Association study between backfat fatty acid composition and SNPs in candidate genes highlights the effect of FASN polymorphism in large white pigs

Association study between backfat fatty acid composition and SNPs in candidate genes highlights the effect of FASN polymorphism in Large White pigs

Zappaterra M., Luise D., Zambonelli P., Mele M., Serra A., Nanni Costa L., Davoli R.

Meat science vol 156, 2019 : 75-84

The final published version is available online at: https://doi.org/10.1016/j.meatsci.2019.05.013

Association study between backfat fatty acid composition and SNPs in candidate genes highlights the effect of *FASN* polymorphism in Large White pigs

Martina Zappaterra[a], Diana Luise[a], Paolo Zambonelli[a], Marcello Mele[b], Andrea Serra[b], Leonardo Nanni Costa[a], Roberta Davoli[a]

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

[b] Department of Agriculture, Food and Environment, University of Pisa, via del Borghetto 80, 56124, Pisa, Italy.

Corresponding author Roberta Davoli: roberta.davoli@unibo.it

Martina Zappaterra: martina.zappaterra2@unibo.it
Diana Luise: diana.luise2@unibo.it
Paolo Zambonelli: paolo.zambonelli@unibo.it
Marcello Mele: marcello.mele@unipi.it
Andrea Serra: andrea.serra@unipi.it
Leonardo Nanni Costa: leonardo.nannicosta@unibo.it

Abstract

The objective of the study was to test the associations between backfat fatty acid composition (FAC) in a sample of 536 Italian Large White pigs and single nucleotide polymorphisms (SNPs) located in candidate genes, using univariate and multivariate approaches. The strongest associations were identified for the SNP AY183428 c.265T>C in *Fatty acid synthase* (*FASN*) gene, with the T allele linked to lower backfat contents of stearic ($P = 0.003$) and arachidic ($P < 0.0001$) acids, and increased amounts of dihomo-γ-linolenic ($P = 0.003$) and arachidonic ($P = 0.009$) acids. These associations were in agreement with the results of a multivariate analysis performed on backfat FAC, and an *in silico* analysis of the sequence flanking *FASN* SNP suggested that the T allele may disrupt a putative exonic splicing enhancer sequence therefore possibly affecting FASN activity. If the results will be further confirmed, the studied *FASN* SNP could be of particular interest for better understanding the regulative patterns controlling porcine backfat FAC.

Keywords: swine; adipose tissue; marker; repeated motif; meat quality; multivariate.

1 Introduction

Fatty acid (FA) composition of meat and meat products, such as dry-cured hams, is influenced by diet, genetics, breed, sex and environmental factors (Lo Fiego et al., 2005; Pena et al., 2016). Backfat FA composition showed to be a moderately heritable trait in several pig populations (Suzuki et al., 2006; Sellier et al., 2010; Davoli et al., 2018), with a heritability (h^2) value in Italian Large White ranging from 0.106 for C12:0 to 0.264 for monounsaturated FA (MUFA) (Davoli et al., 2018), suggesting the existence of a genetic control regulating these phenotypes (Davoli et al., 2018). So far, efforts to change FA composition of pig fat depots have been mainly concerned with the integration of unsaturated fatty acid sources in pig diets (Dugan et al., 2015), but over the last years, the identification of genetic markers related to pork FA composition has gained growing

attention. Indeed, a growing awareness of the association between diet and health has led nutritional quality to become a relevant factor in consumers' food choices. The increasing consumers' demand for healthier food has promoted the production of meats with improved nutritional quality and FA composition, with lower omega-6 (*n*-6) to omega-3 (*n*-3) ratio and enriched in omega- 3 polyunsaturated fatty acids ((*n*-3)PUFA) (Howe et al., 2007; Konieczka et al., 2017). The benefits of *n*-3 long-chain PUFA and their importance in human diet have been proved by a large scientific literature (Kris-Etherton et al., 2003; Nicolaou et al., 2014; Gammone et al., 2019), but at the same time PUFA are more likely to incur in lipolytic and oxidative processes than saturated fatty acids (SFA) (Serra et al., 2018). Therefore high contents in PUFA, and in particular in (*n*-3)PUFA, are detrimental for meat technological quality and in particular for dry-cured hams (Lo Fiego, 1996; Wood et al., 2004; Juárez et al., 2011). For these reasons, dry-cured ham industry requires meat with a limited amount of PUFA (Bosi and Russo, 2004) and adequate contents of SFA, which in turn are generally thought to have a negative impact on human health (Mensink et al., 2003). In this scenario, the knowledge of genes involved in fat deposition and FA composition and the detection of genetic markers could offer new opportunities to drive specific selection schemes aimed to obtain meat products with a nutritional composition that meets both the consumers' and industry requests. Several markers have been associated with intramuscular fat (IMF) deposition and composition (Ballester et al., 2016; Latorre et al., 2016; Ros-Freixedes et al., 2016), but the knowledge on genes and markers with a major effect on backfat thickness and FA composition are still scant. FA synthesis depends on several key enzymes including Acetyl-CoA carboxylase 1 (ACC1), ATP citrate lyase (ACL) and Fatty acid synthase (FAS) (Green et al., 2010) encoded respectively by *ACACA*, *ACLY*, and *FASN* genes. In the cytoplasm, the first rate-limiting step in the synthesis of *de novo* palmitic acid (C16:0) is encoded by ACC1 enzyme, which converts acetyl-CoA to malonyl CoA, the first step of the palmitate synthesis (Gallardo et al., 2009). Then, ACL and FAS enzymes contribute to the synthesis of *de novo* FA, by producing acetyl-CoA from citrate and by catalysing the cyclic condensation of acetyl-CoA, respectively. Palmitic acid may be further elongated to

stearic acid (C18:0) and long chain fatty acids in the endoplasmic reticulum membranes. Moreover, C16:0 and C18:0 SFA are substrates for the *de novo* synthesis of Monounsaturated FA (MUFA), according to the coordinated regulation of elongases and desaturases, such as Stearoyl-CoA desaturase (SCD) (Green et al., 2010). In pigs, mutations of *ACACA*, *ACLY*, *FASN*, and *SCD* genes have been identified but the effect of these polymorphisms on FA metabolism and on backfat FA composition are still poorly known. This research was carried out in a sample of Italian Large White (ILW) heavy pigs and was addressed towards the study of the associations between porcine backfat FA composition and SNPs located in the genes *ACACA*, *ACLY*, *FASN* and *SCD*. To improve the interpretation of results and to account for the complex and multivariate nature of FA synthesis and deposition in porcine backfat, the relations among the measured FA were analysed using different approaches and additional investigations on *FASN* gene sequence were performed. The possible functional link between *FASN* AY183428 c.265T>C variants and backfat FA composition were then analysed and discussed.

2 Materials and Methods

2.1 Animals

The study was carried out on 536 ILW pigs, randomly chosen among a wider population of 950 pigs that has been used in previous work (Davoli et al., 2018). The population considered in the present research was composed of 357 gilts and 169 barrows included in the national selection sib test program carried out in the same environmental conditions at the Central Station of the Italian National Pig Breeders Association (Associazione Nazionale Allevatori Suini, ANAS, http://www.anas.it). Pigs were fed the same diet with a *quasi ad libitum* feeding level from about 30 kg live weight to about 155 kg live weight. Animal care and slaughter of the animals used in this study were performed in compliance with the European rules (Council Regulation (EC) No. 1/2005 and Council Regulation (EC) No. 1099/2009) on the protection of animals during transport and

related operations and at the time of the killing. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health and sampling occurred with the permission of ANAS. Immediately after slaughtering, samples of *Semimembranosus* muscle and backfat were collected, snap-frozen in liquid nitrogen and stored at –80°C until analysis.

2.2 Phenotypes

The phenotypes used for the present research are part of a set of data and measures concerning a bigger sample described in Davoli et al. (2018).

At slaughtering, hot carcass weight (kg), the ultrasonic measure (in mm) taken by Fat-O-Meat'er (FOM - Frontmatec A/S, Kolding, Denmark) between the third and fourth last ribs, 8 cm off the carcass midline of backfat and estimated percentage of lean cuts on carcass were recorded.

Backfat FA composition and intramuscular fat (IMF) content in *Semimembranosus* muscle were processed using an external service. IMF was determined by extraction with petroleum ether from 1 g of fresh *Semimembranosus* muscle by means of an XT15 Ankom apparatus (Macedon, NY, USA), according to Official procedure AOCS Am 5-04 (AOAC, 2005). IMF was expressed in % as g of IMF per 100 g of muscle tissue.

Backfat lipids were extracted according to Serra et al. (2014) as described in Davoli et al. (2018) using a Gas-Chromatograph apparatus (GC 2010 plus, Shimadzu, Columbia, MD, USA) equipped with a flame ionization detector and with a high polar capillary column (SP 2560 100m × 0.25mm, Supelco, Bellefonte, PA, USA).

FA methyl-esters (FAMEs) were identified by comparison with commercial standard mix of FAMEs (GLC-674 Nuchek, Elysian, MN, USA) addicted with single FA (Nuchek, Elysian, MN, US; Larodan, Malmö, Sweden): a complete standard of 105 FAMEs as obtained and compared with the studied samples. Results were expressed as g of FA on 100 g of the total FA after the conversion of FAMEs into FA.

The analysed backfat FA and FA classes are listed in Table 1.

2.3 DNA extraction and genotyping

Genomic DNA was isolated from *Semimembranosus* muscle samples using the Wizard® Genomic DNA Purification Kit (Promega Corp., Madison, WI, USA). Extracted DNA quality and quantity were tested with spectrophotometer NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA, USA). A total of six SNPs reported *in silico* and in the literature were considered for this study (Supplementary Table 1). The polymorphisms *ACLY* c.NM_001105302 g.2956 T>C (Davoli et al., 2014), *ACLY* NM_001105302 g.3923 T>C (Muñoz et al., 2013), *FASN* AY183428 c.265 T>C (rs324640280; Braglia et al., 2014) and *SCD* AY487830 g.2228T>C (rs80912566; Estany et al., 2014) were analyzed by PCR-RFLP technique on PTC-100 thermal cycler (AB Applied Biosystems, Life Technologies, Carlsbad, CA, USA). Two SNPs lying on *ACACA* gene sequence were also tested: the mutation *ACACA* NM001114269 c.5196T>C (rs81434689; Gallardo et al., 2009) in exon 40 and the intron variant NC010454.4:g.38621011G>A (rs81434677; Muñoz et al., 2013). The *ACLY*, *FASN*, *SCD* and *ACACA* NM001114269 c.5196T>C polymorphisms were genotyped using PCR-RFLP: PCR reactions were prepared in a total volume of 20 μl containing 2 μl of 10x buffer, different concentrations of $MgCl_2$ depending on the primer pair, 0.5 pmol of each primer, 160 μM of dNTP, 1 U of Fisher Scientific Taq polymerase (Fisher Molecular Biology, Trevose, PA, USA) and 1 μl of 50–100 ng of DNA. A complete list of the used primer sequences, annealing temperatures, and restriction enzymes is provided in Supplementary Table 1. *ACACA* NC010454.4:g.38621011G>A SNP was genotyped using High-Resolution Melting PCR (HRM) on Rotor-Gene™ 6000 (Corbett Research, Mortlake, New South Wales, Australia), in a total volume of 20 μl containing 2 μl of 10x standard buffer, 3 mM $MgCl_2$, 0.5 pmol of each primer, 160 μM of dNTP, 1 U of Fisher Scientific Taq polymerase (Fisher Molecular Biology, Trevose, PA, USA), 1 U of EvaGreen ™ (Biotium Inc., Hayward, CA, USA) and 50–100 ng of template DNA. Cycling conditions were: initial denaturation at 95°C for 5 min, 40 cycles of 95°C for 20 s, 59°C for 20 s

and 72°C for 25 s, followed by a final extension step of 72°C for 2 min. The HRM data were analysed by Rotor-Gene^{TM} 6000 software.

2.4 Statistical analysis

2.4.1 Investigation of the correlations and Principal Component Analysis

Means and standard deviations were calculated for the measured phenotypes. The Pearson phenotypic correlations between measured traits were obtained using library *Hmisc* on R environment version 3.4.4 (R Core Team, 2018). The correlation matrix was plotted using *corrplot* package on R environment version 3.4.4 (R Core Team, 2018).

Furthermore, with the aim of identifying possible correlation patterns among the measured FA, a Principal Component Analysis (PCA) was performed on the single FA to convert the set of correlated traits to a smaller number of linear combinations of uncorrelated variables, called Principal Components (PC). The PCA was computed with *ropls* package (Thévenot et al., 2015) using mean-centering and unit variance scaling in the R environment (R Core Team, 2018). The *ropls* package implements PCA with the NIPALS-based versions of the algorithms reported in Wold et al. (2001) and Trygg and Wold (2002).

2.4.2 Association study

2.4.2.1 Univariate approach

For the analysed SNPs, the Hardy-Weinberg equilibrium was evaluated by Chi-square test (*P-value* <0.001).

The association study was performed using two approaches. The first approach consisted of using mixed models to test for the existence of associations between each measured phenotype and the SNPs in the studied candidate genes. This analysis was assessed using three different mixed models (proc MIXED) in SAS version 9.4 software (SAS Institute, Cary, NC, USA). The three models were utilized depending on the kind of phenotypes analysed, and were the following:

Model 1: $y_{i,j,k,l} = \mu + G_i + S_j + D_k + b1W + \text{SIRE} + e_{i,j,k}$

Model 2: $y_{i,j,k,l} = \mu + G_i + S_j + D_k + b1W + b2T + \text{SIRE} + e_{i,j,k}$

Model 3: $y_{i,j,k,l} = \mu + G_i + S_j + D_k + b2T + \text{SIRE} + e_{i,j,k}$

The dependent variables *y* were for Model 1 the carcass and meat quality records, for Model 2 the IMF content, for Model 3 the single FA and FA classes; μ is mean of the population; G is the fixed effect of the genotype (*i* = 1, 2, 3) at the studied *loci* (*ACACA, ACLY, FASN, SCD); S* is the fixed effect of sex (*j* = 1, 2); *D* is the fixed effect of slaughter date (*k* = 1… 27); *b1* is the linear regression coefficient of the observation on slaughter weight; *b2* is the linear regression coefficient of the observation on backfat thickness; *W* is the slaughter weight (linear covariate); *T* is the backfat thickness (linear covariate); SIRE is the random effect of sires; $e_{i,j,k}$ is the random error. These three models have been used in order to account for the different variables affecting each trait category. Additive and dominance effects were estimated and tested by *t*-test on significant deviation from zero as reported by Chung and Kim (2005). A first correction of the least-squares mean values for all genotype-trait combinations was computed using the Tukey-Kramer test in SAS (TUKEY option). The *P*-values were adjusted by False Discovery Rate (FDR) using MULTTEST procedure by SAS software 9.4 (SAS Institute Inc., Cary, NC, USA). The results were considered significant for FDR adjusted *P*-value ≤ 0.10 as reported by Macciotta et al. (2017).

2.4.2.2 Association study between the polymorphisms and Principal Component scores

A further association study was performed in order to test if the studied *ACACA*, *ACLY*, *FASN* and *SCD* SNPs were significantly associated with the PC scores obtained by the PCA performed on the single FA reported in section 2.4.1. For this association study, we followed an approach proposed by Bolormaa et al. (2010): the scores of the computed PCs were used as dependent variables and the effect of the genotypes on these new variables were analysed using the mixed Model 3 reported in section 2.4.2.1 (proc MIXED in SAS version 9.4 software). The significant associations (*P*-value <

0.05) between the SNPs fixed effects and the scores of the PCs were further discussed considering the weights of the individual FA in each PC (PC loadings).

2.4.2.3 Association study between arachidonic acid and combined genotypes at FASN *AY183428 c.265T>C and* SCD *AY487830 g.2228T>C loci*

Since both *FASN* AY183428 c.265T>C and *SCD* AY487830 g.2228T>C SNPs showed an effect on backfat arachidonic acid content, a further analysis was computed to test if the different genotype combinations obtained for *FASN* AY183428 c.265T>C and *SCD* AY487830 g.2228T>C SNPs were associated with backfat arachidonic acid content using a mixed model (proc MIXED) in SAS version 9.4 software (SAS Institute, Cary, NC, USA). This model was the same used to perform the association with backfat FA (Model 3) and in this case, the SNP effect G had six possible levels, resulting from the combinations of the three *FASN* genotypes (CC, CT, TT) and the two *SCD* genotypes (C-, TT). For this analysis, only two *SCD* genotypic classes were considered for this *locus*, TT and C-. Both *SCD* CC and CT samples were included in the class *SCD* C-, due to the low frequency of *SCD* C allele.

2.5 Investigation of *FASN* AY183428 c.265T>C flanking sequence

A deeper investigation of *FASN* AY183428 c.265T>C SNP surrounding sequence was carried out with the aim of highlighting possible effects of the polymorphism on *FASN* gene activity. A sequence of about 200 bp length surrounding the *FASN* AY183428 c.265T>C SNP was uploaded in MEME suite version 5.0.4 (Bailey et al., 2009) for the discovery of repeated motifs. The sequences of the motifs obtained by MEME were then uploaded in GOMo (Gene Ontology for Motifs) version 5.0.4 (Buske et al., 2010). GOMo on-line tool analyses the uploaded motifs and scans the sequences of other genes in order to find if the given motifs are maintained. Then, if the considered motifs are found in other genes, GOMo identifies the gene ontologies that are linked to the genes having the searched repeated motif. The latter information could be important for the study of regulative

repeated sequences since significant gene ontology terms could suggest biological roles for the provided motifs. *Mus musculus* database has been used as the reference database for the GOMo analyses since *Sus scrofa* was not comprised among the input reference genomes.

3 Results and Discussion

3.1 Description of the studied population, phenotypic correlations, and multivariate analysis

Table 1 reports mean values and standard deviations for the recorded phenotypes in the analysed samples. The high carcass weight observed in the studied pig population approaches the weight required for heavy pigs grown for the production of high-quality dry-cured hams, and the lean mass deposition is in line with the 40-54.9% of carcass lean meat content required for these productions. In agreement with the literature (Piasentier et al., 2009), the most abundant FA classes in backfat were MUFA and SFA (Table 1).

The phenotypic correlations between measured traits are graphically represented in Figure 1 and the correlation coefficients and the relative *P*-values are reported in Supplementary Table 2. The measured phenotypic correlations are in agreement with the results reported in our previous study on a wider sample of Italian Large White pigs (Davoli et al., 2018). Backfat content of palmitic acid (C16:0) was strongly correlated with medium chain FA % (MCFA; Supplementary Table 2) and was also positively related to the amounts of stearic acid (C18:0), myristic acid (C14:0) and SFA (Supplementary Table 2). The strong phenotypic correlations observed between stearic, palmitic and myristic acids accurately reflect the biological relations linking the abundance of these three FA in backfat tissue, since the endogenous formation of stearic acid is strictly dependent on the elongation of myristic and palmitic, and therefore on their abundance (Inagaki et al., 2002). Other high positive correlations were also noticed between linoleic (C18:2(*n*-6)), dihomo-γ-linolenic

(C20:3(*n*-6)) and arachidonic acids (C20:4(*n*-6); with r ranging from 0.637 to 0.822, and $P <$ 0.0001; Supplementary Table 2). The strong relation existing among these three FA is explained by the fact that linoleic and dihomo-γ-linolenic acids take part in the endogenous synthesis of arachidonic acid. Indeed, linoleic acid is the essential FA precursor for the synthesis of arachidonic acid, which derives from the subsequent desaturation of linoleic acid into dihomo-γ-linolenic acid, and, finally, from the elongation and desaturation (catalysed by Δ-5-desaturase enzyme) of dihomo-γ-linolenic into arachidonic acid (reviewed in Knez et al., 2017). On the other side, linoleic, dihomo-γ-linolenic and arachidonic acids are negatively correlated with stearic, palmitic and arachidic (C20:0) acids with r values ranging from -0.369 to -0.608 and $P < 0.0001$ (Figure 1 and Supplementary Table 2). These results are also in agreement with the genetic correlations reported in Davoli et al. (2018) on a larger sample, where linoleic, dihomo-γ-linolenic and arachidonic acids showed strong negative genetic correlations with stearic and arachidic acids (with $r_g > -0.6$ and $P <$ 0.001). These strong negative genetic correlations indicate these FA share a high proportion of variance, suggesting that a common genetic cause may affect in an opposite way the synthesis of stearic and arachidic acids and the backfat deposition of linoleic, dihomo-γ-linolenic and arachidonic acids.

With the aim of identifying the patterns of correlations among the measured FA characterizing the studied population, a multivariate analysis (PCA) on individual FA was performed. The PCA was conducted on 20 individual FA (reported in Supplementary Table 3) and the loadings plot is reported in Figure 2. The largest amount of explained variance can be achieved considering the first two new variables, PC1 and PC2, explaining 33.9% and 18.7% of the total variance, respectively. The weights of the FA included in these two PC are reported in Supplementary Table 3. The FA with the highest weights in PC1 were stearic, linoleic, dihomo-γ-linolenic and arachidonic acids, while PC2 was mainly related to myristic, palmitic, palmitoleic (C16:1 *cis*-9), eicosadienoic (C20:2(*n*-6)) and C22:1 *cis*-12 acids (Figure 2 and Supplementary Table 3). Interestingly, these two PCs seem to relate FA included into two main FA biosynthetic metabolisms: the arachidonic acid

de novo synthesis in PC1 and the palmitoleic acid biosynthetic process in PC2. Indeed, PC1 comprises linoleic, dihomo-γ-linolenic and arachidonic acids, which, as previously discussed, resulted to be highly correlated with each other (Figure 1 and Supplementary Table 2) and are known to take part in arachidonic acid synthesis (reviewed in Knez et al., 2017). Together with those three FA, also stearic acid entered in PC1, but this FA showed a negative weight, in opposition to the positive weights noticed for linoleic, dihomo-γ-linolenic and arachidonic acids. This evidence could be explained by the negative genetic correlations existing between stearic and the FA related to arachidonic acid synthetic pattern, such as linoleic acid (with $r_g > -0.6$ and $P < 0.001$; Davoli et al., 2018). In the adipose tissue, stearic acid is synthesized by elongase 6 starting from palmitic acid, whereas (*n*-6)PUFA are synthesized by elongase 2 and 5 in coordination with Δ5 and Δ6 desaturases. Long-chain PUFA may regulate PPARa, which, in turn, promote the synthesis of elongase 6 activity (Jakobsson et al., 2006). On the other hand, most of the FA weighting in PC2 were those related to palmitoleic endogenous synthesis, which is catalysed by the elongation of myristic into palmitic acid and by its subsequent step of desaturation into palmitoleic acid (Revilla et al., 2014). The obtained scores of the samples for the two new variables (PC1 and PC2) were then used for the association study.

3.2 Association study

The allele frequencies for the genotyped SNPs are given in Table 2. Among the studied SNPs, *ACACA* NM001114269 c.5196T>C was the only one that did not segregate in the ILW samples. The other SNPs were present in the studied population and, except for *ACLY* NM_001105302 g.2956 T>C, the polymorphisms showed allele frequencies in Hardy–Weinberg equilibrium.
The first step of the association study between SNPs and phenotypic traits was carried out with a mixed model on each measured phenotype, FA and FA class taken individually, and the complete results are reported in Supplementary Table 4.

The associations with a nominal *P*-value < 0.05 found between the tested SNPs and the measured traits are reported in Table 3, together with the least squares means and effects estimated for the genotypes. *ACLY* NM_001105302 g.2956 T>C did not show any significant *P*-values, even when nominal *P*-values were calculated, while *ACACA* NC010454.4 g.38621011G>A, *ACLY* NM_001105302 g.3923 T>C and *SCD* AY487830 g.2228T>C showed some weak effects (nominal *P*-values < 0.05) on the considered phenotypes (Table 3). Despite these weak significances were not retained with FDR adjustment (adjusted *P* > 0.10), the associations found for *ACLY* NM_001105302 g.3923 T>C are in agreement with previous findings in Italian Large White and Italian Duroc pigs, where this polymorphism showed to have an effect on Lean Cut and Backfat Thickness Estimated Breeding Values (Davoli et al., 2014). Interestingly, the polymorphism in *FASN* sequence showed the most consistent associations, with significant FDR adjusted *P*-values. The strongest association was observed for arachidic acid, with an additive effect and the highest estimated mean value associated with *FASN* CC genotype. Together with the arachidic acid content, *FASN* CC genotype was also associated with increased amounts of stearic acid, and lower backfat contents of myristic, palmitoleic, asclepic (C18:1 *cis*-11), dihomo-γ-linolenic and arachidonic acids. The associations observed in the present work between *FASN* mutation and palmitoleic and stearic acids are concordant with the evidence reported by Maharani et al (2012) in Duroc pigs, where CC homozygotes had lower palmitoleic and higher stearic contents than CT pigs. The same polymorphism has also been investigated by Renaville et al. (2018), but few associations with FA composition of backfat and leg fat were observed. The discrepancies between the results of this research and those reported by Renaville et al. (2018) may be due to the different sizes and genetics of the studied populations (536 purebred ILW pigs in this study and 129 crossbred heavy pigs in Renaville et al., 2018). On the whole, the associations observed in the present study for *FASN* gene and myristic and palmitoleic acids are consistent with the effects on backfat FA composition observed for *FASN* gene mutations in other livestock species (Abe et al., 2009).

Furthermore, a second approach has been considered, in order to test the associations of the considered SNPs on sets of FA combined in new variables (the PCs: PC1 and PC2). This approach could be helpful when combined with commonly used univariate models, since the multivariate analysis can help to provide a more accurate interpretation of the biological complexity of metabolisms (Levin, 2014) that are at the basis of quantitative traits, such as FA composition. Additionally, the inclusion of highly correlated FA in a univariate model can be statistically problematic because of multicollinearity among the FA variables, which can inflate the standard errors of the estimates leading to a failure to reject a false null hypothesis (Morris and Tangney, 2014). The use of a multivariate analysis may offer a possible alternative analysis to overcome multicollinearity-related problems (Thévenot et al., 2015). Therefore, the obtained scores of the samples for the two new variables (PC1 and PC2) were then analyzed as dependent variables in a mixed linear model, with the aim of finding associations with the considered SNPs. The results of this analysis are reported in Table 4, where for each polymorphism the estimated least squares means and the genotype effects on the PCs are reported. Consistently with the results found with the univariate approach, *FASN* AY183428 c.265T>C was the only SNP among the tested ones to show significant associations with PC1 scores (Table 4). In particular, the *FASN* CC genotype presented a lower estimated least squares mean for PC1 scores compared with TT samples, suggesting that CC is associated with higher amounts of stearic acid and lower amounts of linoleic, dihomo-γ-linolenic and arachidonic acids when compared to *FASN* TT samples. Therefore, the overall *FASN* SNP effects observed from PCA are in agreement with the univariate results for stearic, dihomo-γ-linolenic and arachidonic acids, suggesting an effect of this gene on stearic acid and on the amounts of FA associated with *de novo* arachidonic acid synthesis.

In addition to the *FASN* AY183428 c.265T>C, the amount of backfat arachidonic acid showed to be linked also to *SCD* AY487830 g.2228T>C. This latter association was weaker than that observed for *FASN* polymorphism but suggested anyway that also *SCD* may have an effect on the amount of backfat arachidonic acid. Therefore, we decided to investigate whether the combination of the

different genotypes for these two *loci* may have enhanced effects on the backfat level of arachidonic acid. To this aim, a mixed model was performed adding as a fixed effect the six different possible combinations of *FASN* and *SCD* genotypes and the estimated means for the six classes of combined genotypes are graphically presented in Figure 2. The mixed model showed a significant effect for the combination of *FASN* and *SCD* genotypes on arachidonic acid content, and significant differences were noticed between the estimated least squares means of the six classes of combinations (Figure 3). The genotypic classes with *SCD* TT genotype showed lower estimated means than the combined genotypic classes with *SCD* C allele, and this effect is particularly evident when *SCD* TT genotype is combined with the CC genotype at the *FASN locus* (Figure 3). This trend may suggest that *SCD locus* could exert with *FASN* SNP a synergistic effect on the arachidonic acid backfat deposition, despite the strongest evidence of association remains for the studied *FASN* SNP.

3.3 Investigation on *FASN* AY183428 c.265T>C flanking sequence

FASN enzyme has a complex homodimeric structure with the major role of regulating the *de novo* synthesis of long-chain FA in mammals through the formation of 16-carbon FA from acetyl-CoA and malonyl-CoA (Chakravarty et al., 2004). This synthesis involves a cyclic-step elongation of activated precursors by 2 carbon units (Smith, 1994) and the growing FA is generally released by a thioesterase when the chain reaches 16 carbon atoms in length. Close to the FASN catalytic center where FA chains are anchored, there is a thioesterase domain with a groove structure (Chakravarty et al., 2004). The thioesterase domain spatial organization is of crucial importance for its function and mutations in this domain of the bovine *FASN* gene were already reported to have significant effects on the fatty acid composition of bovine adipose tissue synthesis (Zhang et al., 2008). These authors suggested that these polymorphisms may influence the structure of FASN thioesterase domain and affect the specific activity of thioesterase towards C-14 acyl carrier protein, resulting in a larger amount of palmitic being elongated to stearic acid, and subsequently converted to oleic acid. Despite *FASN* AY183428 c.265T>C SNP is a synonymous variant and does not bear direct

changes in the aminoacidic sequence, the consistent associations found in the present study with backfat FA composition may suggest that this mutation could be in linkage disequilibrium with other causal *FASN* polymorphisms or that the sequence harbouring the studied polymorphism could have a regulative effect on *FASN* gene. To investigate the latter hypothesis, the sequences flanking the studied *FASN* SNP were tested for the presence of regulative sites. Among the regulative sequences, sequence motifs are short, recurring patterns in DNA that are presumed to have a biological function (reviewed in D'haeseleer, 2006). Interestingly, a hexameric motif comprising the mutation sequence has been identified with the MEME online tool (Supplementary Table 5). The same hexameric motif 5'-AGCTGC-3' (or 3'-GCAGCT-5') has already been indicated in other organisms as a putative Exonic Splicing Enhancer (ESE) hexamer (Pertea et al., 2007). ESE sequences determine whether or not a regulated splice site, usually an upstream 3′ splice site, will be used for RNA splicing, and promote in mature mRNA the inclusion of exons in which they reside (Tacke and Manley, 1999; Zheng, 2004). The *FASN* AY183428 c.265 T allele carries a mismatch at the level of the second nucleotide of the identified 5'->3' hexameric sequence (or at the fifth nucleotide of its reverse complement), thus disrupting the found repeated motif. This result may concur to explain why the analysed synonymous variant in the *FASN* gene has shown strong effects on backfat FA composition. Indeed, polymorphisms in ESE sequences have already been proven to affect the regulation of RNA splicing, leading to the production of different transcripts (Uezato et al., 2015), and thus possibly affecting the coded protein activity. Interestingly Ensembl database reports two transcripts for porcine *FASN* gene: a longer one (ENSSSCT00000029960.2) comprising the exons from 1 to 43, and a shorter one (ENSSSCT00000026033.2) with exons from 28 to 35 (last accessed 7th of May 2019), suggesting the existence of alternative splicing for the first exons. The hypothesis that *FASN* AY183428 c.265 T allele may carry a sequence mismatch influencing *FASN* gene transcription could explain the results reported in Braglia et al. (2014) and in Zambonelli et al. (2009). Both these studies reported for the *FASN* transcript comprising the RNA sequence of the *FASN* first exons (where is located *FASN* AY183428 c.265T>C) significantly

lower levels in muscle and backfat tissues of pigs showing the *FASN* TT genotype compared to CC individuals. These previous results could be explained by the existence of the putative ESE sequence harbouring the *FASN* polymorphism, as hypothesised in the present research, but further dedicated studies are needed to prove this explanation.

Furthermore, aiming to identify possible biological roles for the putative ESE sequence, the repeated motif identified with MEME was analysed with GOMo, and the obtained results are reported in Supplementary Table 6. Among the most specific (specificity = 100%) and significant (adjusted *P*-value < 0.05) gene ontology terms, there are several terms related to calcium ion binding, cation channel activity, and calcium-dependent protein binding, to growth factor activity, to neurofilament and postsynaptic membrane, and to arachidonate 12-lipoxygenase activity. This latter term was of particular interest considering the strong associations of the *FASN* SNP with arachidonic acid noticed in the present study both with univariate and multivariate approaches. Indeed, arachidonate 12-lipoxygenase is an enzyme coded by the gene *Arachidonate 12-lipoxygenase, 12S Type* (*ALOX12*, also referred as *12-LOX*) and catalyzes the peroxidation of free and esterified PUFA generating a spectrum of bioactive lipid mediators. This dioxygenase mainly converts arachidonic acid to (12S)-hydroperoxyeicosatetraenoic acid/(12S)-HPETE, but can also use γ-linolenic, α-linolenic and linoleic acid as substrates (Yokoyama et al., 1986). Interestingly, the scientific literature has already reported a strong association between *FASN* gene and protein activity and arachidonate lipoxygenase family. Indeed, while arachidonic acid has been proven to suppress in different cells *FASN* mRNA synthesis and its enzyme activity (Armstrong et al., 1991; Moon et al., 2002), an increased arachidonate lipoxygenase activity was found to stimulate in human breast cancer cells *FASN* promoter activity through a positive feedback loop (Hu et al., 2011). This positive feedback loop was observed in particular among *FASN*, phosphorylated Extracellular signal-regulated kinase 1/2 (p-ERK1/2) and arachidonate 5-lipoxygenase (*ALOX5* gene, also referred as *5-LOX*) and was found to contribute to the growth of breast cancer cells (Hu et al., 2011). Despite we are not able to claim if the same regulation has taken place in our samples,

the existence of a co-regulation between *FASN* and arachidonate lipoxygenases family could be hypothesized and this may be one of the possible patterns explaining the strong associations noticed between *FASN* AY183428 c.265T>C and the content of arachidonic acid stored in porcine backfat tissue. The reported hypothesis seems to allow explaining several results found in the present study and in the literature, but would need further specific studies.

4 Conclusions

Backfat FA composition is a complex trait and strong correlations exist between individual FA. The complex nature of this trait and the high costs of its measure make it hard to consider FA composition a feasible objective of genetic selection schemes in livestock animals. Anyway, the identification of molecular markers associated with backfat FA composition has gained increasing interest, since understanding the network of genes and biological functions involved in each FA synthesis would make it possible to select pigs with modified backfat FA composition. *FASN* AY183428 c.265T>C mutation evidenced the strongest associations with the backfat FA composition in the studied purebred ILW heavy pigs. The studied *FASN* polymorphism may be a promising marker for pig selection schemes aimed at improving FA composition of meat products since this polymorphism was associated with changes in the backfat contents of myristic, stearic and arachidic and arachidonic acids. Further studies are needed to better understand the involvement of *FASN* AY183428 c.265T>C SNP in the FA composition of porcine tissues.

Acknowledgements

The authors thank Dr Luca Buttazzoni and Dr Maurizio Gallo from CREA and ANAS for providing the porcine backfat and *Semimembranosus* tissue samples and the data of the studied individuals. Funding: This work was supported by PRIN 2015 national project (Grant N. 201549TZXB001) and by AGER – Hepiget project (Grant N. 2011- 0279).

References

Abe, T., Saburi, J., Hasebe, H., Nakagawa, T., Misumi, S., Nade, T., … Kobayashi, E. (2009). Novel mutations of the *FASN* gene and their effect on fatty acid composition in Japanese Black Beef. *Biochemical Genetics*, *47*(5–6), 397–411.

Armstrong, M.K., Blake, W.L., & Clarke, S.D. (1991). Arachidonic acid suppression of fatty acid synthase gene expression in cultured rat hepatocytes. *Biochemical and Biophysical Research Communications*, *177*(3), 1056–1061.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., … Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research, 37*, W202-208.

Ballester, M., Revilla, M., Puig-Oliveras, A., Marchesi, J.A., Castelló, A., Corominas, J., … Folch, J.M. (2016). Analysis of the porcine APOA2 gene expression in liver, polymorphism identification and association with fatty acid composition traits. *Animal Genetics*, *47*(5), 552–559.

Bolormaa, S., Pryce, J.E., Hayes, B. J., & Goddard, M.E. (2010). Multivariate analysis of a genome-wide association study in dairy cattle. *Journal of Dairy Science*, *93*(8), 3818–3833.

Bosi, P., & Russo, V. (2004). The production of the heavy pig for high quality processed products. *Italian Journal of Animal Science, 3*(4), 309–321.

Braglia, S., Zappaterra, M., Zambonelli, P., Comella, M., Dall'Olio, S., & Davoli, R. (2014). Analysis of g.265T>C SNP of fatty acid synthase gene and expression study in skeletal muscle and backfat tissues of Italian Large White and Italian Duroc pigs. *Livestock Science, 162*, 15–22.

Buske, F.A., Bodén, M., Bauer, D.C., & Bailey, T.L. (2010). Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics (Oxford, England)*, *26*(7), 860–866.

Chakravarty, B., Gu, Z., Chirala, S.S., Wakil, S.J., & Quiocho, F.A. (2004). Human fatty acid synthase: Structure and substrate selectivity of the thioesterase domain. *Proceedings of the National Academy of Sciences, 101*(44), 15567–15572.

Chung, E.R. & Kim, W.T. (2005). Association of SNP marker in IGF1 and MYF5 candidate genes with growth traits in Korean cattle. *Asian-Australasian Journal of Animal Science, 18*(8), 1061-1065.

Davoli, R., Braglia, S., Zappaterra, M., Redeghieri, C., Comella, M., & Zambonelli, P. (2014). Association and expression analysis of porcine ACLY gene related to growth and carcass quality traits in Italian Large White and Italian Duroc breeds. *Livestock Science, 165*, 1–7.

Davoli, R., Catillo, G., Serra, A., Zappaterra, M., Zambonelli, P., Meo Zilio, D., ... Russo, V. (2018). Genetic parameters of backfat fatty acids and carcass traits in Large White pigs. *Animal,* Published online: 28 August 2018, DOI:10.1017/S1751731118002082

D'haeseleer, P. (2006). What are DNA sequence motifs? *Nature Biotechnology, 24*(4), 423–425.

Dugan, M.E.R., Vahmani, P., Turner, T.D., Mapiye, C., Juárez, M., Prieto, N., … Aalhus, J.L. (2015). Pork as a Source of Omega-3 (n-3) Fatty Acids. *Journal of Clinical Medicine, 4*(12), 1999–2011.

Estany, J., Ros-Freixedes, R., Tor, M., & Pena, R.N. (2014). A functional variant in the Stearoyl-CoA Desaturase gene promoter enhances fatty acid desaturation in pork. *PLOS ONE, 9*(1), e86177.

Gallardo, D., Quintanilla, R., Varona, L., Díaz, I., Ramírez, O., Pena, R.N., & Amills, M. (2009). Polymorphism of the pig acetyl-coenzyme A carboxylase alpha gene is associated with fatty acid composition in a Duroc commercial line. *Animal Genetics, 40*(4), 410–417.

Gammone, M.A., Riccioni, G., Parrinello, G., & D'Orazio, N. (2019). Omega-3 Polyunsaturated Fatty Acids: Benefits and Endpoints in Sport. *Nutrients, 11*(1), 46.

Green, C.D., Ozguden-Akkoc, C.G., Wang, Y., Jump, D.B., & Olson, L.K. (2010). Role of fatty acid elongases in determination of de novo synthesized monounsaturated fatty acid species.

Journal of Lipid Research, 51, 1871–1877.

Howe, P., Buckley, J., & Meyer, B. (2007). Long-chain omega-3 fatty acids in red meat. *Nutrition & Dietetics, 64*(s4), S135–S139.

Hu, N., Li, Y., Zhao, Y., Wang, Q., You, J., Zhang, X., & Ye, L. (2011). A novel positive feedback loop involving FASN/p-ERK1/2/5-LOX/LTB4/FASN sustains high growth of breast cancer cells. *Acta Pharmacologica Sinica, 32*(7), 921–929.

Inagaki, K., Aki, T., Fukuda, Y., Kawamoto, S., Shigeta, S., Ono, K., & Suzuki, O. (2002). Identification and expression of a rat fatty acid elongase involved in the biosynthesis of C18 fatty acids. *Bioscience, biotechnology, and biochemistry*, *66*(3), 613-621.

Jakobsson, A., Westerberg, R., & Jacobsson, A. (2006). Fatty acid elongases in mammals: Their regulation and roles in metabolism. *Progress in Lipid Research, 45*, 237–249.

Juárez, M., Dugan, M.E.R., Aldai, N., Aalhus, J.L., Patience, J.F., Zijlstra, R.T., & Beaulieu, A.D. (2011). Increasing omega-3 levels through dietary co-extruded flaxseed supplementation negatively affects pork palatability. *Food Chemistry*, *126*(4), 1716–1723.

Knez, M., Stangoulis, J.C.R., Glibetic, M., & Tako, E. (2017). The Linoleic Acid: Dihomo-γ-Linolenic Acid Ratio (LA:DGLA)—An Emerging Biomarker of Zn Status. *Nutrients*, *9*(8), 825.

Konieczka, P., Czauderna, M., & Smulikowska, S. (2017). The enrichment of chicken meat with omega-3 fatty acids by dietary fish oil or its mixture with rapeseed or flaxseed—Effect of feeding duration: Dietary fish oil, flaxseed, and rapeseed and n-3 enriched broiler meat. *Animal Feed Science and Technology*, *223*, 42–52.

Kris-Etherton, P.M., Harris, W.S., & Appel, L.J. (2003). Omega-3 Fatty Acids and Cardiovascular Disease: New Recommendations From the American Heart Association. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *23*(2), 151–152.

Latorre, P., Burgos, C., Hidalgo, J., Varona, L., Carrodeguas, J.A., & López-Buesa, P. (2016). c.A2456C-substitution in Pck1 changes the enzyme kinetic and functional properties

modifying fat distribution in pigs. *Scientific Reports*, *6*, 19617.

Levin, N. (2014). Multivariate statistics and the enactment of metabolic complexity. *Social Studies of Science*, *44*(4), 555–578.

Lo Fiego, D.P. (1996). Carcass fatness and lipid quality in the heavy pig. *Meat Focus International*, *5*(8), 261-263.

Lo Fiego, D.P., Santoro, P., Macchioni, P., & 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*(1), 107–114.

Macciotta, N.P.P., Biffani, S., Bernabucci, U., Lacetera, N., Vitali, A., Ajmone-Marsan, P., & Nardone, A. (2017). Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle. *Journal of Dairy Science*, *100*(6):4683-4697.

Maharani, D., Jung, Y., Jo, C., Jung, W.Y., Nam, K.C., Seo, K.S., … Lee, J.H. (2012). Evaluation of three candidate genes affecting fatty acid composition in pigs. *Korean Journal for Food Science of Animal Resources*, *32*(1), 6–12.

Mensink, R.P., Zock, P.L., Kester, A.D., & Katan, M.B. (2003). Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: a meta-analysis of 60 controlled trials. *The American Journal of Clinical Nutrition, 77*, 1146–1155.

Morris, M.C., & Tangney, C.C. (2014). Dietary fat composition and dementia risk. *Neurobiology of Aging*, *35*, S59–S64.

Moon, Y.S., Latasa, M.J., Griffin, M.J., & Sul, H.S. (2002). Suppression of fatty acid synthase promoter by polyunsaturated fatty acids. *Journal of Lipid Research*, *43*(5), 691–698.

Muñoz, M., Fernández, A.I., Benítez, R., Pena, R.N., Folch, J.M., Rodríguez, M.C., … Alves, E. (2013). Disentangling Two QTL on Porcine Chromosome 12 for Backfat Fatty Acid Composition. *Animal Biotechnology*, *24*(3), 168–186.

Nicolaou, A., Mauro, C., Urquhart, P., & Marelli-Berg, F. (2014). Polyunsaturated Fatty Acid-Derived Lipid Mediators and T Cell Function. *Frontiers in Immunology*, *5*, 75.

Pena, R.N., Ros-Freixedes, R., Tor, M., & Estany, J. (2016). Genetic Marker Discovery in Complex Traits: A Field Example on Fat Content and Composition in Pigs. *International Journal of Molecular Sciences, 17*(12), 2100.

Pertea, M., Mount, S.M., & Salzberg, S.L. (2007). A computational survey of candidate exonic splicing enhancer motifs in the model plant Arabidopsis thaliana. *BMC Bioinformatics*, *8*, 159.

Piasentier, E., Di Bernardo, N., Morgante, M., Sepulcri, A., & Vitale, M. (2009). Fatty acid composition of heavy pig back fat in relationship to some animal factors. *Italian Journal of Animal Science, 8*(2), 531-533.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Renaville, B., Bacciu, N., Lanzoni, M., Mossa, F., & Piasentier, E. (2018). Association of single nucleotide polymorphisms in fat metabolism candidate genes with fatty acid profiles of muscle and subcutaneous fat in heavy pigs. *Meat Science*, *139*, 220–227.

Revilla, M., Ramayo-Caldas, Y., Castelló, A., Corominas, J., Puig-Oliveras, A., Ibáñez-Escriche, N., … Folch, J.M. (2014). New insight into the SSC8 genetic determination of fatty acid composition in pigs. *Genetics Selection Evolution*, *46*(1), 28.

Ros-Freixedes, R., Gol, S., Pena, R.N., Tor, M., Ibáñez-Escriche, N., Dekkers, J.C.M., & 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*(3), e0152496.

Sellier, P., Maignel, L., & Bidanel, J.P. (2010). Genetic parameters for tissue and fatty acid composition of backfat, perirenal fat and longissimus muscle in Large White and Landrace pigs. *Animal*, *4*(4), 497–504.

Serra, A., Buccioni, A., Rodriguez-Estrada, M.T., Conte, G., Cappucci, A., & Mele, M. (2014). Fatty acid composition, oxidation status and volatile organic compounds in "Colonnata" lard from Large White or Cinta Senese pigs as affected by curing time. *Meat Science, 97*, 504–512.

Serra, AS., Conte, G., Giovannetti, M., Casarosa, L., Agnolucci, M., Ciucci, ... Mele, M. (2018). Olive Pomace in Diet Limits Lipid Peroxidation of Sausages from Cinta Senese Swine. *European Journal of Lipid Science and Technology, 120,* 1700236.

Smith, S. (1994). The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *8*(15), 1248–1259.

Suzuki, K., Ishida, M., Kadowaki, H., Shibata, T., Uchida, H., & Nishida, A. (2006). Genetic correlations among fatty acid compositions in different sites of fat tissues, meat production, and meat quality traits in Duroc pigs. *Journal of Animal Science*, *84*(8), 2026–2034.

Tacke, R., & Manley, J.L. (1999). Determinants of SR protein specificity. *Current Opinion in Cell Biology*, *11*(3), 358–362.

Thévenot, E.A., Roux, A., Xu, Y., Ezan, E., & Junot, C. (2015). Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *Journal of Proteome Research*, *14*(8), 3322–3335.

Trygg, J., & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, *16*(3), 119–128.

Uezato, A., Yamamoto, N., Iwayama, Y., Hiraoka, S., Hiraaki, E., Umino, A., … Nishikawa, T. (2015). Reduced cortical expression of a newly identified splicing variant of the DLG1 gene in patients with early-onset schizophrenia. *Translational Psychiatry*, *5*(10), e654.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130.

Wood, J.D., Richardson, R.I., Nute, G.R., Fisher, A.V., Campo, M.M., Kasapidou, E., … Enser, M. (2004). Effects of fatty acids on meat quality: a review. *Meat Science*, *66*(1), 21–32.

Yokoyama, C., Shinjo, F., Yoshimoto, T., Yamamoto, S., Oates, J.A., & Brash, A.R. (1986). Arachidonate 12-lipoxygenase purified from porcine leukocytes by immunoaffinity chromatography and its reactivity with hydroperoxyeicosatetraenoic acids. *Journal of Biological Chemistry*, *261*(35), 16714–16721.

Zambonelli, P., Comella, M., Braglia, S., Buttazzoni, L., & Davoli, R. (2009). Skeletal muscle expression analysis of fat metabolism genes in pig. *Italian Journal of Animal Science*, *8*(2), 171–173.

Zhang, S., Knight, T.J., Reecy, J.M., & Beitz, D.C. (2008). DNA polymorphisms in bovine fatty acid synthase are associated with beef fatty acid composition. *Animal Genetics*, *39*(1), 62–70.

Zheng, Z.M. (2004). Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *Journal of Biomedical Science*, *11*(3), 278–294.

Table 1. Summary statistics of the measured phenotypic traits with the available number of measures (N), means and standard deviations (SD). The common nomenclature for the studied fatty acids is reported between brackets.

Traits	N	Mean	SD
Carcass weight (kg)	517	120.17	8.60
Lean percentage (%)	517	48.56	2.76
Backfat thickness (mm)	527	28.13	5.10
Intramuscular Fat (%)	533	1.99	1.10
Backfat fatty acids (% on the total fatty acids)			
C12:0 (Lauric acid)	528	0.11	0.03
C14:0 (Myristic acid)	528	1.27	0.16
C14:1*cis*-9 (Myristoleic acid)	528	0.01	0.004
C16:0 (Palmitic acid)	528	22.7	1.26
C16:1*cis*-7 (Cis-7 Hexadecenoic acid)	528	0.34	0.06
C16:1*cis*-9 (Palmitoleic acid)	528	1.48	0.27
C18:0 (Stearic acid)	528	13.13	1.80
C18:1*cis*-9 (Oleic acid)	528	38.37	1.57
C18:1*cis*-11 (Asclepic acid)	528	2.19	0.24
C18:2(*n*-6) (Linoleic acid)	528	16.13	1.97
C18:3(*n*-3) (α-linolenic acid)	528	0.78	0.19
C20:0 (Arachidic acid)	528	0.18	0.04
C20:1*cis*-11 (Gadoleic acid)	528	0.79	0.21
C20:2(*n*-6) (Eicosadienoic acid)	528	0.77	0.09
C20:3(*n*-6) (Dihomo-γ-linolenic acid)	528	0.08	0.02
C22:1*cis*-12	528	0.12	0.02

C20:4(*n*-6) (Arachidonic acid)	528	0.22	0.05
C22:4(*n*-6) (Adrenic acid)	528	0.09	0.02
C22:5(*n*-3) (Docosapentaenoic acid- DPA)	528	0.05	0.01
C22:6(*n*-3) (Docosahexaenoic acid- DHA)	528	0.01	0.01
SFA (Saturated fatty acids)	528	37.77	2.57
MUFA (Monounsaturated fatty acids)	528	43.59	1.80
PUFA (Polyunsaturated fatty acids)	528	18.19	2.12
UFA (Unsaturated fatty acids)	528	61.78	2.51
(*n*-6)PUFA (Omega 6 Polyunsaturated fatty acids)	528	16.24	1.98
(*n*-3)PUFA (Omega 3 Polyunsaturated fatty acids)	528	0.84	0.20
SCFA (Short chain fatty acids)[1]	528	0.08	0.02
MCFA (Medium chain fatty acids)[2]	528	26.29	1.36
LCFA (Long chain fatty acids)[3]	528	73.17	1.34

[1] SCFA class comprises C6:0, C8:0 and C10:0 fatty acids. C6:0, C8:0 and C10:0 are not shown individually in table since they were not considered for the association study.

[2] MCFA class comprises C12:0, C14:0, C14:1*cis*-9, C15:0, C16:0, C16:1*cis*-7, C16:1*cis*-9, C17:0 and C17:1*cis*-9. C15:0, C17:0 and C17:1 are not shown individually in table since they were not considered for the association study.

[3] LCFA class comprises all the fatty acids with at least 18 carbons in their tails.

Table 2. The number of Italian Large White pigs showing the different genotypes, the allele frequencies, and the *P*-value for the Hardy Weinberg Equilibrium test.

SNP	N[1]	Allele	Frequencies					HWE	*P*-value
			Genotype[2]			Allele			
			11(N)	12(N)	22(N)	1	2		
ACACA c.5196T>C	50	1 = C 2 = T	1.00 (50)	0.00 (0)	0.00 (0)	1.00	0.00	-	-
ACACA g.38621011G>A	514	1 = A 2 = G	0.05 (23)	0.30 (155)	0.65 (336)	0.20	0.80	68.61	0.329
ACLY g.2956T>C	535	1 = C 2 = T	0.13 (69)	0.29 (156)	0.58 (310)	0.27	0.73	46.06	2.307e-09
ACLY g.3923T>C	535	1 = C 2 = T	0.35(185)	0.49 (263)	0.16 (87)	0.59	0.41	138.45	0.721
FASN c.265T>C	536	1 = C 2 = T	0.13 (68)	0.51 (275)	0.36 (193)	0.38	0.62	175.59	0.055
SCD g.2228T>C	535	1 = C 2 = T	0.03 (17)	0.36 (193)	0.61 (325)	0.21	0.79	122.21	0.071

[1] Number of pigs genotyped for each SNP.

[2] The frequencies and between brackets the number of pigs for each genotype.

Table 3. Associations between *ACACA* g.38621011G>A, *ACLY* g.3923T>C, *FASN* c.265T>C and *SCD* g.2228T>C SNPs, carcass traits and backfat fatty acid composition obtained with mixed models. For the genotypes are reported the estimated least squares means (LSM) ± the standard errors (SE) and the consistencies (N) between brackets.

Associations	N[1]	LSM ± SE (N)			*P*-value[2]	*P*-value A[3]	*P*-value D[4]	FDR adjusted *P*-value
		11	12	22				
ACACA g.38621011G>A								
C16:1*cis*-7 (%)	508	0.357[a] ± 0.011 (23)	0.348[a] ± 0.006 (154)	0.336[a] ± 0.004 (331)	0.049	0.061	0.874	0.465
ACLY g.3923T>C								
Lean percentage (%)	516	48.585[a,b] ± 0.268 (179)	48.689[a] ± 0.236 (257)	47.794[b] ± 0.339 (82)	0.022	0.025	0.030	0.427
Backfat thickness (mm)	516	28.004[a,b] ± 0.484 (179)	28.036[a] ± 0.426 (257)	29.481[b] ± 0.607 (82)	0.033	0.019	0.085	0.427
IMF (%)	512	1.838[a] ± 0.104	1.969[a,b] ± 0.091	2.197[b] ± 0.133	0.040	0.011	0.599	0.427

		(176)	(255)	(81)				
FASN c.265T>C								
IMF (%)	513	$1.694^{a} \pm 0.147$ (63)	$2.075^{b} \pm 0.09$ (262)	$1.894^{a,b} \pm 0.103$ (188)	0.016	0.196	0.0042	0.075
C14:0 (%)	527	$1.228^{A} \pm 0.018$ (68)	$1.270^{A,B} \pm 0.011$ (268)	$1.288^{B} \pm 0.013$ (191)	0.009	0.002	0.327	0.068
C16:1*cis*-9 (%)	527	$1.412^{a} \pm 0.037$ (68)	$1.446^{a} \pm 0.024$ (268)	$1.511^{b} \pm 0.028$ (191)	0.008	0.010	0.540	0.068
C18:0 (%)	527	$13.903^{A} \pm 0.240$ (68)	$13.429^{A,B} \pm 0.152$ (268)	$13.143^{B} \pm 0.173$ (191)	0.010	0.003	0.561	0.068
C18:1*cis*-11 (%)	527	$2.109^{a} \pm 0.032$ (68)	$2.157^{a,b} \pm 0.021$ (268)	$2.201^{b} \pm 0.024$ (191)	0.015	0.006	0.910	0.075
C20:0 (%)	527	$0.206^{a, A} \pm 0.005$ (68)	$0.192^{b, A, B} \pm 0.003$ (268)	$0.182^{c, B} \pm 0.004$ (191)	<.0001	<.0001	0.584	0.003
C20:3(*n*-6) (%)	527	$0.077^{A} \pm 0.002$ (68)	$0.082^{A,B} \pm 0.001$ (268)	$0.084^{B} \pm 0.001$ (191)	0.011	0.003	0.437	0.068

C20:4(*n*-6) (%)	527	$0.205^{a} \pm 0.006$ (68)	$0.214^{a,b} \pm 0.004$ (268)	$0.221^{b} \pm 0.004$ (191)	0.023	0.009	0.819	0.094
SCD g.2228T>C								
C12:0 (%)	526	$0.109^{a,b} \pm 0.008$ (17)	$0.115^{a} \pm 0.003$ (190)	$0.107^{b} \pm 0.002$ (319)	0.030	0.847	0.136	0.350
C16:1*cis*-7 (%)	526	$0.339^{a,b} \pm 0.013$ (17)	$0.344^{a} \pm 0.006$ (190)	$0.331^{b} \pm 0.005$ (319)	0.043	0.528	0.263	0.350
C20:4(*n*-6) (%)	526	$0.211^{a,b} \pm 0.011$ (17)	$0.221^{a} \pm 0.004$ (190)	$0.211^{b} \pm 0.004$ (319)	0.036	0.970	0.106	0.350

For *ACACA* g.38621011G>A: 1 = A allele, 2 = G allele; for *ACLY* g.3923T>C: 1 = C allele, 2 = T allele; for *FASN* c.265T>C: 1 = C allele, 2 = T allele; *SCD* g.2228T>C: 1 = C allele, 2 = T allele.

A,B: means within a row with a different capital letter are different ($P < 0.01$). a,b,c, d: means within a row with a different letter are different ($P < 0.05$).

[1] Total number of observations used in the mixed model.

[2] *P-value* for the polymorphism effect.

[3] *P-value* for the additive effect.

[4] *P-value* for the dominant effect.

Table 4. Associations between *ACACA* g.38621011G>A, *ACLY* g.2956T>C, *ACLY* g.3923T>C, *FASN* c.265T>C and *SCD* g.2228T>C SNPs, and Principal Component (PC) scores of the samples obtained from Principal Component Analysis (PCA) on single fatty acids. For the genotypes are reported the estimated least squares means (LSM) ± the standard errors (SE) and the *P*-values.

Associations	N[1]	LSM ± SE			*P*-value[2]	*P*-value A[3]	*P*-value D[4]
		11	12	22			
ACACA g.38621011G>A							
PC1 scores	502	0.568 ± 0.544	-0.157 ± 0.235	-0.005 ± 0.166	0.420	0.305	0.196
PC2 scores	502	-0.084 ± 0.321	-0.128 ± 0.149	-0.088 ± 0.108	0.965	0.991	0.830
ACLY g.2956T>C							
PC1 scores	521	-0.043 ± 0.362	-0.272 ± 0.260	-0.256 ± 0.206	0.819	0.561	0.651
PC2 scores	521	0.192 ± 0.218	-0.002 ± 0.158	-0.069 ± 0.127	0.482	0.227	0.680
ACLY g.3923T>C							
PC1 scores	522	-0.179 ± 0.243	-0.210 ± 0.214	-0.443 ± 0.320	0.724	0.443	0.663
PC2 scores	522	-0.083 ± 0.150	-0.040 ± 0.134	0.197 ± 0.192	0.352	0.163	0.465

FASN c.265T>C							
PC1 scores	522	$-0.965^{A} \pm 0.352$	$-0.316^{A,B} \pm 0.212$	$0.181^{B} \pm 0.243$	0.007	0.003	0.755
PC2 scores	522	-0.267 ± 0.210	-0.027 ± 0.132	0.107 ± 0.150	0.230	0.092	0.705
SCD g.2228T>C							
PC1 scores	521	-0.877 ± 0.672	-0.033 ± 0.234	-0.382 ± 0.215	0.220	0.463	0.118
PC2 scores	521	0.088 ± 0.390	0.027 ± 0.144	-0.076 ± 0.134	0.737	0.673	0.922

For *ACACA* g.38621011G>A: 1 = A allele, 2 = G allele; for *ACLY* g.3923T>C: 1 = C allele, 2 = T allele; for *FASN* c.265T>C: 1 = C allele, 2 = T allele; *SCD* g.2228T>C: 1 = C allele, 2 = T allele.

A,B: means within a row with a different capital letter are different ($P < 0.01$). a,b,c, d: means within a row with a different letter are different ($P < 0.05$).

[1] Total number of observations used in the mixed model.

[2] *P-value* for the polymorphism effect.

[3] *P-value* for the additive effect.

[4] *P-value* for the dominant effect.

Figure captions

Figure 1. Correlogram of the phenotypic correlations between measured traits. The correlation matrix was reordered according to the correlation coefficient aiming to make the clusters of correlated variables more evident. Positive correlations are displayed in blue and negative correlations in red colour. The colour intensity and size of circles are proportional to the size of the correlation coefficients.

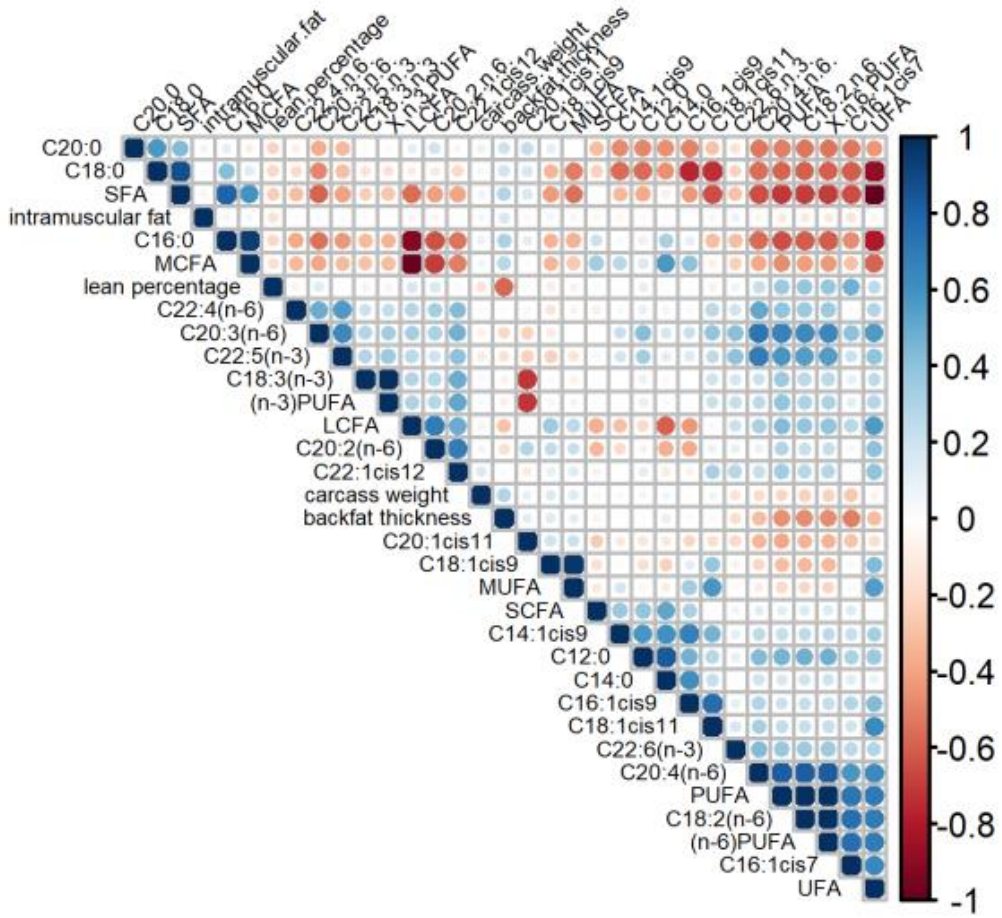

Figure 2. Loadings plot of the two Principal Components (PC1 and PC2) obtained from Principal Component Analysis (PCA)**:** the 3 variables with the most extreme weights (positive and negative) for each loading are labelled in bold characters.

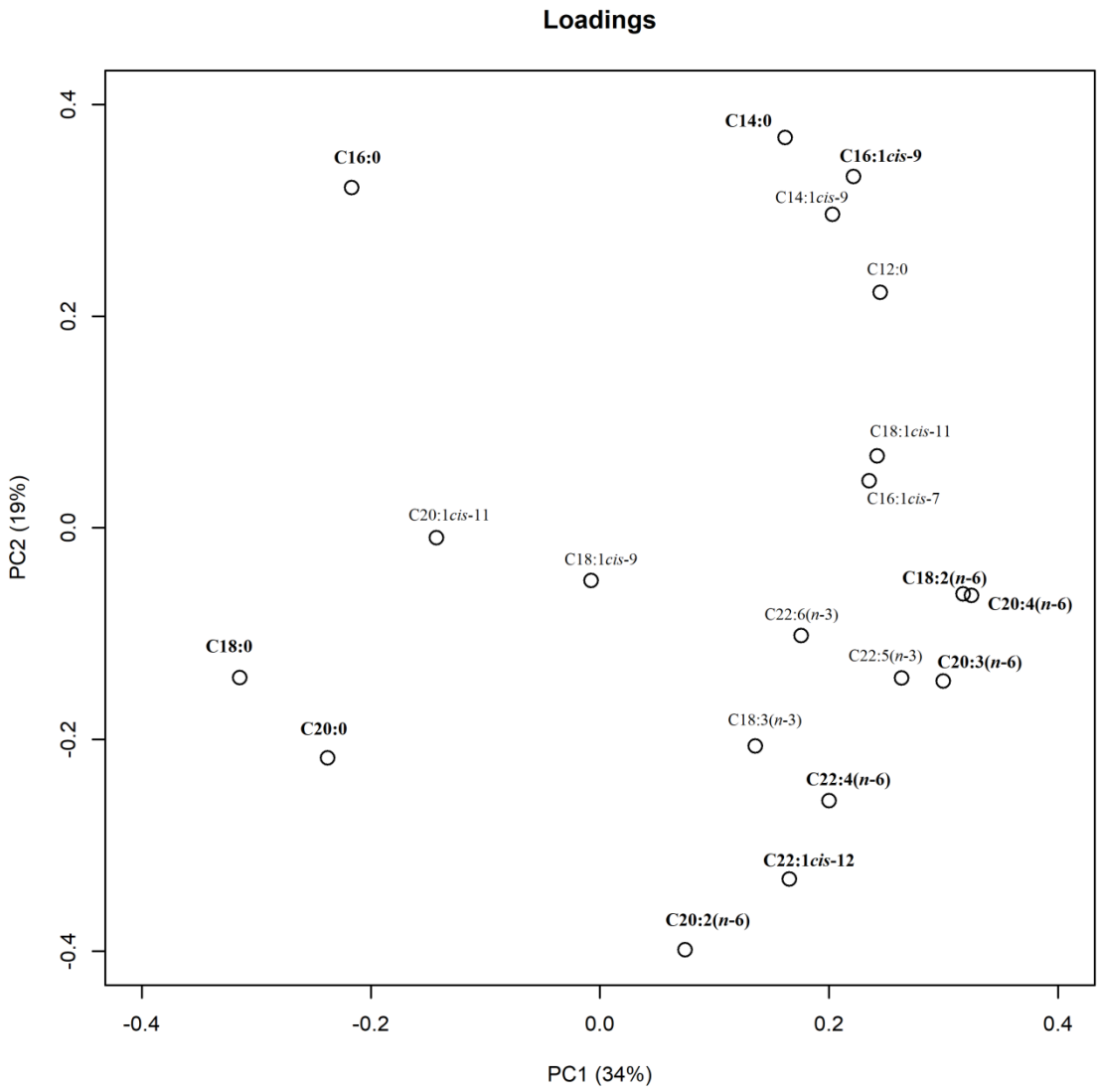

Figure 3. Arachidonic acid least squares means and standard errors estimated for the six different combinations of genotypes for the studied *FASN* and *SCD* loci. The number of samples showing each combination is reported between brackets, and the *P*-values for the significant differences are showed.