Genome-wide association study identifies quantitative trait loci regions involved in muscle acidic profile in Large White heavy pigs

Genome-wide association study identifies quantitative trait loci regions involved in muscle acidic profile in Large White heavy pigs

Catillo G., Zappaterra M., Zambonelli P., Buttazzoni L., Steri R., Minelli G., Davoli R.

Animal, febbraio 2020 : 1-9

Genome-wide association study identifies Quantitative Trait Loci regions involved in muscle acidic profile in Large White heavy pigs

Gennaro Catillo[1†], Martina Zappaterra[2†], Paolo Zambonelli[2], Luca Buttazzoni[1], Roberto Steri[1], Giovanna Minelli[3], Roberta Davoli[2].

[1]*Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA) - Centro di ricerca Zootecnia e Acquacoltura (Research Centre for Animal Production and Aquaculture), Via Salaria 31, I-00015, Monterotondo, Italy.*

[2]*Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Viale Fanin 46, I-40127, Bologna, Italy.*

[3]*Department of Life Sciences, University of Modena and Reggio Emilia, Via G. Amendola 2, I-42122, Reggio Emilia, Italy.*

Corresponding author: Roberta Davoli. E-mail: roberta.davoli@unibo.it.

† These authors contributed equally to this work.

Short title: Markers for porcine muscle fatty acid composition

Abstract

The widespread use of Genome-Wide Association Studies resulted in the discovery of genomic regions associated with Fatty Acid (**FA**) composition in different porcine tissues, but little information exists about the genes involved in FA composition of meat obtained from heavy pigs selected for the production of Italian dry-cured hams. To this objective, we genotyped with a Single Nucleotide Polymorphism (**SNP**) panel 795 Italian Large White heavy-pigs to identify the markers and genomic regions associated with *Semimembranosus* muscle FA profile. Heritability estimates for intramuscular fat FA profile were of low-to-moderate magnitude, suggesting that these traits may be improved with genomic selection. On the whole, 45 SNPs were significantly associated with 14 FA, and four of them (ALGA008109, ALGA0081097, CASI0010164, SIRI0000267) were associated with more than one FA. The palmitoleic/palmitic, and oleic/stearic ratios displayed the highest number of significant markers and the most significant associations (Bonferroni adjusted $P < 5.00E-07$). Of particular interest, the palmitoleic/palmitic ratio was strongly associated with markers located at 111-114 Mb on chromosome 14, in the same chromosomal region where *Stearoyl-CoA desaturase Δ9* (***SCD***) gene is located. Several significant chromosomal regions were found; some of them harbor key genes playing pivotal roles in FA desaturation and elongation, such as *SCD* and some members of the *Elongation Of Very Long Chain FA* (***ELOVL***) gene family. The results suggest that the identification of causal mutations in these regions may provide a set of markers useful for selection schemes aimed at improving FA composition in pork products.

Keywords: swine; intramuscular fat; genetic markers; selection; fatty acid

Implications

To date, the inclusion of fatty acid (**FA**) composition in genomic selection has been limited; however, the consistent advancement of knowledge in livestock genomics and the widespread use of high-throughput genotyping technologies open up new possibilities for the identification of markers to improve the qualitative characteristics of fresh and seasoned pig meat. To our knowledge, this is the first study investigating the genomic regions associated with muscle FA composition in Large White purebred heavy pigs selected for the production of high-quality dry-cured hams.

Introduction

Fat quantity and FA are essential to various technological and nutritional characteristics of meat, both in subcutaneous (backfat) and Intramuscular Fat (**IMF**). Indeed, the relative amounts of Saturated (**SFA**), Monounsaturated (**MUFA**) and Polyunsaturated FA (**PUFA**) in pig carcass and meat are extremely important both for the food processing industry and consumers. A higher degree of lipid unsaturation could be beneficial to human health (Jimenez-Colmenero *et al*., 2001), but unsaturated FA, particularly PUFA, are more likely to incur in oxidative phenomena (Wood *et al*., 2004; Juárez *et al*., 2011) that can worsen the quality of seasoned products such as Parma ham (Lo Fiego *et al*., 2005). Therefore, dry-cured ham industry requires thighs having subcutaneous and IMF with a limited amount of PUFA (Bosi and Russo, 2004). On the contrary, higher percentages of SFA improve fat firmness and oxidative stability (Wood *et al*., 2004), making the meat more desirable by the seasoning industry. These contrasting requirements raise the necessity to elucidate the factors underlying meat FA composition, to find a balance between consumers' and industry demands. IMF FA composition is a complex polygenic trait whose variability depends on feeding aspects (Wood *et al*., 2008; Pena *et al*., 2016; Kim *et al*., 2018) and the genetic background of each pig population (Wood *et al*., 2008; Pena *et al*.,

2016). The widespread use of high-throughput genotyping technologies allowed for the implementation of Genome-Wide Association Studies (**GWAS**) resulting in the discovery of molecular markers and genomic regions associated with FA composition in different porcine tissues (Zhang *et al*., 2016; Zappaterra *et al*., 2018). In this regard, several molecular markers have been reported to be associated with IMF FA composition, but only a small number of genes have been consistently associated with meat acidic profile in different pig populations (Pena *et al*., 2016). Among them, *ELOVL FA elongase 6* (***ELOVL6***) and *Stearoyl-CoA desaturase Δ9* (***SCD***) genes reportedly play a major role in the C16 and C18 FA de novo synthesis. In a previous work of ours, the genomic regions harboring these two genes have been identified as loci of interest for subcutaneous FA composition in Italian Large White pigs belonging to a purebred population selected for the production of high-quality dry-cured hams (Zappaterra *et al.*, 2018). The same population was also considered in a previous study, where the genetic parameters of backfat FA composition were investigated (Davoli *et al*., 2019). Genetics of purebred Italian Large White population has been shaped by the peculiar selection objectives pursued since 1990 by the Italian pig breeders Association. This study aimed at identifying the markers and genomic regions associated with the *Semimembranosus* muscle FA profile in the Italian Large White heavy-pigs sample. To this goal, a GWAS was performed and significant genomic regions for muscle FA were compared with those identified in the literature. The functional analysis of the genes mapped in these regions was also investigated to define the potential candidates involved in the variability of the trait.

Methods

Sampling

The study was carried out on a sample of 795 Italian Large White pigs belonging to sib test population reared by the Italian pig breeders Association and described in Davoli *et al*. (2019). Triplets of full sibs (two females and one castrated male) were fattened in the Italian pig breeders Association testing station, where siblings entered at 30-45 days of age. Pigs were fed the same finishing diet (Supplementary Table S1) at a *quasi ad libitum* feeding level and slaughtered according to Italian and European laws on pig welfare at about 150 kg live weight in 27 different days between 2011 and 2012 at the same commercial abattoir. Siblings were slaughtered in two or three different dates. *Semimembranosus* muscle samples were taken from the thigh of the left carcass side, at the same point in all carcasses. Samples were then wrapped in aluminum foil, immediately put in vacuum-sealed bags, frozen in liquid nitrogen, and then stored at -80°C for further use.

Muscle fatty acid composition analysis

Intramuscular lipids were extracted from each sample of *Semimembranosus* muscle by using a mixture of chloroform: methanol (2:1 v/v) (Carlo Erba Reagents, MI, Italy) according to the modified Folch *et al.* method (Folch *et al*., 1957). Briefly, 6 g of sample was added with methanol (40 mL) and chloroform (80 mL) and homogenized. An aliquot of a KCl 0.88% saline solution equal to 1:4 of total volume was added to purify the samples from contaminants. Extracted lipids were then submitted to methylation before the Gas Chromatographic analysis. Two mL of Hexane (Carlo Erba Reagents, MI, Italy) and 200 µL of a methanolic solution of 2N-potassium hydroxide (Carlo Erba Reagents, MI, Italy) were added to 25 mg of sample. Tridecanoic methyl acid (Larodan AB, Solna, Sweden) was added as internal standard. The FA methyl esters were analyzed using TRACE^{TM}GC Ultra (Thermo Electron Corporation, Rodano, MI, Italy) equipped with a Flame Ionization

Detector, a Programmed Temperature Vaporization injector, and a TR- FA methyl esters Column 30m x 0.25 mm, 0.2 µm film thickness (Thermo Scientific, Rodano, MI, Italy). Helium was used as carrier gas at a flow rate of 1 mL/min. The injection of FA methyl esters sample (1µL) was performed in split mode with a split flow of 10 mL/min, and operating in a constant condition of carrier gas. Injector and detector were kept at 240°C and the initial temperature of the analysis was 140°C. After the first two minutes, the temperature increased by 4°C/min to reach the final temperature of 250°C, kept steady for 5 min. The Chrom-Card software (vers. 2.3.3, Thermo Electron Corporation, Rodano, MI, Italy) was used to record and integrate the peaks of FA methyl esters. Each FA was identified by comparing its retention time with the known retention times of standard solutions of various methyl esters (Larodan AB, Solna, Sweden). The response factor was calculated and the method of internal standard was used for quantification purposes. The amount of each FA was expressed as mg/g of the IMF.

Genotyping

DNA was extracted from blood samples using Wizard Genomic DNA Purification Kit (Promega Corp, Madison, WI, USA) and animals were then genotyped using Illumina PorcineSNP60 v2 BeadChip (Illumina Inc., San Diego, CA, USA) as reported in Zappaterra *et al*. (2018). Quality control of the SNP data was carried out on PLINK (Purcell *et al*., 2007): call rate for SNP > 95%, minor allele frequency (MAF) ≥ 0.01 and Hardy-Weinberg equilibrium with $P \geq 0.001$ were used as quality control thresholds. The call rate was also computed and individuals with more than 10% of missing data were removed. After quality control and the exclusion of the mutations located on sexual chromosomes, 783 pigs and 40 115 SNPs out of the initial 61 565 were retained.

Genome-wide association study

Genome-wide association analysis was performed using the GenABEL package in the R environment and the Genome-wide Association using Mixed Model and Regression – Genomic Control (**GRAMMAR-GC**) approach with the default function gamma (Aulchenko *et al*., 2007).

The following additive polygenic model was fitted with a genomic relationship matrix in GenABEL:

$$Y_i = X_i\beta_i + W_ih_i + Z_ia_i + e_i$$

Where Y_i is the observation vector for the ith trait; β is the vector of effects for three fixed factors (sex: two levels for barrows and gilts; slaughtering date: 27 levels; age at slaughtering as a covariate). The three random factors in the model were litter (h), animal (a) and residuals (e). They were assumed to be normally distributed as $h_i \sim N(0,I\sigma_h^2)$, $a_i \sim N(0,G\sigma_a^2)$ and $e_i \sim N(0,I\sigma_e^2)$, where G is the genomic relationship matrix and σ_a^2, σ_h^2, σ_e^2 the additive genomic, litter and residual variances, respectively. The genomic relationship matrix G was constructed in GenABEL (Aulchenko *et al*., 2007) by calculating the relationship coefficients for every single pair of individuals (for the 795 pigs).

The heritability estimates of FA and the relative standard errors were estimated with GCTA v1.04 software (Yang *et al*., 2011). The heritability was estimated as:

$$h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2)$$

Where σ_a^2 and σ_e^2 were the additive genomic and residual variances, respectively.

Bonferroni correction for multiple testing was implemented by the PROC MULTTEST of SAS software version 9.4 (SAS Institute, Cary, NC, USA). Associations with Bonferroni adjusted $P<0.05$ ($-\log10(P) > 5.904$) were considered significant, while adjusted $P<0.1$ ($-\log10(P) > 5.640$) were treated as suggestive markers. The regions with significant

markers were then compared with previously published results. Linkage Disequilibrium (**LD**) was estimated (r2) using Haploview software (Barrett *et al*., 2005).

Functional characterization of genes mapped in the candidate regions

The genomic regions flanking significant and suggestive markers were further investigated to find candidate genes. The Ensembl BioMart tool was used to obtain the genes located in *Sus scrofa* genome assembly Build 11.1 in the regions 500 kb up- and downstream the significant markers found in this trial. Functional annotation of the genes located in the regions of interest was performed with David online tool version 6.8 (Huang *et al.*, 2009).

Results

Descriptive statistics and heritability estimates

Descriptive statistics and heritability estimates for the 25 individual FA measured in *Semimembranosus* muscle are reported in Table 1. The most abundant FA were oleic (C18:1 *cis*-9), palmitic (C16:0), stearic (C18:0) and linoleic (C18:2 *cis*-9, *cis*-12) acids, accounting for 40.93%, 23.48%, 11.80% and 10.61% of the total FA, respectively. Besides linoleic acid, the other PUFA showed relatively low amounts in *Semimembranosus* muscle. Among them, the most abundant FA were arachidonic (C20:4 *n*-6), α-linolenic (C18:3 *n*-3), adrenic (C22:4 *n*-6) and dihomo-γ-linolenic (C20:3 *n*-6) acids, which together accounted for about 2.5% of total muscle FA.

Heritability estimates ranged from the lowest value (0.041) found for γ-linolenic acid (C18:3 *n*-6) to the highest found for palmitoleic acid (0.389). The majority of FA displayed a moderate heritability (ranging from 0.2 to 0.4), while lower heritabilities (less than 0.20) were found for eight out of the 25 measured FA. Also, the amount of total FA in

Semimembranosus muscle showed a low-to-moderate heritability (0.149), while the heritability of the ratio between *n*-6 PUFA and *n*-3 PUFA was around 0.2.

Significant markers and candidate regions

We performed GWAS for total muscle FA, for the *n*-6/*n*-3 ratio and each of the 25 FA. Additionally, we also tested the associations for some desaturation indexes, as reported by Ros-Freixedes *et al*. (2016). In particular, GWAS was carried out for the oleic/stearic (C18:1 *cis*-9/C18:0), and palmitoleic/palmitic (C16:1 *cis*-9/C16:0) ratios.

The summary of Bonferroni significant markers is reported in Table 2. The list of genes located in the genomic regions neighboring the significant markers is detailed in Supplementary Material S1. On the whole, 45 SNPs were significantly associated with 14 considered traits, and four of them (ALGA008109, ALGA0081097, CASI0010164, SIRI0000267) were associated with more than one FA level. Twenty-one markers out of 45 showed a significant effect on *Semimembranosus* muscle amounts of capric (C10:0), myristic (C14:0), palmitic, stearic, oleic, arachidic (C20:0), α-linolenic (C18:3 *n*-3), gadoleic (C20:1 *cis*-11) and eicosatrienoic (C20:3 *n*-3) acids. The remaining markers were associated with the total FA, the *n*-6/*n*-3 ratio, and the palmitoleic/palmitic, and oleic/stearic ratios (Table 2). Interestingly, these indexes were the traits displaying the highest number and the most significant associations. The most significant SNPs for palmitoleic/palmitic ratio were DRGA0000432 (located on *Sus scrofa* Chromosome- **SSC**-1), ASGA0095534 (SSC8) and ASGA0073484 (SSC16), the latter presenting an effect also on the muscle oleic/stearic ratio (Table 2). Moreover, the oleic/stearic ratio showed associations with eight markers located on SSC1, four on SSC4, two on SSC8 and one SNP on SSC16. Palmitoleic/palmitic index appeared to be linked to four markers located on SSC8, three of which are in very high LD (Supplementary Figure S1), and eight on

SSC14. Additionally, these markers on SSC14 resulted to be highly associated, with a first block constituted by markers in near-complete LD (ALGA0081091, CASI0010164, ALGA0081097, and ASGA0066116) and a second LD block with the SNPs H3GA0042070 and INRA0046731(Supplementary Figure S1). Another interesting region was located on SSC7 between 46.8 and 49.0 Mb, harboring seven markers in LD associated with gadoleic acid (Table 2 and Supplementary Figure S2). Interestingly, we found an SNP (SIRI0000267) located downstream to the SSC7 46.8-49.0 Mb region, showing a likely pleiotropic effect for multiple FA (capric, myristic, palmitic, oleic, muscle total FA and *n*-6/*n*-3 ratio).

Thirteen suggestive mutations were furthermore identified and summarised in Supplementary Material S2. Among these, seven markers (six associated with stearic and one related to palmitoleic/palmitic ratio) were located at 111-114 Mb on SSC14, in the same chromosomal region where other markers displayed a significant effect on palmitoleic/palmitic ratio. Interestingly, the polymorphism H3GA0025321 (on SSC8 at 111684518 bp) with a suggestive effect on palmitoleic acid muscle content was also found in our previous study to be significantly associated with the palmitoleic acid percentage in porcine backfat tissue (Zappaterra *et al*., 2018).

Finally, the genomic regions 500 kb up- and downstream the significant and suggestive markers for each of the 25 FA and FA ratios were further investigated and the genes mapping in these regions were submitted to the functional analysis. Since for some FA only a few markers passed the highly-conservative Bonferroni threshold, we decided to perform the functional analysis clustering together the genes in both the suggestive and significant regions related to the FA based on the number of their double bonds (i.e. all the genes associated with the individual SFA were analyzed together, and the same was done for MUFA and PUFA). The genes located in the regions of interest for the deposition of

capric, myristic, palmitic, stearic and arachidic SFA were analyzed together and the functional annotation clusters are reported in Supplementary Material S3. Significant gene ontology (**GO**) terms for SFA included “GO:0038061~NIK/NF-kappaB signalling” (*P*=0.021); “GO:0031146~SCF-dependent proteasomal ubiquitin-dependent protein catabolic process” (*P*=0.003); and “GO:0019005~SCF ubiquitin ligase complex” (*P*=0.012). These results suggest that genes associated with SFA should not be involved in *de novo* FA synthesis and deposition. The genes located in regions related to oleic, palmitoleic, gadoleic MUFA and desaturation indexes were significantly clustered in terms associated with fat biosynthesis and deposition, such as “GO:0035338~long-chain fatty-acyl-CoA biosynthetic process” (*P*=0.01) comprising the genes *ELOVL FA elongase 5* (***ELOVL5***), *SCD*, *ELOVL6* and *Acyl-CoA Synthetase, Bubblegum Family, member 1* (***ACSBG1***), and the term “GO:0006636~unsaturated FA biosynthetic process” with the genes *ELOVL5*, *SCD* and *ELOVL6* (*P*=0.01; Supplementary Material S4). Finally, the genes located near the markers associated with α-linolenic, eicosadienoic (C20:2 *n*-6), eicosatrienoic (C20:3 *n*-3) and docosahexaenoic (C22:6 *n*-3) PUFA were submitted to functional analysis and the results were summarized in Supplementary Material S5. Only one term was significant: “GO:0006508~proteolysis” (*P*=0.04). Noteworthy, a consistent number of genes were clustered in the UP Keyword “Transport” (*P*=0.03), suggesting that most of the regions found in the present study to be associated with PUFA harbor genes coding for transmembrane transporters and membrane components.

Discussion

The pigs used in this trial displayed high carcass weights (on average 118.7 ± 8.7 kg), approaching the weight of typical heavy pigs grown for the production of high-quality dry-cured hams, such as Parma and San Daniele. For these productions, high values of

PUFA, and in particular *n*-3 PUFA, are undesirable (Bosi and Russo, 2004), while SFA and MUFA are less likely to incur in oxidative and lipolytic processes, and therefore preferable by the ham processing industry. To date, the inclusion of FA composition in genetic improvement programs has been limited; however, rapid advances in technology open up new possibilities for the improvement of this complex trait. In general, IMF FA composition noticed in the present research was consistent with the values reported in previous research on pigs slaughtered at an average live weight of 145 kg (Lo Fiego *et al.*, 2010), where oleic, palmitic, stearic and linoleic acids accounted in *Longissimus lumborum* muscle for about 42.2%, 23.7%, 12.4% and 11.2% of the total IMF FA, respectively. To the best of our knowledge, this is the first study concerning *Semimembranosus* muscle FA composition in purebred Italian Large White pigs reared in the same controlled environmental conditions. The sample size is comparable with the existing literature on similar phenotypes (Corominas *et al*., 2013; Ros-Freixedes *et al*., 2016; van Son *et al*., 2017; Pena *et al.,* 2019). However, it represents nearly all the Italian sib-tested Italian Large White population in 2011 and 2012.

On the whole, the present study identified 45 markers significantly associated with the *Semimembranosus* IMF FA composition in heavy pigs. Of these 45 markers, three SNPs (DRGA0000432, ASGA0095534, and ASGA0073484) associated with oleic/stearic ratio showed highly significant Bonferroni adjusted *P*. The marker DRGA0000432 (SSC1) is an intron variant of the gene *EYA Transcriptional Coactivator And Phosphatase 4* (***EYA4***). The protein encoded by this gene is a transcriptional activator playing a role during eye development, DNA repair, apoptosis, and innate immunity (Schönberger *et al*., 2005). To date, no candidate genes associated with FA composition were indicated in the literature for the SSC1 region where the three mentioned SNPs are located. However, an SNP included in the *EYA4* gene (ALGA0002244) was found to be associated with ham weight

loss at first salting in a GWAS performed on Italian Large White pigs (Fontanesi *et al*., 2017). Weight loss at first salting and during ham seasoning also depends on the amount of fat stored in muscle and on backfat thickness (Bosi and Russo, 2004). Even if the result obtained by Fontanesi *et al.* (2017) was corrected for the subcutaneous fat thickness, the associations found in the two studies may reflect an involvement of the SSC1 region at 30402065-35022863 bp in the oleic storage and muscle fat deposition, with an indirect effect also on hams weight loss at first salting. Anyway, this effect should be further investigated as at present no clear biological evidence for the involvement of this chromosomal region in IMF metabolism or ham seasoning losses has been found. The SNP ASGA0095534, located in *GrpE Like 1, Mitochondrial* (***GRPEL1***) gene on SSC8 was also identified as a highly significant marker. The exact role of its encoded protein has not been completely clarified, neither the differences with its paralog, *GrpE Like 2, Mitochondrial* (***GRPEL2***) gene (Konovalova *et al*., 2018). Anyway, both GRPEL1 and GRPEL2 take part in the mitochondrial presequence translocase-associated motor (**PAM**), a complex driving the translocation of mitochondrial precursor proteins from the intramembrane space to the mitochondria matrix (reviewed in Wiedemann *et al*., 2004). No evidence of direct involvement of *GRPEL1* gene in fat deposition and oxidation exist to date in literature: however, its paralog *GRPEL2* was found to be differentially expressed in the liver of a mice model displaying acute fatty liver disease induced by blocking long-chain FA β-oxidation (van der Leij *et al.*, 2007). The results found by these authors seem to suggest that *GRPEL* paralogs may also be strictly associated with fat deposition and β-oxidation, even though the processes determining this involvement are still unclear. The region flanking the third most significant marker for the oleic/stearic ratio in the present study (ASGA0073484) is a still poorly annotated region located on chromosome 16. The gene closest to the identified marker is ENSSSCG00000036842, a novel gene with an incomplete annotation and no evidence of paralogs or known orthologues in other animal

species since all the identified orthologues have not yet been associated to an official gene symbol in Ensembl database. Anyway, it is worth noting that the same chromosome region was also indicated as associated with average daily gain in Italian Large White pigs (Fontanesi *et al*., 2014), suggesting that candidate genes for growth efficiency and FA composition may be lying in this chromosome sequence. Therefore, on the whole, the three most significant markers (DRGA0000432, ASGA0095534, and ASGA0073484) were not located in genomic regions harboring strong candidate genes for FA composition. However, these results need further validation because these highly-significant markers are rare variants (MAF < 0.1) and this may have biased their estimated *P*. Moreover, the identified effects could be due to genes located in the neighborhood and still unknown due to the incomplete annotation of some porcine genome regions. In this scenario, other markers with higher MAF could be of interest despite their less-significant *P*. We, therefore, decided to discuss also these less-significant markers, and consider their functional roles in FA deposition. Among them, chromosome 8 harbors another interesting region, located at 111.8-113.3 Mb, associated with the palmitoleic/palmitic desaturation index. The four significant markers located in this region are located in the sequence of known genes that are not directly related to fat deposition nor biosynthetic FA processes. Nevertheless, these markers are mapped in the same region where, at 112 Mb, is also located *ELOVL6* gene, a strong candidate that was already reported to be associated with palmitoleic content in porcine backfat tissue in our previous study (Zappaterra *et al*., 2018) and to palmitoleic to palmitic ratio in muscle tissue (Corominas *et al*., 2013). This gene codes for an elongase catalyzing the first and rate-limiting reaction of the four constituting the long-chain FA elongation cycle. This enzyme has a higher activity towards C16:0 acyl-CoAs. These results strongly support the importance of the 112 Mb region on SSC8, suggesting that a consistent candidate gene for the associations found in the present study could be identified in *ELOVL6*. Furthermore, another interesting candidate *locus*

where several markers associated with stearic and palmitoleic/palmitic ratio are mapped is located on SSC14 at 111-114 Mb. This region harbors 25 protein-coding genes, two ribosomal RNA genes, and three noncoding RNAs (Supplementary Material S1), and is known to be a region of particular interest for fat and FA deposition since comprises the known functional candidate gene *SCD*. The enzyme encoded by this gene introduces the first double bond into saturated fatty acyl-CoA substrates and catalyzes the insertion of a cis double bond at the delta-9 position into fatty acyl-CoA substrates including palmitoyl-CoA and stearoyl-CoA. Interestingly, the association with the palmitoleic/palmitic desaturation index noticed for this SSC14 region agrees with the functions of the SCD enzyme. The same *locus* was also detected in our previous study (Zappaterra *et al*., 2018), where the same chromosomal region was significantly associated with backfat proportions of oleic and MUFA in Italian Large White pigs. Furthermore, the same region was also found to be associated with IMF FA composition in Duroc populations (Ros-Freixedes *et al*., 2016; van Son *et al*., 2017), strengthening the hypothesis that this Quantitative Trait Locus (**QTL**) on SSC14 may have a prominent role in FA biosynthesis. A wider QTL region comprising also the sequence harboring *SCD* gene was detected by Park *et al*. (2017), who described a QTL located at 93.4-140.2 Mb on SSC14 associated with *Longissimus lumborum* palmitoleic acid content in crossbred pigs. Also, two other candidate genes are located near *SCD*: *ELOVL FA elongase 3* (***ELOVL3***) and *NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial* (***NDUFB8***). *ELOVL3* gene encodes for another elongase that catalyzes the first and rate-limiting reaction of the four reactions constituting the long-chain FA elongation cycle. This enzyme has higher activity toward C18:0 acyl-CoAs, and it participates in the production of SFA and MUFA of different chain lengths that are involved in multiple biological processes as precursors of membrane lipids and lipid mediators. While *ELOVL3* is directly involved in FA *de novo* biosynthesis, *NDUFB8* gene is related to FA oxidation since it encodes for an

accessory subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I). This subunit is believed to be involved in the electron transfer from NADH to the respiratory chain, and NDUFB8 protein was found to be less expressed in insulin-resistant animal models with increased FA oxidation (Vazquez *et al*., 2015). Therefore, even though *SCD* and *ELOVL3* are considered the most interesting candidate genes lying in this SSC14 region, also *NDUFB8* gene may participate in the complex physiological balance between FA de novo biosynthesis and oxidation.

Gadoleic acid displayed significant associations with markers located at 46.8-49.0 Mb on SSC7. This region contains 25 protein-coding genes and four noncoding RNA genes. From the data reported in Supplementary Material S1, it is possible to observe that two putative functional candidate genes are mapped in this *locus*: *ELOVL5* and *ACSBG1*. Both these genes are involved in long-chain FA biosynthesis. In particular, the *ELOVL5* gene encodes for another elongase that catalyzes the first and rate-limiting reaction of the four constituting the long-chain FA elongation cycle. The enzyme allows for the addition of 2 carbons to the chain of long and very-long-chain FA per cycle. On the other side, ACSBG1 mediates the activation of long-chain FA for both the synthesis of cellular lipids and their degradation via β-oxidation (Steinberg *et al*., 2000). Interestingly, the same chromosomal region was also found significantly associated with the content of gadoleic acid stored in abdominal fat and muscle in a Duroc x Eurhalian pig population (Yang *et al.,* 2013), thus highlighting the evidence of an important role of this *locus* on C20:1 synthesis. Similar results for the same SSC7 genomic sequence were reported by Pena *et al.* (2016), Zhang *et al.* (2016), and Zhang *et al.* (2017) in different pig populations, demonstrating the existence of a strong association linking the SSC7 46.8-49.0 Mb region and IMF FA composition.

To our knowledge, this is the first study investigating the genomic regions associated with muscle FA composition in Italian Large White purebred heavy pigs selected for the production of high-quality dry-cured hams. Several significant chromosomal regions were found to be involved in the *Semimembranosus* muscle FA composition. Among them, key genes with prominent roles in FA desaturation and elongation, such as *SCD* and several members of the *ELOVL* gene family were observed. If confirmed by further studies, the obtained results suggest that the identification of causal mutations in these regions may be useful for selection schemes aimed at improving FA composition in pork products.

Acknowledgments

We thank the Italian pig breeders Association and Dr. Maurizio Gallo for providing the samples and the information about the studied individuals. The authors are deeply grateful to Professor Vincenzo Russo for the support and valuable advice. This work was supported by PRIN2015 national project (N. 201549TZXB001).

Declaration of interest

The authors declare that they have no competing interests.

Ethics statement

Sampling occurred with the permission of the Italian pig breeders Association. Animal care and slaughter of the animals used in this study were performed in compliance with the European rules (Council Regulation (EC) No. 1/2 005 and Council Regulation (EC) No. 1 099/2 009) on the protection of animals during transport and related operations and at the

time of killing. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health.

Software and data repository resources

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

References

Aulchenko YS, Ripke S, Isaacs A and van Duijn CM 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23, 1294-1296.

Barrett JC, Fry B, Maller J and Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263-265.

Bosi P and Russo V 2004. The production of the heavy pig for high quality processed products. Italian Journal of Animal Science 3, 309–321.

Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM and 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.

Davoli R, Catillo G, Serra A, Zappaterra M, Zambonelli P, Zilio DM, Steri R, Mele M, Buttazzoni L and Russo V 2019. Genetic parameters of backfat fatty acids and carcass traits in Large White pigs. Animal 13, 924–932.

Folch J, Lees M and Sloane Stanley GH 1957. A simple method for the isolation and purification of total lipides from animal tissues. Journal of Biological Chemistry 226, 497–509.

Fontanesi L, Schiavo G, Galimberti G, Calò DG and Russo V 2014. A genome wide association study for average daily gain in Italian Large White pigs. Journal of Animal Science 92, 1385–1394.

Fontanesi L, Schiavo G, Gallo M, Baiocco C, Galimberti G, Bovo S, Russo V and Buttazzoni L 2017. Genome-wide association study for ham weight loss at first salting in Italian Large White pigs: towards the genetic dissection of a key trait for dry-cured ham production. Animal Genetics 48, 103–107.

Huang DW, Sherman BT and Lempicki RA 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols 4, 44–57.

Jiménez-Colmenero F, Carballo J and Cofrades S 2001. Healthier meat and meat products: their role as functional foods. Meat Science 59, 5–13.

Juárez M, Dugan MER, Aldai N, Aalhus JL, Patience JF, Zijlstra RT and Beaulieu AD 2011. Increasing omega-3 levels through dietary co-extruded flaxseed supplementation negatively affects pork palatability. Food Chemistry 126,1716–1723.

Kim YM, Choi TJ, Ho Cho K, Cho ES, Lee JJ, Chung HJ, Baek SY and Jeong YD 2018. Effects of sex and breed on meat quality and sensory properties in three-way crossbred pigs sired by Duroc or by a synthetic breed based on a Korean native breed. Korean Journal of Food Science and Animal Research 38, 544-553.

Konovalova S, Liu X, Manjunath P, Baral S, Neupane N, Hilander T, Yang Y, Balboa D, Terzioglu M, Euro L, Varjosalo M and Tyynismaa H 2018. Redox regulation of GRPEL2 nucleotide exchange factor for mitochondrial HSP70 chaperone. Redox Biology 19, 37–45.

Lo Fiego DP, Macchioni P, Minelli G and Santoro P 2010. Lipid composition of covering and intramuscular fat in pigs at different slaughter age. Italian Journal of Animal Science 9, 200-205.

Lo Fiego DP, Santoro P, Macchioni P and De Leonibus E 2005. Influence of genetic type, live weight at slaughter and carcass fatness on fatty acid composition of subcutaneous adipose tissue of raw ham in the heavy pig. Meat Science 69, 107–114.

Park HB, Han SH, Yoo CK, Lee JB, Kim JH, Baek KS, Son JK, Shin SM, Lim HT and Cho IC 2017. Genome scan linkage analysis identifies a major quantitative trait loci for fatty acid composition in longissimus dorsi muscle in an F2 intercross between Landrace and Korean native pigs. Asian-Australasian Journal of Animal Science 30, 1061–1065.

Pena RN, Ros-Freixedes R, Tor M and Estany J 2016. Genetic marker discovery in complex traits: a field example on fat content and composition in pigs. International Journal of Molecular Sciences 17, 2100.

Pena RN, Noguera JL, García-Santana MJ, González E, Tejeda JF, Ros-Freixedes R and Ibáñez-Escriche N 2019. Five genomic regions have a major impact on fat composition in Iberian pigs. Scientific Reports 9, 2031.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ and Sham PC 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. America Journal of Human Genetics 81, 559–575.

Ros-Freixedes R, Gol S, Pena RN, Tor M, Ibáñez-Escriche N, Dekkers JCM and 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.

Schönberger J, Wang L, Shin JT, Kim SD, Depreux FFS, Zhu H, Zon L, Pizard A, Kim JB, Macrae CA, Mungall AJ, Seidman JG and Seidman CE 2005. Mutation in the transcriptional coactivator EYA4 causes dilated cardiomyopathy and sensorineural hearing loss. Nature Genetics 37, 418.

Steinberg SJ, Morgenthaler J, Heinzer AK, Smith KD and Watkins PA 2000. Very long-chain Acyl-CoA synthetases human “Bubblegum” represents a new family of proteins capable of activating very long-chain fatty acids. Journal of Biological Chemistry 275, 35162–35169.

van der Leij FR, Bloks VW, Grefhorst A, Hoekstra J, Gerding A, Kooi K, Gerbens F, te Meerman G and Kuipers F 2007. Gene expression profiling in livers of mice after acute inhibition of β-oxidation. Genomics 90, 680–689.

van Son M, Enger EG, Grove H, Ros-Freixedes R, Kent MP, Lien S and Grindflek E 2017. Genome-wide association study confirm major QTL for backfat fatty acid composition on SSC14 in Duroc pigs. BMC Genomics 18, 369.

Vazquez EJ, Berthiaume JM, Kamath V, Achike O, Buchanan E, Montano MM, Chandler MP, Miyagi M and Rosca MG 2015. Mitochondrial complex I defect and increased fatty acid

oxidation enhance protein lysine acetylation in the diabetic heart. Cardiovascular Research 107, 453–465.

Wiedemann N, Frazier AE and Pfanner N 2004. The protein import machinery of mitochondria. Journal of Biological Chemistry 279, 14473–14476.

Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, Hughes SI and Whittington FM 2008. Fat deposition, fatty acid composition and meat quality: A review. Meat Science 78, 343–358.

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

Yang B, Zhang W, Zhang Z, Fan Y, Xie X, Ai H, Ma J, Xiao S, Huang L and Ren J 2013. Genome-wide association analyses for fatty acid composition in porcine muscle and abdominal fat tissues. PLoS ONE 8, e65554.

Yang J, Lee SH, Goddard ME and Visscher PM 2011, GCTA: a tool for genome-wide complex trait analysis. American Journal of Human Genetics, 88, 76–82.

Zappaterra M, Ros-Freixedes R, Estany J and Davoli R 2018. Association study highlights the influence of ELOVL fatty acid elongase 6 gene region on backfat fatty acid composition in Large White pig breed. Animal 12, 2443–2452.

Zhang J, Zhang Y, Gong H, Cui L, Huang T, Ai H, Ren J, Huang L and Yang B 2017. Genetic mapping using 1.4M SNP array refined loci for fatty acid composition traits in Chinese Erhualian and Bamaxiang pigs. Journal of Animal Breeding and Genetics 134, 472–483.

Zhang W, Zhang J, Cui L, Ma J, Chen C, Ai H, Xie X, Li L, Xiao S, Huang L, Ren J and Yang B 2016. Genetic architecture of fatty acid composition in the longissimus dorsi muscle revealed by genome-wide association studies on diverse pig populations. Genetics Selection Evolution 48, 5.

Table 1. *Descriptive statistics, heritability (h^2) estimates and standard errors of the heritability for the fatty acids measured in porcine* Semimembranosus *muscle.*

Traits	N[1]	Minimum	Maximum	Mean	SD	h^2	SE h^2
C10:0	783	0.209	1.698	0.833	0.213	0.196	0.058
C12:0	783	0.201	1.340	0.670	0.181	0.289	0.065
C14:0	783	2.718	14.120	8.784	2.046	0.310	0.065
C16:0	783	67.900	235.531	151.196	29.010	0.225	0.061
C16:1 *cis*-9	783	4.643	36.268	18.883	4.857	0.389	0.063
C17:0	783	0.510	1.864	0.983	0.209	0.378	0.065
C17:1 *cis*-9	783	0.719	2.773	1.469	0.337	0.090	0.048
C18:0	783	0.786	136.904	75.974	14.762	0.282	0.061
C18:1 *cis*-9	783	100.206	419.702	263.525	58.864	0.160	0.057
C18:1 *cis*-11	783	11.339	44.131	25.152	4.784	0.301	0.063
C18:2 *cis*-9, *cis*-12	783	35.079	128.232	68.332	12.930	0.137	0.052
C18:3 *n*-6	783	0.093	1.929	0.710	0.199	0.041	0.039
C18:3 *n*-3	783	0.109	5.140	2.400	0.702	0.188	0.058
C20:0	783	0.080	1.998	1.001	0.263	0.255	0.058
C20:1 *cis*-11	783	0.168	9.000	4.440	1.188	0.348	0.062
C20:2 *n*-6	783	1.446	5.108	2.863	0.595	0.307	0.061
C20:3 *n*-6	748	0.630	2.964	1.453	0.386	0.105	0.051
C20:4 *n*-6	783	3.155	27.573	11.265	4.116	0.148	0.050
C20:3 *n*-3	783	0.013	1.945	0.563	0.130	0.326	0.063
C20:5 *n*-3	783	0.003	0.258	0.080	0.025	0.110	0.048
C22:1	748	0.016	0.493	0.122	0.046	0.139	0.055
C22:2 *n*-6	783	0.006	1.697	0.572	0.294	0.215	0.057
C22:4 *n*-6	783	0.644	3.698	1.764	0.501	0.238	0.057
C22:5 *n*-3	783	0.166	2.461	0.671	0.220	0.161	0.055

C22:6 *n*-3	783	0.026	0.738	0.177	0.076	0.247	0.067
Total FA	783	315.941	918.311	643.811	112.783	0.149	0.056
n-6/*n*-3	783	12.530	53.270	22.856	3.555	0.229	0.064

[1]Number of analysed samples. The C20:3 *n*-6 and C22:1 fatty acids amounts were available on a smaller number of samples because they were included in the list of fatty acids to be analysed only later.

Table 2. *The genome-wide significant markers associated with pig* Semimembranosus *muscle fatty acid composition.*

Trait	Marker	Marker position on assembly 11.1	Type of variant	Nearest gene[1]	MAF[2]	Bonferroni corrected *P*	-log10(*P*)
C10:0	ALGA0035942	6:89758315	downstream gene variant	*ZSCAN20*	0.048	2.93E-03	7.136
C10:0	MARC0036862	6:89829920	intron variant	*CSMD2*	0.042	4.68E-02	5.933
C10:0	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	7.92E-03	6.705
C14:0	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	1.18E-02	6.532
C16:0	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	1.16E-02	6.539
C16:1 *cis*-9/C16:0	ALGA0049219	8:112372903	synonymous variant	*LRIT3*	0.240	3.07E-03	7.116
C16:1 *cis*-9/C16:0	ASGA0039646	8:112530303	intron variant	*CASP6*	0.124	2.39E-03	7.224
C16:1 *cis*-9/C16:0	M1GA0012034	8:112566311	intron variant	*MCUB*	0.123	3.48E-03	7.062
C16:1 *cis*-9/C16:0	ASGA0039666	8:112601001	intron variant	*MCUB*	0.124	2.96E-02	6.132
C16:1 *cis*-	ALGA0081091	14:111483985	intergenic variant	*SCD*	0.222	4.42E-07	10.958

9/C16:0							
C16:1 *cis*-9/C16:0	CASI0010164	14:111646452	intron variant	*HIF1AN*	0.224	4.42E-07	10.958
C16:1 *cis*-9/C16:0	ALGA0081097	14:111671437	intergenic variant	*HIF1AN*	0.224	4.42E-07	10.958
C16:1 *cis*-9/C16:0	ASGA0066116	14:111809833	intergenic variant	*PAX2*	0.199	1.81E-05	9.345
C16:1 *cis*-9/C16:0	MARC0063250	14:111842298	intron variant	*PAX2*	0.193	1.12E-02	6.552
C16:1 *cis*-9/C16:0	H3GA0042070	14:111950280	intergenic variant	*PAX2*	0.179	2.89E-02	6.143
C16:1 *cis*-9/C16:0	ASGA0066144	14:112329198	intergenic variant	*LBX1*	0.185	1.72E-02	6.367
C16:1 *cis*-9/C16:0	INRA0046731	14:112536527	intron variant	*BTRC*	0.179	1.72E-02	6.367
C18:0	ALGA0081091	14:111483985	intergenic variant	*SCD*	0.222	2.02E-02	6.299
C18:0	CASI0010164	14:111646452	intron variant	*HIF1AN*	0.224	2.02E-02	6.299
C18:0	ALGA0081097	14:111671437	intergenic variant	*HIF1AN*	0.224	2.02E-02	6.299

C18:0	MARC0031817	14:113426095	intron variant	*FBXL15, PSD*	0.212	3.93E-02	6.009
C18:1 *cis*-9	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	3.05E-02	6.119
C18:1 *cis*-9/C18:0	DRGA0000432	1:30402065	intron variant	*EYA4*	0.010	4.78E-10	13.924
C18:1 *cis*-9/C18:0	ALGA0002592	1:35022863	intron variant	*PTPRK*	0.034	3.82E-02	6.021
C18:1 *cis*-9/C18:0	INRA0001776	1:36325329	intergenic variant	*RSPO3*	0.031	1.53E-03	7.417
C18:1 *cis*-9/C18:0	ASGA0002314	1:36345917	intergenic variant	*RSPO3*	0.022	2.55E-05	9.197
C18:1 *cis*-9/C18:0	INRA0001968	1:42521905	intergenic variant	ENSSSCG00000004243	0.031	7.13E-03	6.750
C18:1 *cis*-9/C18:0	ALGA0003145	1:47954023	intergenic variant	-	0.037	2.27E-03	7.248
C18:1 *cis*-9/C18:0	DIAS0000342	1:53037420	intron variant	*DDX43*	0.018	4.33E-07	10.966
C18:1 *cis*-9/C18:0	MARC0063377	1:55613197	upstream gene variant	*ZNF292*	0.010	2.73E-07	11.167

C18:1 *cis*-9/C18:0	ALGA0025221	4:56843082	intergenic variant	*HEY1*	0.020	4.84E-04	7.918
C18:1 *cis*-9/C18:0	M1GA0005902	4:59353020	intergenic variant	*ZFHX4*	0.018	1.29E-04	8.493
C18:1 *cis*-9/C18:0	ALGA0025326	4:59505125	intron variant	*ZFHX4*	0.019	1.53E-03	7.418
C18:1 *cis*-9/C18:0	H3GA0012852	4:64930317	intergenic variant	*NCOA2*	0.018	1.68E-04	8.377
C18:1 *cis*-9/C18:0	ASGA0095534	8:3792767	upstream gene variant	*GRPEL1*	0.015	1.33E-09	11.281
C18:1 *cis*-9/C18:0	CADI0000659	8:43571149	missense variant	*CPE*	0.016	2.99E-05	9.128
C18:1 *cis*-9/C18:0	ASGA0073484	16:54746828	intergenic variant	ENSSSCG00000036842	0.018	1.33E-09	13.481
C18:3 *n*-3	INRA0022476	6:145288074	intergenic variant	-	0.494	4.62E-02	5.938
C18:3 *n*-3	MARC0061838	6:145718706	synonymous variant	*WDR78*	0.419	4.89E-02	5.914
C20:0	ASGA0000671	1:6702897	intron variant	*PRKN*	0.013	1.59E-02	6.403
C20:1 *cis*-11	H3GA0021494	7:47364951	intron variant	*CHRNA3, CHRNA5*	0.226	1.47E-02	6.435

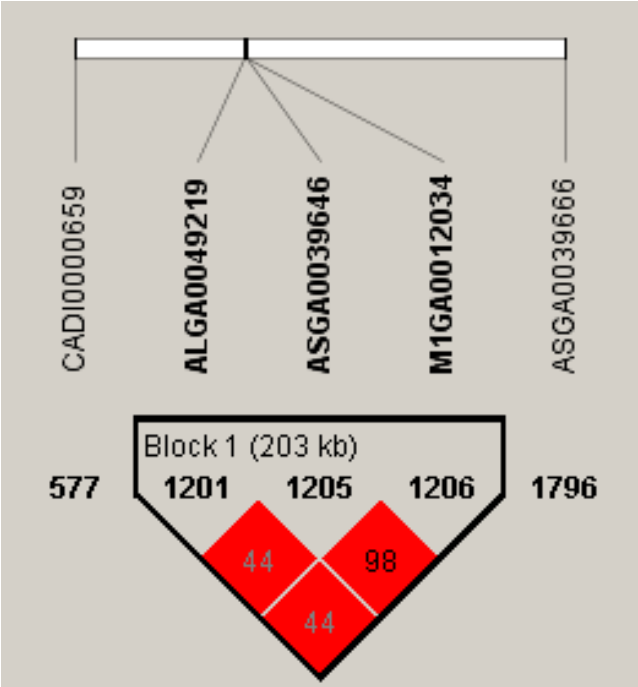
C20:1 *cis*-11	INRA0025702	7:47440163	downstream gene variant	*PSMA4*	0.309	1.66E-02	6.382
C20:1 *cis*-11	DRGA0007600	7:47526194	intron variant	*IREB2*	0.232	1.71E-02	6.369
C20:1 *cis*-11	H3GA0021497	7:47656708	intron variant, intron variant	*DNAJA4, WDR61*	0.238	2.35E-02	6.233
C20:1 *cis*-11	DIAS0002055	7:47774266	missense variant	*IDH3A*	0.218	2.10E-02	6.282
C20:1 *cis*-11	ALGA0041580	7:48022423	intergenic variant	*MORF4L1, ADAMTS7*	0.226	1.66E-02	6.384
C20:1 *cis*-11	ALGA0041669	7:48384635	intergenic variant	*ANKRD34C*	0.347	1.63E-02	6.390
C20:3 *n*-3	CASI0001986	5:67629097	downstream gene variant	*KDM5A*	0.030	2.27E-05	9.247
C20:3 *n*-3	ALGA0032824	5:71968480	intron variant	ENSSSCG00000027998	0.050	4.95E-05	8.909
C20:3 *n*-3	ALGA0057745	10:20170277	intron variant	*CRB1*	0.050	1.78E-02	6.353
C22-6 *n*-3	ALGA0095537	17:48453260	intron variant	*TP53RK*	0.013	4.65E-03	6.936
n-6/*n*-3	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	3.05E-02	6.119
total fatty acids	SIRI0000267	7:87516360	intergenic variant	*SV2B*	0.127	7.87E-03	6.707

[1] - means that no genes were found in the region 500 kb up- and downstream the significant marker.

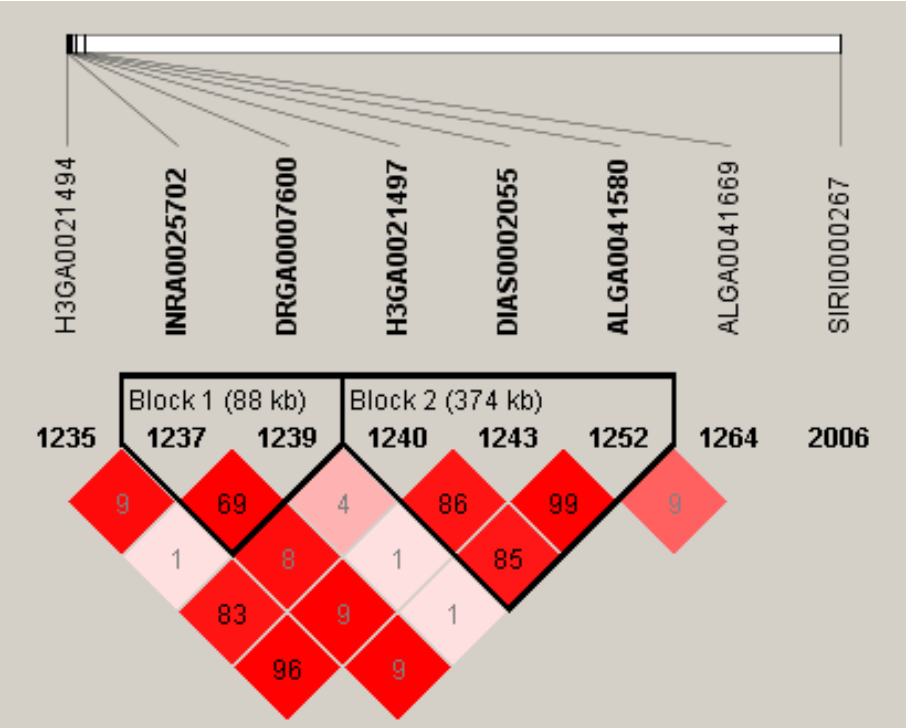
[2] Minor allele frequency.

Supplementary Table S1. *The composition of the diet for growing-finishing pigs.*

Components	Raw Diet
Digestible Energy (MJ/ration)	13 339.738
Composition (%)	
Water	11.642
Crude Protein	14.485
Crude Fat	4.068
Crude Fiber	4.498
Starch	45.274
Lysine	0.775
Methionine	0.230
Methionine + Cysteine	0.485
Tryptophan	0.161
Threonine	0.501
Calcium	0.754
Phosphorus	0.526
Digestible Phosphorus	0.303
Sodium	0.117
Linoleic acid	1.829



Supplementary Figure S1. *Linkage disequilibrium blocks of markers on Sus scrofa chromosome 8 between 112.3 and 112.7 Mb.*



Supplementary Figure S2. *Linkage disequilibrium blocks of markers on Sus scrofa chromosome 7 between 46.8 and 49.0 Mb.*