–83.

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

13. Wimmers K, Murani E, Ponsuksili S. Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. Brief. Funct. Genomics 2010;9:251–8.

14. Steibel JP, Bates RO, Rosa GJM, Tempelman RJ, Rilington VD, Ragavendran A, et al. Genome-wide linkage analysis of global gene expression in loin muscle tissue identifies candidate genes in pigs. PloS One. 2011;6:e16766.

15. Cánovas A, Pena RN, Gallardo D, Ramírez O, Amills M, Quintanilla R. Segregation of regulatory polymorphisms with effects on the gluteus medius transcriptome in a purebred pig population. PLoS One. 2012;7:e35583.

16. Heidt H, Cinar MU, Uddin MJ, Looft C, Jüngst H, Tesfaye D, et al. A genetical genomics approach reveals new candidates and confirms known candidate genes for drip loss in a porcine resource population. Mamm. Genome 2013;24:416–26.

17. Ponsuksili S, Murani E, Brand B, Schwerin M, Wimmers K. Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. J. Lipid Res. 2011;52:668–78.

18. Cinar MU, Kayan A, Uddin MJ, Jonas E, Tesfaye D, Phatsara C, et al. Association and expression quantitative trait loci (eQTL) analysis of porcine *AMBP*, *GC* and *PPP1R3B* genes with meat quality traits. Mol. Biol. Rep. 2012;39:4809–21.

19. Gallardo D, Pena RN, Amills M, Varona L, Ramírez O, Reixach J, et al. Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs. Physiol. Genomics 2008;35:199–9.

20. Gallardo D, Quintanilla R, Varona L, Díaz I, Ramírez O, Pena RN, et al. Polymorphism of the pig acetyl-coenzyme A carboxylase α gene is associated with fatty acid composition in a Duroc commercial line. Anim. Genet. 2009;40:410–17.

21. Mach N, Devant M, Díaz I, Font-Furnols M, Oliver MA, García JA, et al. Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. J. Anim. Sci.

2006;84:3039–48.

22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 2007;81:559–75.

23. Cánovas A, Quintanilla R, Amills M, Pena RN. Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. BMC Genomics. 2010;11:372.

24. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 1987;162:156–9.

25. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004;5:R80.

26. Xu X, Zhao Y, Simon R. Gene set expression comparison kit for BRB-ArrayTools. Bioinforma. Oxf. Engl. 2008;24:137–9.

27. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. Ensembl 2016. Nucleic Acids Res. 2016;44:D710-16.

28. Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J, et al. The BioMart community portal: an innovative alternative to large, centralized data repositories. Nucleic Acids Res. 2015;43:W589-98.

29. González-Prendes R, Quintanilla R, Cánovas A, Manunza A, Figueiredo Cardoso T, Jordana

J, et al. Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. Sci. Rep. 2017;7:39830.

30. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 2012;44:821–4.

31. 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. J. Anim. Sci. 2010;88:2246–54.

32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 1995;57:289–300.

33. Fontanesi L, Schiavo G, Galimberti G, Calò DG, Scotti E, Martelli PL, et al. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. BMC Genomics 2012;13:583.

34. Becker D, Wimmers K, Luther H, Hofer A, Leeb T. A genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars. PloS One. 2013;8:e55951.

35. Paton CM, Ntambi JM. Biochemical and physiological function of stearoyl-CoA desaturase. Am. J. Physiol. Endocrinol. Metab. 2009;297:E28-7.

36. Estany J, Ros-Freixedes R, Tor M, Pena RN. A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. PloS One. 2014;9:e86177.

37. Quintanilla R, Pena RN, Gallardo D, Cánovas A, Ramírez O, Díaz I, et al. Porcine

intramuscular fat content and composition are regulated by quantitative trait loci with muscle-specific effects. J. Anim. Sci. 2011;89:2963–71.

38. Lodhi IJ, Semenkovich CF. Peroxisomes: a nexus for lipid metabolism and cellular signaling. Cell Metab. 2014;19:380–92.

39. Huang TY, Zheng D, Park S, Houmard J,Hickner R, Cortright R. Intramyocellular triacylglycerol is associated with peroxisomal biogenesis in skeletal muscle from lean and obese humans. FASEB J. 2015;28 (Supplement 1): LB708.

40. Nishizuka M, Hayashi T, Asano M, Osada S, Imagawa M. *KCNK10*, a tandem pore domain potassium channel, is a regulator of mitotic clonal expansion during the early stage of adipocyte differentiation. Int. J. Mol. Sci. 2014;15:22743–56.

41. Fontanesi L, Galimberti G, Calò DG, Fronza R, Martelli PL, Scotti E, et al. Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. J. Anim. Sci. 2012;90:2450–64.

42. Murani E, Te Pas MF, Chang KC, Davoli R, Merks JW, Henne H, Wörner R, et al. Analysis of effects of genes differentially expressed during myogenesis on pork quality; 56th Annual Meeting of the European Association for Animal Production; June 5–8; Uppsala, Sweden. 2005. (Abstr).

43. Wimmers K, Murani E, Te Pas MF, Chang KC, Davoli R, Merks JW, et al. Associations of functional candidate genes derived from gene-expression profiles of prenatal porcine muscle

tissue with meat quality and muscle deposition. Anim. Genet. 2007;38:474–84.

44. Puig-Oliveras A, Ramayo-Caldas Y, Corominas J, Estellé J, Pérez-Montarelo D, Hudson NJ, et al. Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. PloS One. 2014;9:e99720.

45. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Estellé J, Castelló A, Alves E, et al. 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;14:843.

46. Verardo LL, Nascimento CS, Silva FF, Gasparino E, Martins MF, Toriyama E, et al. Identification and validation of differentially expressed genes from pig skeletal muscle. J. Anim. Breed. Genet. 2013;130:372–81.

47. Kypreos KE, Karagiannides I, Fotiadou EH, Karavia EA, Brinkmeier MS, Giakoumi SM, et al. Mechanisms of obesity and related pathologies: role of apolipoprotein E in the development of obesity. FEBS J. 2009;276:5720–8.

48. Jethva R, Bennett MJ, Vockley J. Mini-Review: Short-Chain Acyl-Coenzyme A dehydrogenase deficiency. Mol. Genet. Metab. 2008;95:195–200.

Chapter V.

Investigating the genetic regulation of the expression of 63 lipid metabolism genes in the pig skeletal muscle

R. González-Prendes, R. Quintanilla, M. Amills

(Pending of minor revision in Animal Genetics).

Abstract

A comprehensive and systematic view about the genetic regulation of lipid metabolism genes is still lacking in pigs. Herewith, we have investigated the genetic regulation of 63 porcine genes with crucial roles in the uptake, transport, synthesis and catabolism of lipids. With this aim, we have performed an expression QTL (eQTL) scan in 104 pigs with available genotypes for the Illumina Porcine SNP60 BeadChip and microarray measurements of gene expression in the *gluteus medius* muscle. Analysis of the data with the GEMMA software revealed 13 *cis*- and 18 *trans*-eQTL modulating the expression of 19 loci. Genes regulated by eQTL participated in a wide array of lipid metabolism pathways such as the β-oxidation of fatty acids, lipid biosynthesis and lipolysis, fatty acid activation and desaturation, lipoprotein uptake, apolipoprotein assembly and cholesterol trafficking. These data provide a first picture about the genetic regulation of loci involved in porcine lipid metabolism.

Keywords: pigs, *gluteus medius*, muscle gene expression, quantitative trait loci, lipid metabolism

The search of regulatory variants with causal effects on the expression of genes with

important metabolic roles is fundamental to elucidate the genetic basis of multiple physiological and pathological phenotypes (Nica & Dermitzakis 2013). In humans, thousands of expression QTL (eQTL) have been detected so far (Nica & Dermitzakis 2013; GTEx Consortium 2015), and the majority of them appear to act locally (*cis*-eQTL) rather than influencing the expression of genes located at distant genomic regions or chromosomes (*trans*-eQTL). Moreover, around 50% of human *cis*-eQTL are shared across distinct tissues, though the consistency in the magnitude and the direction of these regulatory effects may be variable (GTEx Consortium 2015).

The genetic regulation of lipid metabolism genes has been poorly studied in pigs in spite of the fact that it may have a potential impact on the phenotypic variation of fatness traits. Indeed, the majority of eQTL studies performed in pigs have targeted either genes whose expression correlates with lipid phenotypes or loci comprised within the confidence intervals of fatness quantitative trait loci (Wimmers *et al*. 2010; Steibel *et al*. 2011; Cánovas *et al*. 2012; Heidt *et al*. 2013; Manunza *et al*. 2014). At present, we do not know if porcine lipid genes are predominantly regulated in *cis*- or *trans*- and if such regulation is featured by single or multiple polymorphisms. The goal of the current work was to shed light into these issues by identifying eQTL with effects on the muscle expression of 63 genes with an established role in the uptake, transport, synthesis and catabolism of lipids.

We used 104 barrows from a commercial Duroc porcine line (Lipgen population) distributed in five half-sib families. After weaning, this pig population was transferred to the experimental test station at the Centre de Control Porcí (CCP) of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). A detailed description of the experimental population and management conditions has been reported (Gallardo *et al*. 2008, 2009). Barrows were

slaughtered at an approximate age of 190 days. *Gluteus medius* (GM) muscle biopsies were obtained in the abattoir and they were immediately frozen in liquid nitrogen, being subsequently stored at -80 ºC. All animal care and management procedures followed the ARRIVE guidelines (Kilkenny *et al*. 2010) and they were approved by the Ethical Committee of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

GeneChip Porcine Genomic arrays (Affymetrix, Inc., Santa Clara, CA) were used to measure gene expression in *gluteus medius* samples from the 104 Duroc pigs mentioned above (data are available in the Gene Expression Omnibus public repository of the National Center for Biotechnology Information, accession number: GSE19275). Total RNA isolation and microarray hybridization procedures have been fully reported by González-Prendes et al. (2017). Microarray data were generated at two distinct laboratories (*i.e.* Vall d'Hebron University Hospital and Center for Research in Agricultural Genomics). Data pre-processing and normalization were carried out with the BRB-ArrayTools software version 3.7.1 (Xu *et al*. 2008). Genes displaying more than 20% of expression values over ± 1.5 times the median expression of all arrays were retained for further analysis (Cánovas *et al.* 2010). A detailed description of the techniques and methods used to perform RNA purification and microarray hybridization can be found in Cánovas *et al*. (2010). Finally, sixty three loci annotated in the Ensembl (*S.scrofa 10.2*) database and having a well established role in lipid metabolism (**Supplementary Table 1**) were selected for further analysis.

The Porcine SNP60 BeadChip (Illumina, San Diego, CA) was employed to genotype 62,163 single nucleotide polymorphisms (SNPs) in the 104 Duroc pigs by following a previously reported protocol (Manunza *et al*. 2014). Pig 60K genotypes have been deposited

in the Figshare public repository (https://dx.doi.org/10.6084/ m9.figshare.4263317). The GenomeStudio software (Illumina) was employed to evaluate the quality of the typing data. By using PLINK (Purcell *et al*. 2007), we discarded SNPs with rates of missing genotypes above 10%, minor allele frequencies (MAF) below 5%, as well as those that did not conform to Hardy-Weinberg expectations (threshold set at a *P*-value of 0.001). Markers that did not map to the porcine reference genome (*Sscrofa10.2* assembly) and those located in sex chromosomes were also eliminated from the data set. We also eliminated SNPs that were in strong linkage disequilibrium ($r^2 > 0.98$). After these filtering steps, a total of 28,571 SNPs were used to carry out a GWAS analysis for gene expression phenotypes.

Statistical analyses were performed with the GEMMA software (Zhou & Stephens 2012) by using a previously reported methodology (González-Prendes *et al*. 2017) . The GEMMA software uses a standard linear mixed model and an exact test of significance to identify associations between genotypes and gene expression phenotypes. The existence of population structure was taken into account by considering a relatedness matrix (Zhou & Stephens 2012). The model assumed in the statistical analysis was:

$$\mathbf{y} = \mathbf{W}\alpha + \mathbf{x}\beta + \mathbf{Z}\mathbf{u} + \varepsilon$$

where **y** is the vector of trait values for all individuals; **W** is a matrix of covariates *i.e.* "batch of fattening" (with 4 categories) and "laboratory" (microarray data were generated in two different laboratories); **α** is a vector of the corresponding coefficients including the intercept; **x** is a vector of genotypes of a marker; β is the effect size of the

marker; **u** is a vector of random individual effects with a n-dimensional multivariate normal distribution MVN_n (0, $\lambda \tau^{-1}$ K), where τ^{-1} is the variance of the residual errors; λ is the ratio between the two variance components and **K** is a known relatedness matrix derived from SNPs; and **ε** is a vector of errors. Correction for multiple testing was implemented with a false discovery rate approach (Benjamini & Hochberg 1995) and SNPs with a *q*-value ≤ 0.05 were considered as significantly associated with gene expression. As previously reported (González-Prendes *et al*. 2017), in the analysis of *cis*-eQTL we corrected for multiple testing by taking into consideration the number of SNPs contained within 2 Mb windows around each gene, while in the *trans*-eQTL analysis we took into account the whole set of 28,571 SNPs.

The eQTL scan for lipid-related genes identified 13 *cis*-eQTL and 18 *trans*-eQTL influencing the mRNA levels of 19 loci (**Tables 1** and **2, Supplementary Figure 1**). As shown in **Table 1**, the two *cis*-eQTL detected for the *ACOX3* (SSC8: 2.7-3.7 Mb and 4.4 Mb) and *NPC2* (SSC7: 102.5-103.1 Mb and 104.1-104.4 Mb) genes were located in adjacent positions and they might correspond to two genetic determinants (instead of 4). In a previous study, Chen *et al*. (2013) identified 120 *cis*-eQTLs and 523 *trans*-eQTLs with effects on porcine hepatic gene expression. However, they focused their study on a dataset of 300-400 genes that showed significant correlations with traits under study and their sample size was larger than ours.

Table 1. *Cis*-eQTLs regulating the expression of 11 genes involved in porcine lipid metabolism[1].

Genes			***Cis*-eQTL**									
Symbol	**SSC**	**Location (Mb)**	**SSC**	**N**	**SNP**	**Region (Mb)**	***P*-value**	***q*-value**	***B***	**δ ± SE**	**A_1**	**MAF**
ACADS	14	43.1	14	34	MARC0094155	42.6-45.9	0.00	0.00	0.00	-0.62 ± 0.08	G	0.21
ACOX3	8	4.3-4.4	8	2	M1GA0025674	2.7-3.7	0.00	0.02	0.03	-0.34 ± 0.10	G	0.36
				1	ALGA0118448	4.4	0.00	0.00	0.00	-0.75 ± 0.18	G	0.08
ACSF2	12	26.8	12	8	ALGA0065780	24.0- 26.9	0.00	0.00	0.00	0.85 ± 0.11	A	0.16
CITED2	1	28.2	1	4	MARC0028659	26.5-27.4	0.01	0.02	0.07	0.27 ± 0.10	G	0.38
HMGCS1	16	29.4	16	18	ALGA0089927	28.0-29.8	0.01	0.02	0.14	0.23 ± 0.07	A	0.35
LRP6	5	63.5-63.6	5	20	ASGA0025668	62.2-63.8	0.00	0.02	0.13	0.42 ± 0.13	G	0.40
LIPA	14	110.1	14	10	ASGA0065584	108.8-109.9	0.01	0.04	0.14	0.27 ± 0.11	A	0.19
NCOA1	3	121.2-121.3	3	3	MARC0003746	120.0-120.4	0.00	0.02	0.05	-0.28 ± 0.08	G	0.27
NPC2	7	103.5	7	6	ALGA0043923	102.5-103.1	0.00	0.00	0.00	0.33 ± 0.07	G	0.26
				3	INRA0027651	104.1-104.4	0.00	0.01	0.04	0.28 ± 0.09	G	0.28
SLC25A17	5	4.8	5	27	H3GA0015347	2.7-5.9	0.00	0.00	0.00	-0.79 ± 0.15	G	0.30
VLDLR	1	245.0	1	1	ASGA0005756	244.9	0.00	0.04	0.04	-0.33 ± 0.13	G	0.20

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with traits under study, SNP: SNPs displaying the most significant associations with traits under study, Region (Mb): regions containing SNPs significantly associated with traits under study, *P*-value: nominal *P*-value, *q*-value: q-value calculated with a false discovery rate approach, *B* : Bonferroni-corrected *P*-value, δ: allelic effect and its standard error (SE), A_1: minority allele, MAF: frequency of the minority allele.

Table 2. *Trans*-eQTLs regulating the expression of 12 genes involved in porcine lipid metabolism[1].

Genes			*Trans*-eQTLs									
Symbol	**SSC**	**Location (Mb)**	**SSC**	**N**	**SNP**	**Region (Mb)**	***P*-value**	***q*-value**	**B**	**δ ± SE**	**A_1**	**MAF**
ACADL	15	124.7	3	1	MARC0017993	144.3	0.00	0.03	0.08	-0.58 ± 0.13	C	0.18
				2	ALGA0123606	21.7-21.8	0.00	0.03	0.07	-0.58 ± 0.13	G	0.18
ACADM	6	127.5	9	1	MARC0004327	29.5	0.00	0.04	0.34	-0.56 ± 0.12	C	0.22
			13	7	DIAS0003141	141.6-144.1	0.00	0.04	0.21	-0.72 ± 0.15	G	0.12
ACADS	14	43.1	3	1	MARC0039787	134.6	0.00	0.01	0.19	-0.47 ± 0.11	A	0.18
			12	3	M1GA0017106	58.9-59.4	0.00	0.02	0.24	-0.54 ± 0.12	A	0.12
			14	3	H3GA0040210	53.7-55.5	0.00	0.00	0.02	-0.48 ± 0.08	A	0.28
			17	1	INRA0053259	28.7	0.00	0.05	0.97	-0.46 ± 0.09	G	0.45
ACSF2	12	26.8	12	4	MARC0030253	33.2-34.0	0.00	0.01	0.09	-0.49 ± 0.10	G	0.50
APOA1	9	49.2	1	3	MARC0004843	181.0-183.7	0.00	0.01	0.01	1.02 ± 0.19	A	0.07
CEBPD	4	87.3	7	1	ALGA0045624	128.5	0.00	0.04	0.04	1.30 ± 0.25	A	0.04
CMIP	6	7.1-7.2	5	2	ASGA0103424	12.4-12.7	0.00	0.04	0.36	0.59 ± 0.13	A	0.06
			13	7	DIAS0003141	141.6-144.1	0.00	0.00	0.00	0.58 ± 0.09	G	0.12
ECI2	7	2.5	12	1	MARC0021670	37.0	0.00	0.03	0.03	-0.62 ± 0.13	A	0.16
GPAT3	8	144.2	17	1	H3GA0049617	61.6	0.00	0.05	0.18	0.65 ± 0.14	A	0.22
LACTB	1	120.1	15	2	MARC0020666	3.2-3.4	0.00	0.05	0.08	-0.62 ± 0.14	G	0.13
LRP6	5	63.5-63.6	4	1	MARC0056621	134.9	0.00	0.03	0.03	-0.42 ± 0.08	A	0.47
SLC25A17	5	4.8	1	7	SIRI0000355	129.2-138.3	0.00	0.03	0.26	-0.64 ± 0.13	G	0.18

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with traits under study, SNP: SNPs displaying the most significant associations with traits under study, Region (Mb): regions containing SNPs significantly associated with traits under study, *P*-value: nominal *P*-value, *q*-value: q-value calculated with a false discovery rate approach, *B* : Bonferroni-corrected *P*-value, δ: allelic effect and its standard error (SE), A_1: minority allele, MAF: frequency of the minority allele.

In the current work, the numbers of *cis*- and *trans*-eQTL for lipid genes were quite similar (**Tables 1** and **2**). In contrast, Cánovas *et al*. (2012) performed a genome scan for porcine muscle expression phenotypes and observed a predominance of *trans*- *vs cis*-eQTL. The most likely reason for this discrepancy is that we have used different thresholds of significance to correct for multiple testing in the *cis*- and *trans*-eQTL analyses. Indeed, in humans the majority of eQTL identified so far act in *cis*-. For instance, a recent eQTL scan in 869 lymphoblastoid cell lines revealed that 3,534 and 48 genes were affected by eQTL in *cis*- and *trans*-, respectively (Bryois *et al*. 2014). Similarly, a global analysis of 53 human datasets demonstrated the existence of 116,563 high confidence eQTL. Around 91% and 9% of these eQTL acted in *cis*- and *trans*-, respectively (Zhang *et al*. 2014), and there was an average of 1.8 eQTL per gene.

The majority of *trans*-eQTL detected by us resided in chromosomes different than the one containing the targeted gene, suggesting that they may exert their effects through SNPs that alter the synthesis of a diffusible factor. We also observed the existence of several genes (*e.g. ACADS, ACSF2* and *SLC25A17*) simultaneously regulated by *cis*- and *trans*-eQTL (**Tables 1** and **2** and **Supplementary Figures 1** and **2**). For instance, the expression of the *ACADS* gene is regulated by one *cis*- and four *trans*-eQTL on SSC3, SSC12, SSC14, and SSC17 (**Supplementary Figure 1**). Of note, the *trans*-eQTL on SSC3 is defined by just one isolated SNP, so this result needs to be taken with caution. It might be argued that the *trans*-eQTL regulating the expression of *ACSF2* on SSC12 (33.2-34 Mb) is a "phantom" eQTL produced by the existence of a neighboring *cis*-eQTL on SSC12 (24-26.9 Mb). However, when we introduced the most significant SNP of the *cis*-eQTL in the statistical model as a fixed effect and repeated the analysis, the *trans*-eQTL on SSC12

(33.2-34 Mb) lost part of its significance (q-value > 0.05) but remained significant at the nominal level (**Supplementary Figure 3**).

Our findings illustrate that even simple phenotypes, such as gene expression, can be regulated in a highly complex manner. From a functional point of view, this set of 13 *cis*- and 18 *trans*-eQTL regulated the expression of genes integrated in distinct metabolic pathways. In this way, the acyl-coenzyme A dehydrogenases for short-chain (*ACADS*), medium-chain (*ACADM*) and long-chain (*ACADL*) FA catalyse the first step in the FA β-oxidation pathway (Kim & Miura 2004), and the enoyl-CoA delta isomerase 2 (*ECI2*) gene plays an essential role in the β-oxidation of unsaturated FA (Palosaari *et al*. 1990). Moreover, the solute carrier family 25 member 17 (*SLC25A17*) gene encodes a peroxisomal transporter of coenzyme-A, FAD and NAD^+ cofactors (Agrimi *et al*. 2012) and it could have a role in the α-oxidation of FA (Van Veldhoven 2010). We have also detected eQTL for genes comprised in lipid biosynthetic pathways (**Tables 1** and **2**). For instance, the glycerol-3-phosphate acyltransferase 3 (*GPAT3*) is involved in the synthesis of triacylglycerols (Yamashita *et al*. 2014), and the 3-hydroxy-3-methylglutaryl-CoA synthase 1 (*HMGCS1*) enzyme is a component of the cholesterol biosynthetic pathway (Medina & Krauss 2013). Other relevant loci are the acyl-CoA synthetase family member 2 (*ACSF2*) gene, which may participate in FA activation (Yang *et al*. 2009), the *LACTB* gene that affects adiposity in mice females (Yang. *et al*. 2009), the CCAAT/enhancer binding protein (C/EBP), δ (*CEBPD*) gene that has a key role in the regulation of adipogenesis (Hishida *et al*. 2009) and the Cbp/P300 interacting transactivator with Glu/Asp rich carboxy-terminal domain 2 (*CITED2*) locus that is involved in the regulation of hepatic gluconeogenesis (Sakai *et al*. 2012).

Our results demonstrate that around 30% of the lipid-related genes analysed in the current work are regulated by *cis*- and/or *trans*-eQTL with significant effects on their mRNA levels. In our data set, we have not detected a clear predominance of either *cis*- or *trans*-regulatory factors in the determination of gene expression, a result that contrasts with what has been obtained in humans where gene regulation is mostly exerted by *cis*-factors. In the next future, it would be worth to investigate if the set of eQTL detected herewith displays significant associations with the phenotypic variation of porcine traits of economic interest.

Acknowledgements

The authors are indebted to *Selección Batallé S.A.* for providing the animal material. Thanks also to Angela Cánovas for her advice in the bioinformatic analyses of gene expression. Part of the research presented in this publication was funded by grants AGL2013-48742-C2-1-R and AGL2013-48742-C2-2-R awarded by the Spanish Ministry of Economy and Competitivity and grant 2014 SGR 1528 from the Agency for Management of University and Research Grants of the Generalitat de Catalunya. We also acknowledge the support of the Spanish Ministry of Economy and Competitivity for the *Center of Excellence Severo Ochoa 2016-2019* (SEV-2015-0533) grant awarded to the Center for Research in Agricultural Genomics. Gonzalez-Prendes R. was funded by a FPU Ph.D. grant from the Spanish Ministerio de Educación (FPU12/00860). Thanks also to the CERCA Programme of the Generalitat de Catalunya.

References

Agrimi G., Russo A., Scarcia P. & Palmieri F. (2012) The human gene *SLC25A17* encodes a peroxisomal transporter of coenzyme A, FAD and NAD^+. *The Biochemical Journal* **443**, 241–7.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300.

Bryois J., Buil A., Evans D. M., Kemp J. P., Montgomery S. B., Conrad D. F., Ho K. M., Ring S., Hurles M., Deloukas P., Davey Smith G. & Dermitzakis E. T. (2014) *Cis* and *trans* effects of human genomic variants on gene expression. *PLoS Genetics* **10**, e1004461.

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

Cánovas A., Pena R. N., Gallardo D., Ramírez O., Amills M. & Quintanilla R. (2012) Segregation of regulatory polymorphisms with effects on the *gluteus medius* transcriptome in a purebred pig population. *PLoS One* **7**, e35583.

Chen C., Yang B., Zeng Z., Yang H., Liu C., Ren J. & Huang L. (2013) Genetic dissection of blood lipid traits by integrating genome-wide association study and gene expression profiling in a porcine model. *BMC Genomics* **14**, 848.

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

Gallardo D., Quintanilla R., Varona L., Díaz I., Ramírez O., Pena R. N. & Amills M. (2009) Polymorphism of the pig acetyl-coenzyme A carboxylase α gene is associated with fatty acid composition in a Duroc commercial line. *Animal Genetics* **40**, 410–7.

González-Prendes R., Quintanilla R., Cánovas A. Manunza A., Figueiredo Cardoso T., Jordana J., Noguera J. L., Pena R. N. & Amills M. (2017) Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. *Scientific Reports* 7, 39830.

GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–60.

Heidt H., Cinar M. U., Uddin M. J., Looft C., Jüngst H., Tesfaye D., Becker A., Zimmer A., Ponsuksili S., Wimmers K., Tholen E., Schellander K. & Große-Brinkhaus C. (2013) A genetical genomics approach reveals new candidates and confirms known candidate genes for drip loss in a porcine resource population. *Mammalian Genome* **24**, 416–26.

Hishida T., Nishizuka M., Osada S. & Imagawa M. (2009) The role of C/EBPδ in the early stages of adipogenesis. *Biochimie* **91**, 654–7.

Kilkenny C., Browne W. J., Cuthill I. C., Emerson M. & Altman D. G. (2010) Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biology* **8**, e1000412.

Kim J. J. & Miura R. (2004) Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences. *European Journal of Biochemistry* **271**, 483–93.

Manunza A., Casellas J., Quintanilla R., González-Prendes R., Pena R. N., Tibau J.,

Mercadé A., Castelló A., Aznárez N., Hernández-Sánchez J. & Amills M. (2014) A genome-wide association analysis for porcine serum lipid traits reveals the existence of age-specific genetic determinants. *BMC Genomics* **15**, 758.

Medina M. W. & Krauss R. M. (2013) Alternative splicing in the regulation of cholesterol homeostasis. *Current Opinion in Lipidology* **24**, 147–52.

Nica A. C. & Dermitzakis E. T. (2013) Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **368**, 20120362.

Palosaari P. M., Kilponen J. M., Sormunen R. T., Hassinen I. E. & Hiltunen J. K. (1990) Δ3, Δ2-enoyl-CoA isomerases. Characterization of the mitochondrial isoenzyme in the rat. Journal of Biological Chemistry 265, 3347–53.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A. R., Bender D., Maller J., Sklar P., de Bakker P. I. W., Daly M. J. & Sham P. C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Sakai M., Matsumoto M., Tujimura T., Yongheng C., Noguchi T., Inagaki K., Inoue H., Hosooka T., Takazawa K., Kido Y., Yasuda K., Hiramatsu R., Matsuki Y. & Kasuga M. (2012) *CITED2* links hormonal signaling to PGC-1α acetylation in the regulation of gluconeogenesis. *Nature Medicine* **18**, 612–7.

Steibel J. P., Bates R. O., Rosa G. J. M., Tempelman R. J., Rilington V. D., Ragavendran A., Raney N. E., Ramos A. M., Cardoso F. F., Edwards D. B. & Ernst C. W. (2011) Genome-wide linkage analysis of global gene expression in loin muscle tissue identifies candidate genes in pigs. *PloS One* **6**, e16766.

Van Veldhoven P. P. (2010) Biochemistry and genetics of inherited disorders of peroxisomal fatty acid metabolism. *Journal of Lipid Research* **51**, 2863–95.

Wimmers K., Murani E. & Ponsuksili S. (2010) Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. *Briefings in Functional Genomics* **9**, 251–8.

Xu X., Zhao Y. & Simon R. (2008) Gene set expression comparison kit for BRB-ArrayTools. *Bioinformatics* **24**, 137–9.

Yamashita A., Hayashi Y., Matsumoto N., Nemoto-Sasaki Y., Oka S., Tanikawa T. & Sugiura T. (2014) Glycerophosphate/Acylglycerophosphate Acyltransferases. *Biology* **3**, 801–30.

Yang X., Deignan J. L., Qi H., Zhu J., Qian S., Zhong J., Torosyan G., Majid S., Falkard B., Kleinhanz R. R., Karlsson J., Castellani L. W., Mumick S., Wang K., Xie T., Coon M., Zhang C., Estrada-Smith D., Farber C. R., Wang S. S., van Nas A., Ghazalpour A., Zhang B., Macneil D. J., Lamb J. R., Dipple K. M., Reitman M. L., Mehrabian M., Lum P. Y., Schadt E. E., Lusis A. J. & Drake T. A. (2009) Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nature Genetics* **41**, 415–23.

Zhang X., Gierman H. J., Levy D., Plump A., Dobrin R., Goring H. H., Curran J. E., Johnson M. P., Blangero J., Kim S. K., O'Donnell C. J., Emilsson V. & Johnson A. D. (2014) Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC Genomics* **15**, 532.

Zhou X. & Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44**, 821–4.

Chapter VI.

A genome-wide eQTL scan in the porcine muscle and liver reveals the existence of a significant tissue-specific genetic regulation

Rayner González-Prendes, Raquel Quintanilla, Anna Castelló, Ali Zidi,Yuliaxis Ramayo-Caldas, Tainã Figueiredo-Cardoso, Arianna Manunza, Ángela Cánovas Marcel Amills.

(Manuscript in revision)

Abstract

Background

The identification of expression QTL (eQTL) in different tissues is an essential step to understand how gene expression is genetically regulated in a context-dependent manner. In humans, the performance of eQTL scans across tissues has revealed that 50% of them are shared by multiple tissues. In pigs, most eQTL studies have been focused on genes whose expression correlates with phenotypic variation of traits of economic interest or that are located within QTL regions. In the current work, we aimed to compare the genome-wide eQTL landscape of the porcine skeletal muscle and liver as well as to investigate the co-localization of eQTL with copy number variant regions (CNVR).

Results

By performing genome scans in 104 Duroc pigs with available expression and genotypic data, markers associated with the mRNA levels of loci expressed in the *gluteus medius* muscle (436 *cis*-eQTLs and 450 *trans*-eQTLs) and liver (504 *cis*-eQTLs and 3,228 *trans*-eQTLs) have been found. Only 66 *cis*-eQTL (and none *trans*-eQTL) were shared between the GM muscle and liver tissues. This high proportion of tissue-specific QTL contrasts with data obtained in humans, where 50% of eQTL are shared across tissues. This could be due, at least in part, to technical factors related with the low density of the Porcine SNP60 BeadChip and the poor annotation of porcine microarrays. In addition, 104 CNVRs have been identified in 350 Duroc pigs typed with the Porcine SNP60 BeadChip. Approximately 39% of these CNVR co-localized with *cis*-eQTL signals, whilst the co-localization of

CNVR and *trans*-eQTL was somewhat higher (≈60%). In general, these co-localizations happened to be tissue-specific.

Conclusion

The effects of eQTL on porcine gene expression appears to be predominantly tissue-specific, but this result might be also the consequence of technical factors and needs to be confirmed in a broader sample of tissues. Moreover, the relative contributions of *cis*- and *trans*-eQTL to the regulation of gene expression is clearly different in the muscle and liver, possibly reflecting differences in transcript and functional complexity amongst tissues. The co-localization of CNVR and eQTL can be a first step to understand the role of structural variation on gene expression.

Background

The performance of GWAS in humans has revealed that most of the regions that display significant associations with complex traits are not exonic, meaning that causal polymorphisms probably have regulatory effects on gene expression [1]. This realization has prompted the mapping of expression QTL (eQTL) *i.e.* single nucleotide polymorphisms (SNP), indels or copy number variants (CNV) that explain part of the variance of gene expression phenotypes [1]. Such studies have revealed that the majority of eQTL exert their effects in *cis*- (*i.e.* on neighboring genes), though it is unclear if this is a biological reality or a statistical artifact [2]. There are also evidences that CNV-tagging SNPs are enriched in *cis*-eQTL and that they often modulate multiple expression traits [3]. By examining the patterns of expression of 22,286 genes in 9 human tissues, the GTEx Consortium has shown that approximately 50% of eQTL are shared by nine tissues and that most of them display consistent effects across tissues [4].

In pigs, hundreds of eQTL with effects on muscle [5–9], liver [10,11] and backfat [12] gene expression have been mapped. Often, these pig eQTL studies have targeted subsets of genes either mapping to QTL [12] or displaying significant expression-phenotype correlations [5,6,10,11]. There are conflicting results about the predominant role of either *cis*-eQTL [7] or *trans*-eQTL [8,13] on gene regulation. The broad majority of porcine eQTL studies have targeted single anatomic locations and, in consequence, they do not provide clues about the differential genetic regulation of distinct tissues and organs. Moreover, these genome scans have explored the association of gene expression with

allelic variation of SNPs or microsatellites, neglecting the potential effect of CNVs on expression phenotypes. The goals of the current work were to compare the genomic distribution of eQTL in the pig *gluteus medius* (GM) skeletal muscle and liver, two tissues with highly differentiated patterns of expression [14], as well as to investigate their co-localization with CNVs.

Materials and Methods

Phenotyping and genotyping of a commercial Duroc population

As animal material, we have used a commercial Duroc line of 350 Duroc pigs that were slaughtered at an age of 190 days, with an approximate live weight of 122 kg. This population was generated by crossing 5 boars with ~400 sows, and it was raised in the experimental testing station at the Centre de Control Porcí (CCP) of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). The specific conditions of management and feeding have been previously reported [15,16]. At slaughter, GM muscle and liver biopsies were obtained for 104 pigs. Total RNA purification, measurement of gene expression with GeneChip Porcine Genome microarrays and data pre-processing and normalization have been fully reported [8,17]. All experimental procedures were approved by the Ethical Committee of the IRTA.

Genomic DNA was extracted from blood samples by following a standard phenol-cloroform protocol. Each pig was genotyped for 62,163 single nucleotide polymorphisms (SNPs) with the Porcine SNP60 BeadChip (Illumina, SanDiego, CA). The quality of the genotyping results was evaluated with the GenomeStudio software (Illumina). The PLINK software [18] was used to filter SNP markers with minor allele frequencies below 5%, rates of missing genotypes above 10% as well as those did not conform Hardy-Weinberg expectations (threshold set at a *P*-value of 0.001). Markers that did not map to the porcine reference genome (Sscrofa10.2 assembly) and those located in sex chromosomes were also eliminated from the data set. Single nucleotide polymorphisms that were in complete linkage disequilibrium ($r^2 > 0.98$) were also discarded from further analyses. After these filtering steps, a subset of 28,571 SNPs were used as markers for eQTLs analysis. The filtering criteria in the analysis of CNV were different than those reported above *i.e.* only SNPs that did not map to the Sscrofa10.2 assembly or that were located in sex chromosomes were removed, so the final marker data set contained 46,537 SNPs.

Genome scan for expression QTL

The genome scan for eQTL regulating gene expression in the muscle and liver was carried out with the GEMMA software [19] following the methods described by Gonzalez-Prendes et al.[17]. Fixed effects and parameters assumed in the statistical model were:

$$\boldsymbol{y} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{x}\boldsymbol{\beta} + \mathbf{u} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{y}$ is the vector that defines the expression of each gene in the GM muscle and liver of the i^{th} individual; **W** is the matrix with a column of 1s and the fixed effects *i.e.* "batch of fattening" (with 4 categories) and "laboratory" (microarray data were generated in two different laboratories); $\boldsymbol{\alpha}$ is a c-vector of the corresponding coefficients including the intercept; **x** is an n-vector of marker genotypes; $\boldsymbol{\beta}$ is the SNP allelic effect estimated as a regression coefficient on the corresponding x genotype (values -1, 0, 1); **u** is an n-vector of random effects with a n-dimensional multivariate normal distribution MVN_n (0, $\lambda\ \tau^{-1}$ K) where τ^{-1} is the variance of the residual errors; $\boldsymbol{\lambda}$ is the ratio between the two variance components; **K** is a known relatedness matrix derived from SNPs and $\boldsymbol{\varepsilon}$ is the vector of errors with an MVN_n (0, $\tau^{-1}\ I_n$) being I_n the identity matrix. Correction for multiple testing was implemented with a false discovery rate approach [27]. In the case of *cis*-eQTL the correction for multiple testing was carried out by taking into account the number of SNPs located in a 2 Mb window around the targeted gene [17]. In contrast, the analysis of *trans*-eQTL took into consideration the whole set of SNPs mapping to QTL regions [17].

Detection of copy number variation

The PennCNV software [20] was employed to detect copy number variants (CNV) on the basis of the information provided by 46,537 autosomal SNPs. We did not employ multiple softwares to detect CNVs (and select as true those CNVs identified by two or more softwares) because such practice decreases type 1 error at the expense of substantially increasing type 2 error. The PennCNV software implements a hidden Markov model to infer CNV calls for each genotyped sample using as input the intensity signal Log R Ratio and the B Allele Frequency information generated with the BeadStudio (Illumina) software. Samples with a standard deviation of LRR > 0.30 and BAF drift > 0.01 were discarded. Besides, a wave adjustment procedure for genomic waves was carried out [20]. With these filtering steps, 20 samples were eliminated from the data set. Only CNVs spanning three or more consecutive SNPs were taken into account. Copy number variant regions (CNVR) were created by merging CNVs with an overlap of 80% or more.

Validation of copy number variant regions by quantitative PCR

Quantitative real time PCR (qPCR) assays were used to validate eight CNVR (positive controls) and three putative single-copy genomic regions (negative controls). The

relative quantification (RQ) of the CNVRs was done as previously described [21]. Primers (**Supplementary Table 1**) were designed with the Primer Express Software (Applied Biosystems). Copy number variant regions were quantified in 384-well plates using SYBR Select Master Mix in a QuantStudio 12K Flex Real-Time PCR System platform (Applied Biosystems, Inc., Foster City, CA). Reactions were performed in triplicate, and they contained 7.5 ng genomic DNA and primers at 300 nM in a final volume of 15 μl. The thermocycling profile was: one cycle at 95 °C for 10 min plus 40 cycles of 15 sec at 95 °C and 1 min at 60 °C. Moreover, a melting curve profile (95 ºC for 15 sec, 60 ºC for 15 sec and a gradual increase in temperature with a ramp rate of 1% up to 95 ºC) was implemented to maximize the specificity of the amplification reactions. Relative expression values (RQ) were calculated with the Qbase+ software (Biogazelle, Ghent, Belgium) by applying the $2^{-\Delta\Delta Ct}$ method, after verifying that its assumptions were adequately fulfilled [22]. Relative expression values were calibrated using the arithmetic mean of 3-5 samples showing the lowest number of copies for each specific assay. In the specific case of CNVR87, which encompasses a deletion, the 5 samples chosen for calibration were those with RQ values around 2. Normalization of expression data was done by using a previously reported assay based on the glucagon gene [23].

Results

Detection of expression QTL in the porcine skeletal muscle and liver

A total of 436 *cis*-eQTLs and 450 *trans*-eQTLs regulating the expression of 449 and 319 genes in the GM muscle were identified, respectively (**Table 1, Supplementary Tables 2** and **3**). The proportion of *cis*- and *trans*-eQTLs in the liver was much more unbalanced, with 504 *cis*-eQTLs and 3,228 *trans*-eQTLs modulating the transcription of 525 and 1,902 genes, respectively (**Table 2, Supplementary Tables 4** and **5**). Global analysis of the data showed that the number of *cis*-eQTL per gene (0.96-0.97) was lower than the number of *trans*-eQTL per gene (1.41-1.69). Besides, 90 and 199 loci were simultaneously regulated by *cis*- and *trans*-eQTLs in muscle and liver, respectively (**Supplementary Table 6**, **Figure 1**). The level of overlap between the muscle and liver eQTL data sets was modest, with 66 *cis*-eQTLs (**Figure 2**) and none *trans*-eQTL shared by both tissues. Several examples of muscle- and liver-specific eQTL are shown in **Figure 3**.

Table 1. List of the most significant *cis*- and *trans*-eQTLs detected in the gluteus *medius* (GM) muscle.

Muscle *cis*-eQTL										Gene		
SSC	**N**	**SNP**	**Region (Mb)**	**P-value**	***q*-value**	***B***	**δ±SE**	**A_1**	**MAF**	**Acronym**	**SSC**	**Region (Mb)**
1	27	DRGA0000100	13.2-15.9	0.00	0.00	0.00	0.79±0.12	G	0.39	*CNKSR3*	1	14.4-14.4
1	32	SIRI0000064	246.1-249.8	0.00	0.00	0.00	-1.07±0.12	G	0.06	*FXN*	1	248.6-248.8
2	13	INRA0008800	65.1- 67.7	0.00	0.00	0.00	0.77±0.11	C	0.15	*CALR*	2	66.2-66.2
2	13	ASGA0008809	6.7- 7.8	0.00	0.00	0.00	-0.56±0.08	A	0.49	*TRPT1*	2	6.9-6.9
3	12	MARC0026232	140.0-141.8	0.00	0.00	0.00	0.60±0.06	A	0.37	*ADI1*	3	140.4-140.4
7	20	MARC0077571	46.3- 47.9	0.00	0.00	0.00	-0.69±0.08	C	0.09	*CLIC5*	7	46.9-47.0
8	11	ASGA0101414	122.3-124.1	0.00	0.00	0.00	0.87±0.09	C	0.18	*AIMP1*	8	123.8-123.8
11	20	ASGA0050460	24.0- 27.7	0.00	0.00	0.00	0.73±0.10	G	0.16	*AKAP11*	11	25.4-25.4
14	34	MARC0094155	42.6- 45.9	0.00	0.00	0.00	-0.62±0.08	G	0.21	*ACADS*	14	43.1-43.1
18	18	ASGA0078711	4.3- 5.8	0.00	0.00	0.00	0.96±0.13	G	0.42	*ACTR3B*	18	4.9-4.9
Muscle *trans*-eQTL										**Gene**		
SSC	**N**	**SNP**	**Region(Mb)**	**P-value**	***q*-value**	***B***	**δ±SE**	**A_1**	**MAF**	**Acronym**	**SSC**	**Region (Mb)**
1	5	ASGA0089146	308.9-309.1	0.00	0.00	0.00	0.56±0.08	A	0.48	*RASIP1*	6	49.8-49.8
3	3	ALGA0020311	40.1-42.9	0.00	0.00	0.00	-0.72±0.10	A	0.33	*PREPL*	3	102.2-102.3
4	12	MARC0027501	104.4-109.0	0.00	0.00	0.00	0.62±0.07	G	0.36	*CU207250.1*	4	47.6-47.6
4	1	ASGA0083736	106.3-106.3	0.00	0.00	0.00	0.60±0.08	G	0.45	*UGPP*	7	131.2-131.2
5	1	INRA0019588	46.7-46.7	0.00	0.00	0.00	0.53±0.08	C	0.42	*EMG1*	5	66.1-66.1
5	1	MARC0056503	34.5-34.5	0.00	0.00	0.00	0.63±0.07	G	0.45	*UGPP*	7	131.2-131.2
8	1	ALGA0048824	49.7-49.7	0.00	0.00	0.00	0.51±0.06	A	0.43	*NUDT6*	8	108.4-108.4
8	17	ALGA0115175	93.0-99.7	0.00	0.00	0.00	0.48±0.06	A	0.45	*NUDT6*	8	108.4-108.4
8	8	MARC0049164	114.3-118.8	0.00	0.00	0.00	0.51±0.06	G	0.39	*NUDT6*	8	108.4-108.4
10	1	MARC0112823	14.9-14.9	0.00	0.00	0.00	-1.42±0.15	A	0.36	*IGKC*	3	59.8-59.8

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with the trait under study, SNP: SNP displaying the most significant association with the trait under study, Region (Mb): region containing SNPs significantly associated with the trait under study, P-value: nominal P-value, q-value: q-value calculated with a false discovery rate approach, *B*: Bonferroni corrected P-values, δ: allelic effect and its standard error (SE), A1: minority allele, MAF: frequency of the minority allele.

Table 2. List of the most significant *cis*- and *trans*-eQTLs detected in the liver tissue.

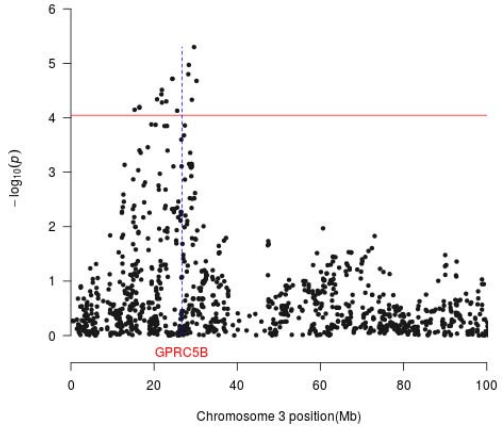
Hepatic *cis*-eQTLs										**Gene**		
SSC	**N**	**SNP**	**Region (Mb)**	**P-value**	**q-value**	***B***	**δ±SE**	**A_1**	**MAF**	**Acronym**	**SSC**	**Region (Mb)**
1	1	MARC0014312	146.3-146.3	0.00	0.00	0.00	2.62±0.58	A	0.46	*IVD*	1	145.9-145.9
1	8	ASGA0006349	266.9-269.2	0.00	0.03	0.12	-0.12±0.04	A	0.43	*ANP32B*	1	268.0-268.0
1	12	MARC0030335	270.3-271.9	0.00	0.00	0.00	1.49±0.29	A	0.33	*TMEFF1*	1	270.5-270.6
2	3	ASGA0103706	11.8-13.3	0.00	0.00	0.00	-2.32±0.40	A	0.05	*MED19*	2	12.8-12.8
3	21	MARC0055489	30.0-33.8	0.00	0.00	0.00	-1.46±0.38	A	0.19	*LITAF*	3	32.3-32.3
5	9	H3GA0016069	22.5-24.0	0.00	0.00	0.00	0.45±0.08	G	0.13	*PTGES3*	5	23.6-23.6
6	10	MARC0030904	84.2-86.9	0.00	0.00	0.00	-1.03±0.24	A	0.13	*INPP5B*	6	86.6-86.7
7	13	H3GA0019664	4.2-5.8	0.00	0.00	0.00	-1.70±0.38	G	0.21	*BLOC1S5*	7	5.4-5.4
13	18	ALGA0071453	96.0-99.5	0.00	0.00	0.00	1.06±0.21	C	0.19	*TSC22D2*	13	98.7-98.7
14	6	MARC0069598	24.1-25.9	0.00	0.00	0.00	0.38±0.08	A	0.50	*ADGRD1*	14	25.5-25.7
Hepatic *trans*-eQTLs										**Gene**		
SSC	**N**	**SNP**	**Region (Mb)**	**P-value**	**q-value**	***B***	**δ±SE**	**A_1**	**MAF**	**Acronym**	**SSC**	**Region (Mb)**
1	7	MARC0058870	34.9-36.6	0.00	0.00	0.00	-3.77±0.32	A	0.01	*FAM63A*	4	107.4-107.4
	4	DRGA0001450	120.2-121.8	0.00	0.00	0.00	-0.67±0.07	G	0.01	*APOA1*	9	49.2-49.2
	3	DRGA0001450	120.2-120.8	0.00	0.00	0.00	4.32±0.50	G	0.01	*COX6A1*	14	42.8-42.8
	4	DRGA0001450	120.2-122.4	0.00	0.00	0.00	-1.02±0.11	G	0.01	*BCAS4*	17	58.6-58.7
	4	DRGA0001450	120.2-121.8	0.00	0.00	0.00	3.90±0.43	G	0.01	*INSIG1*	18	2.8-3.0
2	4	ASGA0084177	0.16-0.36	0.00	0.00	0.00	-2.18±0.25	G	0.01	*CALCOCO1*	5	19.3-19.3
	4	ALGA0111915	162.0-162.2	0.00	0.00	0.00	-2.18±0.25	G	0.01	*CALCOCO1*	5	19.3-19.3
	4	ALGA0111915	162.0-162.2	0.00	0.00	0.00	-3.87±0.43	G	0.01	*LEPROTL1*	15	62.2-62.2
3	6	MARC0001269	22.8-27.9	0.00	0.00	0.00	-3.44±0.56	G	0.16	*CBX1*	12	24.2-24.2
14	13	H3GA0042707	140.7-141.7	0.00	0.00	0.00	-0.33±0.04	G	0.06	*ESD*	11	20.9-21.0

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with the trait under study, SNP: SNP displaying the most significant association with the trait under study, Region (Mb): region containing SNPs significantly associated with the trait under study, P-value: nominal P-value, q-value: q-value calculated with a false discovery rate approach, *B*: Bonferroni corrected P-values, δ: allelic effect and its standard error (**SE**), A_1: minority allele, MAF: frequency of the minority allele.

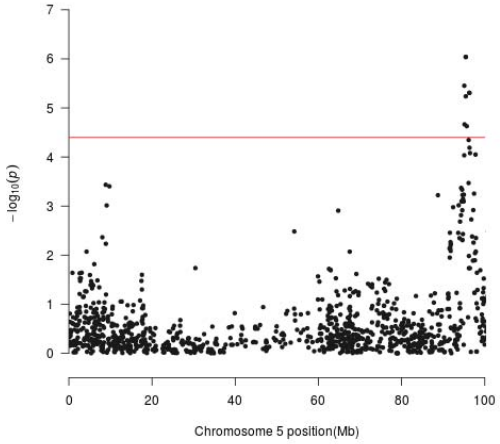
Figure 1a.

***gluteus medius* muscle**

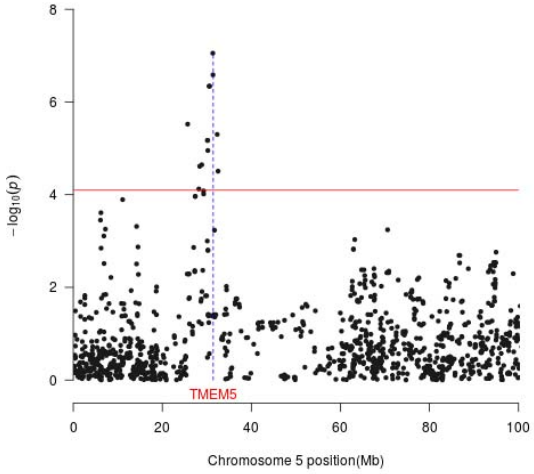
***cis*-eQTL for *GPRC5B* mRNA**



***trans*-eQTL for *GPRC5B* mRNA on SSC5**



***cis*-eQTL for *TMEM5* mRNA**



***trans*-eQTL for *TMEM5* mRNA on SSC15**

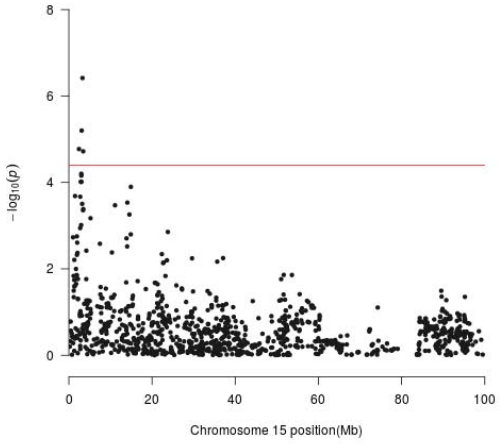
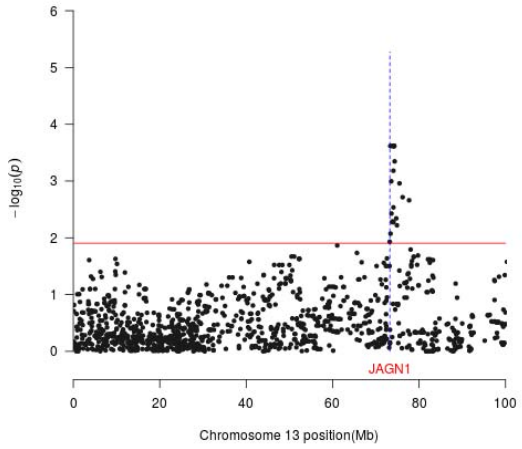


Figure 1a. Examples of genes that are simultaneously regulated by *cis*- and *trans*-eQTL in the *gluteus medius* muscle. In the Manhattan plots, the horizontal line indicates the threshold of significance (q-value ≤ 0.05) whilst the vertical line depicts the genomic location of the genes (*GPRC5B*, *TMEM5*) under consideration.

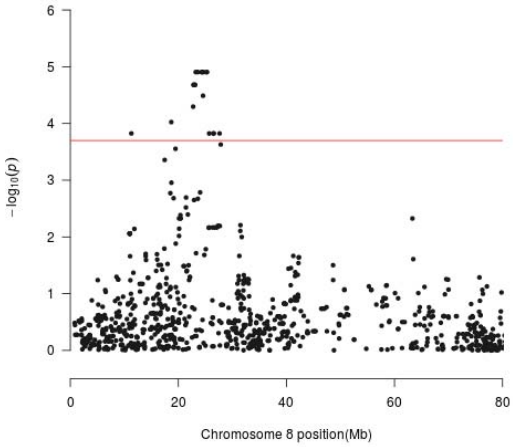
Figure 1b.

Hepatic tissue

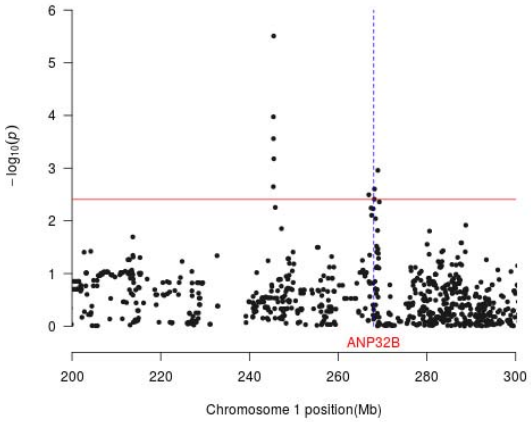
***cis*-eQTL for *JAGN1* mRNA**



***trans*-eQTL for *JAGN1* mRNA on SSC8**



***cis*-eQTL for *ANP32B* mRNA**



***trans*-eQTL for *ANP32B* mRNA on SSC18**

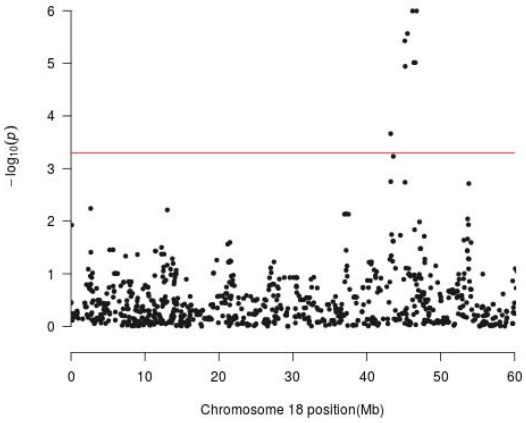


Figure 1b. Examples of genes that are simultaneously regulated by *cis*- and *trans*-eQTL in the liver. In the Manhattan plots, the horizontal line indicates the threshold of significance (q-value ≤ 0.05) whilst the vertical line depicts the genomic location of the genes (*JAGN1*, *ANP32B*) under consideration.

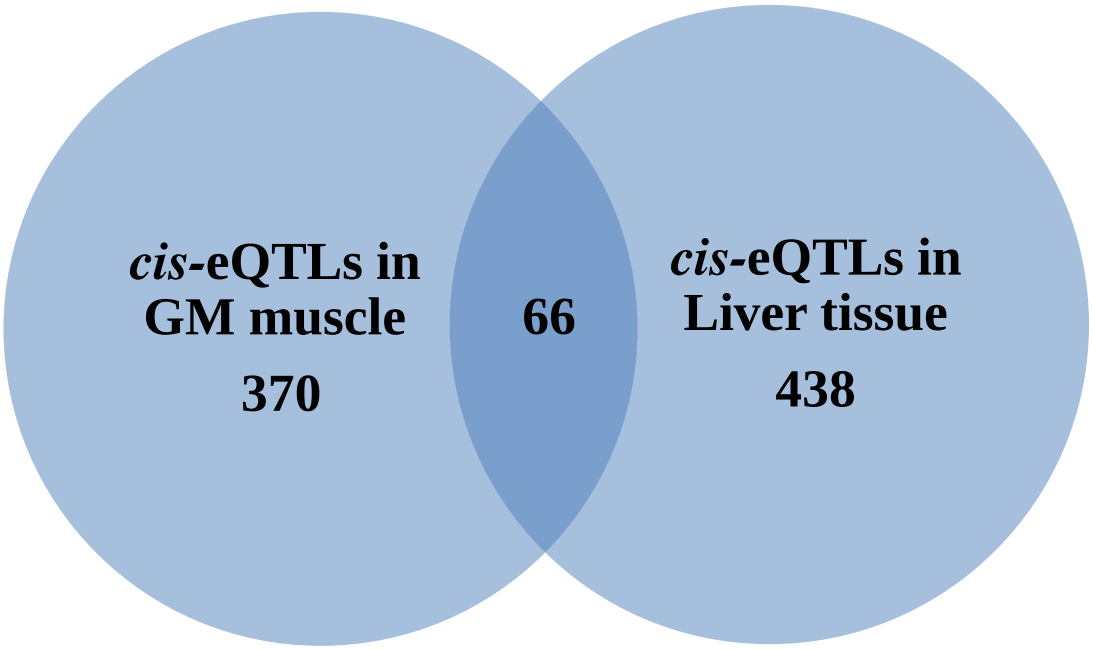
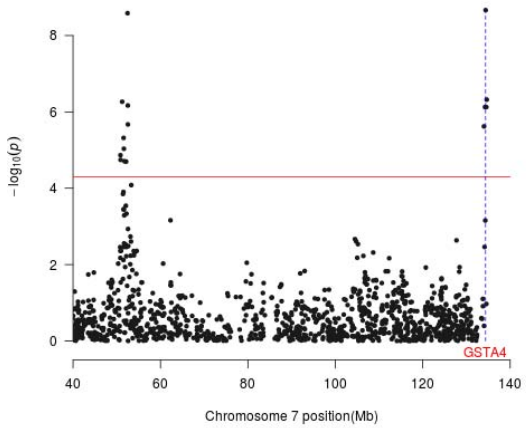


Figure 2. Venn diagram showing the number of *cis*-eQTLs shared by the *gluteus medius* muscle and liver.

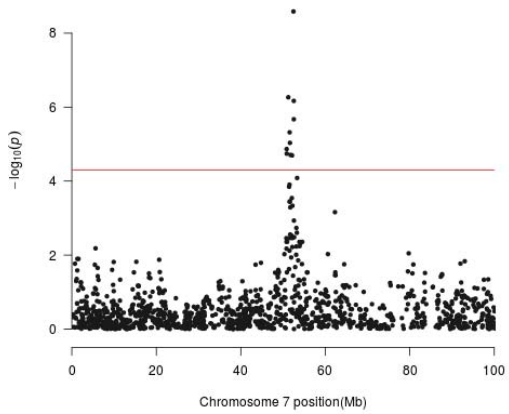
Figure 3a.

***gluteus medius* muscle**

***cis*-eQTL for *GSTA4* mRNA**

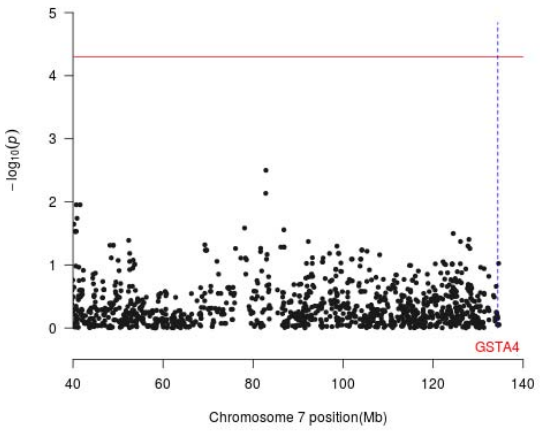


***trans*-eQTL for *GSTA4* mRNA on SSC7**



Hepatic tissue

***no cis-eQTL for GSTA4* mRNA**



no trans-eQTL for GSTA4 mRNA

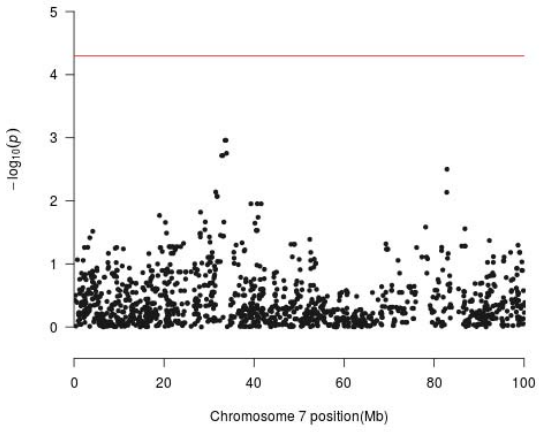
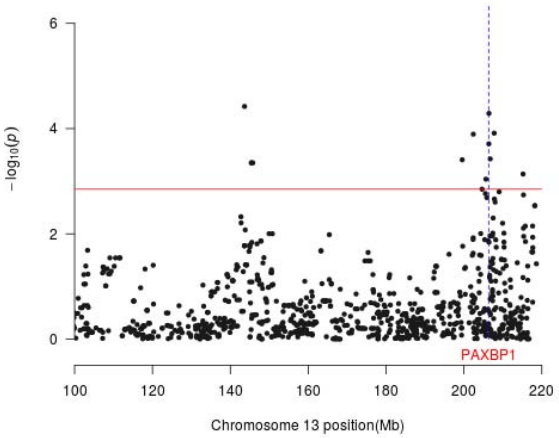


Figure 3a. Examples of *cis*- and *trans*-eQTL that are found in the muscle but not in the liver (a). In the Manhattan plots, the horizontal line indicates the threshold of significance (q-value ≤ 0.05) whilst the vertical line depicts the genomic location of the *GST4* gene.

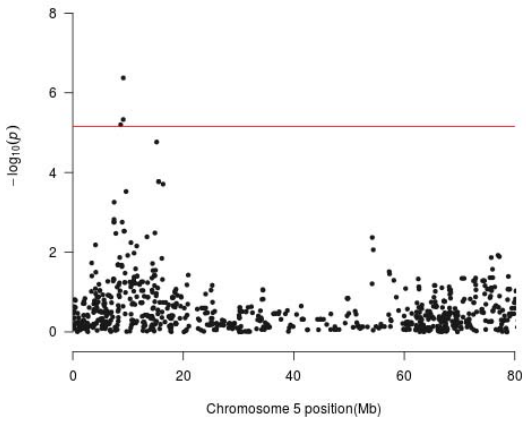
Figure 3b.

Hepatic tissue

***cis*-eQTL for *PAXBP1* mRNA**

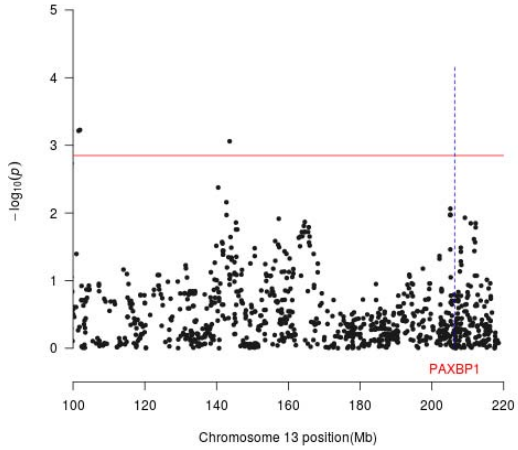


***trans*-eQTL for *PAXBP1* mRNA on SSC5**



***gluteus medius* muscle**

***no cis-eQTL for PAXBP1* mRNA**



***no trans-eQTL for PAXBP1* mRNA**

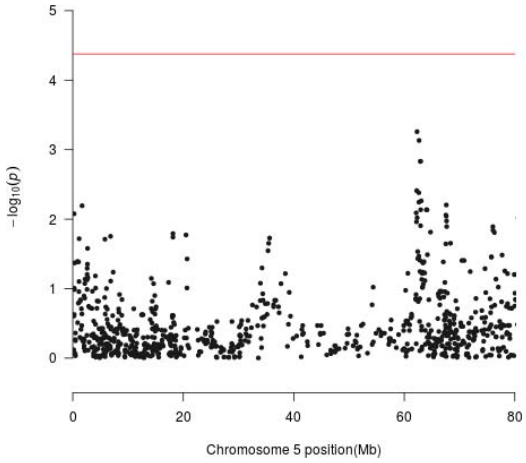


Figure 3b. Examples of *cis*- and *trans*-eQTL that are found in liver tissue but not in the muscle (b). In the Manhattan plots, the horizontal line indicates the threshold of significance (q-value ≤ 0.05) whilst the vertical line depicts the genomic location of the *PAXBP1* gene.

Cytoscape package [24] and Reactome [25] database were used to identify pathways to which genes regulated in *cis*- and *trans*- belong to. The number of pathways detected for eQTL regulated genes expressed in the muscle tissue was lower than those detected for liver genes (**Supplementary Tables 7-10**). This result could be anticipated because the number of genes regulated by eQTL was higher in the liver and, besides, the functions of this organ are more diverse and complex than those of skeletal muscle [26]. The biological functions of eQTL regulated genes were quite diverse, but several of them were integrated in metabolic pathways with a potential impact on phenotypes that are important for the pork industry (**Supplementary Tables 7-10**) *e.g.* amino acid metabolism (*cis*-eQTL in the muscle and *trans*-eQTL in the liver), β-oxidation of saturated and unsaturated fatty acids (*trans*-eQTL in the muscle), metabolism of lipids and lipoproteins (*cis*- and *trans*-eQTL in the liver), transcriptional regulation of white adipocyte differentiation (*cis*-eQTL in the liver), regulation of lipid metabolism by peroxisome proliferator-activated receptor α (*cis*-eQTL in the liver), fatty acid, triacylglycerol and ketone body metabolism (*cis*-eQTL in the liver), metabolism of proteins (*trans*-eQTL in the liver) and sphingolipid metabolism (*trans*-eQTL in the liver).

Co-localization of copy number variants and eQTL

The analysis of structural variation with PennCNV revealed the existence of 1,126 CNVs distributed in 16 pig chromosomes that were assembled into 104 CNVR regions (**Supplementary Table 11**). These CNVRs covered 21% of the pig genome. The

proportions of *copy gain*, *loss* and *loss and gain* CNVRs were 39%, 49% and 12%, respectively. The size of the CNVR ranged from 10.1 kb to 5.5 Mb, with a mean of 331.6 kb. When we compared our CNVR dataset with CNVs previously reported in pigs [27–36] we found that 47% of the CNVR detected by us had been previously reported in the literature (**Supplementary Table 12**). Real time quantitative assays were designed for validating 8 CNVRs (CNVR 21, 23, 32, 54,74,81,87 and 97) and 3 single-copy regions (negative control) in 39 porcine samples. The validation of the 8 CNVR by qPCR showed clear evidence of the existence of copy number variation for 4 of them (CNVR21, 74, 87 and 97), while for the remaining 4 evidence was less conclusive (**Figure 4**, **Supplementary Figure 1**). More unexpectedly, one of the putative single copy regions chosen as negative controls showed strong evidence of containing a CNV, while the status of the remaining two was less obvious (**Supplementary Figure 2**).

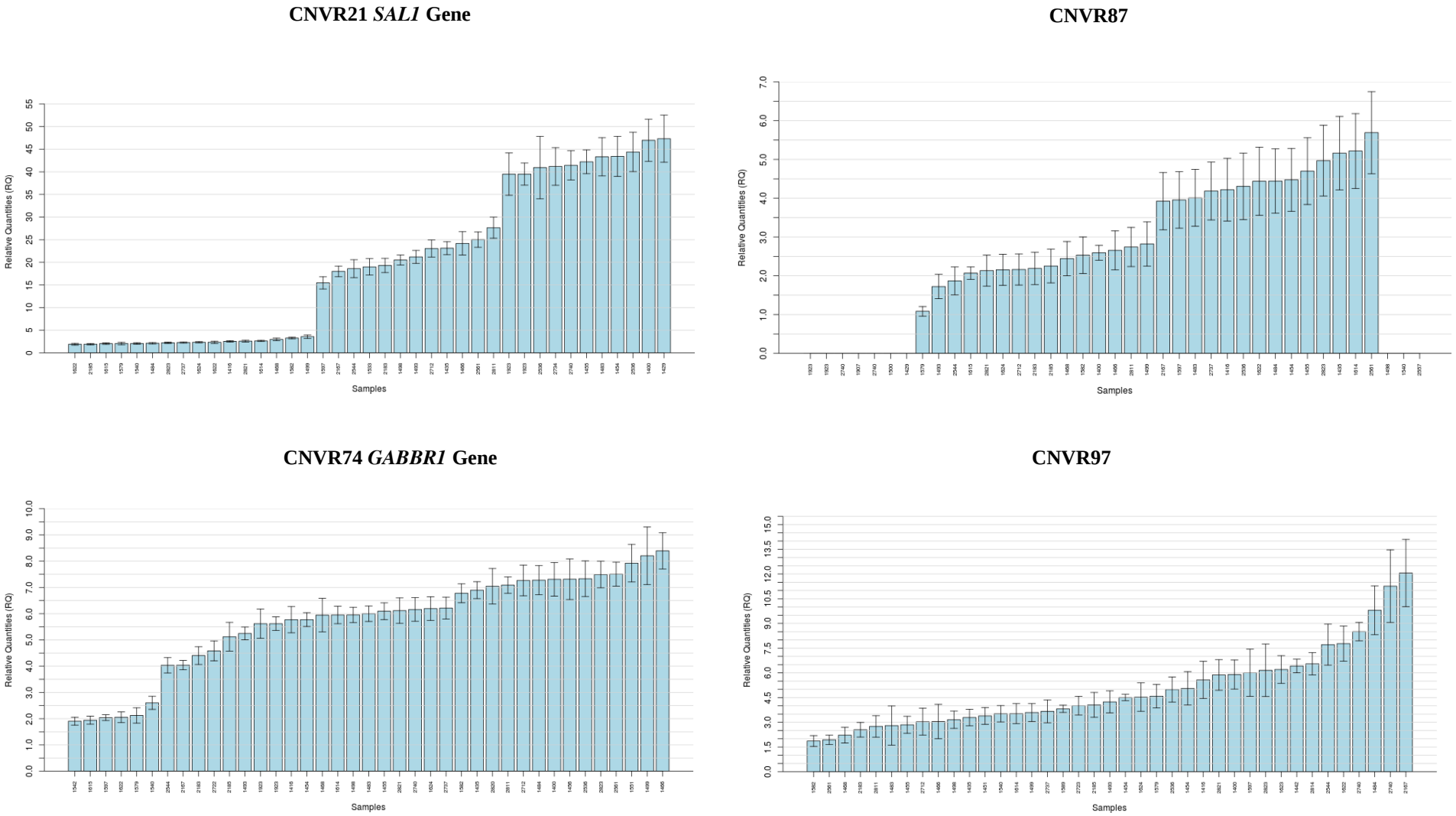


Figure 4. Relative expression values of four high-confidence copy number variation regions validated by quantitative real-time PCR analysis. Each analysed individual is represented in the x-axis, while the y-axis shows the corresponding relative quantification (RQ) value. We have assigned a value of 2 to the arithmetic mean of the samples used as calibrators.

The co-localization of CNVR and eQTLs was also analyzed (**Supplementary Tables 13-16**). In the GM muscle, 28 and 16 CNVR co-localized with 45 *cis*-eQTLs and 16 *trans*-eQTL, respectively. In the liver, 28 and 59 CNVR co-localized with 42 *cis*-eQTLs and 139 *trans*-eQTLs. When the co-localizations between CNVR and *cis*-eQTL were compared in the muscle and the liver tissues, both data sets had 13% of CNVRs in common, but these CNVRs co-localized with different sets of *cis*-eQTL (**Table 3, Supplementary Tables 13-16**). Only six genes (*HSDL2*, *PREPL*, *PACSIN2*, *EMG1*, *RASIP1* and *KCTD20*) located within CNVRs were consistently *cis*-regulated both in the GM muscle and liver (**Table 3**). No genes mapping to CNVR and being *trans*-regulated by the same eQTL in the liver and GM muscle were found.

Table 3. List of CNVRs that co-localize with *cis*-regulated genes in the muscle and the liver (genes shared by both tissues are shown in bold).

CNVR	Genes regulated by muscle *cis*-eQTL	Genes regulated by liver *cis*-eQTL
21	*PTBP3*, ***HSDL2***, *SLC31A2*	***HSDL2***
26	*CST6*, *PACS1*	*PPP1CA*
27	*SNORD26*, *SNORD27*, *snR56*, *METTL12*, *WDR74*	*RAB3IL1*, *TKFC*
46	***PREPL***	*PKDCC*, ***PREPL***
47	*ZNF7*, *ZNF251*	*SHARPIN*
53	*SLC25A17*, ***PACSIN2***, *PARVB*	*ACO2*, ***PACSIN2***
55	***EMG1***, *GABARAPL1*	*KLRK1*, ***EMG1***, *CHD4*
67	*CH242-204P3.4*, ***RASIP1***, *PTOV1*, *BCAT2*, *GYS1*	***RASIP1***, *SULT2A1*, *POLD1*
69	*FUCA1*	*TRNP1*, *GALE*
75	***KCTD20***	***KCTD20***
76	*RCN2*, *ANPEP*	*SIN3A*, *SCAMP5*, *STOML1*, *SCAMP5*, *EDC3*
89	*SLC25A39*, *EFTUD2*	*NMT1*, *HEXIM1*
101	*ATIC*	*FN1*
103	*ZNF622*	*FAM134B*

Discussion

Skeletal muscle and liver have been selected as target tissues to perform a comparative eQTL analysis because they have a key role in body energy homeostasis, a parameter that has a strong impact on growth and fat deposition in pigs. Moreover, the patterns of expression of these two tissues are highly differentiated in pigs, probably as consequence of their distinct embryonic origin and physiological function [14]. The number of *trans*-eQTL happened to be much higher in the liver than in the GM muscle, while the number of *cis*-eQTL were similar in both tissues. In this regard, skeletal muscle, as well as blood and heart, is a good example of a tissue whose transcription is dominated by the expression of a reduced set of genes [37]. This circumstance decreases to a significant extent the power to detect eQTL [37]. Moreover, the biological roles of the liver (detoxification, bile production, metabolism, red blood cell destruction, etc.) are more diverse than those of the skeletal muscle (locomotion and metabolism); a feature that often is associated with a higher transcript complexity and a tighter genetic regulation.

In pigs, hundreds of eQTL with effects on the muscle transcriptome have been detected but the majority of these studies have targeted subsets of genes either mapping to QTL [12] or displaying significant expression-phenotype correlations [5,6,10,11]. By analyzing global gene expression, we have found that the numbers of *cis*- and *trans*-eQTL in the pig GM muscle are remarkably similar, while hepatic *trans*-eQTL are five times more abundant than liver *cis*-eQTL. Liaubet et al. [13] made a genome scan based on microarray measurements of *longissimus lumborum* gene expression in 57 pigs and

identified 335 eQTL, of which only 18 had *cis*-regulatory effects. Similarly, Cánovas et al. [8] identified 613 skeletal muscle genome-wide significant eQTL and only 13% acted in *cis*-. In humans, conversely, the broad majority of eQTL detected so far modulate gene expression in *cis*-. For instance, Bryois et al. [38] performed an association analysis of SNPs and CNVs with gene expression measured in 869 lymphoblastoid cell lines and found that 3,534 and 48 genes are affected by *cis*- and *trans*-eQTL, respectively. Zhang et al. [39] analyzed 53 eQTL data sets and identified 116,563 high-confidence eQTL, but only 9% acted in *trans*. Paradoxically, 60-75% of the heritability of gene expression in humans is explained by *trans*-effects [2], so the low abundance of *trans*-eQTL in humans might be due to the fact that the majority of studies have a low statistical power to detect them (correction for multiple testing is much more stringent in *trans*-eQTL than in *cis*-eQTL and usually *trans*-eQTL have subtler effects than *cis*-eQTL). In yeast, *trans*-regulation predominates over *cis*-regulation [40] and in Drosophila both types of eQTL have a similar prevalence [41]. Differences amongst species related with the density of genotyping platforms, functional annotation of genes and linkage disequilibrium patterns might also explain these discrepancies.

Our results indicate that the sharing of *cis*- and *trans*-eQTLs between pig liver and muscle is low (**Figures 2 and 3**). This would imply that the genetic regulation of gene expression in pigs has a strong tissue-specific component. However, if we had examined more tissues and we had used a chip with a higher SNP density and RNA-seq instead of microarrays, we would have probably found a higher level of concordance between tissues. In the Genotype-Tissue Expression (GTEx) pilot experiment [4], about 50% of eQTL were shared by the nine human tissues under analysis. Moreover, there were two main types of

eQTL *i.e.* those that regulate gene expression in a single tissue and those that are ubiquitously detected in all tissues. Interestingly, the GTEx pilot analysis also showed that eQTL affecting gene expression in the skeletal muscle show a limited replicability in other tissues [4]. In other studies also performed in humans, the tissue-specificity of eQTL took values between 50% [42] and 60-80% [43], implying that the effects of many regulatory mutations are modulated by tissue-associated factors.

Another goal of our study was to investigate the co-localization of eQTL and CNVR in order to make an initial assessment of the potential impact of structural variation on gene expression in pigs. We were able to detect 104 CNVR with an average size of 331 kb, probably reflecting the inability of the Porcine SNP60 BeadChip to detect small CNVs. About 47% of the CNVR detected by us showed positional concordance with porcine CNVRs reported in previous publications [21,27–29,31,33,33–36] (**Supplementary Table 12**). A moderate agreement of CNVR locations amongst studies and populations has been evidenced in several reports [28,36,44–46]. Discrepancies could be due to differences in the genetic background of populations under analysis, filtering criteria (correction factors, criteria to define CNV and CNVRs, etc), genotyping methods, CNV calling algorithms and the use of family information [36,47].

The validation of our CNVR data set by qPCR revealed that 50% of them could be classified as high confidence CNVR (**Figure 4, Supplementary Figure 1)**. On the other hand, one of the three negative control (non-CNVR) regions also happened to be a high-confidence CNVR (**Supplementary Figure 2)**, indicating that our rate of false negatives could be high. Zhang et al. [39] compared the ability of the Birdsuite, Partek, HelixTree, and PennCNV-Affy to call 893 CNVs detected with the Affymetrix Genome-Wide Human

SNP Array 6.0 and validated, in eight HapMap samples, by paired-end sequencing of whole-genome fosmid clones. The averaged recovery rates of these four programs fluctuated between 7.5% (CNVs defined by 2-5 markers) to 54% (CNVs defined by more than 20 markers). The sensitivity of the Porcine SNP60 BeadChip is probably lower than that of chips employed in humans because of its modest SNP density, a feature that decreases very substantially the power to detect small CNVs that in the pig genome are more abundant than the large ones [28,47]. Moreover, SNPs residing in CNVs are often excluded from genotyping platforms owing to Mendelian inconsistency among families, lack of Hardy–Weinberg equilibrium and high missing genotype rates [48]. In summary, the identification of structural variants in domestic species is still very challenging and results need to be interpreted with caution.

Keeping these limitations in mind, we have examined the co-localization of CNVR and eQTL in the muscle and liver. In both tissues, approximately 39% of CNVRs co-localized with *cis*-eQTL signals, whilst the co-localization of CNVR and *trans*-eQTL was somewhat higher (≈60%). The lack of positional concordance between CNVRs and eQTL could be due to technical reasons dealing with the relatively high false positive and negative rates of CNV detection with SNP arrays and the limited sensitivity and poor annotation of porcine microarrays. Besides, the search of CNVs was performed in a population of 350 individuals, while only 104 pigs had microarray measurements available. On the other hand, variations in copy number do not necessarily involve changes in gene expression *e.g.* in heterozygous individuals the loss of one allele can be compensated by an increase in the expression of the remaining one, and duplication can generate additional copies whose expression is silenced [47]. Indeed, the two high confidence CNVRs detected

in our study and containing the *SAL1* and *GABBR1* genes (**Figure 4** and **Supplementary Figure 1**) did not co-localize with *cis*-eQTL regulating the expression of these two loci (**Supplementary Tables 13-16**), suggesting that these CNVRs do not have observable effects on gene expression.

The positional coincidence of porcine CNVR and *cis*- and *trans*-eQTL is not unexpected. Indeed, SNPs tagging CNVs tend to be enriched in *cis*-eQTL and often affect multiple expression traits [3]. In a study performed in mouse, 83% of genes mapping to 19 high-confidence CNVs showed evidence of some level of correlation between expression levels and copy number [49]. When we compared the muscle and liver data sets of CNVRs co-localizing with *cis*-eQTL, we observed that they share a substantial number of CNVRs (**Supplementary Tables 13-16**), but these CNVRs co-localize with different sets of *cis*-eQTL (**Table 3**). The patterns of expression in the GM muscle and liver are very different, a feature that affects the impact and consequences of structural variation on gene expression. There were, however, six genes (*HSDL2*, *PREPL*, *PACSIN2*, *EMG1*, *RASIP1* and *KCTD20*) that were shared by both data sets (**Table 3**), suggesting that variation in copy number could be the causal factor modulating the expression of these loci. A future objective would be to investigate the correlation between the number of copies of genes and their expression levels, but a high throughput method allowing the accurate determination of CNV genotypes will be needed to achieve such goal (genotype determination based on Porcine SNP60 BeadChip data is too imprecise to carry out correlation analyses).

Conclusions

The majority of eQTL detected by us seem to affect gene expression in a tissue-specific manner. However, this result could be explained, at least in part, by technical factors such as the low marker density of the Porcine SNP60 BeadChip and the poor annotation of porcine microarrays. Interestingly, a predominance of *trans*-eQTL vs *cis*-eQTL was observed in the liver, while in the muscle tissue the relative contributions of *cis*- and *trans*-eQTL to the regulation of gene expression were similar. The higher transcript and functional complexity of the liver may partly explain this result. We have also observed a substantial number of CNVR and eQTL co-localizations, the majority of which are not simultaneously found in muscle and liver.

Declarations

Ethics approval

Animal care, management procedures and blood sampling were performed following national guidelines for the Good Experimental Practices and they were approved by the Ethical Committee of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Consent for publication

Not applicable

Availability of data and material

Upon acceptance of the paper, data will be available at Dryad.

Competing interests

The authors declare no competing interests.

Funding

Part of the research presented in this publication was funded by grants AGL2013-48742-C2-1-R and AGL2013-48742-C2-2-R awarded by the Spanish Ministry of Economy and Competitivity. We also acknowledge the support of the Spanish Ministry of Economy and Competitivity for the *Center of Excellence Severo Ochoa 2016-2019* (SEV-2015-0533) grant awarded to the Center for Research in Agricultural Genomics. Angela Canovas was funded under the Juan de la Cierva program awarded by the Spanish Ministry of Science. Taina F Cardoso was funded with a fellowship from the CAPES Foundation-Coordination of Improvement of Higher Education, Ministry of Education (MEC) of the Federal Government of Brazil. Gonzalez-Prendes R. was funded by a FPU Ph.D. grant from the Spanish Ministerio de Educación (FPU12/00860).

Authors' contributions

MA and RQ designed the experiment; ACas and AM contributed to generate SNP data; RGP and ACan analyzed the eQTL data; RGP analyzed the CNV data; AZ and YRC provided support in the analysis of CNV data; ACas and RGP validated CNV data by quantitative PCR; TFC and ACan provided bioinformatic support in the analysis of gene expression; MA and RGP wrote the paper; all authors read and approved the manuscript

Acknowledgements

The authors are indebted to Selección Batallé S.A. for providing the animal material. We gratefully acknowledge to J. Reixach (Selección Batallé) and J. Soler (IRTA) for their collaboration in the experimental protocols.

References

1. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2013;368:20120362.

2. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat. Rev. Genet. 2015;16:197–12.

3. Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. PLoS Genet. 2011;7:e1001292.

4. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348:648–60.

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

6. Wimmers K, Murani E, Ponsuksili S. Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. Brief. Funct. Genomics. 2010;9:251–8.

7. Steibel JP, Bates RO, Rosa GJM, Tempelman RJ, Rilington VD, Ragavendran A, et al. Genome-wide linkage analysis of global gene expression in loin muscle tissue identifies candidate genes in pigs. PloS One. 2011;6:e16766.

8. Cánovas A, Pena RN, Gallardo D, Ramírez O, Amills M, Quintanilla R. Segregation of regulatory polymorphisms with effects on the gluteus medius transcriptome in a purebred pig population. PloS One. 2012;7:e35583.

9. Heidt H, Cinar MU, Uddin MJ, Looft C, Jüngst H, Tesfaye D, et al. A genetical genomics approach reveals new candidates and confirms known candidate genes for drip loss in a porcine resource population. Mamm. Genome. 2013;24:416–26.

10. Ponsuksili S, Murani E, Brand B, Schwerin M, Wimmers K. Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. J. Lipid Res. 2011;52:668–78.

11. Chen C, Yang B, Zeng Z, Yang H, Liu C, Ren J, et al. Genetic dissection of blood lipid traits by integrating genome-wide association study and gene expression profiling in a porcine model. BMC Genomics. 2013;14:848.

12. Muñoz M, Rodríguez MC, Alves E, Folch JM, Ibañez-Escriche N, Silió L, et al. Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. BMC Genomics. 2013;14:845.

13. Liaubet L, Lobjois V, Faraut T, Tircazes A, Benne F, Iannuccelli N, et al. Genetic variability of transcript abundance in pig peri-mortem skeletal muscle: eQTL localized genes involved in stress response, cell death, muscle disorders and metabolism. BMC Genomics. 2011;12:548.

14. Ferraz ALJ, Ojeda A, López-Béjar M, Fernandes LT, Castelló A, Folch JM, et al. Transcriptome architecture across tissues in the pig. BMC Genomics. 2008;9:173.

15. Gallardo D, Pena RN, Amills M, Varona L, Ramírez O, Reixach J, et al. Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs. Physiol. Genomics. 2008;35:199–9.

16. Gallardo D, Quintanilla R, Varona L, Díaz I, Ramírez O, Pena RN, et al. Polymorphism of the pig acetyl-coenzyme A carboxylase α gene is associated with fatty acid composition in a Duroc commercial line. Anim. Genet. 2009;40:410–17.

17. González-Prendes R, Quintanilla R, Cánovas A, Manunza A, Figueiredo Cardoso T, Jordana J, et al. Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. Sci Rep. 2017;7:39830.

18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 2007;81:559–75.

19. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 2012;44:821–4.

20. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 2007;17:1665–74.

21. Ramayo-Caldas Y, Castelló A, Pena RN, Alves E, Mercadé A, Souza CA, et al. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. BMC Genomics. 2010;11:593.

22. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. 2001;25:402–8.

23. Revilla M, Puig-Oliveras A, Castelló A, Crespo-Piazuelo D, Paludo E, Fernández AI, Ballester M, Folch JM. Identificación y validación de CNVs en porcino. XVI Jornada sobre producción animal. 2015. Available from: http://acteon.webs.upv.es/CONGRESOS/AIDA%202015/15-%20Revilla%20et%20al.pdf. Accessed: 27 Feb 2017.

24. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

25. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 2016;44:D481-7.

26. Berg JM, Tymoczko JL, Stryer L. Each organ has a unique metabolic profile. In: Biochemistry. 5th edition. New York: W H Freeman; 2002. Section 30.2. Available from: http://www.ncbi.nlm.nih.gov/books/NBK22436/.

27. Fadista J, Nygaard M, Holm L-E, Thomsen B, Bendixen C. A snapshot of CNVs in the pig genome. PLoS One. 2008;3:e3916.

28. Paudel Y, Madsen O, Megens H-J, Frantz LAF, Bosse M, Bastiaansen JWM, et al. Evolutionary dynamics of copy number variation in pig genomes in the context of adaptation and domestication. BMC Genomics. 2013;14:449.

29. Fowler KE, Pong-Wong R, Bauer J, Clemente EJ, Reitter CP, Affara NA, et al. Genome wide analysis reveals single nucleotide polymorphisms associated with fatness and putative novel copy number variants in three pig breeds. BMC Genomics. 2013;14:784.

30. Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, et al. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. Anim. Sci. 2012;90:2883–93.

31. Schiavo G, Dolezal MA, Scotti E, Bertolini F, Calò DG, Galimberti G, et al. Copy number variants in Italian Large White pigs detected using high-density single nucleotide polymorphisms and their association with back fat thickness. Anim. Genet. 2014;45:745–9.

32. Wang Y, Tang Z, Sun Y, Wang H, Wang C, Yu S, et al. Analysis of genome-wide copy number variations in Chinese Indigenous and Western pig breeds by 60 K SNP genotyping arrays. PLoS One. 2014;9:e106780.

33. Jiang J, Wang J, Wang H, Zhang Y, Kang H, Feng X, et al. Global copy number analyses by next generation sequencing provide insight into pig genome variation. BMC Genomics. 2014;15:593.

34. Li Y, Mei S, Zhang X, Peng X, Liu G, Tao H, et al. Identification of genome-wide copy number variations among diverse pig breeds by array CGH. BMC Genomics. 2012;13:725.

35. Chen C, Qiao R, Wei R, Guo Y, Ai H, Ma J, et al. 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. 2012;13:733.

36. Fernández AI, Barragán C, Fernández A, Rodríguez MC, Villanueva B. Copy number variants in a highly inbred Iberian porcine strain. Anim. Genet. 2014;45:357–66.

37. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. Human genomics. The human transcriptome across tissues and individuals. Science. 2015;348:660–5.

38. Bryois J, Buil A, Evans DM, Kemp JP, Montgomery SB, Conrad DF, et al. Cis and trans effects of human genomic variants on gene expression. PLoS Genet. 2014;10:e1004461.

39. Zhang X, Gierman HJ, Levy D, Plump A, Dobrin R, Goring HH, et al. Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. BMC Genomics. 2014;15:532.

40. Emerson JJ, Li W-H. The genetic basis of evolutionary change in gene expression levels. Philos. Trans. R. Soc. B Biol. Sci. 2010;365:2581–90.

41. Huang W, Carbone MA, Magwire MM, Peiffer JA, Lyman RF, Stone EA, et al. Genetic basis of transcriptome diversity in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 2015;112:E6010-19.

42. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. Nature. 2008;452:423–8.

43. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science. 2009;325:1246–50.

44. Xie J, Li R, Li S, Ran X, Wang J, Jiang J, et al. Identification of copy number variations in Xiang and Kele pigs. PLoS One. 2016; 11: e0148565.

45. Wiedmann RT, Nonneman DJ, Rohrer GA. Genome-wide copy number variations using snp genotyping in a mixed breed swine population. PLoS One. 2015; 10: e0133529.

46. Paudel Y, Madsen O, Megens H-J, Frantz LAF, Bosse M, Crooijmans RPMA, et al. Copy number variation in the speciation of pigs: a possible prominent role for olfactory receptors. BMC Genomics. 2015;16:330.

47. Clop A, Vidal O, Amills M. Copy number variation in the genomes of domestic animals. Anim. Genet. 2012;43:503–17.

48. Bae JS, Cheong HS, Kim J-O, Lee SO, Kim EM, Lee HW, et al. Identification of SNP markers for common CNV regions and association analysis of risk of subarachnoid aneurysmal hemorrhage in Japanese population. Biochem. Biophys. Res. Commun. 2008;373:593–6.

49. Orozco LD, Cokus SJ, Ghazalpour A, Ingram-Drake L, Wang S, van Nas A, et al. Copy number variation influences gene expression and metabolic traits in mice. Hum. Mol. Genet. 2009;18:4118–29.

Chapter VII. General Discussion

7.1 The markers contained in the Porcine SNP60 BeadChip explain a limited amount of the phenotypic variance of meat quality traits recorded in Duroc pigs

As shown in **Table 7.1** the h^2_{snp} estimated for porcine meat quality traits in this thesis ranged between 0.00-0.46 and 0.00-0.51 for the GM and LD muscles, respectively. Besides, we observed clear differences between h^2_{snp} estimated for traits measured in these two muscles *e.g.* h^2_{snp} values for omega-6 to -3 ratio (LD = 0.07, GM = 0.12), and for Minolta b* (LD=0.29, GM=0.00). Similar discrepancies were reported by Larzul et al. (1999) who found h^2 of 0.03 and 0.23 for L* measured in the *gluteus profundus* and *longissimus* muscles, respectively. Moreover, these authors demonstrated that h^2 for pH_{24} traits measured in 4 different muscles ranged between 0.17 (*longissimus*) and 0.39 (*biceps femoris*). These differences suggest that the genetic determinism of meat quality traits varies from muscle to muscle. This finding was previously highlighted by Quintanilla et al. (2011) and Gallardo et al. (2012) who showed that QTL maps for traits recorded in the GM and LD muscles are clearly different.

We also observed important departures from the h^2_{snp} calculated by us and genealogic heritabilities estimated by Casellas et al. (2010) in the same population (**Table 7.2**). For instance, h^2_{snp} for C16:0 in the GM and LD muscles were 0.26 and 0.13, respectively, while genealogic heritabilities calculated by Casellas et al. (2010) were 0.32 (GM) and 0.23 (LD). Similarly, h^2_{snp} and genealogic heritability estimates obtained for C18:1 in the GM and LD muscles were dramatically different with regard to the genealogic heritabilities estimated by Casellas et al. (2010) in the same population. In contrast, h^2_{snp} and genealogic heritability values were similar for intramuscular fat content. These results showed that a considerable amount of phenotypic variance for fatty acid traits cannot be detected by the SNPs contained in the Porcine SNP60 BeadChip. This

"missing heritability" could be due to multiple causes (Manolio et al., 2009) *i.e.* 1) alleles with small effects are hard to detect, especially in small populations as ours 2) causal variants with low frequency in the population are also poorly captured by the chip markers 3) genealogic heritability may be overestimated because of gene-gene or gene-environment interactions that may contribute to similarity between related individuals (Zuk et al., 2012).

Table 7.1. Proportion of the phenotypic variance (and its standard error) for meat quality traits explained by the 36,710 SNP markers.

Intramuscular fat content and composition traits			
Phenotype (%)	**Symbol**	***longissimus dorsi* muscle**	***gluteus medius* muscle**
		$h^2_{snp} \pm SE$	$h^2_{snp} \pm SE$
Intramuscular fat	IMF	0.51 ± 0.10	0.46 ± 0.09
Saturated FA	SFA	0.14 ± 0.07	0.37 ± 0.12
Capric	C10:0	0.03 ± 0.05	0.06 ± 0.06
Lauric	C12:0	0.08 ± 0.09	0.00 ± 0.00
Myristic	C14:0	0.20 ± 0.09	0.00 ± 0.03
Palmitic	C16:0	0.13 ± 0.06	0.26 ± 0.12
Margaric	C17:0	0.14 ± 0.08	0.05 ± 0.05
Stearic	C18:0	0.26 ± 0.12	0.29 ± 0.10
Arachidic	C20:0	0.00 ± 0.06	0.02 ± 0.07
Unsaturated FA	UFA	0.14 ± 0.07	0.37 ± 0.12
Monounsaturated FA	MUFA	0.03 ± 0.05	0.00 ± 0.04
Palmitoleic	C16:1(n-7)	0.20 ± 0.08	0.25 ± 0.08
Palmitelaidic	C16:1(n-9)	0.35 ± 0.11	0.18 ± 0.10
Heptadecenoic	C17:1	0.12 ± 0.07	0.15 ± 0.07
Oleic	C18:1(n-9)	0.05 ± 0.05	0.00 ± 0.04
Gondoic	C20:1	0.02 ± 0.05	0.12 ± 0.10
Polyunsaturated FA	PUFA	0.00 ± 0.04	0.01 ± 0.03
Linoleic	C18:2	0.00 ± 0.03	0.01 ± 0.04
α-Linolenic	C18:3 (n-3)	0.10 ± 0.06	0.05 ± 0.05
Eicosadienoic	C20:2 (n-6)	0.07 ± 0.06	0.01 ± 0.05
Eicosatrienoic	C20:3 (n-3)	0.00 ± 0.04	0.02 ± 0.04
Arachidonic	C20:4	0.00 ± 0.03	0.00 ± 0.03
Eicosapentaenoic	C20:5	0.00 ± 0.00	0.04 ± 0.05
Docosahexaenoic	C22:6	0.00 ± 0.03	0.00 ± 0.03
Omega-3 FA	FA n-3	0.00 ± 0.05	0.06 ± 0.05
Omega-6 FA	FA n-6	0.00 ± 0.04	0.00 ± 0.03
Omega-6 to -3 ratio	n-6/n-3	0.07 ± 0.07	0.12 ± 0.10
Processing meat quality traits			
Electric conductivity	CE	0.20 ± 0.07	0.11 ± 0.08
pH at 24 hours	pH_{24}	0.17 ± 0.10	0.12 ± 0.09
Minolta redness	Colour a*	0.41 ± 0.11	0.45 ± 0.11
Minolta yellowness	Colour b*	0.29 ± 0.12	0.00 ± 0.14
Minolta lightness	Colour L*	0.00 ± 0.25	0.00 ± 0.05

Table 7.2. Proportion of the phenotypic variance for IMF content and composition traits explained by the 36,710 SNP markers (h^2_{snp}) and heritabilities estimated with genealogic information by Casellas et al (2010).

	GEMMA	Casellas et al. (2010)		
Phenotype (%)	**h^2_{snp} ± SE**	**Bayes factor[1]**	**Mean**	**HPD95[2]**
GM IMF	0.46 ± 0.09	992.9	0.47	0.18 a 0.86
LD IMF	0.51 ± 0.10	1,152.3	0.55	0.18 a 0.91
GM C16:0	0.26 ± 0.12	40.4	0.44	0.06 a 0.90
LD C16:0	0.13 ± 0.06	15.6	0.47	0.08 a 0.88
GM C18:0	0.29 ± 0.10	1,575.0	0.43	0.09 a 0.81
LD C18:0	0.26 ± 0.12	883.0	0.45	0.09 a 0.86
GM C18:1 (n-9)	0.00 ± 0.04	1.6	0.32	0.00 a 0.73
LD C18:1 (n-9)	0.05 ± 0.05	1.3	0.30	0.00 a 0.69
GM C18:2	0.01 ± 0.04	8.9	0.37	0.01 a 0.77
LD C18:2	0.00 ± 0.03	0.6	0.25	0.00 a 0.68

GM: *gluteus medius*, **LD**: *longissimus dorsi*, [1]Bayes factor of the model with additive polygenic effects against the same model without additive polygenic effects following (García-Cortés et al., 2001), [2]**HPD95**: highest posterior density region at 95%, **SE**: standard error.

Probably, one of the main limiting factors in our study is the small size of the resource population. In humans, populations used for GWAS are much larger and often encompass hundreds of thousands of individuals (Mackay et al., 2009).

7.2 Genome wide association analysis indicates that the majority of QTL for meat qualiy traits have muscle-specific effects

In this thesis we have generated QTL maps for meat quality and composition traits recorded in two different muscles (GM and LD). By comparing the LD and GM QTL maps, we have confirmed previous results obtained by Quintanilla et al. (2011) and Gallardo et al. (2012) showing that the majority (≈ 90%) of QTL are muscle-specific. Interestingly, Quintanilla et al. (2011) identified remarkable differences in the gene expression patterns of the LD and GM muscles. In this way, many genes related with cell proliferation, tissue development, and muscle contraction showed differential expression in the GM and LD muscles. Li et al. (2010) identified 159 mRNAs) involved in a variety of functions (*i.e.* extracellular matrix, muscle contraction, energy homeostasis and metabolism, etc.) that were differentially expressed in the LD and soleus muscles. More recently, Herault et al. (2014) studied the trascriptomes of two glycolytic muscles, LD and *semimembranosus*, with similar myofiber composition and metabolic characteristics. They found that 3,823 genes were differentially expressed in these two muscles and concluded that this result might be explained by their distinct post-natal myogenic activities. Functional annotation revealed that these genes were mostly related with energy metabolism, cell cycle, gene expression, anatomical structure development, and signal transduction/immune response (Herault et al., 2014). These differences in mRNA expression across muscles could be due to the differential expression of regulatory non-coding RNAs. In this regard, Liu et al. (2013) identified 173 microRNAs that are differentially expressed between glycolytic and oxidative muscles. Recently, Guo et al. (2016) built gene regulatory networks for 13 human tissues by integrating large-scale transcription factor-gene regulations with gene and protein expression data. They found that tissue-specific transcription factors are found to regulate more genes than transcription factors expressed in multiple tissues, and the processes regulated by these

tissue-specific transcription factors are closely related to tissue functions (Guo et al., 2016). As suggested by Quintanilla et al. (2011), differences in the transcriptomic landscape amongst muscles might have consequences on the penetrance of causal SNPs. For instance, Musunuru et al. (2010) demonstrated that a mutation in the sortilin (*SORT1*) gene has a tissue specific penetrance because its effect depends on the expression of a transcription factor that is transcribed in the liver but not in adipose tissue or lymphocytes.

Although the majority of QTL identified by us had muscle-specific effects, we have also identified QTL with consistent effects across muscles. The clearest case is that of the QTL in SSC14 (120-124 Mb) with pleiotropic effects on C18:0, C18:1(n-9), SFA and UFA both in the GM and LD muscles. The existence of pleiotropy for quantitative traits has been reported in many model organisms such as *Drosophila melanogaster*, mice, yeast, and *Arabidopsis thaliana* (Flint and Mackay, 2009). According to Solovieff et al. (2013), almost 17% of genes have pleiotropic effects. In a study encompassing one thousand F_{16} mice belonging to an advanced intercross line, 23 QTLs with pleiotropic effects on obesity, plasma lipid concentrations and susceptibility to diabetes were detected (Lawson et al., 2011). In pigs, Corominas et al. (2013) reported one SNP in the promoter of the *ELOVL6* gene with simultaneous effects on palmitic and palmitoleic fatty acids. Interestingly, Hernández-Sánchez et al. (2013), in the same Duroc population used by us, carried out a bivariate GWAS study which demonstrated that SNPs with pleiotropic effects on intramuscular fat and backfat thickness are scarce.

7.3 Weak positional concordance of QTL detected in distinct studies

In general we have observed a low positional concordance between QTLs maps built with microsatellites (Quintanilla et al 2011, Gallardo et al 2012) and SNPs in the same resource Duroc population. The main concordances were found for the QTL detected for

SFA on SSC12 (≈42.0 Mb), CE (LD SSC4, ≈132 Mb) and for color a* in GM SSC10 (≈72 Mb). However, it is difficult to establish comparisons between microsatellite and SNP based QTL maps because a number of microsatellites are not mapped in the Ensembl data base version 87 (Yates et al., 2016) and also because the confidence intervals of QTL mapped with microsatellite markers can be very large. Other studies have also highlighted that QTL detected with microsatellites cannot be always replicated in GWAS studies *e.g.* Yang et al. (2013) performed a GWAS for IMF composition and abdominal fat tissues and only confirmed 23 out of 63 QTL detected with microsatellite markers in the same population. In the same way, Qiao et al. (2015) performed a GWAS for growth and fatness traits in Sutai and F_2 pig breeds and consistency with a previous linkage analisis study in the same populations was weak. This can be due to multiple factors *i.e.* 1) QTL maps based on microsatellites have a low resolution and many regions of the pig genome are not well covered by microsatellites, 2) factors included in the statistical model or multiple testing correction procedures may affect QTL detection. Indeed, Manunza et al. (2014) performed a GWAS for serum lipid traits in pigs and found that the results generated with different mixed-model methods (*i.e.* GEMMA, GenABEL, EMMAX) were not completely coincident. More recently, Guo et al. (2017) carried out a GWAS for growth and fatness traits in Sutai and F_2 pig breeds and stated that consistency with a previous GWAS study targeting the same populations was quite low (they only confirmed 4 out of 15 previously reported QTL).

Although several meat quality QTL detected in this thesis have been previously reported in other GWAS studies (Ramayo-Caldas et al., 2012; Yang et al., 2013; Zhang et al., 2016) and one of them (SSC14, 120-124 Mb) has been widely described in different porcine populations (Yang et al., 2013; Ros-Freixedes et al., 2016; Zhang et al., 2016), in general our GWAS results showed low concordance with GWAS results published in other pig populations (**Table 7.3**). Yang et al. (2015) performed a GWAS for blood lipid

traits in 2,400 Laiwu, Erhualian and Duroc × (Landrace × Yorkshire) pigs and they identified a total of 22 QTL. Notably, only six regions were identified in more than one population, and 16 were detected in a single population. Guo et al. (2017) analysed fatness traits in 4 pig breeds and showed that the broad majority of detected QTL were breed-specific. In **Table 7.3**, we show a comparative analysis of GWAS results obtained in 3 porcine populations: it can be seen that the majority of QTL segregate in just one population. This could be due to methodological reasons, but likely there is a high genetic heterogeneity in the determinism of meat quality traits in pigs. This could be particularly true for Duroc pigs, since their genetic background is very different from those present in other pig breeds. For instance, Traspov et al. (2016) made a population genetics analysis of 26 porcine breeds from Russia, Belorussia, Kazakhstan, Ukraine, China, and Europe; and found that Duroc pigs display a strong genetic differentiation with regard to other European pig populations (**Figure 7.1**).

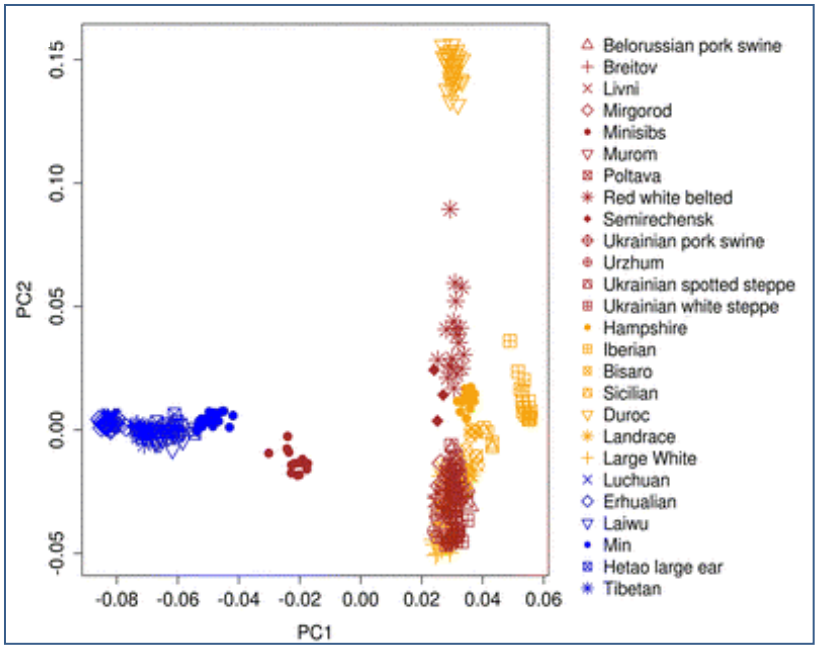


Figure 7.1 A principal coordinate plot showing that Duroc pigs are highly differentiated from their European counterparts. Figure adapted from Traspov et al. (2016) that was published under creative commons attribution 4.0 international license (http://creativecommons.org/licenses/by/4.0/).

The Duroc breed was formed in the United States, during the 19th century, by crossing Red Durocs from New York and Jersey Reds from New Jersey, but the origin of this breed is still unclear (Megens et al., 2008). A low number of founders might explain the high genetic differentiation observed nowadays. Megens et al. (2008) stated that Red Guinea hogs might have participated in the foundation of the Duroc breed but this remains to be demonstrated. This would explain why regions associated with production traits in Duroc are difficult to replicate in other porcine breeds.

Table 7.3 Differences between QTLs maps of four fatty acid traits recorded in three different pig populations.

Traits	[1]Duroc	[2]Landrace × Iberian	[3]White Duroc × Erhualian
	QTLs (Mb)	**QTLs (Mb)**	**QTLs (Mb)**
C16:0	SSC5(71-80)	SSC1(50.3;63,81.3); SSC2(25.7);SSC3(126.3);SSC4(95.1; 136.2);SSC8(88.4; 119.7; 130.6);SSC13(112.7);SSC15(29.3; 60.5; 124.0);SSC17(22.5);SSC18(15.4)	SSC7(31.4)
C18:0	**SSC4 (63.9)**; **SSC14** (58.6-59.4; 65.5-67.9; 87.8-87.9; 92.9-97.0; **120.4-124.4**)	-	**SSC4(63.8)**;**SSC14(120-124)**
C18:1(n-9)	SSC4(63.9;134.9); SSC10(27.0-27.1;54.6;54.6-54.6); **SSC14** (64.7-68.2; 92.9-99.1; **118.7-124.4**; 136.2-136.2; 144.7-148.1;150.3-153.5)	SSC1(147.9);SSC4(13.8; 54.6,107.2);SSC11(7.7);SSC13(112.7)	SSC7(34.8);**SSC4(1 23.6)**
SFA	SSC4 (63.9) SSC12 (6.5-6.5;20.6-20.6;42.4-48.8;50.6-50.6 SSC14 (29.2-29.2;31.8; 92.9-99.1; 119.9-124.4)	SSC1(34.0; 50.3; 63.0; 81.3);SSC4(118.4);SSC8(13.9; 88.4; 119.7);SSC13(112.7);SSC14(38.9)	-
PUFA	SSC5 (71-80)	SSC1(147.9);SSC4(42.8; 136.2); SSC14(5.4)	-

[1]Current study. [2]The genome positions (Mb) representing the QTLs detected in Landrace × Iberian population were extracted from Ramayo-Caldas et al. (2012). We used the SNPs names to update each position from assembly S.scrofa 9 to 10.2 in Ensembl data base version 87 (Yates et al., 2016). [2]The QTLs positions in the White Duroc × Erhualian population were obtained from Yang et al. (2013).

7.4. Analysing the genetic regulation of gene expression through a GWAS approach

One of the goals of the current work was to investigate the positional coincidence of QTL for meat quality traits and eQTL regulating the mRNA levels of genes mapping to QTL regions and having biological functions with a potential impact on muscle physiology. Since microarray measurements of gene expression were exclusively done in the GM muscle, co-localization studies only took into account QTL for traits recorded in the GM muscle. In the analysis of pH, CE, and colour traits we used two approaches to identify *cis*-eQTL. The first of them consisted of doing a genome-wide analysis of *cis*-eQTL and then identifying those that co-localize with QTL regions. In this case, the threshold of correction for multiple testing was established by considering the whole set of markers used in the analysis. In the second analysis, we just considered the set of genes comprised within QTL regions and the threshold of significance for correcting for multiple testing only took into account those SNP comprised within 2 Mb windows around each one of the genes under consideration. In our view, this second analysis is more appropiate than the first one and, in consequence, it has been also used in the GWAS of IMF content and composition traits. Moreover, the study of pH, CE, and color traits only targeted *cis*-eQTL because we considered the information published in humans (GTEx Consortium, 2015) stating that the effects of *trans*-eQTL are much smaller than those of *cis*-eQTL and that they cannot be reliably detected in populations with small sample sizes. In contrast, the genetic analysis of IMF traits targeted both *cis*- and *trans*-eQTL mapping to QTL regions. We did this because we considered that observations made in humans do not necessarily apply to pigs and that excluding *trans*-eQTL from our analysis may imply the loss of valuable genetic information.

The number of *cis*-eQTL detected in QTL regions for pH, CE, and colour traits was quite modest (6 genes regulated by *cis*-eQTL) if compared with those identified for IMF traits (20 genes), but it is also true that we detected much more QTL for IMF traits than for meat technological phenotypes. The *cis*-eQTLs detected in QTL regions for pH, CE, and colour traits were regulating the expression of genes that might have an effect on muscle physiology or metabolism, though this is difficult to establish because the biological mechanisms that regulate meat pH, electrical conductance or colour are poorly known (at least if compared with those that determine fat deposition, a process that has been intensively studied from a biochemical point of view). The *ADCY3* gene was particularly interesting because it has been shown that the partial deletion of this gene in mouse impairs glucose tolerance and insulin sensitivity (Tong et al., 2016). In this regard, it is worth to highlight that meat pH, CE, and color strongly depend on the breakdown of glycogen into glucose and the anaerobic degradation of the glucose into lactic acid (Lee et al., 2000; Warriss et al., 1989). In contrast, the positional coincidence of a *cis*-eQTL regulating the expression of the *IGKC* gene and a QTL for GM redness is probably stochastic because this gene synthesizes the constant region of immunoglobulin kappa, a molecule with immune functions. In other cases (*SLP1*, *CTSA*, *GUSB*), *cis*-regulated genes may have consequences on meat quality but discussing about their potential involvement is still quite speculative because we do not have a detailed knowledge about their biological functions.

The performance of an eQTL scan for 66 genes, located within GM QTL regions made possible to identify 20 *cis*-eQTL regulating the expression of 20 loci. The majority of these *cis*-eQTL mapped to chromosome 14. Several of these genes might have an important role in lipid metabolism *e.g. PEX19*, that plays a key role in the biogenesis of peroxisomes, cellular organelles involved in lipid catabolism (Lodhi and Semenkovich, 2014) and *KCNIP2*, that shows a significant differential expression in the muscle before

and after feeding (our unpublished results). Unexpectedly, we did not observe a *cis*-eQTL for the *SCD* gene that maps to the SSC14 (120-124 Mb) region containing the most significant associations with FA composition traits in both muscles. Recently, Ros-Freixedes et al. (2016) showed that one SNP in the promoter of the *SCD* gene is associated with the composition of IMF and *SCD* mRNA levels. Moreover, Gol et al. (2016) demonstrated that this SNP has causal effects on *SCD* expression in cell cultures. Our inability to detect a *cis*-eQTL regulating the expression of the *SCD* gene might be due to the fact that *SCD* mRNA levels are affected by sex-related and nutritional factors (Joan Estany, personal communication), so the penetrance of the promoter polymorphism reported by Ros-Freixedes et al. (2016) may vary depending on environmental conditions.

We also identified SNPs mapping to GM QTL regions that *trans*-regulated the expression of 103 genes mapping to other chromosomal locations. Several *trans*-regulated loci may have effects on lipid metabolism *e.g.* apolipoprotein E (*APOE*) which plays a key role in the transportation and storage of lipids (Kypreos, 2009) and acyl-CoA dehydrogenase, C-2 to C-3 short chain (*ACADS*), which catalyses the first step of the FA β-oxidation pathway (Jethva et al., 2008). Though the co-localization of QTL and eQTL cannot be interpreted as evidence of a causal relationship, our data suggest that *trans*-eQTL could be an important source of variation for IMF composition traits in pigs. These *trans*-regulatory SNPs could act through a variety of mechanisms *e.g.* altering the sequence of a regulatory RNA or a transcription factor. Westra et al. (2013) carried out a *trans*-eQTL analysis based on 5,311 individuals with measurements of gene expression in blood cells and identified 233 SNPs displaying significant associations. Interestingly, many of these SNPs played an important role in susceptibility to systemic lupus erythematosus, cholesterol metabolism, and type 1 diabetes (Westra et al., 2013).

We have also investigated the factors that regulate the expression of 63 genes related with lipid metabolism. In this way, we have identified 13 *cis*- and 18 *trans*-eQTL modulating the expression of 19 loci with key roles in lipid metabolism. Paradoxically, none of these eQTL mapped to a QTL for lipid traits, suggesting that the detected eQTLs have effects on gene expression but not on phenotypes. There are growing evidences that even mutations abolishing gene function do not always translate into an observable phenotype. For instance, extensive sequencing of Pakistani individuals revealed a woman with homozygous inactivating mutations in the *PRDM9* gene that plays a key role in meiotic recombination (Narasimhan et al., 2016). In mice, the abrogation of *PRDM9* function results in infertility but, paradoxically, this woman had viable offspring (Narasimhan et al., 2016). The most likely explanation for this result is the existence of functional redundancy and compensation mechanisms that mitigate the effects of mutations that might be harmful (Narasimhan et al., 2016). In another study, Tchernitchko et al. (2004) noted that many missense mutations that are predicted to be deleterious do not have an impact on susceptibility to haemolytic anemia.

According to our results, around one third of the 63 lipid genes were regulated by at least one eQTL. The statistical power of our study is quite limited because we only have microarray data from 104 individuals, so we are probably unable to detect eQTL with subtle effects on gene expression. Moreover, it would be interesting to target a higher number of lipid-related genes. The annotation of the porcine microarrays and the pig genome are quite deficient, thus limiting the number of genes that can be analysed. Moreover, in the porcine microarrays gene expression is not measured in a reliable manner when mRNA levels are very low or very high. The optimal situation would be to perform eQTL scans based on RNAseq data obtained in a sufficient number of individuals (100 pigs at least). These large-scale studies can be done in the context of international initiatives such as the FAANG project (Andersson et al., 2015).

Finally, we have demonstrated that eQTL sharing in two distinct porcine tissues, muscle and liver, is quite low. We have analysed gene expression in these two tissues because they play an important role in energy homeostasis and also because, according to Ferraz et al. (2008), their profiles of expression are highly differentiated, probably due to the fact that they derive from different embryonic layers. In this way, we detected 436 *cis*- and 450 *trans*-eQTLs for genes expressed in the *gluteus medius* muscle and 504 *cis*-eQTLs and 3,228 *trans*-eQTLs for hepatic genes. One important observation made by us is that the numbers of *cis*- and *trans*-eQTL in the muscle are similar, while in the liver *trans*-eQTL are clearly predominant, maybe reflecting the functional complexity of this organ. Our results are closely aligned with those presented by other authors indicating that porcine *trans*-eQTL are more abundant than their *cis*- counterparts (Liaubet et al., 2011; Cánovas et al., 2012). In strong contrast, in studies performed in humans, *cis*-eQTL tends to be much more abundant than *trans*-eQTL. For instance, Zhang et al. (2014) analyzed 53 eQTL data sets and identified 116,563 high-confidence eQTL, but only 9% acted in *trans*-. Paradoxically, the analysis of the contributions of *cis*- and *trans*-eQTL to the heritability of gene expression revealed that the latter represent 60-75% of the value of such parameter (Albert and Kruglyak, 2015). This result indicates that gene expression is essentially regulated in *trans*-, but probably these distant genetic determinants are quite scattered across the human genome and their effects are small and depend on the environment, meaning that they are difficult to detect, particularly when resource populations are small (Albert and Kruglyak, 2015). Another important difference with humans is that we have observed a weak sharing of eQTL in the pig muscle and liver, while the GTEx study (GTEx Consortium, 2015) has revealed that 50% of *cis*-eQTL are shared by the nine human tissues under analysis. In contrast, a study comparing the regulatory landscape in three tissues (fibroblasts, lymphoblastoid cell lines and primary T cells) derived from the same set of 75 European individuals reported that

69-80% of *cis*-eQTLs were cell-type specific (Dimas et al., 2009). In consequence, it is difficult to evaluate if discrepancies in the magnitude of eQTL sharing across tissues reveals biological differences between species or if it has technical reasons. Moreover, muscle is a difficult tissue to work with because its transcriptomic profile is overwhelmingly dominated by the expression of a few genes encoding myofibrilar proteins. This means that the expression of other genes can be "masked" to some extent, making difficult to identify eQTL regulating their expression. Thus, it is possible that by using RNAseq data with a high sequencing depth we might be able to find a higher level of QTL sharing between muscle and liver.

7.5. Identification of copy number variants in Duroc pigs and their co-localization with eQTL

We were interested in characterizing copy number variation in our Duroc resource population because such structural polymorphisms might have important consequences on gene expression and phenotypes (Clop et al., 2012). For instance, it is well known that the white coat of pigs is explained by the segregation of CNV with duplicated or triplicated copies of the *KIT* gene that plays a key role in melanocyte survival and differentiation (Giuffra et al., 2002). Unfortunately, the reliable identification of CNV based on Porcine SNP60 BeadChip data has proven to be a difficult task. To infer the copy status of a given individual based on B allele frequency and log R ratio is not straightforward at all (Clop et al., 2012) and it strongly depends on the software used to do so (Ramayo-Caldas et al., 2010). When we compared the putative CNVs identified with the Porcine SNP60 BeadChip and estimates of copy numbers inferred with quantitative real-time PCR assays, only 50% of the CNVs were successfully confirmed. This rate is lower than (75%) that reported by Ramayo-Caldas et al. (2010), but these

authors only took into consideration CNVs identified simultaneously with three softwares, a practice that is expected to decrease substantially the rate of false positives. Indeed, our analysis showed that negative controls represented by putative single-copy regions contained CNVs. This finding makes clear that CNV analyses based on Porcine SNP60 BeadChip data miss many small CNVs that, to make things worse, encompass the broad majority of structural variation in pigs *i.e.* 80% of CNVR are smaller than 15 kb (Paudel et al., 2015). As a matter of fact, many SNPs residing in CNVs are often excluded from genotyping platforms owing to Mendelian inconsistency among families, lack of Hardy–Weinberg equilibrium and high missing genotype rates (Bae et al., 2008). In humans, Cooper et al. (2008) have shown that in old SNP arrays (which happen to have a higher resolution than the Porcine SNP60 BeadChip) 75% of deletions are not covered by probes contained in the chip and that regions containing segmental duplications are often confounded with CNVs. In pigs, Fernández et al. (2014) were able to confirm by high-throughput sequencing only 24 % of CNVR detected with the Porcine SNP60 BeadChip .

Although there was a considerable overlap (47%) between our data set of CNVs and those published in our studies, such co-localizations may have emerged by chance due to the fact that a substantial fraction of the pig genome is covered by CNVs. Another problem that we faced is that we lacked Porcine SNP60 BeadChip data from the mothers of the 350 F_1 barrows, making impossible to evaluate the consistency of CNV segregation with parents-offspring trios. We also obtained evidence that many of the CNVs were present in a few individuals making impossible to assess their effects on gene expression or production phenotypes. A similar observation was made by Bae et al. (2010), who identified 368 CNVR in cattle, of which only 76 had a frequency > 1%.

We have examined the level of co-localization between CNVRs and eQTL detected in the muscle and liver, which happened to be quite low. However, this is a rough approach, so results need to be interpreted with caution. Importantly, eQTL were identified in 104 pigs with microarray measurements while CNVs were detected in 350 pigs. Second, and as previously said, many of the CNVs were harboured by a reduced number of individuals. And third, the duplication or deletion of a genomic region does not necessarily imply changes in gene expression *i.e.* the loss of one allele can be compensated by the increased expression of the remaining allele, and a duplicated copy of one gene might be transcriptionally silent (Clop et al., 2012). To investigate the impact of structural variation on gene and phenotype expression in pigs it will be necessary to use a technique clearly ascertaining copy number status (*e.g.* whole genome sequencing, maybe complemented with some other approach), something that the current Porcine SNP60 BeadChip cannot do in an accurate manner. Subsequently, copy number status should be correlated with gene expression levels measured with RNAseq.

Chapter VIII. Conclusions

Conclusions

1. The segregation of quantitative trait loci (QTL) for meat quality traits in the Duroc population under analysis has been demonstrated. The GWAS analysis for technological meat quality traits recorded in Duroc pigs revealed genome-wide significant associations between electric conductivity of the *longissimus dorsi* muscle and phenotypic variation mapping to SSC4 (104 Mb), SSC5 (15 Mb) and SSC13 (137 Mb), while several additional regions were significantly associated at the chromosome-wide level. The GWAS analysis for intramuscular fat content and composition traits allowed us identifying 37 genome-wide and 83 chromosome-wide QTL. The main genomic regions associated with intramuscular fat composition mapped to SSC2 (9-11 Mb), SSC4 (63.9-64 Mb), SSC5 (71-79 Mb), and SSC14 (87-99 Mb and 120-124 Mb). Particularly relevant was the genomic region on SSC14 (120-124 Mb), which displayed significant associations with C16:1, C18:0, C18:1(n-9), saturated, and unsaturated fatty acids contents.

2. The QTL maps for meat technological and lipid composition traits recorded in the *gluteus medius* and *longissimus dorsi* muscles showed a low positional concordance, indicating that the genetic determinism of these phenotypes has an important muscle-specific component.

3. Several QTL regions for meat quality traits co-localized with expression QTL (eQTL) regulating the mRNA levels of genes with a potential impact on muscle metabolism, though a causal relationship cannot be established yet. In this way, the performance of an eQTL search for SNPs mapping to regions associated with meat technological traits demonstrated that *gluteus medius* a* SSC3 and pH_{24} SSC17 QTL display positional concordance with *cis*-eQTL regulating the expression of several genes (*ADCY3*, *CTSA*, *GUSB*, and *SLP1*) with a potential role on muscle metabolism. Moreover, 20 *cis*-eQTLs affecting the expression of 20 loci mapping to QTL regions for intramuscular fat traits recorded in the *gluteus medius* muscle have been detected in the Duroc population. In addition, 116 *trans*-eQTL mapping to QTL for intramuscular fat phenotypes measured in the *gluteus medius* muscle have been found. These *trans*-eQTL regulate the expression of 103 genes, part of which are involved in lipid metabolism.

4. An eQTLs scan for 63 lipid metabolism genes revealed the existence of 13 *cis*- and 18 *trans*-eQTL regulating the expression of 19 genes with a broad variety of functions such as β-oxidation of fatty acids, lipogenesis, lipolysis, fatty acid

activation and desaturation, and lipoprotein uptake. These results evidence that regulatory mutations are a significant source of variation in the expression of lipid-related genes.

5. A total of 436 *cis*-eQTL and 450 *trans*-eQTL regulating the expression of *gluteus medius* muscle genes have been found, while 504 *cis*-eQTL and 3,228 *trans*-eQTL influencing mRNA levels in the liver have been detected. The weak positional concordance between the eQTL modulating the expression of muscle and hepatic genes indicates that the genetic regulation of gene expression has an important tissue-specific component in pigs.

6. A total of 104 copy number variation regions (CNVR) have been identified in the Duroc pig population under study. Approximately 39% of these CNVR co-localized with *cis*-eQTL, whilst the co-localization of CNVR and *trans*-eQTL was somewhat higher (≈60%). The consistency of these co-localizations in the liver and muscle was weak.

References

Albert, F.W., Kruglyak, L., 2015. The role of regulatory variation in complex traits and disease. Nat. Rev. Genet. 16, 197–212.

Andersson, L., Archibald, A.L., Bottema, C.D., Brauning, R., Burgess, S.C., Burt, D.W., Casas, E., et al., 2015. Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. Genome Biol. 16.

Bae, J.S., Cheong, H.S., Kim, J.O., Lee, S.O., Kim, E.M., Lee, H.W., Kim, S., et al., 2008. Identification of SNP markers for common CNV regions and association analysis of risk of subarachnoid aneurysmal hemorrhage in Japanese population. Biochem. Biophys. Res. Commun. 373, 593–596.

Bae, J.S., Cheong, H.S., Kim, L.H., NamGung, S., Park, T.J., Chun, J.Y., Kim, J.Y., et al., 2010. Identification of copy number variations and common deletion polymorphisms in cattle. BMC Genomics 11, 232.

Cánovas, A., Pena, R.N., Gallardo, D., Ramírez, O., Amills, M., Quintanilla, R., 2012. Segregation of regulatory polymorphisms with effects on the gluteus medius transcriptome in a purebred pig population. PloS One 7, e35583.

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. J. Anim. Sci. 88, 2246–2254.

Clop, A., Vidal, O., Amills, M., 2012. Copy number variation in the genomes of domestic animals. Anim. Genet. 43, 503–517.

Cooper, G.M., Johnson, J.A., Langaee, T.Y., Feng, H., Stanaway, I.B., Schwarz, U.I., Ritchie, M.D., et al., 2008. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. Blood 112, 1022–1027.

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+.

Dimas, A.S., Deutsch, S., Stranger, B.E., Montgomery, S.B., Borel, C., Attar-Cohen, H., Ingle, C., et al., 2009. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 325, 1246–1250.

Fernández, A.I., Barragán, C., Fernández, A., Rodríguez, M.C., Villanueva, B., 2014. Copy number variants in a highly inbred Iberian porcine strain. Anim. Genet. 45, 357–366.

Ferraz, A.L., Ojeda, A., López-Béjar, M., Fernandes, L.T., Castelló, A., Folch, J.M., Pérez-Enciso, M., 2008. Transcriptome architecture across tissues in the pig. BMC Genomics 9, 173.

Flint, J., Mackay, T.F., 2009. Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. 19, 723–733.

Gallardo, D., Pena, R.N., Quintanilla, R., Ramírez, O., Almuzara, D., Noguera, J.L., Amills, M., 2012. Quantitative trait loci analysis of a Duroc commercial population highlights differences in the genetic determination of meat quality traits at two different muscles. Anim. Genet. 43, 800–804.

García-Cortés, L.A., Cabrillo, C., Moreno, C., Varona, L., 2001. Hypothesis testing for the genetic background of quantitative traits. Genet. Sel. Evol. 33, 3–16.

Giuffra, E., Törnsten, A., Marklund, S., Bongcam-Rudloff, E., Chardon, P., Kijas, J.M., Anderson, S.I., et al., 2002. A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking *KIT*. Mamm. Genome 13, 569–577.

Gol, S., Tor, M., Estany, J y Pena, R. N., 2016. Clonación molecular y caracterización de la región promotora del gen *SCD* porcino. Available from: http://acteon.webs.upv.es/. Accessed: 27 Feb 2017

GTEx Consortium, 2015. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348, 648–660.

Guo, Y., Alexander, K., Clark, A.G., Grimson, A., Yu, H., 2016. Integrated network analysis reveals distinct regulatory roles of transcription factors and microRNAs. RNA. 22, 1663–1672.

Guo, Y., Huang, Y., Hou, L., Ma, J., Chen, C., Ai, H., Huang, L., et al., 2017. Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. Genet. Sel. Evol. 49, 21.

Herault, F., Vincent, A., Dameron, O., Le Roy, P., Cherel, P., Damon, M., 2014. The *longissimus* and *semimembranosus* muscles display marked differences in their gene expression profiles in pig. PloS One 9, e96491.

Hernández-Sánchez, J., Amills, M., Pena, R.N., Mercadé, A., Manunza, A., Quintanilla, R., 2013. Genomic architecture of heritability and genetic correlations for intramuscular and back fat contents in Duroc pigs. J. Anim. Sci. 91, 623–632.

Jethva, R., Bennett, M.J., Vockley, J., 2008. Mini-review: Short-chain acyl-coenzyme a dehydrogenase deficiency. Mol. Genet. Metab. 95, 195–200.

Kypreos, K.E., 2009. Mechanisms of obesity and related pathologies. FEBS J. 276, 5719.

Larzul, C., Roy, P., Gogué, J., Talmant, A., Jacquet, B., Lefaucheur, L., Ecolan, P., et al., 1999. Selection for reduced muscle glycolytic potential in Large White pigs. II. Correlated responses in meat quality and muscle compositional traits. Genet. Sel. Evol. 31, 1–16.

Lawson, H.A., Cady, J.E., Partridge, C., Wolf, J.B., Semenkovich, C.F., Cheverud, J.M., 2011. Genetic effects at pleiotropic loci are context-dependent with consequences

for the maintenance of genetic variation in populations. PLoS Genet. 7, e1002256+.

Lee, S., Norman, J.M., Gunasekaran, S., van Laack, R.L., Kim, B.C., Kauffman, R.G., 2000. Use of electrical conductivity to predict water-holding capacity in post-rigor pork. Meat Sci. 55, 385–389.

Li, Y., Xu, Z., Li, H., Xiong, Y., Zuo, B., 2010. Differential transcriptional analysis between red and white skeletal muscle of Chinese Meishan pigs. Int. J. Biol. Sci. 6, 350–360.

Liaubet, L., Lobjois, V., Faraut, T., Tircazes, A., Benne, F., Iannuccelli, N., Pires, J., et al., 2011. Genetic variability of transcript abundance in pig peri-mortem skeletal muscle: eQTL localized genes involved in stress response, cell death, muscle disorders and metabolism. BMC Genomics 12, 548.

Liu, Y., Li, M., Ma, J., Zhang, J., Zhou, C., Wang, T., Gao, X., et al., 2013. Identification of differences in microRNA transcriptomes between porcine oxidative and glycolytic skeletal muscles. BMC Mol. Biol. 14, 7.

Lodhi, I.J., Semenkovich, C.F., 2014. Peroxisomes: a nexus for lipid metabolism and cellular signaling. Cell Metab. 19, 380–392.

Mackay, T.F., Stone, E.A., Ayroles, J.F., 2009. The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. 10, 565–577.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., et al., 2009. Finding the missing heritability of complex diseases. Nature 461, 747–753.

Manunza, A., Casellas, J., Quintanilla, R., González-Prendes, R., Pena, R.N., Tibau, J., Mercadé, A., et al., 2014. A genome-wide association analysis for porcine serum lipid traits reveals the existence of age-specific genetic determinants. BMC Genomics 15, 758.

Megens, H.J., Crooijmans, R.P., Cristobal, M.S., Hui, X., Li, N., Groenen, M.A., 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genet. Sel. Evol. 40, 103–128.

Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., et al., 2010. From noncoding variant to phenotype via *SORT1* at the 1p13 cholesterol locus. Nature 466, 714–719.

Narasimhan, V.M., Hunt, K.A., Mason, D., Baker, C.L., Karczewski, K.J., Barnes, M.R., Barnett, A.H., et al., 2016. Health and population effects of rare gene knockouts in adult humans with related parents. Science 352,474-477

Paudel, Y., Madsen, O., Megens, H.J., Frantz, L.A., Bosse, M., Crooijmans, R.P., Groenen, M.A., 2015. Copy number variation in the speciation of pigs: a possible prominent role for olfactory receptors. BMC Genomics 16, 330.

Qiao, R., Gao, J., Zhang, Z., Li, L., Xie, X., Fan, Y., Cui, L., Ma, J., et al., 2015. Genome-wide association analyses reveal significant loci and strong candidate genes for growth and fatness traits in two pig populations. Genet. Sel. Evol. 47, 17.

Quintanilla, R., Pena, R.N., Gallardo, D., Cánovas, A., Ramírez, O., Díaz, I., Noguera, J.L., et al., 2011. Porcine intramuscular fat content and composition are regulated by quantitative trait loci with muscle-specific effects. J. Anim. Sci. 89, 2963–2971.

Ramayo-Caldas, Y., Castelló, A., Pena, R.N., Alves, E., Mercadé, A., Souza, C.A., Fernández, A.I., et al., 2010. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. BMC Genomics 11, 593.

Ramayo-Caldas, Y., Mercadé, A., Castelló, A., Yang, B., Rodríguez, C., Alves, E., Díaz, I., et al., 2012. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. Anim. Sci. 90, 2883–2893.

Ros-Freixedes, R., Gol, S., Pena, R.N., Tor, M., Ibáñez-Escriche, N., Dekkers, J.C, Estany, J., 2016. Genome-wide association study singles out *SCD* and *LEPR* as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. PloS One 11, e0152496.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., Smoller, J.W., 2013. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 14, 483–495.

Tchernitchko, D., Goossens, M., Wajcman, H., 2004. In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. Clin. Chem. 50, 1974–1978.

Tong, T., Shen, Y., Lee, H.W., Yu, R., Park, T., 2016. Adenylyl cyclase 3 haploinsufficiency confers susceptibility to diet-induced obesity and insulin resistance in mice. Sci. Rep. 6, 34179.

Traspov, A., Deng, W., Kostyunina, O., Ji, J., Shatokhin, K., Lugovoy, S., Zinovieva, N., et al., 2016. Population structure and genome characterization of local pig breeds in Russia, Belorussia, Kazakhstan and Ukraine. Genet. Sel. Evol. 48, 16.

Warriss, P.D., Bevis, E.A., Ekins, P.J., 1989. The relationships between glycogen stores and muscle ultimate pH in commercially slaughtered pigs. Br. Vet. J. 145, 378–383.

Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., et al., 2013. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. 45, 1238–1243.

Yang, B., Zhang, W., Zhang, Z., Fan, Y., Xie, X., Ai, H., Ma, J., et al., 2013. Genome-wide association analyses for fatty acid composition in porcine muscle and abdominal fat tissues. PloS One 8, e65554+.

Yang, H., Huang, X., Zeng, Z., Zhang, W., Liu, C., Fang, S., Huang, L., et al., 2015. Genome-wide association analysis for blood lipid traits measured in three pig populations reveals a substantial level of genetic heterogeneity. PloS One 10, e0131667.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., et al., 2016. Ensembl 2016. Nucleic Acids Res. 44, D710-716.

Zhang, W., Yang, B., Zhang, J., Cui, L., Ma, J., Chen, C., Ai, H., et al., 2016. Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. Sci. Rep. 6, 24718.

Zhang, X., Gierman, H.J., Levy, D., Plump, A., Dobrin, R., Goring, H.H., Curran, J.E., et al., 2014. Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. BMC Genomics 15, 532.

Zuk, O., Hechter, E., Sunyaev, S.R., Lander, E.S., 2012. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc. Natl. Acad. Sci. 109, 1193–1198.

Chapter IX. Annexes

9.1. Supplementary material of *Chapter IV*: *Identifying genomic regions associated with muscle lipid phenotypes and gene expression traits in Duroc pigs*

9.1.1 Supplementary Tables of Chapter IV.

Chapter IV. **Supplementary Table 1**. Mean percentages and standard deviations (SD) of intramuscular fat and composition traits recorded in two muscles of 350 Duroc pigs.

Phenotype	Symbol	*longissimus dorsi* muscle	*gluteus medius* muscle
		Mean ± SD (%)	Mean ± SD (%)
Intramuscular fat	**IMF**	**3.91±1.53**	**5.2±2.05**
Saturated FA	**SFA**	**37.11 ± 2.41**	**36.47 ± 2.07**
Capric	C10:0	0.10 ± 0.07	0.11 ± 0.06
Lauric	C12:0	0.09 ± 0.04	0.09 ± 0.04
Myristic	C14:0	1.37 ± 0.27	1.39 ± 0.23
Palmitic	C16:0	23.47 ± 1.64	23.23 ± 1.42
Margaric	C17:0	0.21 ± 0.08	0.27 ± 0.16
Stearic	C18:0	11.71 ± 1.22	11.21 ± 1.13
Arachidic	C20:0	0.18 ± 0.07	0.17 ± 0.12
Unsaturated FA	**UFA**	**62.89 ± 2.41**	**63.53 ± 2.07**
Monounsaturated FA	**MUFA**	**43.29 ± 5.65**	**43.19 ± 4.9**
Palmitoleic	C16:1 (n-7)	2.98 ± 0.59	2.82 ± 0.49
Palmitelaidic	C16:1 (n-9)	0.46 ± 0.04	0.59 ± 0.05
Heptadecenoic	C17:1	0.17 ± 0.06	0.21 ± 0.06
Oleic	C18:1 (n-9)	34.96 ± 5.16	35.13 ± 4.48
Gondoic	C20:1	0.66 ± 0.16	0.68 ± 0.14
Polyunsaturated FA	**PUFA**	**19.6 ± 7.35**	**20.35 ± 6.04**
Linoleic	C18:2	14.12 ± 5.09	14.93 ± 4.1
α-Linolenic	C18:3 (n-3)	0.48 ± 0.09	0.62 ± 0.1
Eicosadienoic	C20:2 (n-6)	0.41 ± 0.1	0.53 ± 0.14
Eicosatrienoic	C20:3 (n-3)	0.49 ± 0.22	0.45 ± 0.2
Arachidonic	C20:4	3.52 ± 1.8	3.18 ± 1.55
Eicosapentaenoic	C20:5	0.16 ± 0.12	0.18 ± 0.14
Docosahexaenoic	C22:6	0.12 ± 0.1	0.12 ± 0.12
Omega-3 FA	**FA n-3**	**0.91 ± 0.29**	**1.10 ± 0.34**
Omega-6 FA	**FA n-6**	**18.69 ± 7.12**	**19.25 ± 5.81**
Omega-6 to -3 ratio	**n-6/n-3**	**20.43 ± 4.71**	**17.82 ± 3.86**

Chapter IV. **Supplementary Table 2**. QTL displaying chromosome-wide significant associations with intramuscular fat and composition traits recorded in the *gluteus medius* (GM) and *longissimus dorsi* (LD) muscles of 350 Duroc pigs[1].

Traits	SSC	N	SNP	Region (Mb)	*P*-value	*q*-value	*B*	δ	A_1	MAF
GM IMF	13	6	MARC0046697	26.9-27.1	0.00	0.04	0.18	-0.67 ± 0.17	A	0.37
Majority FA (g*luteus medius*)										
GM C18:0	14	1	ALGA0074770	7.5-7.5	0.00	0.05	1.00	-0.28 ± 0.09	G	0.46
		3	ASGA0063465	55-58.6	0.00	0.02	1.00	-0.44 ± 0.11	A	0.23
		1	ASGA0066212	119.9-119.9	0.00	0.03	1.00	-0.42 ± 0.12	G	0.16
		7	MARC0019623	131.2-132.3	0.00	0.03	1.00	-0.47 ± 0.14	A	0.10
GM C18:1 (n-9)	14	1	ALGA0074874	9.3- 9.3	0.00	0.05	1.00	-0.11 ± 0.03	C	0.20
GM PUFA	5	6	ASGA0085283	71.7-79.8	0.00	0.04	0.26	2.99 ± 0.80	A	0.07
GM n-6 FA	5	6	ASGA0085283	71.7-79.8	0.00	0.04	0.21	2.9 ± 0.76	A	0.07
GM n-6/n-3	4	2	DIAS0002565	130.5-130.6	0.00	0.02	0.04	-1.39 ± 0.29	G	0.35
Majority FA (*longissimus dorsi*)										
LD C18:0	14	5	H3GA0041168	87.8-88.2	0.00	0.01	0.73	0.41 ± 0.11	A	0.44
LD C18:1 (n-9)	10	2	DBMA0000150	54.6-54.6	0.00	0.05	0.14	-0.11 ± 0.03	A	0.45
	14	16	ALGA0078299	64.7-68.2	0.00	0.02	1.00	0.12 ± 0.03	G	0.23
		9	ASGA0064951	93.2-99.1	0.00	0.02	1.00	-0.11 ± 0.03	G	0.41
		1	ASGA0091963	118.7-118.7	0.00	0.02	1.00	0.11 ± 0.03	A	0.38
		1	DRGA0014640	136.2-136.2	0.00	0.04	1.00	-0.12 ± 0.04	A	0.14

Traits	SSC	N	SNP	Region (Mb)	*P*-value	*q*-value	*B*	δ	A_1	MAF
		20	MARC0062790	150.3-153.5	0.00	0.01	0.41	0.11 ± 0.03	C	0.42
LD SFA	12	1	ASGA0082934	6.5- 6.5	0.00	0.04	0.45	0.70 ± 0.18	G	0.32
		2	ASGA0054485	20.6-20.6	0.00	0.03	0.25	0.73 ± 0.18	A	0.31
		1	ALGA0066768	50.6-50.6	0.00	0.04	0.46	0.78 ± 0.22	G	0.17
	14	1	ASGA0062403	29.2-29.2	0.00	0.03	1.00	0.64 ± 0.18	A	0.32
		1	DRGA0013769	31.8-31.8	0.00	0.04	1.00	0.73 ± 0.21	A	0.24
LD UFA	12	1	ASGA0082934	6.5- 6.5	0.00	0.04	0.48	-0.69 ± 0.18	G	0.32
		2	ASGA0054485	20.6-20.6	0.00	0.03	0.26	-0.73 ± 0.18	A	0.31
		1	ALGA0066768	50.6-50.6	0.00	0.04	0.42	-0.78 ± 0.22	G	0.17
	14	1	ASGA0062403	29.2-29.2	0.00	0.03	1.00	-0.63 ± 0.18	A	0.32
Minority FA (g*luteus medius*)										
Traits	**SSC**	**N**	**SNP**	**Region (Mb)**	***P*-value**	***q*-value**	***B***	**δ**	**A_1**	**MAF**
GM C10:0	12	2	DIAS0001577	26.8-27.4	0.00	0.01	0.02	-0.03 ± 0.01	A	0.14
GM C14:0	5	15	ASGA0085283	71.7-79.8	0.00	0.01	0.03	-0.13 ± 0.03	A	0.07
		12	ALGA0103880	80.0-80.5	0.00	0.02	0.18	-0.07 ± 0.02	G	0.21
GM C16:1 (n-7)	14	42	ALGA0081091	120.4-124.3	0.00	0.01	0.06	0.19 ± 0.04	C	0.35
		2	ALGA0104036	143.0-144.4	0.00	0.02	0.84	0.19 ± 0.05	G	0.13
GM C17:0	5	4	ASGA0025952	68.0-68.1	0.00	0.02	0.08	0.08 ± 0.02	A	0.11
	6	1	ASGA0092684	20.5-20.5	0.00	0.03	0.67	0.10 ± 0.03	A	0.05
		4	M1GA0008595	63.3-63.5	0.00	0.02	0.28	0.06 ± 0.01	G	0.28
		15	ALGA0036836	124.3-126.0	0.00	0.01	0.04	0.07 ± 0.02	A	0.17

Traits	SSC	N	SNP	Region (Mb)	*P*-value	*q*-value	*B*	δ	A_1	MAF
		2	DIAS0003128	133.2-133.3	0.00	0.02	0.17	0.07 ± 0.02	A	0.17
		1	MARC0038340	146.8-146.8	0.00	0.02	0.20	0.11 ± 0.03	G	0.05
	11	1	H3GA0032297	73.9-73.9	0.00	0.00	0.00	0.08 ± 0.02	G	0.19
	13	1	ALGA0074006	214.6-214.6	0.00	0.01	0.01	0.12 ± 0.03	G	0.06
GM C17:1	2	1	MARC0050503	11.1-11.1	0.00	0.04	0.04	-0.02 ± 0.01	G	0.30
GM C18:2	5	6	ASGA0085283	71.7-79.8	0.00	0.05	0.30	2.02 ± 0.54	A	0.07
GM C18:3	10	2	ALGA0057399	18.7-18.8	0.00	0.03	0.07	0.06 ± 0.01	G	0.11
GM C20:3 (n-3)	13	1	ASGA0083465	2.2- 2.2	0.00	0.05	0.19	0.06 ± 0.02	G	0.45
		1	ALGA0067910	11.1-11.1	0.00	0.04	0.13	0.08 ± 0.02	G	0.13
		1	MARC0023520	189.8-189.8	0.00	0.04	0.05	0.08 ± 0.02	C	0.12
		1	H3GA0037916	207.4-207.4	0.00	0.04	0.13	0.07 ± 0.02	C	0.18
	18	2	ASGA0095005	37.0-37.5	0.00	0.03	0.49	0.10 ± 0.03	A	0.06
GM C20:4	4	1	ASGA0020884	99.0-99.0	0.00	0.04	0.04	0.92 ± 0.21	G	0.06
Minority FA (*longissimus dorsi*)										
LD C10:0	18	3	ASGA0099368	5.1- 5.8	0.00	0.03	0.09	-0.03 ± 0.01	G	0.19
LD C14:0	9	3	ASGA0041280	7.7- 9.5	0.00	0.00	0.09	-0.09 ± 0.02	G	0.30
		3	ASGA0043433	65.2-65.7	0.00	0.02	0.78	0.08 ± 0.02	A	0.23
LD C16:1 (n-9)	6	1	H3GA0017791	28.3-28.3	0.00	0.02	0.05	-0.02 ± 0.00	A	0.26
	7	8	ASGA0037068	128.8-129.9	0.00	0.01	0.10	0.02 ± 0.01	A	0.16
		2	M1GA0011374	130.2-130.2	0.00	0.01	0.05	-0.02 ± 0.00	A	0.48
	12	7	MARC0034121	42.4-48.8	0.00	0.03	0.30	-0.01 ± 0.00	A	0.30
		3	ASGA0055013	52.6-54.6	0.00	0.04	0.90	0.01 ± 0.00	G	0.29

Traits	SSC	N	SNP	Region (Mb)	*P*-value	*q*-value	*B*	δ	A_1	MAF
LD C16:1 (n-7)	4	3	MARC0050687	63.9-63.9	0.00	0.03	0.09	0.24 ± 0.06	C	0.29
		1	H3GA0014326	125.6-125.6	0.00	0.03	0.13	0.19 ± 0.05	C	0.32
	9	3	ASGA0043433	65.2-65.7	0.00	0.04	0.13	0.23 ± 0.06	A	0.23
	10	2	ALGA0057868	27.0-27.1	0.00	0.02	0.06	0.23 ± 0.05	A	0.26
		4	H3GA0029887	34.2-38.3	0.00	0.02	0.05	0.25 ± 0.06	A	0.18
		1	ALGA0106385	52.8-52.8	0.00	0.04	0.30	-0.18 ± 0.05	A	0.36
		1	MARC0089740	77.7-77.7	0.00	0.04	0.29	-0.17 ± 0.05	A	0.40
	14	40	ALGA0081091	120.9-124.3	0.00	0.01	0.17	0.22 ± 0.05	C	0.35
LD C17:0	2	1	ASGA0008934	9.4- 9.4	0.00	0.04	0.46	0.04 ± 0.01	A	0.12
	5	16	INRA0019528	60.6-68.1	0.00	0.02	0.14	0.04 ± 0.01	G	0.13
		3	ALGA0107635	103.8-104.0	0.00	0.03	0.39	0.02 ± 0.01	A	0.42
	8	1	MARC0036242	27.4-27.4	0.00	0.04	0.28	0.03 ± 0.01	A	0.12
		7	ALGA0047393	31.7-32.2	0.00	0.04	0.20	0.05 ± 0.01	A	0.07
	10	2	ALGA0058366	37.8-38.0	0.00	0.04	0.14	-0.03 ± 0.01	C	0.33
		3	ALGA0058457	42.2-42.3	0.00	0.04	0.22	0.03 ± 0.01	A	0.14
LD C17:1	2	9	MARC0088806	3.2- 9.5	0.00	0.03	0.06	-0.04 ± 0.01	C	0.11
		2	ALGA0011938	10.2-10.9	0.00	0.05	0.53	0.02 ± 0.01	G	0.48
		1	DRGA0002832	25.6-25.6	0.00	0.05	0.58	-0.03 ± 0.01	A	0.12
	11	3	ASGA0050304	23.4-23.5	0.00	0.03	0.10	-0.03 ± 0.01	A	0.29
LD C18:3	2	1	ALGA0117101	44.7-44.7	0.00	0.03	0.03	0.06 ± 0.01	C	0.06
	9	1	H3GA0027297	51.8-51.8	0.00	0.01	0.06	-0.05 ± 0.01	A	0.09
		8	ALGA0053212	60.9-64.2	0.00	0.01	0.06	-0.07 ± 0.02	A	0.05

Traits	SSC	N	SNP	Region (Mb)	*P*-value	*q*-value	*B*	δ	A_1	MAF
		1	ASGA0095645	147.5-147.5	0.00	0.04	0.40	0.06 ± 0.02	A	0.05
LD C20:2 (n-6)	5	1	MARC0074262	98.8-98.8	0.00	0.04	0.04	0.03 ± 0.01	G	0.41
LD C20:3 (n-3)	11	4	ASGA0088784	20.8-20.8	0.00	0.01	0.03	0.05 ± 0.01	G	0.14
LD C20:4	3	1	H3GA0010420	116.3-116.3	0.00	0.03	0.03	-0.56 ± 0.13	A	0.34
LD n-3 FA	1	2	MARC0013135	65.3-65.8	0.00	0.05	0.52	0.09 ± 0.02	G	0.45
		9	ALGA0005329	115.2-117.0	0.00	0.05	0.54	0.09 ± 0.02	G	0.39

[1]SSC: porcine chromosome, N: Number of SNPs significantly associated with the trait under study , SNP: SNP displaying the most significant association with the trait under study, Region (Mb): region containing SNPs significantly associated with the trait under study, *P*-value: nominal *P*-value, *q*-value: *q*-value calculated with a false discovery rate approach, *B* : Bonferroni-corrected *P*-value, δ: allelic effect and its standard error (SE), A_1: minority allele, MAF: frequency of the minority allele.

9.2. Supplementary material of Chapter V: *Investigating the genetic regulation of the expression of 63 lipid metabolism genes in the pig skeletal muscle*

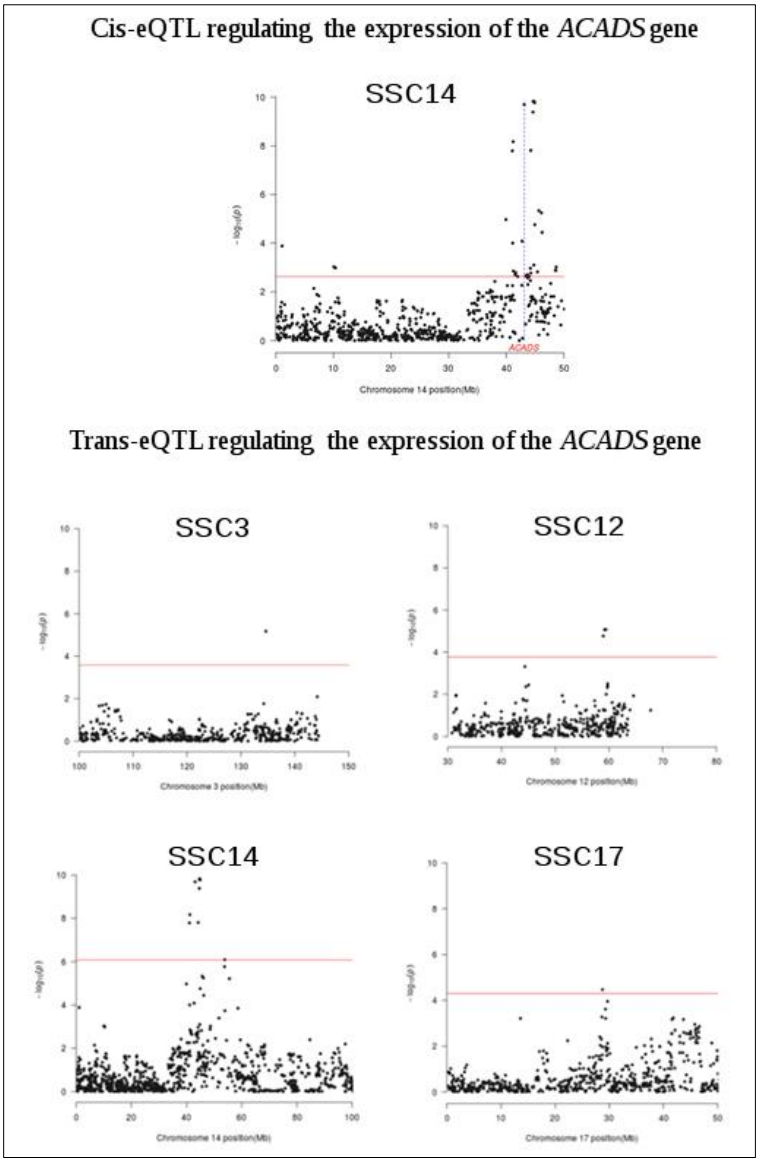
9.2.1 Supplementary Tables of Chapter V.

Chapter V. **Supplementary Table 1.** List of 63 genes involved in lipid metabolism and annotated in the Ensembl database (release 84).

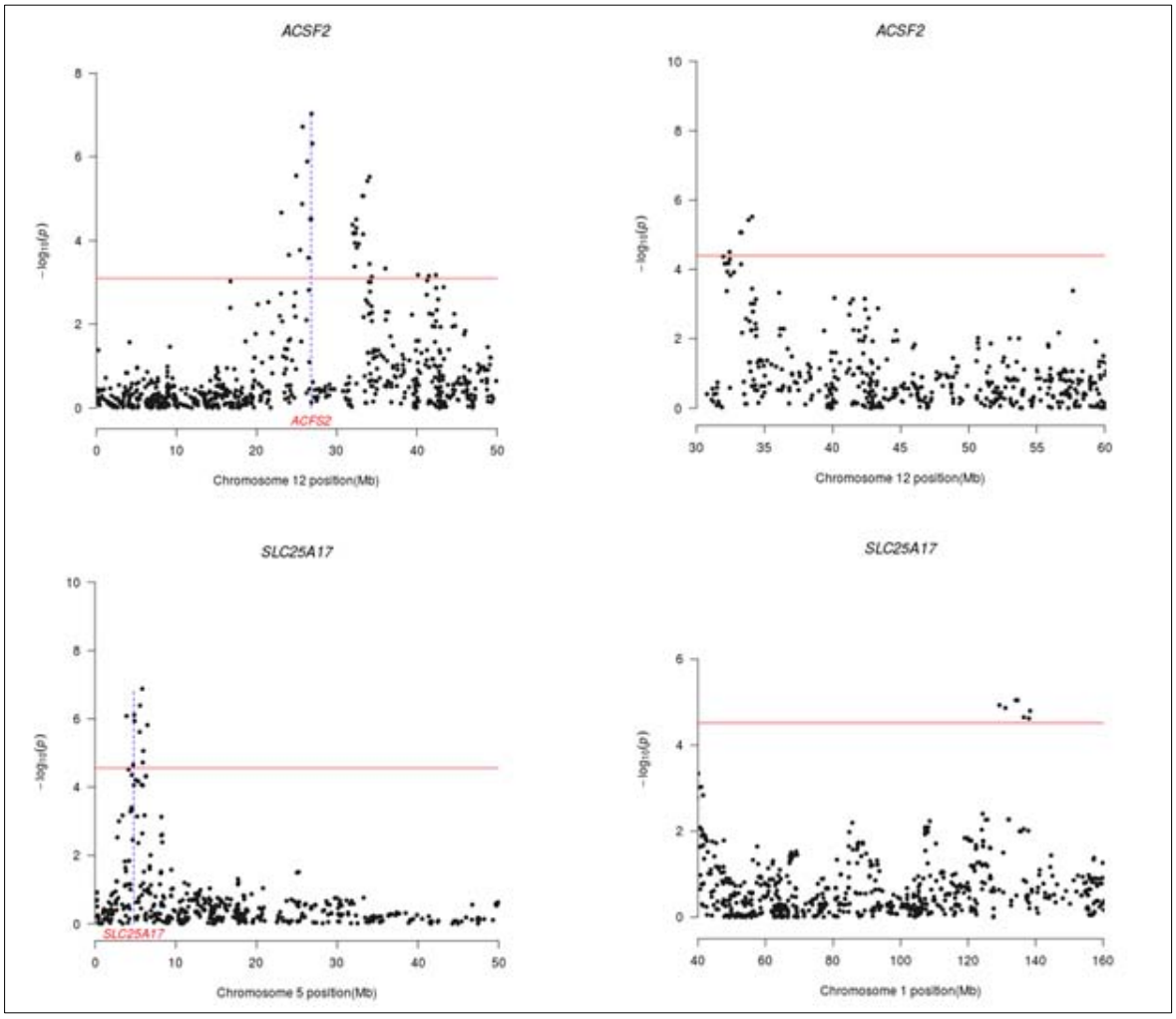
Ensembl ID	Name	Acronym
ENSSSCG00000026173	ATP-binding cassette, sub-family A (ABC1), member 1	*ABCA1*
ENSSSCG00000028620	ATP-binding cassette, sub-family D (ALD), member 3	*ABCD3*
ENSSSCG00000016156	acyl-CoA dehydrogenase, long chain	*ACADL*
ENSSSCG00000003776	acyl-CoA dehydrogenase, C-4 to C-12 straight chain	*ACADM*
ENSSSCG00000009916	acyl-CoA dehydrogenase, C-2 to C-3 short chain	*ACADS*
ENSSSCG00000008724	acyl-CoA oxidase 3, pristanoyl	*ACOX3*
ENSSSCG00000017566	acyl-CoA synthetase family member 2	*ACSF2*
ENSSSCG00000015784	acyl-CoA synthetase long-chain family member 1	*ACSL1*
ENSSSCG00000016223	acyl-CoA synthetase long-chain family member 3	*ACSL3*
ENSSSCG00000012583	acyl-CoA synthetase long-chain family member 4	*ACSL4*
ENSSSCG00000000757	adiponectin receptor 2	*ADIPOR2*
ENSSSCG00000005829	1-acylglycerol-3-phosphate O-acyltransferase 2	*AGPAT2*
ENSSSCG00000015755	1-acylglycerol-3-phosphate O-acyltransferase 5	*AGPAT5*
ENSSSCG00000013599	angiopoietin like 4	*ANGPTL4*
ENSSSCG00000030921	apolipoprotein A1	*APOA1*
ENSSSCG00000003088	apolipoprotein E	*APOE*
ENSSSCG00000016634	caveolin 1, caveolae protein, 22kDa	*CAV1*
ENSSSCG00000016635	caveolin 2	*CAV2*
ENSSSCG00000006276	CCAAT/enhancer binding protein (C/EBP), delta	*CEBPD*
ENSSSCG00000010449	cholesterol 25-hydroxylase	*CH25H*
ENSSSCG00000004142	Cbp/p300-interacting transactivator, with Glu/Asp rich carboxy-terminal domain, 2	*CITED2*
ENSSSCG00000002689	c-Maf inducing protein	*CMIP*
ENSSSCG00000015391	carnitine O-octanoyltransferase	*CROT*
ENSSSCG00000006126	2,4-dienoyl CoA reductase 1, mitochondrial	*DECR1*
ENSSSCG00000003854	enoyl CoA hydratase domain containing 2	*ECHDC2*
ENSSSCG00000001000	enoyl-CoA delta isomerase 2	*ECI2*
ENSSSCG00000026044	farnesyl-diphosphate farnesyltransferase 1	*FDFT1*
ENSSSCG00000010631	glycerol-3-phosphate acyltransferase, mitochondrial	*GPAM*
ENSSSCG00000008569	hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase , beta subunit	*HADHB*
ENSSSCG00000016379	high density lipoprotein binding protein	*HDLBP*
ENSSSCG00000016872	3-hydroxy-3-methylglutaryl-CoA synthase 1 (soluble)	*HMGCS1*
ENSSSCG00000026025	3-hydroxymethyl-3-methylglutaryl-CoA lyase	*HMGCL*

Ensembl ID	Name	Acronym
ENSSSCG00000016420	insulin induced gene 1	*INSIG1*
ENSSSCG00000010226	jumonji domain containing 1C	*JMJD1C*
ENSSSCG00000004569	lactamase beta	*LACTB*
ENSSSCG00000010450	lipase A, lysosomal acid, cholesterol esterase	*LIPA*
ENSSSCG00000003018	lipase, hormone-sensitive	*LIPE*
ENSSSCG00000004509	lipase, endothelial	*LIPG*
ENSSSCG00000000625	low density lipoprotein receptor-related protein 6	*LRP6*
ENSSSCG00000028960	lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	*LSS*
ENSSSCG00000016918	mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase	*MAP3K1*
ENSSSCG00000004454	malic enzyme 1, NADP(+)-dependent, cytosolic	*ME1*
ENSSSCG00000024134	monoglyceride lipase	*MGLL*
ENSSSCG00000025447	MID1 interacting protein 1	*MID1IP1*
ENSSSCG00000001063	myosin regulatory light chain interacting protein	*MYLIP*
ENSSSCG00000008581	nuclear receptor coactivator 1	*NCOA1*
ENSSSCG00000003707	Niemann-Pick disease, type C1	*NPC1*
ENSSSCG00000002366	Niemann-Pick disease, type C2	*NPC2*
ENSSSCG00000016863	3-oxoacid CoA transferase 1	*OXCT1*
ENSSSCG00000011215	3-oxoacyl-ACP synthase, mitochondrial	*OXSM*
ENSSSCG00000001539	peroxisome proliferator-activated receptor delta	*PPARD*
ENSSSCG00000011579	peroxisome proliferator-activated receptor gamma	*PPARG*
ENSSSCG00000003837	protein kinase, AMP-activated, alpha 2 catalytic subunit	*PRKAA2*
ENSSSCG00000000185	protein kinase, AMP-activated, gamma 1 non-catalytic subunit	*PRKAG1*
ENSSSCG00000016432	protein kinase, AMP-activated, gamma 2 non-catalytic subunit	*PRKAG2*
ENSSSCG00000026281	SREBF chaperone	*SCAP*
ENSSSCG00000009759	scavenger receptor class B, member 1	*SCARB1*
ENSSSCG00000010554	stearoyl-CoA desaturase (delta-9-desaturase)	*SCD*
ENSSSCG00000010116	solute carrier family 25 (mitochondrial carrier	*SLC25A1*
ENSSSCG00000000072	solute carrier family 25 (mitochondrial carrier	*SLC25A17*
ENSSSCG00000015232	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	*ST3GAL4*
ENSSSCG00000017402	signal transducer and activator of transcription 5A	*STAT5A*
ENSSSCG00000005229	very low density lipoprotein receptor	*VLDLR*

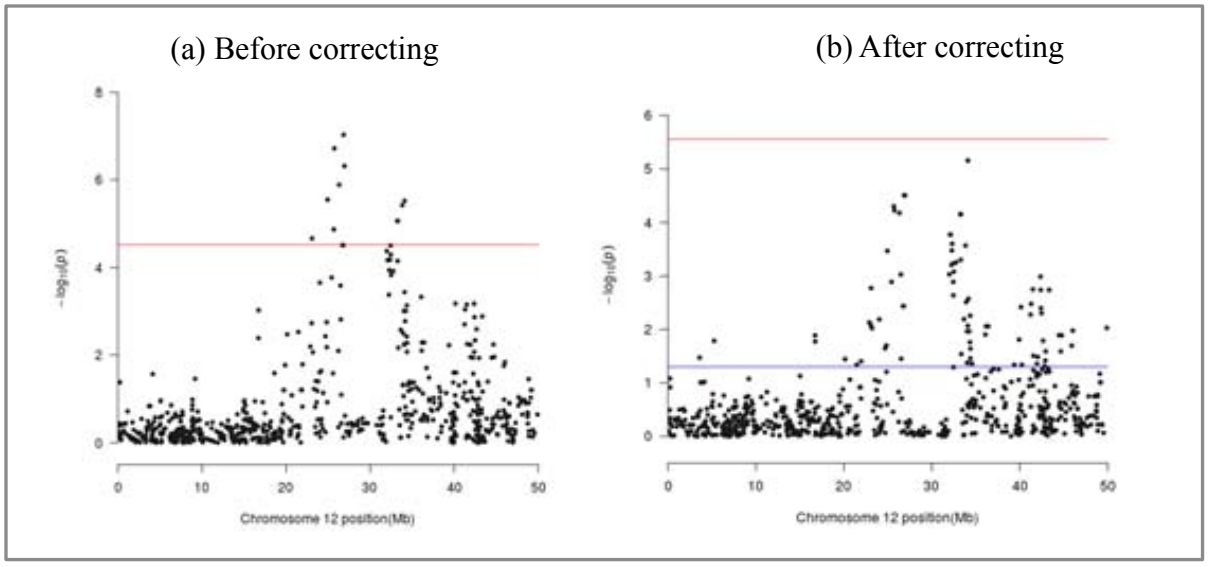
9.2.2 Supplementary Figures of Chapter V.



Chapter V. **Supplementary Figure 1.** Manhattan plots of *cis*- and *trans*-eQTLs regulating the expression of the porcine *ACADS* gene. The x-axis represents the chromosomal region containing the eQTL (measured in Mb), and the y-axis shows the -$\log_{10}$ (*P*-value) of the associations found. The horizontal line indicates the threshold of significance (q-value ≤ 0.05). The vertical blue line depicts the genomic location of the *ACADS* gene



Chapter V. **Supplementary Figure 2.** Manhattan plots of *cis*- and *trans*-eQTLs regulating the expression of the porcine *ACSF2* and *SLC25A17* genes. The x-axis represents the chromosomal region containing the eQTL (measured in Mb), and the y-axis shows the -log10 (*P*-value) of the associations found. The horizontal red line indicates the threshold of significance (*q*-value ≤ 0.05). The vertical blue line depicts the genomic location of the *ACSF2* and *SLC25A17* genes



Chapter V. **Supplementary Figure 3.** Comparison of the significance of a *trans*-eQTL (SSC12, 33.2-34 Mb) regulating the expression of the *ACSF2* gene (SSC12, 26.8 Mb) before (a) and after (b) correcting the data for the most significant SNP of a neighboring *cis*-eQTL (SSC12, 24-26.9 Mb). The x-axis represents the chromosomal region containing the eQTLs (measured in Mb), and the y-axis shows the $-\log_{10}$ (*P*-value) of the associations found. The red and blue lines indicate the thresholds of significance after and before correcting for multiple testing, respectively.

The present Ph.D. thesis is based on the list of studies below:

González-Prendes, R., Quintanilla, R., Cánovas, A., Manunza, A., Figueiredo Cardoso, T., Jordana, J., Noguera, J.L., Pena, R.N., Amills, M., 2017. Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. Sci Rep 7, 39830.

González-Prendes, R., Quintanilla, R., Figueiredo Cardoso, T., Manunza, A., Casellas, J., Cánovas, A., Díaz, I., Noguera, J.L., Castelló, A., Mercadé, A., Amills, M., 2017. Identifying genomic regions associated with muscle lipid phenotypes and gene expression traits in Duroc pigs. BMC Genomics. (*submitted*)

González-Prendes, R., Quintanilla, R., Amills, M., 2016. Investigating the genetic regulation of the expression of 63 lipid metabolism genes in the pig skeletal muscle. (Pending of minor revision in Animal Genetics).

González-Prendes, R., Quintanilla, R., Castelló, A., Zidi, A., Ramayo-Caldas, Y., Figueiredo Cardoso, T., Manunza, A., Cánovas, A., Amills, M., 2017. A genome-wide eQTL scan in the porcine muscle and liver reveals the existence of a significant tissue-specific genetic regulation. (*In preparation*)

Related contribution performed by the author and not included in this thesis:

Cardoso, T.F., Cánovas, A., Canela-Xandri, O., **González-Prendes, R.**, Amills, M., Quintanilla, R., 2017. RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. Sci Rep 7, 40005.

Eusebi, P.G., **González-Prendes, R.**, Quintanilla, R., Tibau, J., Cardoso, T.F., Clop, A., Amills, M., 2017. A genome-wide association analysis for carcass traits in a commercial Duroc pig population. Anim Genet. doi:10.1111/age.12545

Chański, W., **González-Prendes, R.**, Castelló, A., Jordana, J., Manunza, A., Quintanilla, R., Amills, M., 2016. An association analysis between a missense polymorphism at the pig PCSK9 gene and serum lipid and meat quality traits in Duroc pigs. Livest. Sci. 190, 27–30.

Manunza, A., Casellas, J., Quintanilla, R., **González-Prendes, R**., Pena, R.N., Tibau, J., Mercadé, A., Castelló, A., Aznárez, N., Hernández-Sánchez, J., Amills, M., 2014. A genome-wide association analysis for porcine serum lipid traits reveals the existence of age-specific genetic determinants. BMC Genomics 15, 758.

Cardoso, T.F., Quintanilla, R., Castelló, A., **González-Prendes, R.**, Amills, M., Cánovas, A., 2017. Differential expression of splicing mRNA isoforms in the skeletal muscle of pigs with distinct fatness profiles". BMC Genomics. (*submitted*)

Cardoso, T.F., Quintanilla, R., Tibau, J.,Gil, M., Mármol-Sánchez, E., González-Rodríguez, O., **González-Prendes, R.**, Amills, M., 2017. The ingestion of food modulates the expression of genes integrated in the porcine muscle circadian clock. BMC Genomics. (*submitted*)

Cardoso, T.F., Quintanilla, R., Mármol-Sánchez, E., **González-Prendes, R.**, Amills, M., 2017. "The ingestion of food drives changes in the expression of long non-coding RNAs in the pig skeletal muscle". (*In preparation*)