About the existence of common determinants of gene expression in the porcine liver and skeletal muscle

Rayner González-Prendes1,6†, Emilio Mármol-Sánchez1†, Raquel Quintanilla2, Anna Castelló1,3, Ali Zidi1, Yulaxis Ramayo-Caldas1, Tainã Figueiredo Cardoso1,4, Arianna Manunza1, Ángela Cánovas1,5 and Marcel Amills1,3*

Abstract

Background: The comparison of expression QTL (eQTL) maps obtained in different tissues is an essential step to understand how gene expression is genetically regulated in a context-dependent manner. In the current work, we have compared the transcriptomic and eQTL profiles of two porcine tissues (skeletal muscle and liver) which typically show highly divergent expression profiles, in 103 Duroc pigs genotyped with the Porcine SNP60 BeadChip (Illumina) and with available microarray-based measurements of hepatic and muscle mRNA levels. Since structural variation could have effects on gene expression, we have also investigated the co-localization of cis-eQTLs with copy number variant regions (CNVR) segregating in this Duroc population.

Results: The analysis of differential expression revealed the existence of 1204 and 1490 probes that were overexpressed and underexpressed in the glutus medius muscle when compared to liver, respectively (|fold-change| > 1.5, q-value < 0.05). By performing genome scans in 103 Duroc pigs with available expression and genotypic data, we identified 76 and 28 genome-wide significant cis-eQTLs regulating gene expression in the glutus medius muscle and liver, respectively. Twelve of these cis-eQTLs were shared by both tissues (i.e. 42.8% of the cis-eQTLs identified in the liver were replicated in the glutus medius muscle). These results are consistent with previous studies performed in humans, where 50% of eQTLs were shared across tissues. Moreover, we have identified 41 CNVRs in a set of 350 pigs from the same Duroc population, which had been genotyped with the Porcine SNP60 BeadChip by using the PennCNV and GADA softwares, but only a small proportion of these CNVRs co-localized with the cis-eQTL signals.

Conclusion: Despite the fact that there are considerable differences in the gene expression patterns of the porcine liver and skeletal muscle, we have identified a substantial proportion of common cis-eQTLs regulating gene expression in both tissues. Several of these cis-eQTLs influence the mRNA levels of genes with important roles in meat (CTSF) and carcass quality (TAPT1), lipid metabolism (TMEM97) and obesity (MAR2C), thus evidencing the practical importance of dissecting the genetic mechanisms involved in their expression.

© The Author(s). 2019 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
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 regulatory polymorphisms might have important effects on phenotypic variation [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) [2]. There are also evidences that CNV-tagging SNPs are enriched in cis-eQTLs 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 the nine tissues and that most of them display consistent effects across tissues [4].

In pigs, hundreds of eQTLs 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, 13] or displaying significant expression-phenotype correlations [5, 6, 10, 11]. 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 the 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 gene expression patterns and cis-eQTL landscapes of the pig gluteus medius (GM) skeletal muscle and liver, two tissues with highly differentiated patterns of expression [13], as well as to investigate the co-localization of cis-eQTLs and CNVRs.

Results
Differential expression analysis
A total of 1204 probes were found to be overexpressed in the GM muscle (|FC| > 1.5; q-value < 0.05), whereas 1490 showed overexpression in the liver (Fig. 1 and Additional file 1). The list of genes showing the highest levels of differential expression between muscle and liver, included the bridging integrator 1 (BIN1, log₂FC = 5.34, q-value = 2.21E-198), the sorbitol dehydrogenase (SORD, log₂FC = −5.82, q-value = 3.47E-185), the protein phosphatase 1 regulatory inhibitor subunit 1A (PPP1R1A, log₂FC = 5.67, q-value = 5.62E-184), the zinc
finger protein 106 (ZNF106, log₂FC = 4.94, q-value = 1.60E-183), the hydroxysteroid 17-beta dehydrogenase 13 (HSD17B13, log₂FC = −5.54, q-value = 3.04E-180) and apolipoprotein A1 (APOA1, log₂FC = −5.88, q-value = 1.25E-179). Noteworthy, the log₂FC values are very high and the q-values highly significant, thereby evidencing that the expression profiles of the porcine liver and the skeletal muscle are strongly divergent.

Detection of expression QTL in the porcine skeletal muscle and liver
A total of 76 and 28 genome-wide significant cis-eQTLs were identified in the GM muscle and liver, respectively (Tables 1 and 2). Several genes that are cis-regulated in the muscle participate in lipid or carbohydrate metabolism, e.g. HIV-1 Tat Interactive Protein 2 (HTATIP2) [14], exocyst complex component 7 (EXOC7) [15] acyl-CoA dehydrogenase short chain (ACADS) [16], insulin degrading enzyme (IDE) [17], solute carrier family 38 member 9 (SLC38A9) [18], and family with sequence similarity 3 member C (FAM3C) [19]. With regard to the liver, hydroxysteroid dehydrogenase like 2 (HSDL2) [20] and lipase A (LIPA) [21] genes have also important roles in lipid metabolism. The level of overlap between the muscle and the liver cis-eQTL data sets was considerable, with 12 cis-eQTLs shared by both tissues (42.8% of the cis-eQTLs detected in the liver were also detected in the GM muscle). Four examples of shared cis-eQTLs are shown in Fig. 2. There was a strong consistency in the direction of the effects of the shared cis-eQTLs in both tissues. For instance, the cis-eQTL regulating ARL2BP and TMEM97 showed positive allelic effects in both tissues, whilst consistent negative allelic effects were estimated for the cis-eQTL modulating the hepatic and muscular expression of BSDC1, POLR3D and TAP1 (Tables 1 and 2). We also detected a number of cis-eQTLs that had effects on gene expression either in the muscle or in the liver (Figs. 3 and 4). RNA-Seq data from the GM muscle from 52 of the 103 pigs with GM muscle microarray was also available. Comparison of the genotypic means obtained with both platforms and corresponding to genes that are cis-regulated in the GM muscle provided highly consistent results, i.e. the direction of the difference between the two homozygous genotypes was concordant in 82% of the comparisons and the same ordering of genotypes (from highest to lowest expression) was obtained in 70% of the comparisons (Additional file 2). However, both methods did not always yield significant results (about 50% of coincidence with regard to statistical significance). The most probable reason for this is that the sample size of the RNA-Seq data set is half the size when compared to the microarray data set, which leads to larger standard deviations for the RNA-Seq datasets. This would consequently reduce the power to detect significant differences between the genotypic means.

Co-localization of copy number variants and eQTLs
The analysis of structural variation was performed with two softwares: PennCNV and GADA. We detected 93 and 103 CNVRs with PennCNV and GADA, respectively, and 44.08% of the CNVRs detected with PennCNV were also called by GADA. These 41 common CNVRs were distributed along 13 pig chromosomes (Additional file 3). The proportions of copy gain, loss and loss/gain CNVRs were ~48.78%, ~39.02% and ~12.19% respectively. The size of the CNVRs ranged between 31.4 kb and 5.2 Mb, with a mean of 457.4 kb. We compared our CNVR dataset with other CNVRs previously reported in pigs [22–29], and found that 60.97% of our CNVRs had been previously reported (Additional file 4). Real time quantitative assays were designed and used to validate 4 CNVRs (CNVR 9, 15, 32 and 38) in 39 porcine samples. According to D’haene [30], estimates of copy number between 1.414 and 2.449 most likely correspond to a normal copy number of 2, whilst anything below or above these thresholds might represent a deletion (CN = 1) or a duplication (CN = 3), respectively. Following these criteria, the four regions under analysis showed evidence of structural variation (Fig. 5). The co-localization of CNVRs and eQTLs was also analyzed (Additional file 5). In the GM muscle, 2 CNVRs co-localized with 3 cis-eQTLs, while 2 CNVRs co-localized with 2 hepatic cis-eQTLs. This low concordance between CNVRs and cis-eQTLs was not surprising as only 10 CNVRs co-localized with the genomic coordinates of the gene data set analyzed in our experiment.

Discussion
The profiles of mRNA expression of the porcine liver and skeletal muscle are highly divergent
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. The analysis of differential expression revealed that the patterns of expression of these two tissues are highly differentiated in pigs (Fig. 1), probably as a consequence of their distinct embryonic origins and physiological functions [13]. Skeletal muscle and liver also show different patterns of gene expression in mouse [31] and humans [4]. The strongly up-regulated genes in muscle included BIN1, which mediates calcium signaling and excitation–contraction coupling [32], tropomyosin (TPM2), which is fundamental for muscle contraction [33], and ZNF106, which is required for the postnatal maintenance of myofiber innervation by motor neurons [34]. In the liver, we detected a
| Muscle c-eQTL | Gene | Acronym | Region (Mb) | P-value | q-value | B | δ ± SE | A | MAF | Region (Mb) |
|--------------|------|---------|------------|---------|---------|---|-------|---|-----|------------|
| 1 | 51 | ASGA0000817 | 0.2–9.9 | 6.69E-08 | 2.12E-04 | 1.91E-03 | −0.37 ± 0.06 | A | 0.388 | ARID1B | 9.7–10.2 |
| 1 | 15 | ALGA0002427 | 32.2–36.6 | 2.74E-11 | 5.90E-07 | 7.80E-07 | 0.58 ± 0.06 | C | 0.365 | AKAP7 | 32.2–32.3 |
| 1 | 10 | INRA003613 | 101.9–108.9 | 3.29E-12 | 2.00E-08 | 9.00E-08 | −0.29 ± 0.03 | C | 0.433 | WDR7 | 106.0–106.4 |
| 1 | 2 | MARC0039560 | 125.9–128.7 | 5.67E-06 | 3.24E-02 | 1.62E-01 | −0.16 ± 0.03 | G | 0.481 | HYPK | 127.7–127.7 |
| 1 | 5 | DIAS000210 | 183.3–183.9 | 5.36E-07 | 7.63E-03 | 1.53E-02 | 0.28 ± 0.05 | C | 0.466 | CNH1 | 183.8–183.8 |
| 1 | 9 | ASGA006202 | 233.0–239.7 | 3.49E-11 | 8.10E-07 | 9.09E-07 | 0.50 ± 0.07 | A | 0.115 | GHRPR | 238.1–238.1 |
| 1 | 18 | M1GA0001375 | 240.2–249.6 | 1.92E-13 | 1.00E-08 | 1.00E-08 | 0.27 ± 0.03 | A | 0.399 | ALG2 | 240.9–240.9 |
| 1 | 1 | H3GA0004649 | 260.9–260.9 | 2.36E-06 | 3.24E-02 | 6.73E-02 | −0.36 ± 0.07 | A | 0.17 | PSMD5 | 260.8–260.8 |

Table 1 List of the genome-wide significant c-eQTLs detected in the gluteus medius (GM) muscle (those that have been consistently detected in the liver are shown in bold)
strong overexpression of *SORD*, an enzyme that converts sorbitol into fructose [35], *HSD17B13*, whose inactivation leads to chronic liver disease [36], and *APOAI*, the major structural component of high-density lipoproteins [37].

Identification of *cis*-eQTLs with consistent effects on the mRNA levels of genes expressed in the liver and skeletal muscle

In pigs, hundreds of eQTLs with effects on the muscle transcriptome have been detected, but the majority of
these studies focused on target subsets of genes either mapping to QTLs [12] or displaying significant expression-phenotype correlations [5, 6, 10, 11]. Herewith, we have detected 76 muscle cis-eQTLs and 28 liver cis-eQTLs which attained genome-wide significance. These numbers are consistent with previous reports. Liabet et al. [38] made a genome scan based on microarray measurements of *longissimus lumborum* gene expression in 57 pigs and identified 335 eQTLs. Of these, only 18 had cis-regulatory effects. Similarly, Cánovas et al. [8] identified 478 skeletal muscle genome-wide significant eQTLs but only 13% acted in cis-. Although modest sample size (N = 103) limits our ability to identify eQTLs, this sample size ought to allow detecting eQTLs with moderate to large effects on gene expression. As a valid reference, the Genotype-Tissue Expression project detected thousands of eQTLs in 44 tissues characterized by RNA-Seq, and 18 of these tissues were represented by sample sizes equal or below the one used in the current experiment [2]. However, it is also true that sample size largely affects the statistical power of eQTL mapping [4] and thus, the analysis of a larger number of pigs would most likely uncover additional eQTLs that remained undetected in our study.

Our results indicate that almost 42.8% of the *cis-*eQTLs regulating gene expression in the liver also
a  

**gluteus medius muscle**

---

**cis-eQTL for ARL2BP mRNA**

---

**cis-eQTL for TAPI1 mRNA**

---

**cis-eQTL for POLR3D mRNA**

---

**cis-eQTL for PNPO mRNA**

---

b  

**liver tissue**

---

**cis-eQTL for ARL2BP mRNA**

---

**cis-eQTL for TAPI1 mRNA**

---

**cis-eQTL for POLR3D mRNA**

---

**cis-eQTL for PNPO mRNA**

---

Fig. 2 (See legend on next page.)
display significant associations with the muscle mRNA levels of the very same genes (Tables 1 and 2, Fig. 2). The liver and muscle common cis-eQTLs encompass genes previously related to the function of these two organs, such as the cathepsin F (CTSF) gene, that harbors a polymorphism that has been associated with longissimus dorsi tenderness, ham weight and fatness in Italian crossbred pigs [39], the transmembrane anterior transformation 1 (TAPT1) gene, which was linked to carcass and eviscerated weight in a GWAS in chicken [40], the mitochondrial amidoxime reducing component 2 (MAR2) gene, that appears to be involved in porcine fatness [41], and the transmembrane protein 97 (TMEM97) gene, which contributes to the regulation of cholesterol levels [42].

There is also a relevant fraction of porcine cis-eQTLs that display significant effects only in one of the two tissues (Figs. 3 and 4). In the Genotype-Tissue Expression (GTEx) pilot experiment [4], approximately 50% of eQTLs were shared by the nine human tissues under analysis. Moreover, two main types of eQTLs were especially prevalent, 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 eQTLs affecting gene expression in the skeletal muscle show a limited replicability in other tissues [4]. In other studies, also performed in humans, the eQTL tissue-specificity ranged between 50% [43] and 60–80% [44], which implies that the effects of many regulatory mutations are modulated by tissue-associated factors.

Several of the cis-eQTLs detected in our experiment affected the expression of genes involved in lipid or carbohydrate metabolism. For instance, the down-regulation of HTATIP2 leads to elevated fatty acid synthesis and enhanced levels of lipogenic enzymes [14]. The EXOC7 gene regulates the uptake of fatty acids by adipocytes and ACADS is involved in the β-oxidation of fatty acids [16], while FAM3C can suppress hepatic gluconeogenesis [45]. It would be interesting to investigate whether polymorphisms associated with the expression of lipid genes also display associations with fatness traits. Two of the muscle cis-eQTLs detected in our study have been previously reported by other authors. An eQTL for the BSIDCI gene was detected by Ponsuksili et al. [46] and the expression of this gene was also correlated with the percentage of weight loss of the longissimus dorsi muscle. Moreover, a local eQTL that regulates the expression of PNPO and which co-localizes with several meat quality retail traits (such as the percentage of fat and moisture in meat) was described by Steibel et al. [7].

A remarkable level of heterogeneity has been observed in the genetic determinism of production traits in different porcine breeds [47]. In consequence, we anticipated a limited positional concordance amongst eQTLs detected in different breeds. Indeed, a joint analysis of eQTLs across five human populations revealed that varying linkage disequilibrium patterns across populations results in the detection of large numbers of eQTLs with heterogeneous effects [48].

**Limited positional concordance between cis-eQTLs and copy number variation regions**

Another goal of our study was to investigate the co-localization of cis-eQTLs and CNVRs in order to make an initial assessment of the potential impact of structural variation on gene expression in pigs. With PennCNV and GADA, we consistently detected 41 CNVRs with a mean size of 457.4 kb. When we compared our CNVR dataset with CNVRs previously reported in pigs [22–29], we found that 60.97% of the CNVRs detected by us had been previously reported in the literature (Additional file 4). A moderate agreement of CNVR locations amongst studies and populations has been evidenced in several reports [27, 49–52]. Discrepancies could be due to differences in the genetic background of the 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 [27, 53]. In the GM muscle, 2 CNVRs (14 and 38) co-localized with 2 cis-eQTLs (SSC2, 32.6–39.4 Mb and SSC14, 102.7–106.9 Mb), whilst the co-localizations between CNVRs and cis-eQTLs in the liver tissue were somewhat similar, i.e. 2 cis-eQTLs (SSC2, 5.83–5.90 Mb and SSC13, 148.90–149.17 Mb) co-localized with CNVR 11 and 36. This limited positional concordance is probably due to the fact that only 10 CNVRs co-localized with the set of probes analyzed in the current experiment. The low positional concordance between CNVRs and eQTLs could be also due to technical reasons related to the difficulties of detecting CNVs 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 103 pigs had available gene-expression data. 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...
Fig. 3 (See legend on next page.)

**Gluteus medius muscle**

**a**

cis-eQTL for CTSC mRNA  
cis-eQTL for ACADS mRNA

**Liver**

no cis-eQTL for CTSC mRNA  
no cis-eQTL for ACADS mRNA

**b**

cis-eQTL for HSDL2 mRNA  
cis-eQTL for HMOX1 mRNA

**Gluteus medius muscle**

no cis-eQTL for HSDL2 mRNA  
no cis-eQTL for HMOX1 mRNA
allele can be compensated by an increase in the expression of the other allele, and duplication can generate additional copies which expression is silenced [53]. A future objective would be to investigate the association 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 this goal (genotype determination based on 60 K SNP array data is too imprecise to carry out such analyses).

Conclusions
Despite the fact that there are considerable differences in the gene expression patterns of the porcine liver and skeletal muscle, we have identified a substantial proportion of common cis-eQTLs regulating gene expression in both tissues.

Materials and methods
Phenotyping and genotyping of a commercial Duroc population
We have used a commercial Duroc line of 350 Duroc barrows that were slaughtered at the age of ~190 days, with an approximate live weight of 122 kg. This population was generated by crossing 5 boars with ~400 sows and selecting one offspring/sow (50 piglets did not complete their productive cycle so we ended up with 350 barrows with valid records). The 350 barrows were bred in the experimental farm of the Pig Control Center of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). The specific conditions of management and feeding have been previously reported [54, 55]. Pigs were slaughtered in a commercial abattoir following the guidelines of the Spanish Royal Decree 54/1995 (January 20, 1995) to preserve animal welfare. In this way, swine were stunned with carbon dioxide (70% or higher) until they lost consciousness and they were subsequently bled by making an incision in, at least, one carotid artery. At slaughter, GM muscle and liver biopsies were obtained for 103 pigs. Total RNA purification procedures have been previously reported [8, 56]. All experimental procedures were approved by the Ethical Committee of IRTA.

Genomic DNA was extracted from blood samples by following a standard phenol-chloroform protocol. Each pig was genotyped for 62,163 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 [57] was used to filter out the SNP markers with minor allele frequency below 5%, missing genotype rate above 10%, and displaying significant departures (P-value < 0.001) from the Hardy-Weinberg equilibrium. Markers which could not be mapped to the Sscrofa11.1 assembly or that mapped to either the X or Y chromosomes were also eliminated from the data set. Markers in complete linkage disequilibrium (r² > 0.98) were also purged. The final data set used for performing eQTL analyses contained 28,493 SNPs. The filtering criteria used for the CNV analysis was different and only SNPs that did not map to the Sscrofa11.1 assembly or that were located in sex
chromosomes were removed, leading to a final number of 46,537 used SNPs.

Differential expression analysis between muscle and liver tissue

Gluteus medius muscle and liver samples were collected from 103 Duroc pigs (Lipgen population) after slaughter, and immediately frozen in liquid nitrogen. These 103 pigs were selected on the basis of a principal component analysis focused on 13 lipid and growth related traits [58]. We chose individuals representing two different metabolic types, i.e. (i) fat pigs with high intramuscular fat (high saturated and monounsaturated fatty acid content) and also high serum lipid levels, and (ii) pigs that were lean and displayed a low level of intramuscular fat (high polyunsaturated fatty acid content) and circulating lipids [58]. Total RNA was extracted from both GM and liver samples, and mRNA expression profiles were characterized by hybridization to the GeneChip Porcine arrays (Affymetrix Inc., Santa Clara, CA), as previously reported by Cánovas et al. [58]. Hepatic and muscular microarray expression data were deposited in the Gene Expression Omnibus (GEO) public repository, and are accessible through GEO Series accession number GSE115484. The Robust Multi-array Average (RMA) algorithm [59] was employed for carrying out data preprocessing, background correction, normalization and log-transformation of expression values. Gene Intensity level of significance for detecting expressed probes was calculated with the MAS 5.0 algorithm [60]. Control probes and those probes that did not show expression levels above the detection threshold in all samples were filtered out. Differential expression analysis between the GM muscle and liver followed the guidelines of the limma-trend pipeline [61, 62]. The limma’s empirical Bayes method incorporated a mean-variance trend, thus making possible to adequately model the relationship between variance and gene signal intensity. In our study, fold-change (FC) values reflect the mean expression levels in the GM muscle vs liver. The correspondence between differentially expressed probes (|FC| > 1.5; q-value < 0.05) and genes was based on the Affymetrix porcine annotation data assembled database [63] and the Biomart database (https://www.ensembl.org/biomart/martview).

Genome scan for expression QTLs

The genome scan for eQTLs targeted a total of 3326 probes that are simultaneously expressed in both tissues. These genes were selected using the following criteria: The MAS 5.0 [60] algorithm was used to detect the probes with expression levels above the Gene Intensity level of significance. Probes were retained when expressed in all the samples analyzed in both tissues. Expressed
probes where then annotated according to Biomart database available at Ensembl repositories (https://www.ensembl.org/biomart/martview) and those with no gene correspondence were removed from further analyses. We used the GEMMA software [64] and followed the methods described by González-Prendes et al. [56] to carry out the association analyses. The fixed effects and parameters assumed in the statistical model were:

\[ y = W \alpha + x \beta + u + \epsilon \]

where \( y \) 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); \( \alpha \) is a c-vector of the corresponding coefficients including the intercept; \( x \) is an n-vector of marker genotypes; \( \beta \) is the SNP allelic effect estimated as a regression coefficient on the corresponding x genotype (values \(-1, 0, 1\)); \( u \) is a n-vector of random effects with a n-dimensional multivariate normal distribution \( \text{MVN}_n(0, \lambda^{-1} \tau^{-1} K) \) where \( \tau^{-1} \) is the variance of the residual errors; \( \lambda \) is the ratio between the two variance components; \( K \) is a known relatedness matrix derived from SNPs and \( \epsilon \) is the vector of errors with an \( \text{MVN}_n(0, \tau^{-1} I_n) \) being \( I_n \) the identity matrix. The false discovery rate approach reported by Benjamini and Hochberg [65] was used to correct for multiple testing at the genomewide level. In this study, cis-eQTLs were defined as these genomic regions with SNPs significantly associated with a probe located at a maximum distance of \( \pm 1 \text{Mb} \) of the associated SNPs. Co-localization was defined as an overlap of at least 1 base pair between the genomic locations of two eQTLs. The GM muscle cis-eQTLs detected with microarrays were validated by considering an RNA-seq data set corresponding to the same tissue and comprising 52 Duroc pigs [66]. These 52 pigs were a subset of the 103 pigs investigated in the current work and the 52 Duroc pigs [66]. These 52 pigs were a subset of the 103 pigs investigated in the current work and the 52 Duroc pigs [66].

Detection of copy number variation

The PennCNV software [67] was employed to detect copy number variants (CNVs) on the basis of the information provided by 46,537 autosomal SNPs. The PennCNV software implements a hidden Markov model (HMM) to infer CNVs calls for each genotyped sample using as input the intensity signal Log R Ratio (LRR) and the B Allele Frequency (BAF) information generated with the BeadStudio software (Illumina). 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 [67]. With these filtering steps, 20 samples were eliminated from the data set. Only CNVs spanning three or more consecutive SNPs were taken into account. The CNV calling was performed using the default parameters of the HMM model by assuming a UF factor of 0.01. Copy number variant regions (CNVRs) were created by merging CNVs with an overlap of 80% or more. Additionally we used the Genome Alteration Detection Analysis (GADA) package [68] to further validate the CNVR detected with PennCNV and minimize the rate of false positives. The GADA software uses a sparse Bayesian algorithm, based on LRR and BAF values obtained along the genome with the BeadStudio software (Illumina), to identify genomic locations where copy number changes occur. Briefly, a LRR average of markers in a chromosome segment is computed and compared to the median of the respective chromosome. Finally, the copy number class is classified as gain, loss or gain/loss. In our study, copy number variations spanning three or more consecutive SNPs were taken into account and the multiple array analysis option was employed. The parameters defined for the Bayesian learning model and the backward elimination were: 0.8 for the sparseness hyperparameter and 8 for the critical value of the backward elimination.

Validation of copy number variant regions by quantitative PCR

Quantitative real time PCR (qPCR) assays were used to validate four CNVRs in 39 porcine samples randomly selected from the Lippen population. The relative quantification (RQ) of the CNVRs was done as previously described [24]. Primers (Additional file 6) 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 12 K Flex Real-Time PCR System platform (Applied Biosystems, Inc., Foster City, CA). Reactions were performed in triplicate, and they contained 7.5 ng of 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 s at 95 °C and 1 min at 60 °C. Moreover, a melting curve profile (95 °C for 15 s, 60 °C for 15 s and a gradual increase in temperature with a ramp rate of 1 °C up to 95 °C) was implemented to maximize the specificity of the amplification reactions. Relative expression values 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) [69]. 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 CNVR32, which encompasses a deletion, the five samples chosen for calibration were those with RQ values around 2. Normalization of the expression data was done by using a previously reported assay based on the glucagon gene [70].
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.

Authors’ contributions
MA and RQ designed the experiment; AC1 and AM contributed to generate SNP data; EMS made the analyses of differential expression and selected the genes tested in the eQTL analysis; RGP and AC2 analyzed the eQTL data; RGP analyzed the CNV data; AZ and YRC provided support in the analysis of CNV data; AC1 and RGP validated CNV data by quantitative PCR; TFC and SNP data; EMS made the analyses of differential expression and selected the CNVRs by dPCR.

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. We also acknowledge the support of the Spanish Ministry of Economy and Competitiveness for the Center of Excellence Severo Ochoa 2016. We also acknowledge the support of the Spanish Ministry of Economy and Competitiveness for the Center of Excellence Severo Ochoa 2016.

Additional files

Additional file 1: Table S1. List of genes that are differentially expressed (DE) in the gluteus medius muscle and the liver (q-value < 0.05). (XLSX 335 kb)

Additional file 2: Figure S1. Boxplots depicting the mRNA expression levels of cis-eQTL regulated genes measured with RNA-Seq and microarrays in the gluteus medius muscle of 52 and 103 Duroc pigs, respectively. Means were compared with a Student’s t-test: P-value > 0.05 (ns), P-value ≤ 0.05 (*), P-value ≤ 0.01 (**), P-value ≤ 0.001 (***), and P-value ≤ 0.0001 (****). (DOCX 807 kb)

Additional file 3: Table S2. Locations, absolute frequencies, lengths and status of CNVRs detected in 350 Duroc pigs with the PennCNV and GADA softwares. (XLSX 12 kb)

Additional file 4. Table S3. Positional coincidence of the copy number variant regions (CNVR) detected in the current work with CNVRs reported in other pig populations. (XLSX 10 kb)

Additional file 5: Table S4. Co-localization of CNVRs and cis-eQTLs in the gluteus medius muscle (eQTL and gene positions are expressed in Mb). (XLSX 13 kb)

Additional file 6: Table S6. Sequences of the primers used to validate CNVRs by dPCR. (XLSX 9 kb)

Abbreviations
[FC]: absolute value of the fold-change; bp: base pair; CNV: Copy number variant; CNVR: Copy number variant regions; DE: Differentially expressed; e.g.: for example; eQTL: expression quantitative trait loci; GEMMA: Geno- wide Efficient Mixed-Model Association; GEO: Gene Expression Omnibus; GM: gluteus medius; GWAS: genome-wide association analysis; i.e.: in other words; Kb: Kilobase pairs; Mb: Megabase pairs; mRNA: messenger RNA; SNP: Single nucleotide polymorphism; SSC: porcine chromosome

References
1. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. Philos Trans R Soc Lond Ser B Biol Sci. 2013;368. https://doi.org/10.1098/rstb.2012.0362.
2. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat Rev Genet. 2015;16:197–212.
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. Ardlie KG, Deluca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET, et al. The genotype-tissue expression (GTEx) pilot analysis: multisite gene regulation in humans. Science. 2015;348:648–60.
5. Ponsusilsi S, Jonas E, Murani E, Phatthara C, Srirakchal 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, Ponsusilsi S. Functional genetics and genetic approaches to 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 VN, 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:e161766.
8. Cánovas A, Pena RN, Gallardo D, Ramírez Q, Amills M, Quintanilla R. Segregation of regulatory polymorphisms with effects on the gluteus medius transcriptome in a purebred pig population. PLoS One. 2012;7:e33593.
9. Heidt H, Cinar MU, Uddin MJ, Locht C, Jungst 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;24:146–26.
10. Ponsusilsi S, Murani E, Brand B, Schwerin M, Wimmers K. Integrating expression profiling and whole-genome association for dissection of fat traits in a pig 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, Barbe As-Esch N, Silid 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;14:84.
13. Ferraz AL, Ojeda A, López-Belar M, Fernandes LT, Castello A, Folch JM, et al. Transcriptome architecture across tissues in the pig. BMC Genomics. 2008;9:173.

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.

Competing interests
The authors declare that they have no competing interests.

Author details
1Department of Animal Genetics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UIAB-UB, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain. 2Animal Breeding and Genetics Program, Institute for Research and Technology in Food and Agriculture (IRTA), Torre Marimon, 08140 Caldes de Montbui, Spain. 3Department of Ciência Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. 4CAPES Foundation, Ministry of Education of Brazil, Brasília D, F. Zip Code 70.049-020, Brazil. 5Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, 50 Stone Road East, Guelph, Ontario N1G 2W1, Canada. 6Departament de Producció Animal-Agroetecnio Center, Universitat de Lleida, 191 Rovira Roure, 25198 Lleida, Spain.

Received: 27 July 2018 Accepted: 7 June 2019

Published online: 24 June 2019
14. Yin F, Sharen G, Yuan F, Peng Y, Chen R, Zhou X, et al. TIP30 regulates lipid metabolism in hepatocellular carcinoma by regulating SREBP1 through the Akt/mTOR signaling pathway. Oncogene. 2017;36:e547.

15. Inoue M, Akama T, Jiang Y, Chun TH. The exocyst complex regulates free fatty acid uptake by adipocytes. PLoS One. 2015;10:e0120289.

16. Ghosh S, Kruger C, Wicks S, Simon J, Kumar KG, Johnson WD, et al. Short chain acyl-CoA dehydrogenase deficiency and short-term high-fat diet perturb mitochondrial energy metabolism and transcriptional control of lipid-handling in liver. Nutr Metab. 2016;13:17.

17. Farris W, Mansouri S, Chang Y, Lindsley L, Eckman EA, Frosch MP, et al. Insulin-degrading enzyme regulates the levels of insulin, amyloid β-protein, and the β-amyloid precursor protein intracellular domain in vivo. Proc Natl Acad Sci U S A. 2003;100:4162–7.

18. Castellano BM, Thelen AM, Moldavski O, Feltes M, van der Welle RE, González-Prendes B. The cardioprotective protein apolipoprotein A1 promotes potent anti-tumorigenic effects. J Biol Chem. 2013;288:21237–52.

19. Chen Z, Ding L, Yang W, Wang J, Chen L, Chang Y, et al. Hepatic activation of the FAM5C-HSF1-CaM pathway attenuates hyperglycemia of obese diabetic mice. Diabetes. 2017;66:1185–97.

20. Skogberg J, Lundström J, Kovacs A, Nilsson R, Noori P, Maleki S, et al. Transcriptional profiling uncovers a network of cholesterol-responsive atherosclerosis target genes. PLoS Genet. 2008;4:e1000336.

21. Dubiland JA, Francis GA. Lysosomal acid lipase: at the crossroads of normal and atherogenic cholesterol metabolism. Front Cell Dev Biol. 2015;3:53.

22. Fadista J, Nygaard M, Holm L-E, Thomsen B, Bendixen C. A snapshot of transcriptomic bodymap across 17 tissues by RNA-seq. Sci Rep. 2017;7:4200.

23. Fowler KE, Pong-Wong R, Bauer J, Clemente EJ, Reitter CP, Affara NA, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science. 2009;325:1246–50.

24. Chen Z, Wang J, Yang W, Chen J, Meng Y, Feng B, et al. FAM3C activates HSF1 to suppress hepatic gluconeogenesis and attenuate hyperglycemia of type 1 diabetic mice. Oncotarget. 2017;8:100388–94.

25. Ponsuksili S, Murani E, Schwerin M, Schellander K, Wimmen K. Identification of expression QTLs (eQTLs) of genes expressed in porcine M. longissimus dorsi and associated with meat quality traits. BMC Genomics. 2010;11:572.

26. Zhang W, Yang B, Zhang J, Liu 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.

27. Nogués-Fontcuberta I, Fenster P-L, Bartolome J, Cervantes A, Jørensen C. Evolutionary dynamics of copy number variation in pig genomes in the context of adaptation and domestication. BMC Genomics. 2013;14:449.

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

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

30. Paudel Y, Madsen O, Megens H-J, Frantz LAF, Bosse M, Bastiaansen JW, et al. Evolutionary dynamics of copy number variation in pig genomes. PLoS Genet. 2015;11:e1005176.

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

32. Clop A, Vidal O, Amills M. Copy number variation in the genomes of domestic animals. Anim Genet. 2015;46:303–17.

33. Gallardo D, Pena RN, Amills M, Varona L, Ramirez 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.

34. Gallardo D, Quintanilla R, Varona L, Díaz I, Ramírez O, Pena RN, et al. Polyposis of the pig acetyl-coenzyme a carboxylase α gene is associated with fatty acid composition in a Duroc commercial line. Anim Genet. 2009;40:410–7.
62. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014;15:R29.
63. Carlson M. Affymetrix porcine annotation data (chip porcine). R package version 3.2.3; 2016.
64. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44:821–4.
65. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57:289–300.
66. Cardoso TF, Cánovas A, Canela-Xandri O, González-Prendes R, Arnills M, Quintanilla R. RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. Sci Rep. 2017;7:40005.
67. 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.
68. Pique-Regí R, Monso-Varona J, Ortega A, Seeger RC, Triche TJ, Asgharzadeh S. Sparse representation and Bayesian detection of genome copy number alterations from microarray data. Bioinformatics. 2008;24:309–18.
69. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^(-ΔΔCt) method. Methods. 2001;25:402–8.
70. Revilla M, Puig-Oliveras A, Castelló A, Crespo-Piazzuelo D, Paludo E, Fernández AI, et al. A global analysis of CNVs in swine using whole genome sequence data and association analysis with fatty acid composition and growth traits. PLoS One. 2017;12:e0177014.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.