Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Curriculum: Scienze Zootecniche

Ciclo XXIX

Settore Concorsuale di afferenza: 07/G1

Settore Scientifico disciplinare: AGR17/ AGR19

GENOMICS AND NEW APPROACHES TO STUDY COMPLEX TRAITS IN PIGS AND OTHER LIVESTOCK SPECIES. A FOCUS ON THE INVESTIGATION OF GENE NETWORKS RELATED TO FAT QUALITY AND DEPOSITION IN PIGS AND PRELIMINARY RESEARCH TO STUDY FACTORS RELATED TO PERFORMANCES IN PIGLETS AND POULTRY.

Presentata da: **Martina Zappaterra**

Coordinatore Dottorato **Relatore**

Prof. Giovanni Dinelli **Prof.ssa Roberta Davoli**

Esame finale anno 2017

Summary

General introduction

General introduction

Livestock breeding programs: modern genetic selection and its challenges

Since thousands of years ago, man has always tried to model the environment to obtain better food supply and living conditions. The first livestock breeding started thousands of years ago with the domestication of most animal species used today. For most of the human history, the selection of the best breeders was exclusively based on phenotypic evaluations with little awareness of the underlying causes of different productivities and appearances (Jonas and de Koning, 2015). In the last century, the increasing information on the genetic background of productive traits in livestock species has changed substantially breeding selection. The birth of quantitative genetics and the combination of theoretical approaches and experimental achievements led to the inclusion of multifactorial models in breeding schemes (reviewed in Hill, 2014; Jonas and de Koning, 2015). In livestock species, selection based on the recording of pedigree and performance information has been effective for the improvement of many traits, in particular for animals with short generation intervals and for productive traits with high heritability. However, the selection process for some traits has been slowly accelerating: this delay is due to the complexity of some production traits (sometimes hardly recordable or expressed in a sex-dependent manner), the long generation intervals characterising some livestock species and the existence of negative correlations between traits. Furthermore, it is worth remembering that quite all the production traits are quantitative, that is to say that have a measurable phenotype that depends on the cumulative actions of many genes and the environment. The polygenic nature of almost all the production traits participate in slowing down and hardening the improvement of these quantitative traits through breeding schemes. Over the last ten years, the development of new technologies and the progress of genetics and genomics have opened new horizons in breeding programs and have offered new solutions to these problems. In particular, genetic information is highly useful and informative when selection is addressed towards the improvement of productive traits that are linked to each other through negative correlations, which show low heritability, or whose expression depends from sex and is hardly measurable.

The first breeding approach combining genetic information and traditional selection schemes has been marker assisted selection (MAS). MAS aimed at finding molecular markers highly associated to productive traits: the assumption on which this approach was based is that rather than selecting animals exclusively on the phenotypic expression of the considered trait, it could be more efficient to select animals bearing favourable genetic variants for the genes and regions associated to quantitative traits (regions defined quantitative trait *loci*, QTL). Anyway, there has been a limited number of traits for which MAS used direct markers coding for the functional mutation. Indeed, direct markers (causal mutations) are extremely difficult to detect as it is particularly complex to prove causality between single markers and quantitative traits (Andersson, 2001). For this reason, MAS has been mainly performed using molecular markers (such as single nucleotide polymorphisms-SNPs- and microsatellites) in linkage disequilibrium with the unknown causal mutation. The identification of molecular markers and genetic regions associated to quantitative traits could be possible through the application of two main strategies: the genome scan approach (Andersson, 2001) and the candidate gene approach (Rothschild and Soller, 1997). The first strategy aims at finding the associations between a trait and anonymous DNA markers across the genome, in order to identify chromosomal locations involved in the expression of the trait. The genome scan has represented a powerful tool for the identification of QTL with a major effect, but provides reliable results only on condition that the markers considered are widely distributed across the genome, the postulated genetic model is accurate and the size of the used sample is reasonable. Furthermore, this approach has a low sensitivity for the identification of *loci* with small effects, and the obtained regions, spanning on average from 5 to 30 cM, need to be further investigated through subsequent fine mapping. On the other side, the candidate gene approach is addressed directly to investigate associations between a trait of interest and variations within known genes, which have been selected for the functional role of the coded proteins and/or for the location of the gene sequence inside an already identified QTL. Anyway, despite MAS was welcomed in a greatly optimistic way, it has not reached the initial proposed achievements due to the fact that the identification of reliable markers is difficult, especially when working with complex polygenic traits (Jonas and de Koning, 2015).

More recently, the advancements of molecular genetics facilitated large scale genotyping, allowing for each individual the study of hundreds of thousands SNPs at the same time through the use of SNP arrays. This technological progress has led to the birth of a new approach named genomic selection (GS), combining large scale population genotyping and selection schemes (Meuwissen et al., 2001). GS is based on the use of SNP arrays (high-density SNP panels), which are genome-wide marker panels; the large scale population genotyping obtained by SNP arrays is then used to derive animals breeding values based on their genomic information. With this approach, the traditional estimated breeding values (EBVs) of a candidate breeder become Genomic EBVs (GEBVs). Traditional EBVs are estimated using pedigree information and measures obtained from evaluation schemes considering the breeder performances and the sib and/or the progeny productions. Generally all this information is processed using statistical models based on best linear unbiased predictions (BLUP; Henderson, 1975). In GS the GEBVs are the predictions of the genetic merit of an individual based on its genome. Anyway, to be accurate, GEBVs need to be validated on a large training population (or reference population). To date, GS can be implemented in practice for all main livestock species since genome-wide SNP panels, and sometimes also full sequence information, are available. Genomic selection has already been introduced into dairy cattle breeding programs leading to a consistent lowering in generation intervals, passing from 5-6 years in traditional selection schemes to 1.5 years in GS programs (Pryce and Daetwyler, 2012). The integration of genomic tools in dairy cattle had a substantial success, but the same can not be observed for other livestock species. The reasons of this difference between dairy cows and other livestock animals (such as beef cattle, pig and chicken) has to be found in the particular attributes characterising dairy cattle breeding programs. The generation interval is quite long and the offspring born from each insemination is generally constituted of one or two individuals, unlike in pigs and chickens where the generation interval is shorter and the offspring numerous. Additionally, beef cattle populations compared to dairy cattle are less uniform, as each country has different beef cattle breeds and asks for peculiar meat characteristics; this situation makes it difficult to obtain unique results from the genotypic data and perform cross-comparisons with large training populations in beef cattle (Garrick, 2011). The scenery of pig selection schemes is completely different from what can be observed in dairy cattle. First of all, the generation interval in pig species is noticeably shorter than that of dairy cattle, which makes less cost-

effective the application of GS to pig selection. Furthermore, most pig breeding schemes use crossbreeding between paternal production-oriented breeds and maternal reproduction-oriented breeds (Visscher et al., 2000). Because of this specific breeding program, it is particularly complex to set up GS in pig: any production traits are measured in crossbred pigs, while selection is based on the purebred lines they derived from, hindering the possibility of distinguishing environmental and genetic effects on the phenotypes. In the particular situation of pig breeding programs, the existence of different environmental effects between the purebred breeders (nucleus) and the crosbred pigs makes it difficult to obtain accurate genomic selection breeding values (GEBVs) using nucleus records (Nirea and Meuwissen, 2017). For all these reasons, GS has not yet been used in pig as a selection tool with the same reliability as in dairy cattle (Jonas and de Koning, 2015). Lillehammer et al. (2011) proposed GS as a potential tool to select efficiently maternal traits, which in traditional breeding programs are quite arduous to improve due to their low heritability. Dekkers (2007) suggested that a feasible way to overcome the problems of low genetic correlations between purebred and commercial crossbred performances could be the application of MAS to pig selection schemes. However, MAS to be effective need to rely on consistent knowledge of the genetic variations associated to the considered traits. In a recent paper, Lopes et al. (2017) propose for pig selection schemes to reconsider MAS in a new approach that combines also high-density SNP panels. In particular, the use of SNP arrays and association studies performed on the results of large scale population genotyping (genome wide association studies- GWAS) has increased the power and precision of the detection of markers linked to QTL. This strategy may represent a viable alternative to GS for the livestock breeding programs where the latter is still not economically sustainable. The suggested approach consists in performing a GWAS for the trait of interest on different populations genotyped through high-density SNP chips; then, the most associated markers derived from the GWAS are used for the MAS on the entire population and are accounted as fixed effects in BLUP prediction models (such as marker-assisted BLUP and marker-assisted genomic BLUP). This approach was tested in four distinct pig populations for the maternal trait "number of teats”: the most significant SNP resulting from all the four population was used for MAS, and the obtained prediction accuracy was compared to the accuracies obtained for the same trait using other prediction strategies (for example the traditional BLUP, the genomic BLUP and the Bayesian variable selection model). The application of MAS

using as markers the SNPs accounting for the strongest associations led to better prediction accuracies respect to BLUP and genomic BLUP (Lopes et al., 2017). This approach seems to be particularly effective respect to GS when it is not possible to have large training populations, such as for maternal traits in pig breeding programs. Nirea and Meuwissen (2017) proposed that a possible solution to the thorny issue of small reference populations could be represented by the use of a mixed training population constituted of both nucleus and crossbred animals. This strategy could be an answer for the application of GS to traits particularly affected by environmental effects, such as feed efficiency ratio (Nirea and Meuwissen, 2017).

During the last four years, one world-wide running pig breeding company has announced to use GS in crossbred, while another pig breeding corporation reported to integrate pedigree information with genomic information obtained from large scale population genotyping; however, no details about the prediction methods and accuracies have been disseminated.

Selection in chickens is characterised by features that make it quite similar to the pig selection scenario: such as in pigs, chickens are selected through pyramidal breeding programs, cross-breeding schemes are often used, the generation interval is short and the offspring is numerous (Jonas and de Koning, 2015). The studies performed until now are not able to provide a clear answer about whether or not genomic selection could bring advantages in chicken selection breeding programs. The theoretical studies about GS application in chicken breeding programs returned contradictory results, highlighting also the possibility of unwanted side-effects (Sitzenstock et al., 2013).

A particular case: pig breeding and selection in Italy

In this international scenario, a particular case is represented by Italian pig selection. Italian pig selection goals differ significantly from those demanded from the international market: in addition to the feed efficiency, Italian market demands meats suitable for the production of high quality meat products, such as dry-cured hams with protected origin designation (Parma and San Daniele hams). To obtain meat with these characteristics pig are slaughtered

at higher live weights (at least 160 kg for Italian heavy pigs or 130 kg for pigs slaughtered at an intermediate live weight) and older ages (eight or nine months of age). The breeding program is carried out by the Italian pig breeder association (Associazione Nazionale Allevatori Suini, ANAS) and is based on the selection of pigs belonging to three main breeds: Italian Large White, Italian Landrace and Italian Duroc. Italian selection schemes pursue the same efficiency objectives of other countries (for example reproductive efficiency, growth rate, feed efficiency and carcass characteristics), but additionally have a distinctive attention towards meat quality traits (Bosi and Russo, 2004). For such high quality production, meat must have an excellent aptitude for salting and seasoning and a right balance of fat and lean mass deposition is required in pig carcasses and thighs. Thighs must have a sufficient amount of covering fat, which prevents hams from loosing an excessive amount of water during the seasoning period. On the other side, hams should not have a too high backfat thickness and an excessive content of visible intramuscular fat, which are not well perceived by the consumers. Therefore, one of the main objectives for Italian heavy pig selection is represented by the maintenance of a constant backfat thickness while keeping a satisfactory amount of deposited lean mass. For the production of high quality seasoned products it is also mandatory that covering fat does not contain a percentage of linoleic acid beyond the established threshold of the 15% (Bosi and Russo, 2004). Higher contents of linoleic and polyunsaturated fatty acids cause the occurrence of rancidity and unwanted flavours in seasoned products, making them unsuitable for human consumption. Unlike in other countries, the Italian heavy pig selection aims at preserving a balance between fat and leanness without chasing the increase of muscle deposition: indeed, high contents of lean mass negatively affect meat colour, flavour and dry-cured ham firmness, causing also higher seasoning loss. The selection index utilized for the Italian breeding programs includes the estimated breeding values (EBVs) for: Average Daily Gain (ADG) calculated from 30 to 155 kg of live weight with a *quasi ad libitum* feeding level (expressed in grams); Backfat Thickness (BFT) recorded *post mortem* at the level of *gluteus medius* muscle (expressed in mm); Lean Cuts (LC) obtained from the sum of neck and loin weights (expressed in kg); Feed Conversion Ratio (FCR) obtained from feed intake recorded daily and body weight measured bimonthly (expressed in units); and Hams Weight (HW), which is the measure of thighs (expressed in kg). For Italian Duroc pigs a further EBV is considered in the breeding programs: Visible Intermuscular Fat (VIF), a selection criterion set up and utilised exclusively in this pig breed

in order to avoid an excessive presence of inter- and intramuscular fat in Italian Duroc thighs. VIF indicates the genetic value of Italian Duroc boars for the probability of transmitting excessive intramuscular fat depots to the offspring (Bosi and Russo, 2004). The EBVs are calculated using a BLUP Animal Model (Henderson and Quaas, 1976) with different fixed effects depending on the considered trait (Russo et al., 2000).

Italian pig selection shows peculiar features, characterising Italian market from the ones of other countries. The differentiated breeding goals would make it particularly tricky to apply GS in this scenario: Italian heavy pig has been selected since many years for the improvement of different objectives respect to those of other European and international markets, creating a population with different genetics, different abilities towards fat and lean mass storage, and a longer production cycles than the other European pig populations (nine months of age for Italian heavy pig *vs*. on average six months of age for "lean" pigs used for the production of fresh pork products). Samorè et al. (2015) tested the impact of GS in a simulated Italian pig population. The whole experiment was designed to evaluate the possible application of GS in Italian pig selection schemes: to this aim, the authors considered the candidate boars as prediction population, for which was assumed the genotype, and as training population their full-sibs, for which were assumed both genotypic and phenotypic data. The application of GS showed positive results only when applied to traits with low heritability (such as reproduction or health traits), while none of the considered genomic models provided improved predictions for traits with average heritability levels and for pigs without recorded phenotypes.

Genetics, genomics and new approaches to study complex traits in livestock species

In the last 20 years, livestock science has seen an astonishing development in genomic technologies, passing from the QTL mapping in the early 1990s, to the release of the whole genome sequences of major livestock species (Bovine Genome Sequencing and Analysis Consortium et al., 2009; International Sheep Genomics et al., 2010; Groenen et al., 2012), until the availability of high-density SNP panels ranging from 10,000 to 50,000, 100,000 and up to five million SNPs today. Genomics have been successfully implemented in dairy cattle

modern selection schemes, but less consistently in breeding programs of other livestock species. However, these technological advancements offer the possibility for investigating more in deep also in these animal species the genetic and physiological causes affecting economically important traits in breeding populations. Technical progress in the field of large scale genotyping, next-generation sequencing, mass spectrometry and bioinformatics facilitates the study of highly complex traits, offering new approaches for the identification of complex biological patterns. The integration of OMICs technologies (genomics, transcriptomics, proteomics, lipidomics and metabolomics) in livestock science was welcomed in an optimistic way, with the hope that these new tools could shed light of the determinism of quantitative traits. High-density SNP panels are today widely used to perform Genome Wide Association Studies (GWAS) with productive traits in all the livestock species (reviewed in Sharma et al., 2015): this studies permit the identification of single markers, or contiguous regions of markers, associated with the trait of interest. In a second step, the regions highlighted by means of GWAS need to be investigated more in deep through the detection of candidate genes harbouring the causal mutations involved in the phenotypic variability. Similarly, the results obtained from transcriptomics, proteomics, lipidomics and metabolomics can help identifying transcripts and metabolites involved in complex traits, but require also validation studies targeting the identified candidate genes and physiological processes. Therefore, OMICs technologies and targeted studies on particular candidate genes or metabolic processes represent two complementary tools, which are able to offer a more complete knowledge of the genetic and physiological players behind livestock animals' phenotype. Furthermore, the technological progress led to the development of new phenotypes: OMICs sciences provide information that outlines the biological responses of an organism to genetic mutations, diseases and environmental effects. Therefore, transcriptome, proteome or metabolome can also be considered as phenotypes. New tecnologies are also used to supply less expensive methods for the analysis of already known phenotypes: there is, for example, an increasing use of advanced imaging technologies for measuring carcass and meat characteristics (Santos-Garcés et al., 2014; Matika et al., 2016). Taken together, the technological advances and the use of a multidisciplinary approach can help identifying the main factors affecting complex traits, such as carcass and meat qualities, fat deposition and composition (which are central in

particular for pork high quality seasoned products), health and maternal traits (which are known to have low heritability).

In the present research, different approaches are used to study complex traits in livestock species: the studies were mainly oriented towards the investigation of gene networks related to fat traits in pigs; additionally two preliminary researches were also assessed to detect factors related to piglets' survival performances and the occurrence of dystrophic-like defects in chicken breast muscle.

Fat deposition and composition for the improvement of pork quality

The different fat depots in carcass

Due to the growing obesity epidemic over the last twenty years, there has been increasing interest in human medicine towards the investigation of the different functions of fat depots. In particular, the adipocytes showed to be cells with much more complex roles than being the simple lipid storage keepers that were considered until the discovery of fat-secreted hormone leptin (Zhang et al., 1994; Friedman, 2010). From the beginning in 1994, the knowledge on adipocyte roles has increased and evidences about the involvement of these cells in the regulation of major metabolic, physiological, and genetic changes have been added. Even if they are generally referred as “fat”, the different adipose tissues exhibit divergent characteristics depending on their anatomical location. Animals show five main fat depots with different anatomical location and physiology: subcutaneous, intermuscular, intramuscular, visceral fat and bone adipose depots (Hausman et al., 2014). While in human medicine fat depots are mainly studied for their linking with mechanisms leading to obesity, insulin sensitivity and metabolic disregulation (Hausman et al., 2014), in animal science the quantity and composition of lipids stored in carcass and meat represent essential factors influencing the technological and nutritional quality of animal products. In particular, subcutaneous and intramuscular fat content and composition contribute importantly to the nutritional and technological values of fresh and seasoned pork products, and therefore are traits displaying a consistent economical interest.

Intramuscular fat

Intramuscular fat: intramuscular adipocytes and intramyocellular lipid droplets

In livestock science, the term “intramuscular fat” (IMF) is generally used to define both the intramuscular fat depots and the intramyocellular lipid droplets. Intramuscular fat depots are composed of adipocytes infiltrated between and among skeletal muscle fibers (Hausman

et al., 2014)**;** this interspersed adipose tissue gives the muscle a marbled appearance, causing IMF to be called “marbling” in meat animals (Wood et al., 1999; Hausman et al., 2014). Adipocytes develop from pluripotent stem cells in both animals and humans, the potential to generate new fat cells continues throughout the lifespan. Adipocyte differentiation is a highly orchestrated process mediated by a large number of hormones, small non-coding RNA molecules (miRNAs), cytoskeletal proteins and transcription factors. The best understood regulators of adipogenesis are transcription factors (Mota de Sá et al., 2017) such as CCAAT/enhancer binding protein (C/EBP), adipocyte determination and differentiation-dependent factor 1 (ADD1 or SREBP), C/EBPα undifferentiated protein (CUP or AP-2α), peroxisome proliferator-activated receptors (PPARs) and monocyte chemoattractant protein-1 (MCP-1) (Hausman et al., 2009). Intramuscular white adipocytes are cells with the specific role of storing energy: adipocytes contain a unique large lipid vacuole, which forces nucleus and mitochondria to be located into a thin rim at the periphery of the cell (Figure 1). In addition to this role, muscle-interspersed adipocytes may also have an active function in insulin sensitivity: Goodpaster et al. (2000) found in human a positive correlation between insulin resistance and the amount of adipose tissue interspersed in skeletal muscle, suggesting that IMF could be involved in a paracrine mechanism (Komolka et al., 2014) and play an essential role communicating with muscle through adipokines and myokines (Vettor et al., 2009).

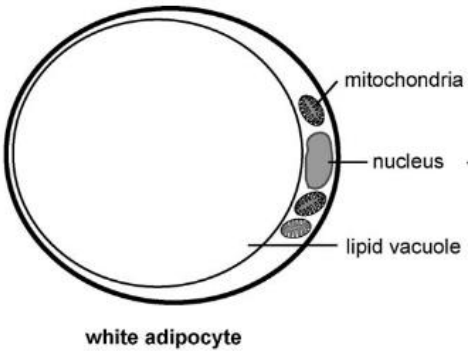


Figure 1: Illustration of a white adipocyte morphology (adapted from Sell et al., 2004).

On the other side, lipid droplets are intracellular organelles present in most cells (including myocytes) with the fundamental role of storing lipids useful as energy reserve and as matrix for the synthesis of cell membrane constituents (Thiam et al., 2013). Intracellular lipid

droplets core is mainly composed of triacylglycerols and cholesteryl esters, surrounded by a phospholipid monolayer with the hydrophobic acyl-chains dissolved in the triacylglycerols core and the hydrophilic head groups interfacing with the aqueous cytosol (Wolins et al., 2006; Thiam et al., 2013). This characteristic composition permits the lipid droplets to remain cohere inside the aqueous cytosol (Thiam et al., 2013). In myofibers, these intracellular organelles are mainly found adjacent to mitochondria and are known to be more present in Oxidative (slow-twitch type I) muscle fibers: the latter are reported to contain larger depots of intracellular triacylglycerols than fast-twitch (type IIA and IIB) fibers (Malenfant et al., 2001). Furthermore, intracellular lipid droplets surface is embedded with specific proteins controlling lipid metabolism and lipid droplets packaging inside cytosol (Figure 2). Among these proteins are Perilipins and lipases (such as Hormone sensitive lipase-HSL, alias LIPE, or Adipose triglyceride lipase- ATGL), which drive lipid storage and lipolysis (Wolins et al., 2006). In Figure 2 is reported the hypothesised general structure of intracellular lipid droplets. Anyway, apparently, there is spreading evidence that intracellular lipid droplets may be more complex and active organelles than thought so far: some lipid droplets were found to bear in their hydrophobic core ribosome-bound membranes (Wan et al., 2007), RNA, RNA binding proteins and ribosomal units (Cermelli et al., 2006; Wan et al., 2007) and eicosanoid-producing enzymes and perilipins (Melo et al., 2011; Robenek et al., 2009; Singh et al., 2009).

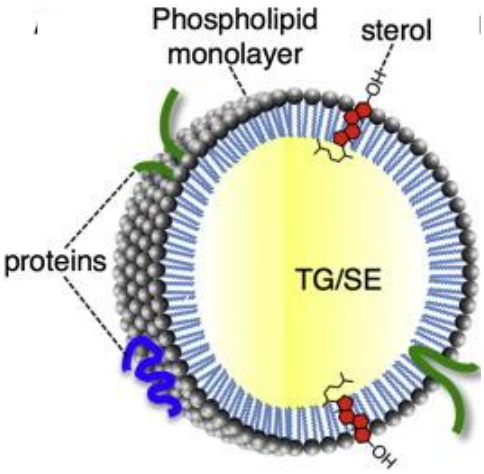


Figure 2: Illustration of the basic structure of lipid droplets (adapted from Ohsaki et al., 2014).

In animal science, despite intramuscular adipocytes and intramyocellular lipid droplets diverge for their anatomical location and physiology, it is not possible to analyse separately

their amount and composition: all the chemical analyses performed on meat return a picture that is the result of both intramuscular adipocytes and intramyocellular lipid droplets (Komolka et al., 2014). In Figure 3 are reported some images showing the differences among intramuscular adipocytes (Figure 3 B) and an example of how are located intracellular lipid droplets inside a myocyte (Figure 3 C).

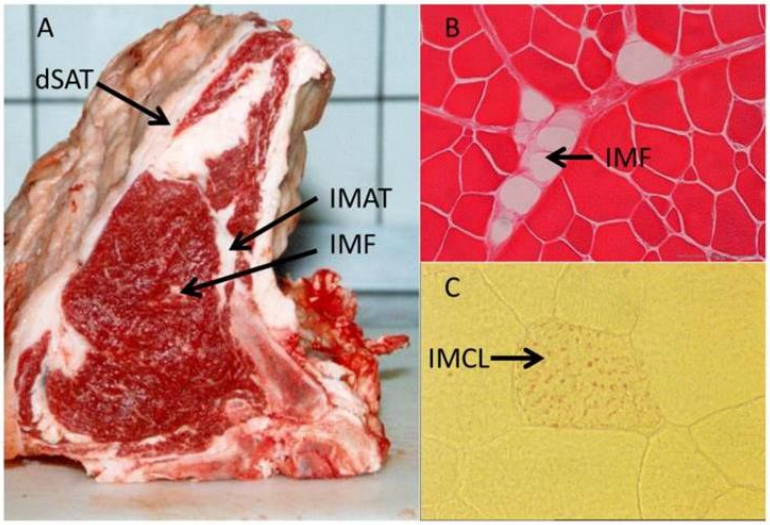


Figure 3: Illustration with the different fat depots in a beef cut. (A) Deep subcutaneous adipose tissue (dSAT) covering *serratus dorsalis* muscle, intermuscular adipose tissue (IMAT) between *intercostalis interni* and *longissimus dorsi* muscle, and intramuscular fat (IMF) within *longissimus dorsi* in cattle. (B) Cellular structure of IMF in M. longissimus dorsi (cattle, Eosin stained). (C) Intramyocellular lipids (IMCL, red dots, Oil-red O stained) in a muscle cell (M. longissimus dorsi, mouse) (adapted from Komolka et al., 2014).

However, changes in IMF content are mainly due to variation in the accumulation of adipocytes between muscle fibers. Both adipocytes hypertrophy and hyperplasy may be involved in IMF variations in pigs (Gerbens, 2004).

Intramuscular fat importance in pig meat quality

In recent years sales of meat in many countries have fallen slightly, where more and more consumers are moved by the slogan “less and healthier”. Today, consumers demand for higher quality meats and are willing to pay more for healthier products obtained from

animals reared in high standard welfare conditions. Anyway, this request for high quality meats sometimes seems to be at odds with consumers' request for leaner products. Even though fat is generally considered unhealthy, it influences meat characteristics: fat and fatty acids are extremely important as they contribute substantially to various aspects of meat quality and are central to the nutritional value of meat (Wood et al., 2008). Meat quality is a complex trait determined by different aspects, such as the microbiological food safety, the nutritional composition and the organoleptic attributes of the product. Meat eating quality is mainly related to tenderness, juiciness and flavour. Interestingly, IMF content seems to be positively correlated to all of these three characteristics in meat (Barton-Gade, 1990; Brewer et al., 2001). Several studies in different livestock species have shown that high IMF depots are correlated with lower resistance to shearing and higher tenderness, maybe because of the dilution of muscle fibrous protein by soft fat (Wood et al., 1999) or because the disposition of adipocytes around perimysial connective tissue forces the muscle structure (Wood, 1990). Fernandez et al (1999) proved that IMF levels higher than 2.5% bear favourable effects on sensory attributes of pork products obtained from different European pig populations, and Schworer et al. (1995) reported that the selection of high IMF in pigs is necessary to improve pork quality. On the whole, the recommended optimum range of IMF stands around 2.5-3%, and percentages below this threshold were associated to diminished eating quality. Furthermore, IMF content also participates in determining meat flavour through the compounds produced by the lipids degradation (such as aldehydes, alcohols and ketones), which participate in Maillard-like reactions (Mottram and Salter, 1989). Due to IMF proved importance for meat quality, there is an increasing interest towards the development of fast and cheap methods for IMF content determination in different muscles. In particular, the technological progress has come to selection aid, bringing more accurate technologies: Newcom et al. (2002) conducted one of the first studies aimed at using real-time ultrasound to predict IMF percentage in a Duroc pig population finding values of correlation around 0.6 between the analysed and the predicted IMF content. Recently, Jung et al. (2015) performed the same study on four pig breeds, identifying a high phenotypic correlation (0.76) between measured and predicted IMF content.

Candidate genes involved in porcine intramuscular fat deposition

The estimated heritability of IMF content in pigs ranges from 0.26 to 0.86, depending on the characteristics of the studied population, with an average of 0.5 (Sellier, 1998; Schwab, 2007). This indicates that this trait could be manipulated through selection, although the conventional selection for IMF resulted poorly effective. The improvement of this trait through traditional selection approaches results particularly difficult, due to the unfavorable genetic correlations linking IMF with leanness deposition and feed efficiency traits (Hermesch et al., 2000) and the limitations in the measurement of this trait (which can be measured exclusively after slaughter). Together with the development of cheaper technologies for the determination of IMF content, there is increased interest towards the deciphering of the molecular mechanisms underlying IMF deposition. The identification of genes and markers that contribute to genetic variation in IMF can provide an alternative and effective strategy to improve IMF in pigs avoiding side-effects on other carcass traits. To date IMF is associated to 244 QTLs all over pig genome (http://www.animalgenome.org/cgi-bin/QTLdb/ last accessed 23rd of March 2017). There are several genes lying in these QTLs that were found to be associated with IMF deposition. Some of them are directly related to fat deposition and metabolism, and due to their involvement in obesity and other related diseases were also investigated in human medicine. Among them is the *Alpha-ketoglutarate Dependent Dioxygenase* (also known as *Fat mass and obesity associated gene- FTO*), which has been associated to IMF deposition in different pig populations (Fan et al., 2009; Fontanesi et al., 2009). Fontanesi et al. (2009) investigated a mutation of *FTO* gene in two groups of Italian Duroc pigs divergent for VIF EBV and found statistically different allele frequencies between the two groups of extreme and divergent pigs. This association was then confirmed on a wider population of Italian Duroc (Fontanesi et al., 2010), suggesting that SNPs on this gene may have consistent effects on IMF deposition. Other genes found associated with IMF are the genes codifying for Fatty Acid Binding Proteins (FABP). Among them, *FABP3* gene expression showed statistically different transcription levels between pigs diverging for IMF deposition (Li et al., 2010; Tyra et al., 2013), and nucleotide variations in its sequence have been related to this trait (Gerbens et al., 2001; Li et al., 2010; Tyra and Ropka-Molik, 2011). With the dawn of OMICs science, the new technological and

computational advances have been also applied to the investigation of IMF molecular bases. Davoli et al. (2016) performed a GWAS on Italian Large White individuals identifying three candidate genes possibly affecting the IMF content of *Semimembranosus* muscle: *Sidekick cell Adhesion Molecule 1* (*SDK1*) gene located to date (Sus scrofa Build 10.2 assembly) on Sus scrofa chromosome (SSC) 3, *Serine Carboxypeptidase 1* (*SCPEP1*) mapped on SSC 12 and *Protein Phosphatase 3 Catalytic Subunit alpha* (*PPP3CA*) lying on SSC 8. In particular, the authors suggested that, among these three genes, the most promising one, which deserves to be more deeply investigated, is *PPP3CA*: this gene encodes for a calcium- and calmodulin-dependent protein phosphatase, also called calcineurin, which is a key enzyme in muscle fibers differentiation (da Costa et al., 2007). On the other hand, another GWAS performed on a Spanish population of Duroc pigs hypothesized the implication of the gene *Leptin Receptor* (LEPR) in IMF deposition (Ros-Freixedes et al., 2016). *LEPR* gene is mapped on SSC6 in the region comprised between 135 and 136 Mb: this region accounted in Ros-Freixedes et al. study for 3.1% of the genetic variance associated to IMF in *Gluteus medius* muscle and corresponded to a chromosomal region where are reported several QTLs associated with feed intake, carcass fatness, backfat thickness and IMF content. *LEPR* appears to be an interesting candidate gene, especially because the coded receptor interacts with Leptin, an adipocytokine that regulates energy intake and expenditure. Anyway, the signal detected by the authors seems to indicate that the whole region where *LEPR* gene lies could have a relevant role in IMF deposition: *LEPR* overlaps another gene, the *Leptin Receptor Overlapping Transcript* (*LEPROT*) and is mapped near *Janus Kinase 1* (*JAK1*) gene. These two other genes may be both involved in the phenotypic variations affecting IMF: on one hand, *LEPROT* encodes a protein that negatively regulates the presence of leptin receptors in the cell surface, while *JAK1* is involved in the adipocytokine signaling pathway, promoting the leptin-induced transactivation of the satiety neuropeptide NPY gene (Muraoka et al., 2003).

On the whole, the results obtained to date are often not concordant among the published studies. This divergence among the investigations reported in literature may be due to the polygenic nature of this trait, whose expression is probably affected by complex gene networks. To complicate matters further, the comparison of literature is also made harder by the fact that for the deciphering of the genetic basis of this trait were used breeds or crossbred pigs with different genetic backgrounds and divergent predispositions for the

deposition of IMF (i.e. Chinese pig breeds are more prone towards IMF deposition than European breeds). All these factors, together with the different environmental effects affecting pig populations studied in literature, contribute to the wide range of results in the current literature. Furthermore, livestock science literature lacks investigations on other possible candidate genes affecting IMF trait in pigs. In particular, among these genes are *Perilipins* (*PLINs*). *PLIN* genes code for proteins coating intracellular lipid droplets surface that are known to control stored lipid hydrolysis through the interaction with lipases. *PLIN* genes family in mammals is composed of five actors, from *PLIN1* to *PLIN5*, sharing high sequence homology and a highly conserved N-terminal sequence (Figure 4) (Kimmel and Sztalryd, 2016).

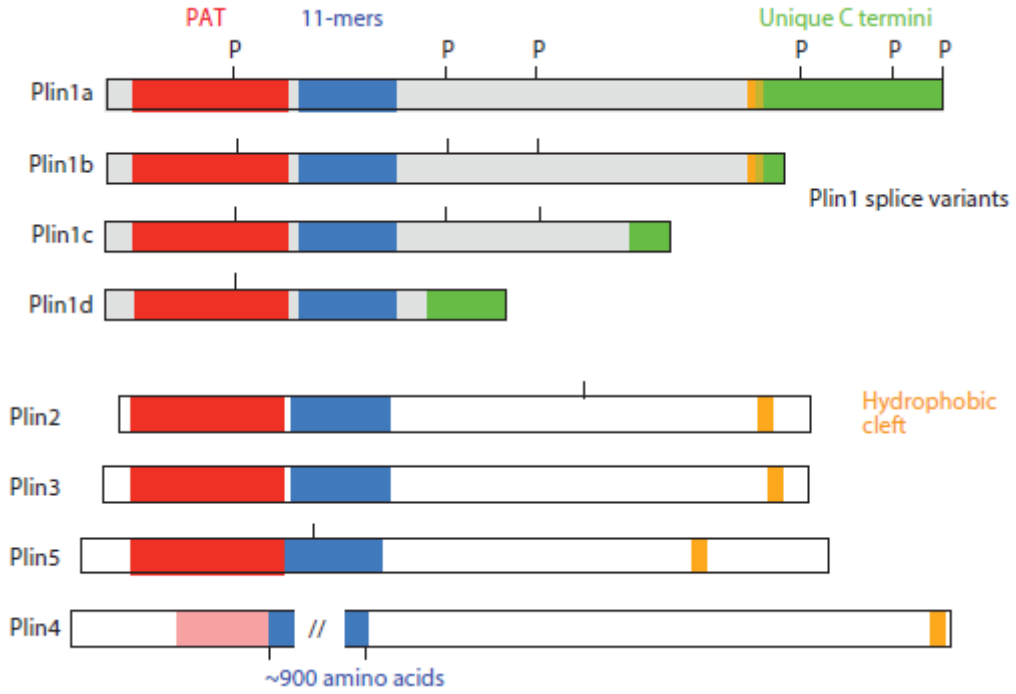


Figure 4: The conserved architecture of *Plin* genes family in mouse. The PAT domain (in red) is maintained among all the *Plin* genes and proteins, followed by 11-mer helical motif of varying lengths (blue). These two motifs are part of the N-terminal, an aminoacidic sequence highly conserved among Plins. The C-terminal (green) is the most divergent aminoacidic sequence among Plins proteins. The murine phosphorylation sites are indicated with a "P". In mouse, four splicing variants for Plin1 have been described (adapted from Kimmel and Sztalryd, 2016).

Each PLIN gene has a peculiar expression tissue: *PLIN1* is mainly found in adipose tissue, *PLIN2* and *PLIN3* seem to be widely distributed, *PLIN4* is characteristic of adipocytes, brain,

heart, and skeletal muscle, while *PLIN5* is predominantly expressed in oxidative tissues, such as heart, oxidative muscle fibers and brown adipose tissue (Wolins et al., 2006; Yamaguchi et al., 2006; Dalen et al., 2007). The expression of all *PLIN* genes is controlled by Peroxisome Proliferator Activated Receptor (PPAR) family (Mandard et al., 2004; Poulsen et al., 2012), except for *PLIN3*, which is the only Perilipin lacking functional PPAR response elements in its promoter sequence. The regulation of *PLINs* transcription probably is also controlled by other transcription factors, apart from the PPARs: *PLIN1* expression in human white adipose tissue is suppressed by the Liver X Receptor α (LXR-α) (Stenson et al., 2011) and Sterol Regulatory Element-binding protein 2 (SREBP2) was found inhibiting *Plin5* in mouse liver (Langhi et al., 2014). Taken together, their expression patterns differentiated among tissues and their different regulation suggest that PLINs have different functional roles. PLIN1 is the most investigated component of this gene and protein family: compared to wild type individuals, *Plin1* knockout mice fed diets enriched in fat content have higher lipolysis in adipocytes, leading to increases in circulating lipids and fatty acids with a consequent insurgence of steatosis in other organs, inflammatory responses, and insulin resistance (reviewed in Kimmel and Sztalryd, 2016). Similar responses were also observed in *Plin2* knockout mice and in *Plin5* knockout mice, but in the latter no systematic effect on insulin sensitivity was observed (reviewed in Kimmel and Sztalryd, 2016). Thus, PLINs promote and regulate lipids storage and hydrolysis, but differences have been observed depending on the tissue considered, the training effort (whether the individuals were at rest or subjected to muscular training) and the fat content level in the diet. For this reason, it is not possible to highlight a unique mode of action for this family, whose members seem to act differently depending on where they are expressed and what is the physiological state and the diet of the animal. On the whole, PLINs seem to control lipases activity on the surface of intracellular lipid droplets: Plin5 is reported to be a scaffolding protein for key lipolytic players, such as Adipose Triglyceride Lipase (ATGL) (Wang et al., 2011; Granneman et al., 2011), Hormone Sensitive Lipase (HSL, also known as LIPE) (Wang et al., 2009), and Comparative Gene Identification-58 (CGI-58) (Granneman et al., 2009; Wang et al., 2011). ATGL, HSL and CGI-58 are the main actors involved in the hydrolysis of the lipid droplets (Figure 5). The increasing number of studies identifying regulative pathways linking PLINs with these proteins suggests for PLINs a pivotal role in the control of intracellular energy fluxes.

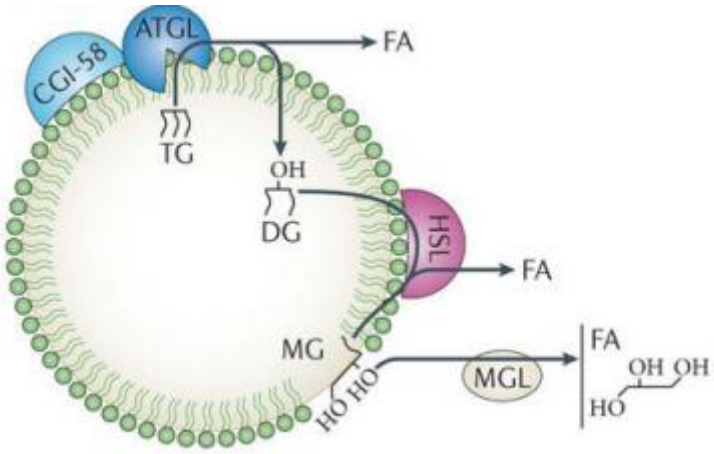


Figure 5: ATGL catalyses the first step of lipolysis and is recruited onto lipid droplets surface by the co-factor CGI-58. ATGL hydrolyses triacylglycerol (TG) into free fatty acids (FA) and 1,3- diacylglycerol (DG). HSL, also bound to lipid droplets, hydrolyses diacylglycerol (DG) into free fatty acids (FA) and monoacylglycerol (MG). The latter is in turn hydrolysed to free fatty acids (FA) and glycerol by Monoglyceride Lipase (MGL or MGLL), which is soluble in the cytosol (adapted from Thiam et al., 2013).

Interestingly, the genes encoding for the above-mentioned lipases have also been identified as candidate genes involved in pig IMF deposition and composition: *ATGL* gene was found less expressed in pigs with high IMF deposition (Zhao et al., 2009; Zhang et al., 2015), similarly to HSL, which has been observed to have a significantly increased expression in lean pig breeds (Zhang et al., 2015); furthermore, SNPs in *Monoglyceride Lipase* (MGLL) and its gene expression levels were associated to IMF composition in Spanish crossbred pigs (Puig-Oliveras et al., 2016). These evidences reported in literature suggest that together with these lipases, also *PLIN* genes (which are known to control the activity of the above-mentioned lipases) can be considered interesting candidate genes for porcine IMF deposition. Anyway, very little information is known about PLINs in livestock species, and most of the knowledge about these genes and proteins derive from studies in mouse, human and cultured cells. Gandolfi et al. (2011) and Davoli et al., (2011) performed the first studies on *PLIN1* and *PLIN2* genes and proteins in pig: Gandolfi et al. identified for the first time that in pig muscle PLIN1 and PLIN2 proteins are localized in correspondence with intramuscular adipocytes and intramyocellular lipids, respectively, and Davoli et al. found a novel SNP in *PLIN2* gene (*GU461317:g.98G>A*). Samples belonging to five different pig breeds were genotyped for *PLIN2* SNP, but intermediate allele frequencies were found

exclusively in the Italian Duroc breed, where the SNP was significantly associated with ADG, FCR, LC and HW EBVs.

A completely different landscape was proposed by Ren et al. (2017) in a recently published research. As IMF is generally recognized as a desirable attribute in pork meat and the selection for the improvement of this trait is particularly difficult, the authors proposed the use of transgenic pigs as a valuable resource for pork industry. The used pigs were genetically engineered individuals, exhibiting ectopic expression of *Phosphoenolpyruvate Carboxykinase* (*PEPCK-C*) driven by an α-skeletal-actin gene promoter. The transgenic pigs showed a consistent increase in IMF deposition respect to their wild type sibs, suggesting that enhancing the transcription of *PEPCK-C* has a direct effect on marbling. By the way, despite this interesting result, the use of engineered animals as meat source is not accepted public opinion, and to date this practice would arise consumers' concerns.

Backfat deposition and fatty acid composition traits in pigs

Adipocytes as already reported in previous subchapters are highly specialized cells, which have different roles from serving the crucial function of energy storage to being hormone secreting cells. In animals and humans, the potential to differentiate new fat cells from pluripotent stem cells continues throughout the lifespan, and the reduction in adipocyte number may occur as a result of adipocyte apoptosis or dedifferentiation. To date, adipose tissue is considered as a major endocrine organ that secretes numerous proteins and lipid hormones (adipokines and lipokines) (reviewed in Galic et al., 2010) and controls several major metabolic, physiological, and genetic changes. Adipose tissue is also the site of *de novo* FA synthesis in pigs (O'Hea and Leveille, 1969). Pig backfat (BF) is composed of subcutaneous adypocytes. Subcutaneous adipocytes display higher capacity for storing fat (lipogenic activity) compared to IMF adipocytes, in addition to increased insulin-induced lipogenic and lipolytic efficiency and higher levels of both genes and enzymes involved in lipid metabolism. This indicates that BF and IMF have different and specific biological features.

Backfat importance in meat products quality

BF thickness and BF fatty acid composition are extremely important traits affecting meat products quality, and play an essential role in particular in Italian heavy pig products: indeed, BF must be thick enough to obtain retailed fresh hams with fat cover ranging from 20 to 30 mm. BF has a prominent effect in reducing seasoning loss in high quality hams: an insufficient fat covering of the thigh causes higher seasoning loss and worse organoleptic characteristic in dry-cured hams (Bosi and Russo, 2004). BF thickness is in fact negatively correlated to seasoning loss, with correlation coefficients ranging from –0.79 to –0.50 (Bosi et al., 1984; Bosi and Russo, 2004) and genetic correlations between BF thickness and curing loss after salting in the same range (Carnier et al.,1999). The positive effect exerted by BF thickness on dry-cured hams is due to the lower amount of water contained in adipose tissue respect to muscle (5- 15% vs 70-75%), fact that turns BF into a barrier hampering the exchanges between the inner muscle and the external environment (Bosi and Russo, 2004). In addition to BF deposition, another important trait for influencing the quality of hams and meat is represented by BF fatty acid composition.

Fatty acids are classified on the basis of the length of their aliphatic chain (short, medium and long chain fatty acids) and on the basis of the type of the carbon-carbon bonds intervening in their chain: if the carbon atoms in their chains are exclusively tied with single bonds fatty acids are defined saturated, if there is a double bond they are known as monounsaturated, and when fatty acids have more than one carbon-carbon double bond are called polyunsaturated. Fatty acids have different characteristics depending on their composition: saturated fatty acids melt at higher temperatures (for example stearic melts at 69°C) compared to unsaturated fatty acids, and in particular with polyunsaturated ones (linoleic acid melts at 5°C) (Wood, 1984). Therefore, fatty acid composition of adipose tissue affects its firmness, as the different fatty acids have different melting points. Wood et al. (1989) found that both the objective and the subjective measures of adipose tissue firmness can be simulated basing on the fat content of stearic and linoleic acids: stearic is positively correlated with the firmness of subcutaneous fat in the shoulder region (r=0.35 with the objectively measured firmness and r=0.40 with the subjectively measured firmness), while linoleic content is negatively correlated with these two parameters (r=-0.75 with the

objectively measured firmness and r=-0.78 with the subjectively measured firmness). In addition to firmness, the biochemical features of fatty acids relevantly affect also fat depots oxidative stability. A higher prevalence of polyunsaturated fatty acids results in an increased rate of rancidity, as polyunsaturated fatty acids are liable to oxidative breakdown, causing a worsening in the organoleptic and nutritional quality of the meat products and an overall impairment of these products suitability for technological processing. For that reasons, Parma and San Daniele consortia accept only thighs with a content of linoleic acid in BF lower than 15% of total fatty acids (Bosi and Russo, 2004; Disciplinare Prosciutto di Parma DOP, reported at http://www.prosciuttodiparma.com/pdf/it_IT/disciplinare.28.11.2013.it.pdf). On the other side, polyunsaturated fatty acids have beneficial effects on human health and are requested for a balanced diet (omega 3 fatty acids especially). This dichotomy rises questions about what tools should be used to reach a favourable balance between saturated and unsaturated fatty acids in meat products. BF composition is affected by the diet lipids composition (Wood et al., 2008), but also by the pig genetic type (Lo Fiego et al., 2005). The latter result suggests that BF composition can be successfully modified through genetic selection.

Candidate genes involved in porcine backfat deposition and fatty acid composition

The identification of markers associated to BF deposition and composition would provide genetic tools useful for the selection of animals displaying a right balance between lean and fat deposition and between unsaturated and saturated fatty acids. To date, pig genome is known to harbour 240 QTLs associated to BF at last rib, 190 to BF at tenth rib, 1,311 for fat composition (http://www.animalgenome.org/cgi-bin/QTLdb/ last accessed 23rd of March 2017). Up to now, several genes have been identified as possibly associated to BF accumulation. One of the first genes found associated with pig fatness was *Insulin Growth Factor 2* (*IGF2*): several SNPs in its sequence have been associated to variations in lean and fat deposition. Since Van Laere et al. (2003) found a SNPs in intron 3 of *IGF2* responsible for

the association between the QTL comprising *IGF2* region and fatness, numerous other authors have investigated mutations in this gene sequence and their associations with porcine fat deposition (Vykoukalová et al., 2006; Burgos et al., 2012). Anyway, despite IGF2 was one of the first genes associated to fatness traits in pigs, its exact location in Sus scrofa Build 10.2 assembly is still unknown. Together with *IGF2* gene, also the *melanocortin-4 receptor* gene (*MC4R*) has been related to changes in pig adipogenic capabilities and has been widely studied in different pig breeds. Kim et al. (2000) identified a missense mutation in *MC4R* sequence with consequences on BF and growth rate of various pig lines, Chen et al. (2005) studied SNPs in both *IGF2* and *MC4R* in different pig population (comprising purebred and crossbred pigs) and obtained significant associations with carcass traits and fat deposition, and Davoli et al. (2012) analysed the effects of a *MC4R* SNP on carcass and growth performances in Italian Duroc and Italian Large White pigs. Anyway, literature is not concordant about the effects of *MC4R* mutations on BF accumulation, as this effect has not always been detected in all the studies, and different pig breeds showed divergent associations. For instance, MC4R showed to affect BF accumulation only in Italian Large White, while no association was found between Italian Duroc and BF thickness (Davoli et al., 2012). The authors suggested that this effect may be caused by breed-specific epistatic gene interactions between *MC4R* and other genes, affecting nutrient utilisation and addressing the growth towards lean or fat deposition depending on the breed different genetic and metabolic backgrounds. Another study carried out on Italian pig populations identified another candidate gene related to adipose tissue deposition: a SNP located in the sequence of the gene *Proprotein Convertase Subtilisin/Kexin Type 1* (*PCSK1*) showed several associations with fatness (in Italian Duroc pigs exclusively) and with growth traits (in both Italian Duroc and Italian Large White pigs) (Fontanesi et al., 2012). Recently, among the genes found associated with BF thickness is present also *Nuclear receptor subfamily 1, group H, member 3* (*NR1H3*, also known as *Liver X receptor α*, *LXRα*) (Zhang et al., 2016a). This member of the LXR nuclear receptor super family is an important regulator of lipid and fatty acid accumulation in different tissues and, among the others, its expression suppresses also *PLIN1* transcription (Stenson et al., 2011). Two other genes associated to pig BF deposition and composition are *Fatty Acid Synthase* (*FASN*) and *Insulin Induced Gene 2* (*INSIG2*): Grzes et al. (2016) found in four commercial pig breeds 12 novel polymorphisms in *FASN* gene sequence and seven novel SNPs in *INSIG2* nucleotide sequence. Surprisingly, the two genes

showed inverted associations with fat traits respect to their respective functional role. Indeed, *FASN*, which is known for its role in fatty acid synthesis (FASN enzyme catalyses the synthesis of palmitate from acetyl-CoA and malonyl-CoA), showed statistical associations with BF thickness, while *INSIG2* resulted to be highly associated with the fatty acid profile of the considered tissues (Grzes et al., 2016). By the way, despite the consistent number of studies, there are discrepancies among the results reported in literature and the understanding of the effects exerted by those candidate genes remains arduous. To shed light on the expression patterns linked to BF accumulation, recent studies have focused on the transcriptome of porcine adipose tissue, with the aim of finding differentially expressed transcripts suggesting new candidate genes related to this trait. Xing et al. (2016) used a high-throughput sequencing approach to identify transcriptomes and whole-genome differences from adipose tissue samples collected from three full-sibling pairs of pigs with divergent BF thickness. The authors obtained 20 differentially expressed genes that were then matched to the QTLs associated with fatness in pigs. This integrated approach permitted to identify two SNPs and one haplotype of *Malic Enzyme 1* (*ME1*) gene significantly associated to fat deposition in pigs (Xing et al., 2016). The results achieved in the same research highlighted also *Stearoyl-CoA Desaturase Δ9* (*SCD*) among the genes up-regulated in the high BF thickness group of pigs. *SCD* is a candidate gene involved in the biosynthesis of unsaturated fatty acids, in particular it encodes a desaturase that catalyses the insertion of a cis double bond at the delta-9 position into fatty acyl-CoA substrates, including palmitoyl-CoA and stearoyl-CoA among its preferential substrates. This gene has already been studied by other authors as a candidate gene for oleic fatty acid content in muscle: Estany et al. (2014) identified different SNPs in its sequence and suggested that the mutation AY487830:g.2228T>C in *SCD* promoter region enhances fat desaturation. These results were confirmed in Spanish Duroc pigs by a GWAS that highlighted the *SCD* chromosomal region as involved in intramuscular fat content of oleic acid (Ros-Freixedes et al., 2016). Another GWAS performed by Zhang et al. (2016b) identified the *SCD* genomic region as involved in fatty acid composition: the authors performed a comprehensive GWAS taking into account five different pig populations, whose genotypes were investigated through high-density SNP panels. This research was addressed towards the deciphering of the main players involved in the fatty acid composition of *Longissimus dorsi* muscle and abdominal fat, but highlighted some candidate genes that can have a prominent role also on

subcutaneous fat composition. Both *SCD* and *FASN* regions showed associations with the contents of the different fatty acids: in agreement with the previous studies (Estany et al., 2014; Ros-Freixedes et al., 2016) *SCD* region affected in particular the contents of palmitoleic, oleic and monounsaturated fatty acids, and markers nearby *FASN* were associated to the content of palmitic fatty acid (Zhang et al., 2016b). In addition, other candidate genes have been suggested by the authors: in particular several genes belonging to the ELOVL elongase family (*ELOVL5*, *ELOVL6* and *ELOVL7*), which take part in the elongation of fatty acid carbon chain. Other loci associated to the fatty acid composition of muscle and abdominal fat were those harbouring the genes *Acyl-CoA Synthetase Bubblegum Family Member 1* (*ACSBG1*, also known as *Very Long-Chain Acyl-CoA Synthetase*), *Fatty Acid Desaturase 2* (*FADS2*) and *Sterol Regulatory Element Binding Transcription Factor 2* (*SREBP2*) (Zhang et al., 2016b). These players seem to affect mainly the synthesis of long and very long fatty acids, while *SCD* and *ELOVL6* are hypothesised to take a more active part in the synthesis of palmitoleic and oleic acids (C16 and C18). On the whole, the candidate genes suggested by Zhang et al. (2016b) may be considered also for their involvement in *de novo* fatty acid synthesis of subcutaneous fat tissue. Anyway, much remains to be understood and further studies are needed to shed light on the complex gene patterns acting in the biosynthesis of fatty acids.

Maternal traits in pig

Maternal traits are extremely important for the pig breeding industry, but at the same time are quite arduous to improve as the prediction of the genetic merit for these traits requires long times and has low accuracy if progeny tests are not performed. In the 1980s and 1990s, litter size trait dominated the maternal index and was considered the main objective of selection schemes in maternal lines (Quinton et al., 2006). However, most studies recently present the economic relevance of a more balanced maternal breeding objective (Knap, 2005; Serenius and Muhonen, 2007; Amer et al., 2014). Several authors proposed to include among the maternal traits that undergo the selection process also traits describing vitality, uniformity, robustness, welfare, and health of animals (Knap, 2005; Merks et al., 2012; Hermesch and Amer, 2013). Knap (2005) defined robustness traits as preweaning survival, growing pig survival, and the number of litters a sow has over a lifetime, which were all shown to contribute in the improvement of pig production profitability.

Colostrum composition: an opportunity for the deciphering of sows' maternal traits

Within the landscape of the on-going research for new traits to be included in selection schemes, sow colostrum constitutes a virgin territory: unlike in dairy cattle, pig *colostrum* is to date poorly studied.

Colostrum starts to be produced by the porcine mammary gland before parturition and this production continues until up to 48 hours after the onset of lactation (Klobasa et al., 1987). Then from around 24 to 36 hours after parturition, *colostrum* is gradually replaced by mature milk (Rooke and Bland, 2002). As in other mammals, nutrients from blood pass inside the mammary gland epithelial cells, where are used to synthesise milk components. Anyway, unlike dairy cattle, the porcine mammary gland does not contain any cisterns and milk is only stored in the alveoli and milk ducts (Hartmann and Holmes, 1989). This anatomical characteristic impairs the possibility of milking sows and makes it difficult to measure the

yield of sow *colostrum* and milk. *Colostrum* production is highly floating among sows (Farmer and Quesnel, 2009; Quesnel, 2011) and can be affected by the farrowing parity, the weight of the sow at the time of parturition, sow health and nutritional state, sow diet and sow genetic type (Klobasa et al., 1987; Le Dividich et al., 2005; Devillers et al., 2007; Farmer and Quesnel, 2009). *Colostrum* is known to have a different macroscopical composition respect to milk: similarly to what has been observed in dairy cattle, porcine *colostrum* has a higher content of protein than milk, and lower levels of lactose, fat and caseins (Klobasa et al., 1987; Le Dividich et al., 2005). This difference in the protein amount has to be mainly imputed to *colostrum* content of immunoglobulins, which provides piglets with passive immunity: during the first day of life piglets intestinal mucosa is able to absorb *colostrum* immunoglobulins; this ability is then lost after the first hours of life, and if ingested immunoglobulins undergo digestion. *Colostrum* is a highly variable secretion, which experiences alterations throughout the whole period of its secretion. As time passes, the percentages of fat and lactose increase, while immunoglobulins consistently decrease (Table 1).

Variable	3 h	6 h	24 h	*P*
% Fat	8.9 (0.6)	9.9 (0.5)	14.2 (0.6)	< 0.001
% Protein	25.2 (0.6)	21.9 (0.5)	11.6 (0.4)	< 0.001
% Lactose	3.1 (0.1)	3.6 (0.1)	5.4 (0.1)	< 0.001
% Dry matter	37.2 (0.7)	35.5 (0.6)	31.2 (0.7)	< 0.001
IgG (mg/ml)	92 (12)	85 (13)	18.3 (3)	< 0.001
IgA (mg/ml)	11 (1.3)	8.1 (0.8)	2.8 (0.3)	< 0.001

IgG: Immunoglobulin G.
IgA: Immunoglobulin A.

Table 1: Composition (%) and mean concentrations of immunoglobulins G and A (mg/ml) in *colostrum* collected at 3, 6 and 24 hours after the birth of the first piglet. The standard error of the mean is shown between brackets. ANOVA was used to analyse the time effect on *colostrum* composition and the *P* values were reported (adapted from Decaluwé et al., 2014).

Decaluwé et al. (2014) demonstrated that piglets' daily weight gain and survival until weaning is positively associated with *colostrum* intake per kg of birth weight, confirming the importance of this secretion for the survival and growth until weaning.

Apart from immunoglobulins, *colostrum* composition in pigs is still poorly known and investigated. Other still unknown compounds could affect piglets' survival and growth in addition to immunoglobulins, and the understanding of their changes could represent a first step towards the knowledge of the genes associated to different colostrum compositions.

Growth and health traits in broiler

Selection in chickens has similar characteristics compared to pig breeding programs for both their architecture (pigs and chickens are selected through pyramidal breeding programs) and the peculiar attributes of the reproductive cycle in these two species (the generation interval is short and the offspring is numerous) (Jonas and de Koning, 2015). Furthermore, chickens are the fastest-growing farmed species, reaching in 40 days of life the slaughter weight. For all these reasons, GS (genomic selection) has not gained a prominent role in the selection of this livestock species, unlike in dairy cattle. Broiler have been selected intensively for growth traits such as body weight and feed efficiency, leading to a dramatic shortening of the production cycle: the National Chicken Council estimates that the average number of days it took to raise a chicken passed from 112 days in 1925 to 47 in 2016 (http://www.nationalchickencouncil.org/about-the-industry/statistics/u-s-broiler-performance/ last accessed on 26th of March 2017). In parallel, the market demand for heavier broiler increased drastically: in 2016 the market weight almost triples the one requested for broilers in 1925. Thus, broilers have been selected to reach heavier weights in less than the half of the time. Furthermore, consumers' demand for breast meat has increased the value of this cut and led poultry producers to look for ways to optimise breast muscle growth (Thiruvenkadan et al., 2011). Selection for higher breast yield resulted in a genetic gain of 277% per generation, whilst keeping the body weight in the range of 2400 to 2450 g and maintaining feed conversion and fertility (Schmidt et al., 2006). However, the selection for these traits also led to a dramatic increase in the incidence of health problems, such as lameness, ascites and several muscle myopathies and abnormalities (Dransfield and Sosnicki, 1999; Sandercock et al., 2009; Petracci and Cavani, 2012). Among the health problems observed in the last decades, ascites have been estimated as one of the pathologies causing the most important losses for poultry industry, accounting for about 1 billion dollars annually around the world and for over 25% of broiler losses (Maxwell and Robertson, 1997; Navarro et al., 2002). This pathology has a complex aetiology: this growth-related disease is a functional hypoxia caused by the high oxygen requirement for rapid growth and the inability of the heart and lungs to deliver sufficient oxygen to muscle tissue. This disorder leads in its final stages to accumulation of fluid in the abdominal cavity

(resulting in death). The measure of this phenotype (accumulation of fluid in abdomen) anyway has a very low heritability (ranging from 0.08 to 0.15) and is extremely difficult to be recorded in live animals. For this reason, Pakdel et al. (2002a, b) suggested the use of other phenotypic measures that could represent effective traits for the genetic selection against ascites. In particular, the haematocrit value showed high heritability (0.46-0.50) and is a trait that can be easily recorded. The heritability of the traits related to ascites is presented in Table 2.

Traits	Heritability estimates (Mean ± s.e.)	Heritability values
Haematocrit value (HCT)	0.46 ± 0.05	0.50
Right ventricle weight (RV)	0.47 ± 0.05	0.41
Total ventricle weight (TV)	0.46 ± 0.05	-
Ratio (RV:TV)	0.45 ± 0.05	0.54.
ABD (Accumulation of fluid in abdomen)	0.08 ± 0.03	0.15
	Pakdel *et al.* (2002a)	Pakdel *et al.* (2002b)

Table 2: Heritability estimates for the traits related to ascites (adapted from Thiruvenkadan et al., 2011).

These results showed that ascites had a genetic root, suggesting that selection against this pathology was possible. Three QTL affecting ascites related traits were detected (Rabie et al., 2005; Pakdel et al., 2005b) and using this information in selection programs resulted in a higher broiler weight (+122 g) without increasing ascites incidence.

The occurrence of dystrophy-like defects in broiler breast muscle

During the last five years, an increasing number of flocks of commercial broiler chickens has showed a myopathy affecting the pectoralis major (and occasionally minor) muscles. This disease is characterised by myodegeneration of breast muscle, where the apoptotic myofibers are replaced with fat and fibrotic tissue (Sihvo et al., 2014; Mazzoni et al., 2015; Mutryn et al., 2015; Soglia et al., 2016). These characteristics strongly impair the nutritional and organoleptic quality of chicken breast meat: this detrimental effect is due to the altered

composition and reduced protein functionality in affected muscles, whose proteins show a lower ability to hold/bind water, causing the formation of gels containing denatured proteins and water (Mudalal et al., 2014; Mudalal et al., 2015). These alterations give birth to abnormal breasts, showing the defects generally called “white striping” (Kuttappan et al., 2009) and “wooden breast” (Sihvo et al., 2014). It is not clear whether these two defects are part of the same myopathy, which worsening cause breasts to pass from white striping to wooden breast, or if white striping and wooden breast must be considered separately: white-striped fillets are characterised by the occurrence of white striations parallel to muscle fibers on the surface of the *Pectoralis major* muscle (Kuttappan et al., 2009) and are sometimes coupled with ‘wooden breast’, which is characterised by macroscopically visible hardened and pale areas in the caudal part of the fillet (Sihvo et al., 2014). Anyway, what is clear is that both these two defects exhibit similar histological changes, consisting of moderate-to-severe polyphasic myodegeneration with regeneration, as well as variable amounts of interstitial connective tissue accumulation or fibrosis (Sihvo et al., 2014). Lorenzi et al. (2014) reported that the incidence of white striping in commercial broiler chickens raised in Italy is quite high, in particular in medium and heavy broilers where it reaches 43.0% of the slaughtered animals, with 6.2% of samples considered severe. Therefore, considering these data it is obvious that the occurrence of these defects is a troubling issue in poultry industry. As in the ascites case, one of the way that could be followed is the application of MAS for the selection of animals resistant to these myopathies. The heritability estimates for these two defects are not concordant among the results reported in literature: Bailey et al. (2015) reported for white striping an estimated heritability of 0.338 and for wooden breast a heritability lower than 0.1. On the other hand, Alnahhas et al. (2016) proposed a stronger genetic determinism at the basis of white striping defect, with an estimated heritability of 0.65. With the aim of identifying genes and biological pathways related to the occurrence of wooden breast, Mutryn et al. (2015) performed a transcriptome analysis on *Pectoralis major* muscle samples of chickens showing a normal appearance of the breast muscle and individuals with affected muscles. On the whole, over 1500 genes were differentially expressed between affected and unaffected birds, and most of them were related to the physiological response against hypoxia and oxidative stress, were involved in cellular repair and in intracellular calcium homeostasis (Mutryn et al., 2015). Literature suggests a strong genetic determinism at the basis of the occurrence of these dystrophy-like

defects, but to date the information is still lacking of markers and candidate genes associated with the incidence of white striping and wooden breast.

Aim of the study

The studies reported in the present thesis are addressed towards the application of different OMICs technologies to the analysis of productive traits in different species of farm animals. The study was mainly focused on the investigation of candidate genes and gene networks associated to porcine fatness traits using genomics, transcriptomics and single gene studies. Moreover, specific additional studies were planned to explore in preliminary analyses the factors affecting sows' maternal traits through metabolomics and chicken breast muscle myopathies through microarray technique.

The main subject developed was the identification and study of markers, genes and transcripts involved in porcine backfat deposition and fatty acid composition. Investigation on genes affecting intramuscular fat content in pig *Semimembranosus* muscle was also faced. For this purpose, different approaches have been considered: the research has been carried out integrating genome wide strategies with investigations focused on candidate genes chosen for their functional roles (such as *Perilipin* and lipases genes). The integration of different powerful and innovative approaches can be considered a promising emergent strategy to adopt, in particular for the traits for which selection process has been slowly accelerating.

The research has been carried out in pigs and chickens with the following peculiar aims:

- to investigate markers, genomic regions and transcripts related to porcine backfat deposition and/or backfat fatty acid composition through the combined approach of genomics, transcriptomics technologies and single gene studies (Chapters 1 to 5) and to investigate candidate genes for porcine intramuscular fat deposition integrating gene expression and protein expression data (Chapter 6); ii) to explore the factors affecting pig colostrum composition using a metabolomics approach and test the existence of compositional differences between three porcine breeds (Chapter 7); iii) to explore the factors affecting pig colostrum composition using a metabolomics approach and test the existence of compositional differences between three porcine breeds (Chapter 7); iv) To perform a microarray study for the detection of gene expression profiles involved in the occurrence of dystrophy-like defects in broilers breast muscle (Chapter 8).

References

Alnahhas N, Berri C, Chabault M, Chartrin P, Boulay M, Bourin MC & Le Bihan-Duval E (2016) Genetic parameters of white striping in relation to body weight, carcass composition, and meat quality traits in two broiler lines divergently selected for the ultimate pH of the pectoralis major muscle. *BMC Genet* 17:61.

Amer PR, Ludemann CI & Hermesch S (2014) Economic weights for maternal traits of sows, including sow longevity. *J Anim Sci* 92:5345–5357.

Andersson L (2001) Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet* 2:130–138.

Bailey RA, Watson KA, Bilgili SF & Avendano S (2015) The genetic basis of pectoralis major myopathies in modern broiler chicken lines. *Poult Sci* 94:2870–2879.

Barton-Gade PA (1990) Pork quality in genetic improvement programmes - the Danish experience. Proceeding of the National Swine Improvement Federation Annual Meeting; Des Moines, IA, USA.

Bosi P & Russo V (2004) The production of the heavy pig for high quality processed products. *I J Anim Sci* 3:309–321.

Bosi P, Russo V & Santoro P (1984) Calo di stagionatura del prosciutto rifilato e qualità della carcassa e della carne. Proc. 38th Nat. Congr. SISVet.,Rimini, Italia, 38: 487-489.

Bovine Genome Sequencing and Analysis Consortium et al. (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 324:522–528.

Brewer MS, Zhu LG & McKeith FK (2001) Marbling effects on quality characteristics of pork loin chops: consumer purchase intent, visual and sensory characteristics. *Meat Sci* 59:153–163.

Burgos C, Galve A, Moreno C, Altarriba J, Reina R, García C & López-Buesa P (2012) The effects of two alleles of IGF2 on fat content in pig carcasses and pork. *Meat Sci* 90:309–313.

Carnier P, Cassandro M, Knol E & Padoan D (1999) Genetic parameters for some carcass and fresh ham traits of crossbred Goland pigs. In: *Proceedings of the ASPA Congress - Recent Progress in Animal Production Science (Italy)*. .

Cermelli S, Guo Y, Gross SP & Welte MA (2006) The lipid-droplet proteome reveals that droplets are a protein-storage depot. *Curr Biol* 16:1783–1795.

Chen JF, Xiong YZ, Zuo B, Zheng R, Li FE, Lei MG, Li JL, Deng CY & Jiang SW (2005) New Evidences of Effect of Melanocortin-4 Receptor and Insulin-like Growth Factor 2 Genes on Fat Deposition and Carcass Traits in Different Pig Populations. *Asian-Australasian Journal of Animal Sciences, Asian-Australas J Anim Sci* 18:1542–1547.

da Costa N, Edgar J, Ooi P-T, Su Y, Meissner JD & Chang K-C (2007) Calcineurin differentially regulates fast myosin heavy chain genes in oxidative muscle fibre type conversion. *Cell Tissue Res* 329:515–527.

Dalen KT, Dahl T, Holter E, Arntsen B, Londos C, Sztalryd C & Nebb HI (2007) LSDP5 is a PAT protein specifically expressed in fatty acid oxidizing tissues. *Biochim Biophys Acta* 1771:210–227.

Davoli R, Braglia S, Valastro V, Annaratone C, Annarratone C, Comella M, Zambonelli P, Nisi I, Gallo M, Buttazzoni L & Russo V (2012) Analysis of MC4R polymorphism in Italian Large White and Italian Duroc pigs: association with carcass traits. *Meat Sci* 90:887–892.

Davoli R, Gandolfi G, Braglia S, Comella M, Zambonelli P, Buttazzoni L & Russo V (2011) New SNP of the porcine perilipin 2 (PLIN2) gene, association with carcass traits and expression analysis in skeletal muscle. *Mol Biol Rep* 38:1575–1583.

Davoli R, Luise D, Mingazzini V, Zambonelli P, Braglia S, Serra A & Russo V (2016) Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *J Anim Breed Genet* 133:277–282.

Decaluwe R, Maes D, Wuyts B, Cools A, Piepers S & Janssens G (2014) Piglets' colostrum intake associates with daily weight gain and survival until weaning. *Livest Sci* 162:185–192.

Dekkers JCM (2007) Marker-assisted selection for commercial crossbred performance. *J Anim Sci* 85:2104–2114.

Devillers N, Farmer C, Le Dividich J. & Prunier A (2007) Variability of colostrum yield and colostrum intake in pigs. *Animal*. 1:1033-1041.

Dransfield E & Sosnicki AA (1999) Relationship between muscle growth and poultry meat quality. *Poultry Science* 78:743–746.

Fan B, Du Z-Q & Rothschild MF (2009) The Fat Mass and Obesity-Associated (FTO) Gene is Associated with Intramuscular Fat Content and Growth Rate in the Pig. *Anim Biotec* 20:58–70.

Farmer C & Quesnel H (2009) Nutritional, hormonal, and environmental effects on colostrum in sows. *J Anim Sci* 87:56-65.

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

Fontanesi L, Bertolini F, Scotti E, Trevisi P, Buttazzoni L, Dall&#39 S, Olio, Davoli R, Bosi P & Russo V (2012) Polymorphisms in an obesity-related gene (PCSK1) are associated with fat deposition and production traits in Italian heavy pigs. *Animal* 6:1913–1924.

Fontanesi L, Scotti E, Buttazzoni L, Dall'Olio S, Bagnato A, Fiego DPL, Davoli R & Russo V (2010) Confirmed association between a single nucleotide polymorphism in the FTO gene and obesity-related traits in heavy pigs. *Mol Biol Rep* 37:461–466.

Friedman JM (2010) A tale of two hormones. *Nat Med* 16:1100–1106.

Galic S, Oakhill JS & Steinberg GR (2010) Adipose tissue as an endocrine organ. *Mol Cell Endocrinol* 316:129–139.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V & Davoli R (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631–637.

Garrick DJ (2011) The nature, scope and impact of genomic prediction in beef cattle in the United States. *Genet Sel Evol* 43:17.

Gerbens F, Verburg FJ, Van Moerkerk HT, Engel B, Buist W, Veerkamp JH & te Pas MF (2001) Associations of heart and adipocyte fatty acid-binding protein gene expression with intramuscular fat content in pigs. *J Anim Sci* 79:347–354.

Goodpaster BH, Thaete FL & Kelley DE (2000) Thigh adipose tissue distribution is associated with insulin resistance in obesity and in type 2 diabetes mellitus. *Am J Clin Nutr* 71:885–892.

Granneman JG, Moore H-PH, Mottillo EP & Zhu Z (2009) Functional interactions between Mldp (LSDP5) and Abhd5 in the control of intracellular lipid accumulation. *J Biol Chem* 284:3049–3057.

Granneman JG, Moore H-PH, Mottillo EP, Zhu Z & Zhou L (2011) Interactions of perilipin-5 (Plin5) with adipose triglyceride lipase. *J Biol Chem* 286:5126–5135.

Groenen MAM et al. (2012) Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491:393–398.

Grzes M, Sadkowski S, Rzewuska K, Szydlowski M & Switonski M (2016) Pig fatness in relation to FASN and INSIG2 genes polymorphism and their transcript level. *Mol Biol Rep* 43:381–389.

Hartmann PE & Holmes M A (1989) Sow lactation. In: Manipulating Pig Production II. Proceedings of the Biennial Conference of the Australasian Pig Science Association (A.P.S.A.), pp 72-97. Albury, NSW on November 27 -29.

Hausman GJ, Basu U, Du M, Fernyhough-Culver M & Dodson MV (2014) Intermuscular and intramuscular adipose tissues: Bad vs. good adipose tissues. *Adipocyte* 3:242–255.

Hausman GJ, Dodson MV, Ajuwon K, Azain M, Barnes KM, Guan LL, Jiang Z, Poulos SP, Sainz RD, Smith S, Spurlock M, Novakofski J, Fernyhough ME & Bergen WG (2009) Board-invited review: the biology and regulation of preadipocytes and adipocytes in meat animals. *J Anim Sci* 87:1218–1246.

Henderson CR & Quaas RL (1976) Multiple Trait Evaluation Using Relatives' Records. *J Anim Sci* 43:1188–1197.

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–447.

Hermesch S & Amer PR (2013) Deriving economic values for reaction norms of growth in pigs. *Proc Assoc Adv Anim Breed Genet* 20:475–478.

Hermesch S, Luxfordb BG & Grasera HU (2000) Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs: 2. Genetic relationships between production, carcase and meat quality traits. *Livest Prod Sci* 65:249–259.

Hill WG (2014) Applications of population genetics to animal breeding, from wright, fisher and lush to genomic prediction. *Genetics* 196:1–16.

International Sheep Genomics Consortium, Archibald AL, Cockett NE, Dalrymple BP, Faraut T, Kijas JW, Maddox JF, McEwan JC, Hutton Oddy V, Raadsma HW, Wade C, Wang J, Wang W & Xun X (2010) The sheep genome reference sequence: a work in progress. *Anim Genet* 41:449–453.

Jonas E & de Koning D-J (2015) Genomic selection needs to be carefully assessed to meet specific requirements in livestock breeding programs. *Front Genet* 6.

Jung J-H, Shim K-S, Na C-S & Choe H-S (2015) Studies on Intramuscular Fat Percentage in Live Swine Using Real-time Ultrasound to Determine Pork Quality. *Asian-Australas J Anim Sci* 28:318–322.

Kim KS, Larsen N, Short T, Plastow G & Rothschild MF (2000) A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm Genome* 11:131–135.

Kimmel AR & Sztalryd C (2016) The Perilipins: Major Cytosolic Lipid Droplet-Associated Proteins and Their Roles in Cellular Lipid Storage, Mobilization, and Systemic Homeostasis. *Annu Rev Nutr* 36:471–509.

Klobasa F, Werhahn E & Butler J E (1987) Composition of Sow Milk During Lactation. *J Anim Sci* 64:1458-1466.

Knap PW (2005) Breeding robust pigs. *Aust J Exp Agric* 45:763–773.

Komolka K, Albrecht E, Wimmers K, Michal JJ & Maak S (2014) Molecular Heterogeneities of Adipose Depots - Potential Effects on Adipose-Muscle Cross-Talk in Humans, Mice and Farm Animals. *J Genomics* 2:31-44.

Kuttappan VA, Brewer VB, Clark FD, McKee SR, Meullenet JF, Emmert JL & Owens CM (2009) Effect of white striping on the histological and meat quality characteristics of broiler fillets. *Poult Sci* 88:136–137.

Langhi C, Marquart TJ, Allen RM & Baldán A (2014) Perilipin-5 is regulated by statins and controls triglyceride contents in the hepatocyte. *J Hepatol* 61:358–365.

Le Dividich J, Rooke JA & Herpin P (2005) Nutritional and immunological importance of colostrum for the new-born pig. *J Agricul Sci* 143:469-485.

Li X, Kim S-W, Choi J-S, Lee Y-M, Lee C-K, Choi B-H, Kim T-H, Choi Y-I, Kim J-J & Kim K-S (2010) Investigation of porcine FABP3 and LEPR gene polymorphisms and mRNA expression for variation in intramuscular fat content. *Mol Biol Rep* 37:3931–3939.

Lillehammer M, Meuwissen THE & Sonesson AK (2011) Genomic selection for maternal traits in pigs. *J Anim Sci* 89:3908–3916.

Lo Fiego DP, 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 Sci* 69:107–114.

Lopes MS, Bovenhuis H, van Son M, Nordbø Ø, Grindflek EH, Knol EF & Bastiaansen JWM (2017) Using markers with large effect in genetic and genomic predictions. *J Anim Sci* 95:59–71.

Lorenzi M, Mudalal S, Cavani C & Petracci M (2014) Incidence of white striping under commercial conditions in medium and heavy broiler chickens in Italy. *J Appl Poult Res* 23:754–758.

Mandard S, Müller M & Kersten S (2004) Peroxisome proliferator-activated receptor alpha target genes. *Cell Mol Life Sci* 61:393–416.

Matika O, Riggio V, Anselme-Moizan M, Law AS, Pong-Wong R, Archibald AL & Bishop SC (2016) Genome-wide association reveals QTL for growth, bone and in vivo carcass traits as assessed by computed tomography in Scottish Blackface lambs. *Genet Sel Evol* 48:11.

Maxwell MH & Robertson GW (1997) World broiler ascites survey 1996. *Poultry International* 36:16-19

Mazzoni M, Petracci M, Meluzzi A, Cavani C, Clavenzani P & Sirri F (2015) Relationship between pectoralis major muscle histology and quality traits of chicken meat. *Poult Sci* 94:123–130.

Melo RCN, D'Avila H, Wan H-C, Bozza PT, Dvorak AM & Weller PF (2011) Lipid bodies in inflammatory cells: structure, function, and current imaging techniques. *J Histochem Cytochem* 59:540–556.

Merks JWM, Mathur PK & Knol EF (2012) New phenotypes for new breeding goals in pigs. *Animal* 6:535–543.

Mota de Sá P, Richard AJ, Hang H & Stephens JM (2017) Transcriptional Regulation of Adipogenesis. *Compr Physiol* 7:635-674.

Mottram DS & Salter LJ (1989) Flavour formation in meat-related Maillard systems containing phospholipids. In Thermal Generation of Aromas, pp. 442–451 [Parliment TH, McGorrin RJ & Ho CT, editors]. Washington, DC: American Chemical Society.

Mudalal S, Babini E, Cavani C & Petracci M (2014) Quantity and functionality of protein fractions in chicken breast fillets affected by white striping. *Poult Sci* 93:2108–2116.

Mudalal S, Lorenzi M, Soglia F, Cavani C & Petracci M (2015) Implications of white striping and wooden breast abnormalities on quality traits of raw and marinated chicken meat. *Animal* 9:728–734.

Muraoka O, Xu B, Tsurumaki T, Akira S, Yamaguchi T & Higuchi H (2003) Leptin-induced transactivation of NPY gene promoter mediated by JAK1, JAK2 and STAT3 in the neural cell lines. *Neurochem Int* 42:591–601.

Muscle Triglyceride and Insulin Resistance (2002) *Annual Rev Nutr* 22:325–346.

Mutryn MF, Brannick EM, Fu W, Lee WR & Abasht B (2015) Characterization of a novel chicken muscle disorder through differential gene expression and pathway analysis using RNA-sequencing. *BMC Genomics* 16.

Navarro P, Koerhuis ANM, Chatziplis D, Visscher PM & Haley CS (2002) Genetic studies of ascites in a broiler population. Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August 19-23.

Newcom DW, Baas TJ & Lampe JF (2002) Prediction of intramuscular fat percentage in live swine using real-time ultrasound. *J Anim Sci* 80:3046–3052.

Nirea K g. & Meuwissen T h. e. (2017a) Improving production efficiency in the presence of genotype by environment interactions in pig genomic selection breeding programmes. *J Anim Breed Genet* 134:119–128.

Nirea K g. & Meuwissen T h. e. (2017b) Improving production efficiency in the presence of genotype by environment interactions in pig genomic selection breeding programmes. *J Anim Breed Genet* 134:119–128.

O'Hea EK & Leveille GA (1969) Significance of adipose tissue and liver as sites of fatty acid synthesis in the pig and the efficiency of utilization of various substrates for lipogenesis. *J Nutr* 99:338–344.

Ohsaki Y, Suzuki M & Fujimoto T (2014) Open Questions in Lipid Droplet Biology. *Chem & Biol* 21:86–96.

Pakdel A, Van Arendonk JAM, Vereijken ALJ & Bovenhuis H (2002a) Genetic and phenotypic correlations for ascites elated traits in broilers. Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August 19-23.

Pakdel A, Van Arendonk L, Vereijken A & Boverhuis H (2002b) Direct and maternal genetic effects for ascites related traits in broilers. *Poult Sci* 81:1273-1279.

Petracci M & Cavani C (2012) Muscle growth and poultry meat quality issues. *Nutrients* 4:1-12.

Poulsen L la C, Siersbæk M & Mandrup S (2012) PPARs: fatty acid sensors controlling metabolism. *Semin Cell Dev Biol* 23:631–639.

Pryce JE & Daetwyler HD (2012) Designing dairy cattle breeding schemes under genomic selection: a review of international research. *Anim Prod Sci* 52:107–114.

Puig-Oliveras A, Revilla M, Castelló A, Fernández AI, Folch JM & Ballester M (2016) Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. *Sci Rep* 6.

Quesnel H (2011) Colostrum production by sows: variability of colostrum yield and immunoglobulin G concentrations. *Animal* 5:1546-1553.

Quinton VM, Wilton JW, Robinson JA & Mathur PK (2006) Economic weights for sow productivity traits in nucleus pig populations. *Livestock Science* 99:69–77.

Rabie TSKM, Crooijmans RPMA, Bovenhuis H, Vereijken ALJ, Veenendaal T, Van Der Poel JJ, Van Arendonk JAM, Pakdel A. & Groenen MAM (2005) Genetic mapping of quantitative trait loci affecting susceptibility in chicken to develop pulmonary hypertension syndrome. *Anim Genet* 36:468-476.

Ren Z, Wang Y, Ren Y, Zhang Z, Gu W, Wu Z, Chen L, Mou L, Li R, Yang H & Dai Y (2017) Enhancement of porcine intramuscular fat content by overexpression of the cytosolic form of phosphoenolpyruvate carboxykinase in skeletal muscle. *Sci Rep* 7:43746.

Robenek H, Buers I, Hofnagel O, Robenek MJ, Troyer D & Severs NJ (2009) Compartmentalization of proteins in lipid droplet biogenesis. *Biochim Biophys Acta* 1791:408–418.

Rooke JA & Bland IM (2002) The acquisition of passive immunity in the new-born piglet. *Livest Prod Sci* 78:13-23.

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

Rothschild MF & Soller M (1997) Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. *Probe* 8:13–20.

Russo V, Buttazzoni L, Baiocco C, Davoli MR, Nanni Costa NL, Schivazappa OC & Virgili PC (2000) Heritability of muscular cathepsin B activity in Italian Large White pigs. *J Anim Breed Genet* 117:37–42.

Samorè AB, Buttazzoni L, Gallo M, Russo V & Fontanesi L (2015) Genomic selection in a pig population including information from slaughtered full sibs of boars within a sib-testing program. *Animal* 9:750–759.

Sandercock DA, Barker Z, Mitchell MA & Hocking PH (2009) Changes in muscle cell cation regulation and meat quality traits are associated with genetic selection for high body weight and meat yield in broiler chickens. *Genet Sel Evol* 41:8.

Santos-Garcés E, Muñoz I, Gou P, Garcia-Gil N & Fulladosa E (2014) Including estimated intramuscular fat content from computed tomography images improves prediction accuracy of dry-cured ham composition. *Meat Sci* 96:943–947.

Schmidt GS, Figueiredo EAP, Ledur M, De Avila VS & Scheuermann GN (2006) Effect of selection for breast size in broilers. Proc. 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, MG, Brasil, August 13-18, 2006

Schwab CR (2007) *Quantitative and Molecular Genetic Components of Selection for Intramuscular Fat in Duroc Swine*. ProQuest.

Schworer DA, Rebsamen A & Lorenz D (1995) Selection of intramuscular fat in Swiss pig breeds and the importance of fatty tissue quality. Proceeding of 2nd Dummerstorf Muscle Workshop on Growth and Meat Quality; Rostock, Germany.

Sell H, Deshaies Y & Richard D (2004) The brown adipocyte: update on its metabolic role. *Int. J. Biochem. Cell Biol.*36:2098–2104.

Serenius T & Muhonen P (2007) Economic values of pork production related traits in Finland. *Agr Food Sci* 16:79–88.

Sharma A, Lee JS, Dang CG, Sudrajad P, Kim HC, Yeon SH, Kang HS & Lee S-H (2015) Stories and Challenges of Genome Wide Association Studies in Livestock — A Review. *Asian-Australas J Anim Sci* 28:1371–1379.

Sihvo H-K, Immonen K & Puolanne E (2014) Myodegeneration with fibrosis and regeneration in the pectoralis major muscle of broilers. *Vet Pathol* 51:619–623.

Singh R, Kaushik S, Wang Y, Xiang Y, Novak I, Komatsu M, Tanaka K, Cuervo AM & Czaja MJ (2009) Autophagy regulates lipid metabolism. *Nature* 458:1131–1135.

Sitzenstock F, Ytournel F, Sharifi AR, Cavero D, Täubert H, Preisinger R & Simianer H (2013) Efficiency of genomic selection in an established commercial layer breeding program. *Genet Sel Evol* 45:29.

Soglia F, Mudalal S, Babini E, Di Nunzio M, Mazzoni M, Sirri F, Cavani C & Petracci M (2016) Histology, composition, and quality traits of chicken Pectoralis major muscle affected by wooden breast abnormality. *Poult Sci* 95:651-659.

Stenson BM, Rydén M, Venteclef N, Dahlman I, Pettersson AML, Mairal A, Aström G, Blomqvist L, Wang V, Jocken JWE, Clément K, Langin D, Arner P & Laurencikiene J (2011) Liver X receptor (LXR) regulates human adipocyte lipolysis. *J Biol Chem* 286:370–379.

Thiam AR, Farese RV & Walther TC (2013) The Biophysics and Cell Biology of Lipid Droplets. *Nat Rev Mol Cell Biol* 14:775–786.

Thiruvenkadan AK, Prabakaran R & Panneerselvam S (2011) Broiler breeding strategies over the decades: an overview. *World's Poultry Science Journal* 67:309–336.

Tyra M & Ropka-Molik K (2011) Effect of the FABP3 and LEPR gene polymorphisms and expression levels on intramuscular fat (IMF) content and fat cover degree in pigs. *Livest Sci* 142:114–120.

Tyra M, Ropka-Molik K, Terman A, Piórkowska K, Oczkowicz M & Bereta A (2013) Association between subcutaneous and intramuscular fat content in porcine ham and loin depending

on age, breed and FABP3 and LEPR genes transcript abundance. *Mol Biol Rep* 40:2301–2308.

Van Laere A-S, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M & Andersson L (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425:832–836.

Vettor R, Milan G, Franzin C, Sanna M, De Coppi P, Rizzuto R & Federspil G (2009) The origin of intermuscular adipose tissue and its pathophysiological implications. *Am J Physiol Endocrinol Metab* 297:E987–998.

Visscher P, Pong-Wong R, Whittemore C & Haley C (2000) Impact of biotechnology on (cross)breeding programmes in pigs. *Livest Prod Sci* 65:57–70.

Vykoukalová Z, Knoll A, Dvorák J & Čepica S (2006) New SNPs in the IGF2 gene and association between this gene and backfat thickness and lean meat content in Large White pigs. *J Anim Breed Genet* 123:204–207.

Wang H, Bell M, Sreenivasan U, Sreenevasan U, Hu H, Liu J, Dalen K, Londos C, Yamaguchi T, Rizzo MA, Coleman R, Gong D, Brasaemle D & Sztalryd C (2011) Unique regulation of adipose triglyceride lipase (ATGL) by perilipin 5, a lipid droplet-associated protein. *J Biol Chem* 286:15707–15715.

Wang H, Hu L, Dalen K, Dorward H, Marcinkiewicz A, Russell D, Gong D, Londos C, Yamaguchi T, Holm C, Rizzo MA, Brasaemle D & Sztalryd C (2009) Activation of hormone-sensitive lipase requires two steps, protein phosphorylation and binding to the PAT-1 domain of lipid droplet coat proteins. *J Biol Chem* 284:32116–32125.

Wolins NE, Quaynor BK, Skinner JR, Tzekov A, Croce MA, Gropler MC, Varma V, Yao-Borengasser A, Rasouli N, Kern PA, Finck BN & Bickel PE (2006) OXPAT/PAT-1 is a PPAR-induced lipid droplet protein that promotes fatty acid utilization. *Diabetes* 55:3418–3428.

Wood JD (1990) Consequences for meat quality of reducing carcass fatness. In Reducing Fat in Meat Animals, pp. 344–397 [JD Wood and AV Fisher, editors]. London: Elsevier Applied Science.

Wood JD (1984). Fat deposition and the quality of fat tissue in meat animals. In Fats in animal nutrition, pp. 407–435 [J Wiseman, editor]. London: Butterworths.

Wood JD, Enser M, Fisher AV, Nute GR, Richardson RI & Sheard PR (1999) Manipulating meat quality and composition. *Proc Nutr Soc* 58:363–370.

Wood JD, Enser M, Whittington FM, Moncrieff CB & Kempster AJ (1989) Backfat composition in pigs: Differences between fat thickness groups and sexes. *Livest Prod Sci* 22:351–362.

Xing K, Zhu F, Zhai L, Chen S, Tan Z, Sun Y, Hou Z & Wang C (2016) Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data. *Sci Rep* 6:23219.

Zhang B, Shang P, Qiangba Y, Xu A, Wang Z & Zhang H (2016a) The association of NR1H3 gene with lipid deposition in the pig. *Lipids Health Dis* 15.

Zhang C, Luo JQ, Zheng P, Yu B, Huang ZQ, Mao XB, He J, Yu J, Chen JL & Chen DW (2015) Differential expression of lipid metabolism-related genes and myosin heavy chain isoform genes in pig muscle tissue leading to different meat quality. *Animal* 9:1073–1080.

Zhang W, Bin Yang null, Zhang J, Cui L, Ma J, Chen C, Ai H, Xiao S, Ren J & Huang L (2016b) Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. *Sci Rep* 6:24718.

Zhang Y, Proenca R, Maffei M, Barone M, Leopold L & Friedman JM (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature* 372:425–432.

Zhao SM, Ren LJ, Chen L, Zhang X, Cheng ML, Li WZ, Zhang YY & Gao SZ (2009) Differential expression of lipid metabolism related genes in porcine muscle tissue leading to different intramuscular fat deposition. *Lipids* 44:1029–1037

Chapter 1: Genome-wide association study confirms the importance of ELOVL6 region on subcutaneous fatty acid composition in Italian Large White pigs.

M. Zappaterra[1], R. Ros-Freixedes[2,3], J. Estany[2], R. Davoli[1]

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

[2]*Departament de Ciència Animal, Universitat de Lleida-Agrotecnio Centre, Lleida, Catalonia (Spain).*

[3]*Present address: The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Scotland (UK).*

In preparation

Summary

Dietary fatty acid composition has an impact on human health. There is an increasing request from consumers for healthier food and pork industry must respond to it without worsening performance and the technological properties of pork products. The inclusion of genetic markers for carcass fatty acid composition in pig selection schemes could be a useful tool to reach the right balance between unsaturated and saturated fatty acids to satisfy market demands. With the aim of finding genomic regions associated with porcine backfat fatty acid composition, a genome-wide association study was performed on 798 Italian Large White pigs genotyped through Illumina PorcineSNP60k. The strongest associations with backfat contents of palmitic, palmitoleic, oleic, medium chain and long chain fatty acids were found for the SSC8 region located at 119-122 Mb, where the gene *ELOVL elongase 6* is mapped. Palmitic, palmitoleic, stearic and oleic acid contents were also found associated with SSC14, in particular with the genomic region at 121-124 Mb, where stearoyl-CoA desaturase Δ9 gene lies. On the other hand, the genomic regions associated with backfat contents of arachidic, arachidonic, omega6 and omega3 fatty acids showed to harbour mainly genes involved in dietary lipids and carbohydrates digestion, absorption and utilisation. To our knowledge, this is the first study performed in Large White pigs identifying markers and genomic regions associated with backfat fatty acid composition. The obtained results indicate the likely involvement of distinct molecular pathways leading to different fatty acid deposition.

Introduction

Over the last decades, the interest in the fatty acid (FA) composition of meat and other animal products has increased. Fat composition of both subcutaneous and intramuscular fat is an important factor influencing organoleptic and nutritional quality of animal products (Wood et al., 2008). In particular, FA composition strongly affects firmness, tenderness (Wood et al., 2008) and the technological quality of pork processed products. While pork rich in polyunsaturated fatty acids (PUFA) is preferable in healthy human diets (WHO, 2003), their higher susceptibility to oxidation (Shackelford et al., 1990) and the lower melting temperatures (Wood, 1984) make PUFA-rich pork cuts unsuitable for seasoning (Bosi et al.,

2000). A higher amount of saturated fatty acids (SFA) is instead preferable for pork industry, but on the other side they may result in undesirable effects on human health (Kris-Etherton and Yu, 1997). The contrasting requests of consumers and pig meat industry raised in the last years the issue of selection schemes aimed at improving fat and meat FA composition. This goal appears to be still far from being reached and therefore a deeper knowledge of genes involved in FA biosynthesis and metabolism is needed (Zhang et al., 2016). Among the instruments available for the modern genomic studies, Genome Wide Association (GWA) analyses are the most used tool for the identification of the molecular drivers underlying animal and human traits and diseases (Sharma et al., 2015; Denny et al., 2016). Several GWA studies aimed at identifying *loci* associated with the different FA composition of pig subcutaneous (Corominas et al., 2013; Ros-Freixedes et al., 2016) and intramuscular fat (Muñoz et al., 2013; Davoli et al., 2015; Ros-Freixedes et al., 2016; Zhang et al., 2016) have been performed since today, and different QTL regions harbouring candidate genes for fat traits have been indicated. However, the knowledge on the genes involved in fat composition is still incomplete, and the results of the different studies do not always agree on the genomic regions associated with fat composition complex trait.

The main objective of this study was to detect genomic regions associated with the most important backfat FA in a population of Italian Large White pigs.

Materials and Methods

Sampling

For the present study, 889 Italian Large White pigs were used. The animals were pure breed pigs included in the Italian sib test genetic evaluation scheme performed by ANAS (Associazione Nazionale Allevatori Suini, ANAS; www.anas.it), reared in the same environmental conditions at the genetic test station with a *quasi ad libitum* feeding level (60% of the pigs were able to ingest the entire supplied ration). At about 150 kg of live weight the animals were transported to a commercial abattoir located at about 25 km from the test station in accordance with Council Rule (EC) No. 1/2005 regarding the protection of animals during transport and related operations. At the slaughterhouse, the pigs were

electrically stunned and bled in a supine position in agreement with Council Regulation (EC) No. 1099/2009 regarding the protection of animals at the time of slaughter. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health. Carcass weight and backfat thickness measured with Fat-O-Meter (FOM) at 8 cm off the midline of the carcass at the level placed between the third and fourth last ribs. Samples of blood and backfat were gathered, immediately frozen in liquid nitrogen and stored at -20°C.

Determination of backfat fatty acid composition

The backfat tissue samples collected after slaughtering were conserved at -20°C until processed. Backfat FA composition was detected by direct trans-esterification, following the protocol reported by Murrieta et al. (2003). For each sample, 50 mg of frozen backfat was used for the total lipid extraction and then in each tube 0.5 mg of C19:0 methyl ester in hexane was added as internal standard. Gas chromatography was performed on GC-2010 Plus High-end Gas Chromatograph (Shimadzu Corporation, Tokyo, Japan), using SP^{TM}-2560 Capillary GC Column (Sigma-Aldrich, Merck, Darmstadt, Germany). Backfat FA composition was expressed as the ratio between each FA and the total. Due to some missing information in the pedigree or in the measured phenotypes, 798 animals were taken into account for the following steps. The means and the standard deviations of the phenotypic traits measured in the 798 animals used for the GWAS are reported in Table 1.

Genotyping data quality control and imputation of the missing genotypes

DNA was extracted from the blood samples. The animals were then genotyped using Illumina PorcineSNP60 v2 BeadChip (Illumina Inc., San Diego, CA, USA), which contains 61,565 SNP markers distributed across the whole genome (Ramos *et al*., 2009). Quality control of the high-density SNP data was carried out on PLINK (Purcell *et al*., 2007): SNPs with more than 10% of missing genotypes, minor allele frequencies below 0.01, or deviations from Hardy-Weinberg equilibrium with p-value below 0.001 were filtered out. After the quality control the data set included 49,662 markers. All individuals had a call rate greater than 0.90 and passed the quality control. Markers were then mapped using the pig

genome assembly *Sus scrofa* build 10.2, and the unmapped SNPs were excluded. Markers located on sexual chromosomes were also excluded from the study. Finally, missing genotypes were imputed using Beagle version 3.3.2 (Browning and Browning, 2009). The final number of SNPs included in the GWAS was 45,704.

Genome-wide association study

Before performing the GWAS, the phenotypes of the 798 animals were adjusted for carcass weight (as a covariate), slaughter day (27 slaughtering batches), sex (castrate or female), animal (using a pedigree with 2,301 individuals) and litter effects (393 litters) using an animal model.

The associations between the genotypes and the adjusted phenotypes were assessed using the Bayes B approach as implemented in the GenSel software (Fernando and Garrick, 2014). The model was the following:

$$\boldsymbol{y} = \sum_{i=1}^{k} \boldsymbol{z}_i \alpha_i \delta_i + \boldsymbol{e}$$

where **y** was the vector of the adjusted phenotypes for each trait; $\mathbf{z}_i$ was the vector of the coded genotypes for a SNP at locus *i* (*i* = 1 to *k*, where *k* is the number of SNPs); α_i was the effect of the allele substitution of the SNP at locus *i*; δ_i was a random 0/1 variable representing the absence (0) or the presence (1) of a SNP *i* in the model for a certain iteration of the Markov chain Monte Carlo procedure; and **e** was the vector of random residuals normally distributed. Alternate homozygous genotypes were coded as -10 and 10, and heterozygotes as 0. The prior probability (π) that the SNPs had no effect ($\delta_i = 0$) on the adjusted phenotypes was fixed at $\pi = 0.985$, and consequently the prior probability of the markers having an effect on the adjusted phenotypes ($\delta_i = 1$) was $1–\pi = 0.015$. Thus, the model fitted approximately 745 SNPs in each iteration. A total of 500,000 iterations were run, with a burn-in of 100,000.

A Bayes Factor was calculated for each locus *i* (BF_i) to evaluate the statistical relevance of the association between each SNP and the adjusted phenotypes. The Bayes Factor was calculated as follows:

$$BF_i = \frac{\frac{\widehat{P}_i}{1-\widehat{P}_i}}{\frac{1-\pi}{\pi}}$$

where $\widehat{P}_i$ is the posterior probability of a marker *i* of being included in the model at a given iteration of the Markov chain Monte Carlo procedure. Generally, for BF above 3.2 the marker association with the trait is considered substantial, strong for BF between 10 and 100, and decisive for BF > 100 (Kass and Raftery, 1995).

Furthermore, for each trait we predicted the collective genetic variance of the SNPs included in consecutive non-overlapping 1-Mb windows based on the markers position in Sus scrofa assembly build 10.2. This approach permitted taking into consideration the combined effects of SNPs which are closely located and could be in linkage disequilibrium (LD). Contiguous 1-Mb windows that explained at least 0.5% of the total genetic variance each were merged and considered together, as in Ros-Freixedes et al. (2016). This approach permitted to take into account also the LD between markers placed in neighbouring regions or spanning more than 1 Mb. LD in candidate regions was evaluated using Haploview software (Barrett et al., 2005).

Functional characterization of the genes mapped in the most relevant regions

Candidate genes in the most associated regions were identified through Ensembl (EMBL-EBI), using the BioMart tool (Guberman et al., 2011) (url: http://www.ensembl.org/biomart/), and posteriorly their functional gene annotation was analysed using Enrichr (Kuleshov et al., 2016) (url: http://amp.pharm.mssm.edu/Enrichr/), with the aim of identifying pathways of genes involved in the genetic determinism of the studied traits. Among the results obtained from Enrichr, the pathways from both Reactome Pathway Database (Febregat et al., 2016) and KEGG PATHWAY database (Kanehisa et al., 2016) were taken into account, with the aim of obtaining more complete information about the pathways related to the most associated regions. One of the regions associated with backfat FA composition comprised microRNAs. Aiming to identify if there were genes involved in adipogenesis and FAs biosynthesis among their target genes, their corresponding

hsa-miRNAs were used on PicTar (url: http://pictar.mdc-berlin.de/), utilising the algorithm for the predictions in vertebrates (Krek et al., 2005).

The obtained associated regions were also compared with the information about the known QTLs reported on QTLdb (Hu et al., 2016) (url: http://www.animalgenome.org/cgi-bin/QTLdb/index).

Estimation of the genotypic effects for the most relevant markers

For the most interesting genes according to the GWAS results and the functional characterization analysis, we selected a tag SNP in order to further evaluate its effect. Estimated means and differences between genotypes were assessed using a model that also included carcass weight (as a covariate), slaughter day, sex, and litter. The program Rabbit (Rabbit programme, 2012) was used.

Results and discussion

GWA studies have become a commonly used analysis tool in genomics, allowing the identification of new regions and markers on a wider level compared with single marker association studies. With the spreading use of GWAS, some issues regarding the most appropriate statistical approach to be used were raised. The use of Bayesian models permits to overcome some of the limits highlighted by the application of linear models to GWA studies (Guo et al., 2016), but on the other side the spread of distinct methods has complicated the comparison of results from different GWA studies. The use of a Bayesian approach for the present study permitted to obtain more than 30 different genomic regions associated with backfat composition and thickness, some of which were consistent with other regions reported in the literature. On the whole, the majority of the regions identified in this study were located near or inside QTLs already reported in literature, demonstrating the consistency of the present results. In particular, among all the 30 regions found herein associated with backfat FA composition, approximately 15 were consistent with (or directly related to) QTLs found in previous studies, 6 regions were new, showing no corresponding QTLs on databases, while the remaining were located in QTLs associated with traits and FA

other than the ones herein considered (Supplementary Table S1). Additionally, all the 4 regions associated with backfat thickness in the present study were in agreement with literature.

The markers most associated with the studied traits are reported in Table 2 and Supplementary TableS 2. As visible from Supplementary Table S2 almost half of the markers that were found associated with backfat FA composition were mapped on chromosomes 7 and 8 (8 markers were located on SSC7 and 9 SNPs on SSC8), but other relevant SNPs lied on SSC1, SSC5, SSC9, SSC10, SSC11, SSC14, SSC16, SSC17 and SSC18. In particular, the region included between 119 and 122 Mb on SSC8 consistently explained a great proportion of the genetic variance associated to medium chain fatty acids (MCFA) and long chain fatty acids (LCFA) (about the 70% of the genetic variance), and a smaller part of the genetic variance associated to palmitic, palmitoleic and oleic acids (10%, 3% and 1.78%, respectively) (Table 2). This region harboured three of the markers most associated with backfat FA composition (H3GA0025321, SIRI0000509 and INRA0030422). Among them, the SNP that showed the highest Bayes Factor (BF_i) was H3GA0025321, which was strongly associated with MCFA (BF_i > 1000), with LCFA (BF_i > 1000), and also with palmitic acid (BF_i > 166) (Table 2). The same marker resulted to have a relevant effect also on palmitoleic and oleic acids (BF_i = 12.02 and BF_i = 10.05, respectively). Although the effects of this marker on MCFA and LCFA could be overestimated due to the phenomenon known as Beavis effect (Xu, 2003), it is worth noting that H3GA0025321 has already been identified in other association studies performed on different pig breeds (Corominas et al., 2013; Zhang et al., 2016). The H3GA0025321 marker has already been identified in a GWA study assessed by Corominas et al. (2013) on Iberian x Landrace crosses, where this polymorphism was found strongly associated with the palmitic and palmitoleic acid content in the intramuscular fat of Longissimus dorsi. Furthermore, in agreement with Corominas et al. (2013), we have found that, together with H3GA0025321, SIRI0000509 and INRA0030422 markers also were significantly associated with palmitic acid. In literature, these three markers were found to be in LD with mutations located in the promoter sequence of the nearby gene *ELOVL elongase 6* (Corominas et al., 2013). This gene codes for an elongase, which catalyses the first and rate-limiting reaction of the four that constitute the long-chain fatty acids elongation cycle. In a recent investigation Corominas et al. (2015) suggested that the causal mutation responsible for the strong association found

for SSC8 119-120 Mb region could be *ELOVL6:c.-394 G>A,* a SNP located in *ELOVL6* gene promoter region. Additionally, the three identified markers (H3GA0025321, SIRI0000509 and INRA0030422) and *ELOVL6* gene mapped in the same SSC8 region where QTLs associated with palmitoleic (Muñoz et al., 2013), palmitic and oleic acid contents (Revilla et al., 2014) have already been identified. The associations observed in the present study for the three markers are completely in agreement with the *ELOVL6* elongase critical role in regulating the length of FA chain: H3GA0025321 C allele was strongly related to higher amounts of stearic, oleic and LCFA, causing at the same time a consistent decrease in MCFA and C16 acids (palmitic and palmitoleic) (Table 3).

In addition to the region where *ELOVL6* gene is located, other regions were associated with the oleic acid content in backfat. In particular, the region included between 116 and 124 Mb on SSC14, which is involved in desaturation, was found to be consistently associated with stearic acid (1.18% of the trait genetic variance), oleic acid (1.23%), monounsaturated fatty acids (MUFA) (1.26%) and UFA content (0.82% of the trait genetic variance) (Table 2). MARC0006531 was the marker showing the highest Bayes Factors for this SSC14 region, and individuals displaying the AA genotype for this *locus* resulted to be more prone towards backfat MUFA and oleic acid deposition (Table 4). The region included between 121 and 123 Mb on SSC14 that was found in the present study to be associated with MUFA and Oleic acid ontent in backfat, in line with previous studies (Ros-Freixedes et al., 2016). Two candidate genes are currently annotated in this region: *stearoyl-CoA desaturase delta 9* (*SCD*) and *ELOVL elongase3* (*ELOVL3*). Ros-Freixedes et al. (2016) pointed out that the signal detected at 123 Mb (where *ELOVL3* gene lies) may be the result of a long LD block containing the markers in the downstream region of *SCD* gene. The results obtained from the present research showed that this region on SSC14 is not only important in Duroc, but also in Large White.

Additionally, three different SSC5 regions were related to porcine backfat FA composition, although the amount of the genetic variance explained by these regions was quite limited. The window located at 65-66 Mb was associated to linolenic acid (explaining the 0.48% of the trait genetic variance), the region ranging from 70 to 75 Mb explained the highest genetic variance for backfat arachidic acid content (2.57%), and the window at 104-105 Mb was linked to linoleic, PUFA and PUFAΩ6 amounts (0.59%, 0.59% and 0.61% of the traits

genetic variance, respectively) (Table 2 and Supplementary Table S3). Interestingly, comparing Table 2 and Supplementary Table S3 results revealed that the genomic regions associated to short chain fatty acids (SCFA), arachidic, arachidonic, PUFA and PUFAΩ6 backfat contents were other than those linked to palmitic, palmitoleic, stearic, oleic and MUFA backfat contents. In particular, this difference was also reflected in the results obtained from the functional characterization of the associated regions (Supplementary Table S4). On one hand, the regions associated to backfat contents of MUFA, UFA, MCFA, LCFA, palmitic, palmitoleic and oleic acids clustered together and resulted to harbour mainly genes related to FA chain elongation, desaturation and *de novo* biosynthesis. Indeed, the genes mapped inside these regions were involved in the "Regulation of cholesterol biosynthesis by SREBP" (*SEC24 Homolog B, COPII Coat Complex Component- SEC24B*, *Insulin Induced Gene 1- INSIG1- and ELOVL6 gene*), the "Synthesis of very-long fatty acyl-CoAs" (*ELOVL6* and *Very Long-Chain Acyl-CoA Synthetase- ACSBG1*), the "Metabolism of lipids and lipoproteins" (*SEC24B, ELOVL6, Phospholipase A2, Group XIIA- PLA2G12A*- and *Ethanolamine-Phosphate Phospho-Lyase- ETNPPL*), the "Biosynthesis of unsaturated fatty acids" (*SCD and ELOVL6* genes) and "Fatty acid elongation" (*ELOVL3* and *ELOVL6*) (Supplementary Table S4).

On the other hand, the genomic regions associated with backfat arachidic, arachidonic acid, PUFA, PUFAΩ6 and SCFA showed to harbour mainly genes involved in dietary lipids digestion, mobilisation and transport, in pancreatic secretion, in carbohydrates metabolism and in the pathway regulating gastrin secretion (Table 2 and Supplementary Table S4). Among these genes, *Amyloid beta precursor protein binding family B member 1 interacting protein* (*APBB1IP*) was located on SSC10 and on SSC14 were located *Phospholipase A2 group III* (*PLA2G3*), *Pancreatic Lipase* (*PNLIP*), *Pancreatic Lipase-Related Protein 1* (*PNLIPRP1*) and *Pancreatic Lipase-Related Protein 2* (*PNLIPRP2*). The gene *APBB1IP* (also called *Rap1-GTP-Interacting Adaptor Molecule- RIAM*) is known for its role in the regulation of gastrin gene transcription in pancreatic islet cells (Simon et al., 1994). *APBB1IP* plays an essential role in mammalian energy metabolism, as *APBB1IP* knockout mice were found to show severe conditions of obesity, glucose intolerance, insulin resistance, liver steatosis and adipose tissue hypertrophy (Yeung et al., 2013). The region at 138-142 Mb on SSC14 associated to arachidic acid harbours the genes encoding for pancreatic lipases (*PNLIP*, *PNLIPRP1* and

PNLIPRP2). These pancreatic-secreted lipases have the essential role of hydrolysing triglycerides and making the dietary fat digestion efficient (Berton et al., 2009). The implication of these lipases in fat deposition was unknown a few years ago; then, in the early 2000s the administration of a pancreatic lipases inhibitor (Orlistat) was noticed to positively affect obesity in treated patients (Finer et al., 2000). Since that moment, the activity of pancreatic lipases and their genes have been studied in medicine for their effect on obesity and diabetes (Drew et al., 2007), and a mutation in *PNLIP* gene sequence was related to human congenital pancreatic lipase deficiency and fat malabsorption (Behar et al., 2014). Hence, mutations in the coding sequence of pancreatic lipase genes may be responsible for different lipids absorption in pigs, influencing also backfat fatty acid synthesis and deposition. The results of the present study suggest that *APBB1IP*, *PLA2G3, PNLIP*, *PNLIPRP1* and *PNLIPRP2* may influence backfat FA composition and represent new candidate genes for backfat FA composition. Interestingly, taken together, the associations found for the different FA are consistent with what is known in literature about FA metabolism in mammals. While palmitic, stearic, SFA and MUFA are known to be mainly *de novo* synthesised, linoleic and α-linoleic acids are essential FA that must be provided by the diet, as mammals are known to be unable to endogenously synthesise them.

Finally, it is worth noting that the windows of markers relevantly associated with backfat thickness harbour genes affecting miRNAs biogenesis (*P*-value = 2.38E-02, with the gene *Dicer 1, Ribonuclease Type III* (*DICER1*)) and hemostasis (*P*-value = 1.49E-02) (Supplementary Table S4). Among the genomic windows associated to backfat thickness, the chromosomal region included between 64 and 67 Mb on SSC11 to date is known to harbour exclusively the sequences of microRNAs taking part in the miR-17-92 cluster. This cluster is composed of ssc-mir-20a, ssc-mir-19a, ssc-mir-18a, ssc-mir-19b-1, ssc-mir-17 and ssc-mir-92a-1, which are known for promoting adipocyte differentiation in mouse preadipocytes by targeting *Rb2/p130* gene (Wang et al., 2008) and have predicted binding sites on several genes related to adipocyte differentiation and lipids metabolism (Table 5). In an unpublished study carried out on porcine backfat tissue transcriptome, we found that microRNAs of the miR-17-92 cluster are expressed in pig adipose tissue (unpublished data). These results suggest that this cluster may have an important role also in swine adipocytes differentiation. The association found between the region harbouring these microRNAs and backfat thickness may indicate

that mutations in the sequence of miR-17-92 cluster could alter their affinity towards the target genes, affecting preadipocytes differentiation and backfat thickness. Further studies are needed to support this hypothesis and to prove that the markers in this region can be related to changes in swine differentiating preadipocytes.

To our knowledge, this is the first study performed in Large White pigs identifying markers and genomic regions associated with backfat fatty acid composition. The results of the present research strengthen the hypothesis that both *ELOVL6* and *SCD* genes play an important role in determining pig backfat fatty acid composition, but emphasise also the need to deepen the knowledge of genes involved in feed digestion, as they too may affect backfat fatty acid composition through changes in lipids absorption efficiency.

Aknowledgements

We thank Dr. Luca Buttazzoni and Dr. Maurizio Gallo from Associazione Nazionale Allevatori Suini (ANAS) for providing the porcine backfat and *Semimembranosus* tissue samples and the information about the studied individuals. This work was supported by Progetto "AGER - Agroalimentare e ricerca": Advanced research in genomics and processing technologies for the Italian heavy pig production – Hepiget (Grant N. 2011- 0279) and by PRIN 2015.

References

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

Behar DM, Basel-Vanagaite L, Glaser F, Kaplan M, Tzur S, Magal N, Eidlitz-Markus T, Haimi-Cohen Y, Sarig G, Bormans C, Shohat M & Zeharia A (2014) Identification of a novel mutation in the PNLIP gene in two brothers with congenital pancreatic lipase deficiency. *J Lipid Res* 55:307–312.

Berton A, Sebban-Kreuzer C, Rouvellac S, Lopez C & Crenon I (2009) Individual and combined action of pancreatic lipase and pancreatic lipase-related proteins 1 and 2 on native versus homogenized milk fat globules. *Mol Nutr Food Res* 53:1592–1602.

Bosi P, Cacciavillani JA, Casini L, Lo Fiego DP, Marchetti M & Mattuzzi S (2000) Effects of dietary high-oleic acid sunflower oil, copper and vitamin E levels on the fatty acid composition and the quality of dry cured Parma ham. *Meat Sci* 54:119–126.

Browning BL & Browning SR (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* 84:210–223.

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

Corominas J, Marchesi JA, Puig-Oliveras A, Revilla M, Estellé J, Alves E, Folch JM & Ballester M (2015) Epigenetic regulation of the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. *Gen Sel Evol* 47:20.

Davoli R, Luise D, Mingazzini V, Zambonelli P, Braglia S, Serra A & Russo V (2015) Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *J Anim Breed Gen* 133:277-82.

Denny JC, Bastarache L & Roden DM (2016) Phenome-Wide Association Studies as a Tool to Advance Precision Medicine. *Ann Rev Genom Hum Genet* 17:353-373.

Drew BS, Dixon AF & Dixon JB (2007) Obesity management: Update on orlistat. *Vasc Health Risk Manag* 3:817–821.

Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, Matthews L, May B, Milacic M, Rothfels K, Shamovsky V, Webber M, Weiser J, Williams M, Wu G, Stein L, Hermjakob H & D'Eustachio P (2016) The Reactome pathway Knowledgebase. *Nucl Acids Res* 44**:**D481-487.

Finer N, James WP, Kopelman PG, Lean ME & Williams G (2000) One-year treatment of obesity: a randomized, double-blind, placebo-controlled, multicentre study of orlistat, a gastrointestinal lipase inhibitor. *International Int J Obes Relat Metab Disord.* 24:306–313.

Guberman JM, Ai J, Arnaiz O, Baran J, Blake A, Baldock R, Chelala C, Croft D, Cros A, Cutts RJ et al. (2011) BioMart Central Portal: an open database network for the biological community. *Database* 2011:bar041.

Guo X, Su G, Christensen OF, Janss L & Lund MS (2016) Genome-wide association analyses using a Bayesian approach for litter size and piglet mortality in Danish Landrace and Yorkshire pigs. *BMC Genomics* 17:468.

Hu ZL, Park CA & Reecy JM (2016) Developmental progress and current status of the Animal QTLdb. *Nucl Acids Res* 44:D827–D833.

Kanehisa M, Sato Y, Kawashima M, Furumichi M & Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucl Acids Res* 44:D457-462.

Kass RE & Raftery AE (1995) Bayes Factors. *J Am Stat Ass* **90**:73–795.

Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M & Rajewsky N (2005) Combinatorial microRNA target predictions. *Nat Genet* 37**:**495–500.

Kris-Etherton PM & Yu S (1997) Individual fatty acid effects on plasma lipids and lipoproteins: human studies. *Am J Clin Nutr* 65:1628S–1644S.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW & Ma'ayan A (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucl Acids Res* 44:W90-97.

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

Murrieta CM, Hess BW & Rule DC (2003) Comparison of acidic and alkaline catalysts for preparation of fatty acid methyl esters from ovine muscle with emphasis on conjugated linoleic acid. *Meat Sci* 65:523–529.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.

Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C et al. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLOS ONE* 4:e6524.

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Revilla M, Ramayo-Caldas Y, Castelló A, Corominas J, Puig-Oliveras A, Ibáñez-Escriche N, Muñoz M, Ballester M & Folch JM (2014) New insight into the SSC8 genetic determination of fatty acid composition in pigs. *Genet Sel Evol* 46:28.

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

Shackelford SD, Miller MF, Haydon KD, Lovegren NV, Lyon CE & Reagan JO (1990) Acceptability of Bacon as influenced by the Feeding of Elevated Levels of Monounsaturated Fats to Growing-finishing Swine. *J Food Sci* 55:621–624.

Sharma A, Lee JS, Dang CG, Sudrajad P, Kim HC, Yeon SH, Kang HS and Lee SH (2015) Stories and Challenges of Genome Wide Association Studies in Livestock — A Review. *Asian Australas J Anim Sci* 28:1371–1379.

Simon B, Tillotson L & Brand SJ (1994) Activation of gastrin gene transcription in islet cells by a RAP1-like cis-acting promoter element. *FEBS Letters* 351:340–344.

Wang Q, Li YC, Wang J, Kong J, Qi Y, Quigg RJ & Li X (2008) miR-17-92 cluster accelerates adipocyte differentiation by negatively regulating tumor-suppressor Rb2/p130. *Proc Natl Acad Sci* 105:2889–2894.

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

Xu S (2003) Theoretical Basis of the Beavis Effect. *Genetics* 165:2259–2268.

Yeung F, Ramírez CM, Mateos-Gomez PA, Pinzaru A, Ceccarini G, Kabir S, Fernández-Hernando C & Sfeir A (2013) Nontelomeric Role for Rap1 in Regulating Metabolism and Protecting against Obesity. *Cell Rep* 3:1847–1856.

Zhang W, Zhang J, Cui L, Ma J, Chen C, Ai H, Xie X, Li L, Xiao S, Huang L, Ren J & Yang B (2016) Genetic architecture of fatty acid composition in the longissimus dorsi muscle revealed by genome-wide association studies on diverse pig populations. *Genet Sel Evol* 48:5.

Table 1. Means and standard deviations of the measured traits in the study population.

Traits	Mean	SD
Palmitic acid (C16)[1]	22.37	1.28
Palmitoleic acid (C16:1, cis-9) [1]	1.43	0.25
Stearic acid (C18) [1]	13.29	1.76
Oleic acid (C18:1, cis-9) [1]	38.43	1.59
Linoleic acid (C18:2, cis-9, 12) [1]	16.28	2.00
Linolenic acid (C18:3Ω3) [1]	0.74	0.17
Arachidic acid (C20) [1]	0.18	0.04
Arachidonic acid (C20:4Ω6) [1]	0.23	0.05
Docosapentaenoic acid (C22:5Ω3) [1]	0.05	0.01
Docosahexaenoic acid (C22:6Ω3) [1]	0.01	0.01
Saturated Fatty Acids (SFA) [1]	37.52	2.60
Monounsaturated Fatty Acids (MUFA) [1]	43.70	1.84
Polyunsaturated Fatty Acids (PUFA) [1]	18.34	2.16
Unsaturated Fatty Acids (UFA) [1]	62.04	2.55
Polyunsaturated Fatty Acids omega 6 (PUFAΩ6) [1]	16.39	2.01
Polyunsaturated Fatty Acids omega 3 (PUFAΩ3) [1]	0.81	0.17
Short Chain Fatty Acids (SCFA) [1]	0.07	0.02
Medium Chain Fatty Acids (MCFA) [1]	25.86	1.38
Long Chain Fatty Acids (LCFA) [1]	73.63	1.41
Backfat thickness (mm)	26.52	5.06
Carcass weight (kg)	118.79	8.57

[1] All the fatty acids or fatty acid categories are expressed as percentage on the total fatty acids amount.

Table 2. List of the 1-Mb windows explaining over 0.5% of the genetic variance associated with pigs backfat fatty acid composition and backfat thickness. In the table are reported also the extended regions, composed of at least two contiguous windows associated with the same trait, and whose explained genetic variances were summed. The top SNP and the candidate genes lying in the associated regions are listed in the right part of the table.

Traits	SSC	Region[1] (Mb)	Genetic Variance[2]	$\hat{P}_1$[3]	Extended region[4]		Top SNP	SNP Bayes Factor	Candidate gene(s)
					Region (Mb)	Genetic Variance[5]			
Palmitic acid (C16)	8	119-120	9.2	0.99	119-122	10.19	H3GA0025321	166.13	*ELOVL6*
	8	9-10	0.68	0.58	8-10	1.15	SIRI0000509	23.01	
	7	52-53	0.59	0.57					*ACSBG1; IDH3A*
Palmitoleic acid (C16:1, cis-9)	8	119-120	1.27	0.37	119-122	3.01	H3GA0025321	12.02	*ELOVL6*
Stearic acid (C18)	1	77-78	0.79	0.43	75-80	2			
	14	117-118	0.68	0.54	116-118	1.18	MARC0006531	11.71	
Oleic acid (C18:1, cis-9)	8	119-120	1.78	0.74			INRA0030422	28.15	*ELOVL6*
	14	122-123	0.45	0.51	121-124	1.23			*SCD; ELOVL3*
Linoleic acid (C18:2, cis-9, 12)	5	104-105	0.97	0.59			DRGA0006379	28.94	
	10	55-56	0.52	0.45			H3GA0030320	13.24	*APBB1IP*
	7	57-58	0.39	0.52	56-58	0.75			
Linolenic acid	17	17-18	1.2	0.52			MARC0013253	19.38	

(C18:3Ω3)									
	9	26-27	0.78	0.59			H3GA0026788	10.76	
	5	65-66	0.67	0.48					
Arachidic acid (C20)	5	72-73	0.9	0.38	70-75	2.57	MARC0070351	7.39	*ADIPOR2*
	14	138-139	0.48	0.34	138-142	1.13			*PNLIPRP1; PNLIPRP2; PNLIP; PNLIPRP3*
Arachidonic acid (C20:4Ω6)	7	53-54	0.72	0.48					*ACSBG1; IDH3A*
	9	22-23	0.65	0.42					*ME3*
Docosahexaenoic acid (C22:6Ω3)	14	50-51	0.4	0.16	47-52	1.65			*MTMR3; INPP5J; PLA2G3; PISD*
Saturated Fatty Acids (SFA)	7	53-54	0.97	0.68	51-54	1.58	ASGA0033717	12.90	*ACSBG1; IDH3A*
	1	77-78	0.74	0.50	75-80	2.04			
Monounsaturated Fatty Acids (MUFA)	7	56-57	0.51	0.62					
	14	122-123	0.5	0.53	121-124	1.26			*SCD; ELOVL3*
Polyunsaturated Fatty Acids (PUFA)	5	104-105	0.95	0.59					
	10	55-56	0.53	0.45			H3GA0030320	10.51	*APBB1IP*
Unsaturated Fatty	7	53-54	0.91	0.68	51-54	1.42	ASGA0033717	11.82	*ACSBG1; IDH3A*

Acids (UFA)									
	1	77-78	0.73	0.49	75-80	1.98			
	1	277-278	0.59	0.39	277-280	0.87	MARC0053473	17.67	
	14	117-118	0.49	0.56	116-118	0.82			
Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	5	104-105	0.99	0.61					
	10	55-56	0.5	0.44			H3GA0030320	10.51	*APBB1IP*
Polyunsaturated Fatty Acids omega 3 (PUFAΩ3)	14	50-51	0.43	0.22	48-52	1.64			*MTMR3; INPP5J; PLA2G3; PISD*
Short Chain Fatty Acids (SCFA)	14	50-52	0.57	0.19	46-52	2.15			*MTMR3; INPP5J; PLA2G3; PISD*
Medium Chain Fatty Acids (MCFA)	8	119-120	70.24	1.00			H3GA0025321	>1000	*ELOVL6*
Long Chain Fatty Acids (LCFA)	8	119-120	70.38	1.00			H3GA0025321	>1000	*ELOVL6*
Backfat thickness	7	15-16	1.38	0.63	10-12; 15-16	2.02	H3GA0020080	42.27	
	11	65-66	17	0.53	0.39				miR-17-92 cluster
	7	122-123	0.91	0.60			ALGA0045097	17.67	*DICER1*

[1] Windows positions referred to *Sus scrofa* assembly Build 10.2, expressed in Mb.

[2] Proportion of the collective genetic variance explained by the markers in the window on the total genetic variance of the trait.

[3] Posterior probability of inclusion in the model for the markers in the window.

[4] Extended region comprising contiguous windows explaining at least 0.5% of the genetic variance.

[5] The proportion of genetic variance explained by the markers placed in the neighbouring windows of the extended regions.

Docosapentaenoic acid (C22:5Ω3) is not reported in the table because no marker window explained at least 0.5% of the genetic variance.

Table 3. Estimated differences for the different genotypes of the marker H3GA0025321 and probability that the difference is lower than zero (P(<0)) for the traits resulted to be associated with the marker from the GWA study.

Traits	AA- AC		AC- CC		AA- CC	
	Mean	P(<0)	Mean	P(<0)	Mean	P(<0)
Palmitic acid (C16)	0.86	0.00	0.11	0.36	0.97	0.00
Stearic acid (C18)	-0.27	0.96	-0.78	0.95	-1.06	0.99
Palmitoleic acid (C16:1, cis-9)	0.14	0.00	0.06	0.18	0.20	0.00
Oleic acid (C18:1, cis-9)	-0.52	1.00	0.33	0.20	-0.18	0.68
Medium chain fatty acids (MCFA)	1.04	0.00	0.11	0.35	1.14	0.00
Long chain fatty acids (LCFA)	-1.00	1.00	-0.24	0.78	-1.24	1.00

Table 4. Estimated differences for the different genotypes of the marker MARC0006531 and probability that the difference is lower than zero (P(<0)) for the traits resulted to be associated with the marker from the GWA study.

Traits	AA- AC		AC- CC		AA- CC	
	Mean	P(<0)	Mean	P(<0)	Mean	P(<0)
Stearic acid (C18)	-0.37	0.99	-0.57	1.00	-0.94	1.00
Oleic acid (C18:1, cis-9)	0.13	0.16	0.30	0.01	0.43	0.00
Saturated fatty acids (SFA)	-0.43	0.98	-0.67	1.00	-1.09	1.00
Monounsaturated fatty acids (MUFA)	0.18	0.10	0.48	0.00	0.67	0.00

Table 5. List of the miRNAs mapped in the region on chromosome 11 associated with backfat thickness.

Region on SSC11 included between 64 and 67 Mb associated with backfat thickness				
miRNAs	Ensembl Gene ID	Ensembl Transcript ID	miRNA predicted target genes related to lipid metabolism	PicTar score
ssc-mir-20a	ENSSSCG00000019883	ENSSSCT00000021478	*Very low density lipoprotein receptor* (*VLDLR*)	7.35
			Monoglyceride lipase (*MGLL*)	2.58
			Factor for adipocyte differentiation 104 (*FAD104*)	1.14
ssc-mir-19a	ENSSSCG00000019897	ENSSSCT00000021492	*Acyl-CoA synthetase long-chain family member 4* (*ACSL4*)	10.95
			Low density lipoprotein-related protein 2 (*LRP2*)	7.44
			Lysocardiolipin acyltransferase (*LYCAT*)	1.62
[#]ssc-mir-18a	ENSSSCG00000019535	ENSSSCT00000021130	*Estrogen receptor 1* (*ESR1*)	6.03
			Prostaglandin F2 receptor negative regulator (*PTGFRN*)	4.18
[§]ssc-mir-19b-1	ENSSSCG00000019902	ENSSSCT00000021497	*Low density lipoprotein-related protein 2* (*LRP2*)	7.44
			Adiponectin receptor 2 (*ADIPOR2*)	2.61
			Lysocardiolipin acyltransferase (*LYCAT*)	1.62
			Factor for adipocyte differentiation 104 (*FAD104*)	1.49
*ssc-mir-17	ENSSSCG00000018617	ENSSSCT00000020212	*Very low density lipoprotein receptor* (*VLDLR*)	7.35
			Factor for adipocyte differentiation 104 (*FAD104*)	1.26
[+]ssc-mir-92a-1	ENSSSCG00000018681	ENSSSCT00000020276	*Factor for adipocyte differentiation 104* (*FAD104*)	4.83

	Monoglyceride lipase (*MGLL*)	3.53

All the predicted sites were obtained using the corresponding hsa-mir, and using on PicTar the algorithm for the predictions in vertebrates and choosing the dataset “target predictions for all human microRNAs based on conservation in mammals”.

[#] Results are referred to hsa-mir-18.

[§] Results are referred to hsa-mir-19b.

* Results are referred to hsa-mir-17-5p, previously known as hsa-mir-17.

[+] Results are referred to hsa-mir-92.

Supplementary Table S1. Comparison between the regions obtained from the present study and the QTLs reported in literature for the same regions or in nearby chromosomal regions.

GWA results			QTLs in literature[1]		
SSC[2]	Region (Mb)	Trait	Trait	Region (Mb)	Literature
1	18-19[§]	Linolenic acid (C18:3Ω3)	-	-	-
1	75-80	Stearic acid (C18), Saturated Fatty Acids (SFA), Unsaturated Fatty Acids (UFA)	Saturated fatty acid content	81.4	Ramayo-Caldas et al., 2012
1	277-280	Unsaturated Fatty Acids (UFA), Palmitic acid (C16)	Oleic acid content	280.5-280.6	Sanchez et al., 2007
2	107-108[§]	Palmitoleic acid (C16:1, cis-9)	-	-	-
3	32-33[§]	Linolenic acid (C18:3Ω3)	-	-	-
5	65-66[§]	Linolenic acid (C18:3Ω3)	-	-	-
5	70-75	Arachidic acid (C20)	Arachidic acid content	69; 77	Yang et al., 2013
5	104-105	Linoleic acid (C18:2, cis-9, 12), Polyunsaturated Fatty Acids (PUFA), Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	Average Feeding rate	107.5	Do et al., 2013
7	5-6	Stearic acid (C18), Unsaturated Fatty Acids (UFA)	Linoleic acid content	8.8-13.8	Kim et al., 2006
7	10-16	Backfat thickness	Average backfat	9.7-14.8	Paszek et al., 2001

			thickness		
7	51-54	Palmitic acid (C16), Stearic acid (C18), Arachidonic acid (C20:4Ω6), Saturated Fatty Acids (SFA), Unsaturated Fatty Acids (UFA)	Arachidonic acid content; Cis-11-Eicosenoic acid content	50.8; 52.2	Yang et al., 2013
7	56-58	Monounsaturated Fatty Acids (MUFA), Linoleic acid (C18:2, cis-9, 12)	Linoleic acid content	49.2-49.4	Guo et al., 2009
7	122-123	Backfat thickness	Average backfat thickness	106.8-120.7	Kim et al., 2005
8	8-10	Palmitic acid (C16), Unsaturated Fatty Acids (UFA)	Palmitic acid content	0.8-6.7	Uemoto et al., 2012
8	33-34	Stearic acid (C18)	Average backfat thickness	33-34	Jiao et al., 2014
8	119-122	Palmitic acid (C16), Palmitoleic acid (C16:1, cis-9), Oleic acid (C18:1, cis-9), Medium Chain Fatty Acids (MCFA), Long Chain Fatty Acids (LCFA)	Palmitoleic acid content; Oleic acid content and Palmitic acid content	114.2-122.3; 117.4	Muñoz et al., 2013; Revilla et al., 2014
8	141-142	Unsaturated Fatty Acids (UFA)	Palmitoleic acid content	72-139	Clop et al., 2003
9	22-23[§]	Arachidonic acid (C20:4Ω6)	-	-	-

9	26-27	Linolenic acid (C18:3Ω3)	Palmitoleic acid content	24.6-71.4	Uemoto et al., 2012
9	28-29	Oleic acid (C18:1, cis-9), Monounsaturated Fatty Acids (MUFA)	Palmitoleic acid content	24.6-71.4	Uemoto et al., 2012
10	10-11	Palmitic acid (C16)	Palmitic acid content	5.9-13.9	Uemoto et al., 2012
10	55-56	Linoleic acid (C18:2, cis-9, 12), Polyunsaturated Fatty Acids (PUFA), Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	Oleic acid content	61.5	Sanchez et al., 2007
11	7-8	Linolenic acid (C18:3Ω3)	Linoleic acid content	7.8	Ramayo-Caldas et al., 2012
11	65-66	Backfat thickness	Backfat at rump	68.5	Fontanesi et al., 2012
13	167-174	Short Chain Fatty Acids (SCFA)	Cholesterol level	185.4-206.7	Yoo et al., 2012
14	10-11[§]	Monounsaturated Fatty Acids (MUFA)	-	-	-
14	46-52	Short Chain Fatty Acids (SCFA), Docosahexaenoic acid (C22:6Ω3), Polyunsaturated Fatty Acids omega 3 (PUFAΩ3)	Homolonolenic acid content (C20:3)	45	Yang et al., 2013
14	116-124	Stearic acid (C18), Oleic acid (C18:1, cis-9), Unsaturated Fatty Acids (UFA), Monounsaturated	Monounsaturated fatty acid content and Stearic acid	116.6-116.7; 120.2	Sanchez et al., 2007; Yang et al., 2013

		Fatty Acids (MUFA)	content; Oleic acid content		
14	138-139	Arachidic acid (C20)	Arachidic acid content	121-121.5	Zhang et al., 2016
16	0-3	Backfat thickness	Backfat above muscle dorsi	0.3-67.6	Liu et al., 2008
16	36-37	Arachidonic acid (C20:4Ω6)	Arachidic acid content	38.8-38.9	Guo et al., 2009
17	17-18	Linolenic acid (C18:3Ω3)	Eicosenoic acid to eicosanoic acid ratio	22.5	Ramayo-Caldas et al., 2012
18	3-4	Palmitic acid (C16)	Palmitic acid content	5.7-22.9	Uemoto et al., 2012
18	41-42	Oleic acid (C18:1, cis-9)	Palmitic acid content	46.7-47.7	Quintanilla et al., 2011

[1] The reported QTLs were obtained from QTLdb (url: http://www.animalgenome.org/cgi-bin/QTLdb/index).

[2] SSC stands for *Sus scrofa* chromosome.

[§] For these regions there are not QTLs located in the neighbourhood or in the same genomic region reported in literarture.

Additional references

Clop A, Ovilo C, Perez-Enciso M, Cercos A, Tomas A, Fernandez A, Coll A, Folch JM, Barragan C, Diaz I, Oliver MA, Varona L, Silio L, Sanchez A & Noguera JL (2003) Detection of QTL affecting fatty acid composition in the pig. *Mamm Genome* 14:650–656.

Fontanesi L, Schiavo G, Galimberti G, Calò DG, Scotti E, Martelli PL, Buttazzoni L, Casadio R & Russo V (2012) A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genomics* 13:583.

Guo T, Ren J, Yang K, Ma J, Zhang Z & Huang L (2009) Quantitative trait loci for fatty acid composition in longissimus dorsi and abdominal fat: results from a White Duroc x Erhualian intercross F2 population. *Anim Genet* 40:185–191.

Jiao S, Maltecca C, Gray KA & Cassady JP (2014) Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genome wide association. *J Anim Sci* 92:2846–2860.

Kim JJ, Zhao H, Thomsen H, Rothschild MF & Dekkers JCM (2005) Combined line-cross and half-sib QTL analysis of crosses between outbred lines. *Genet Res* 85:235–248.

Kim Y, Kong M, Nam YJ & Lee C (2006) A quantitative trait locus for oleic fatty acid content on Sus scrofa chromosome 7. *J Hered* 97:535–537.

Liu G, Kim JJ, Jonas E, Wimmers K, Ponsuksili S, Murani E, Phatsara C, Tholen E, Juengst H, Tesfaye D, Chen JL & Schellander K (2008) Combined line-cross and half-sib QTL analysis in Duroc-Pietrain population. *Mamm Genome* 19:429–438.

Paszek AA, Wilkie PJ, Flickinger GH, Miller LM, Louis CF, Rohrer GA, Alexander LJ, Beattie CW & Schook LB (2001) Interval mapping of carcass and meat quality traits in a divergent swine cross. *Anim Biotechnol* 12:155–165.

Quintanilla R, Pena RN, Gallardo D, Cánovas A, Ramírez O, Díaz I, Noguera JL & Amills M (2011) Porcine intramuscular fat content and composition are regulated by quantitative trait loci with muscle-specific effects. *J Anim Sci* 89:2963–2971.

Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, Díaz I, Ibáñez-Escriche N, Noguera JL, Pérez-Enciso M, Fernández AI & Folch JM (2012) Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. *J Anim Sci* 90:2883–2893.

Sanchez MP, Iannuccelli N, Basso B, Bidanel JP, Billon Y, Gandemer G, Gilbert H, Larzul C, Legault C, Riquet J, Milan D & Le Roy P (2007) Identification of QTL with effects on intramuscular fat content and fatty acid composition in a Duroc x Large White cross. *BMC Genetics* 8:55.

Uemoto Y, Soma Y, Sato S, Ishida M, Shibata T, Kadowaki H, Kobayashi E & Suzuki K (2012) Genome-wide mapping for fatty acid composition and melting point of fat in a purebred Duroc pig population. *Anim Genet* 43:27–34.

Yoo CK, Cho IC, Lee JB, Jung EJ, Lim HT, Han SH, Lee SS, Ko MS, Kang T, Hwang JH, Park YS & Park HB (2012) QTL analysis of clinical-chemical traits in an F_2 intercross between Landrace and Korean native pigs. *Physiol Genom* 44:657–668.

Supplementary Table S2. Markers most relevantly associated with pigs backfat fatty acids composition listed on the basis of their genomic position.

SSC	Nucleotide Position[1]	Marker	Trait	Genetic Variance[2]	$\widehat{P}_I$[3]	BF_i[4]
1	18,577,487	ASGA0101644	Linolenic acid (C18:3Ω3)	1.21E-05	0.21	17.24
1	277,975,238	MARC0053473	Saturated Fatty Acids (SFA)	2.45E-03	0.23	19.36
			Unsaturated Fatty Acids (UFA)	1.96E-03	0.21	17.67
5	65,687,961	MARC0035523	Linolenic acid (C18:3Ω3)	4.29E-06	0.13	10.13
5	71,056,487	MARC0070351	Arachidic acid (C20)	3.24E-07	0.10	7.39
5	104,959,696	DRGA0006379	Linoleic acid (C18:2, cis-9, 12)	4.18E-03	0.35	28.94
			Polyunsaturated Fatty Acids (PUFA)	4.23E-03	0.33	26.99
			Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	4.08E-03	0.34	28.34
7	5,620,893	ALGA0038213	Stearic acid (C18)	6.66E-04	0.20	16.32
			Saturated Fatty Acids (SFA)	2.27E-03	0.24	20.61
			Unsaturated Fatty Acids (UFA)	2.49E-03	0.25	21.87
7	10,506,639	ALGA0038542	Backfat thickness	2.05E-03	0.14	10.56
7	11,385,909	ASGA0031202	Backfat thickness	2.83E-03	0.16	12.35

7	15,108,730	H3GA0020080	Backfat thickness	3.41E-02	0.39	42.27
7	51,426,179	ALGA0041370	Saturated Fatty Acids (SFA)	8.84E-04	0.15	11.74
7	53,685,735	ASGA0033717	Stearic acid (C18)	3.67E-04	0.14	10.83
			Saturated Fatty Acids (SFA)	1.29E-03	0.16	12.90
			Unsaturated Fatty Acids (UFA)	9.76E-04	0.15	11.82
7	122,615,348	ALGA0045097	Backfat thickness	6.58E-03	0.21	17.67
7	122,628,410	ASGA0036491	Backfat thickness	5.27E-03	0.20	16.07
8	8,826,922	ASGA0037719	Palmitic acid (C16)	6.40E-04	0.21	17.61
			Saturated Fatty Acids (SFA)	7.62E-04	0.15	11.45
			Unsaturated Fatty Acids (UFA)	7.85E-04	0.15	11.72
8	9,494,068	DRGA0008295	Palmitic acid (C16)	3.16E-04	0.16	12.34
8	33,282,771	H3GA0024720	Stearic acid (C18)	7.43E-04	0.20	16.55
8	119,727,823	SIRI0000509	Palmitic acid (C16)	2.48E-03	0.26	23.01
			Oleic acid (C18:1, cis-9)	9.47E-04	0.21	14.62
8	119,851,261	INRA0030422	Palmitic acid (C16)	7.93E-04	0.16	12.09
			Oleic acid (C18:1, cis-9)	3.35E-03	0.34	28.15

8	119,887,465	H3GA0025321	Palmitic acid (C16)	3.59E-02	0.72	166.13
			Palmitoleic acid (C16:1, cis-9)	7.96E-05	0.18	12.02
			Oleic acid (C18:1, cis-9)	4.59E-04	0.16	10.05
			Medium Chain Fatty Acids (MCFA)	1.61E-01	1.00	> 1000
			Long Chain Fatty Acids (LCFA)	1.55E-01	1.00	> 1000
8	121,743,169	ALGA0049269	Palmitic acid (C16)	8.69E-04	0.22	18.39
8	121,829,267	ASGA0039683	Palmitic acid (C16)	3.12E-04	0.14	10.73
8	140,400,180	ASGA0102346	Backfat thickness	4.41E-03	0.17	13.84
9	22,748,295	ASGA0099198	Arachidonic acid (C20:4Ω6)	9.83E-07	0.17	13.84
9	26,225,411	H3GA0026788	Linolenic acid (C18:3Ω3)	4.45E-06	0.14	10.76
10	10,036,337	MARC0022071	Palmitic acid (C16)	3.02E-04	0.17	13.01
10	54,960,156	ALGA0105237	Palmitic acid (C16)	1.79E-04	0.14	10.32
10	55,858,372	H3GA0030320	Linoleic acid (C18:2, cis-9, 12)	1.03E-03	0.20	13.24
			Polyunsaturated Fatty Acids (PUFA)	1.03E-03	0.19	12.57
			Unsaturated Fatty Acids (UFA)	5.50E-04	0.14	10.51
			Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	1.00E-03	0.19	13.02

11	7,320,581	ASGA0049556	Linolenic acid (C18:3Ω3)	7.33E-06	0.18	13.98
11	63,662,080	M1GA0015159	Linolenic acid (C18:3Ω3)	6.28E-06	0.16	12.85
11	65,750,596	INRA0036865	Backfat thickness	7.18E-03	0.20	16.79
14	123,733,166	MARC0006531	Stearic acid (C18)	4.50E-04	0.15	11.71
16	665,251	ASGA0071831	Backfat thickness	1.98E-03	0.14	10.43
16	864,816	ASGA0071832	Backfat thickness	7.03E-03	0.22	18.35
16	81,266,338	MARC0078879	Saturated Fatty Acids (SFA)	1.87E-03	0.22	18.35
			Monounsaturated Fatty Acids (MUFA)	3.66E-04	0.15	9.55
			Unsaturated Fatty Acids (UFA)	5.85E-04	0.14	11.12
17	17,548,504	INRA0052808	Linolenic acid (C18:3Ω3)	1.36E-05	0.21	17.63
17	17,686,439	MARC0013253	Linolenic acid (C18:3Ω3)	1.50E-05	0.23	19.38
18	13,833,809	ASGA0078689	Palmitic acid (C16)	1.79E-04	0.15	11.28
18	13,900,864	ASGA0088995	Palmitic acid (C16)	2.46E-04	0.16	12.89
18	141,909,500	ASGA0102347	Oleic acid (C18:1, cis-9)	6.46E-04	0.18	11.93

[1] Marker positions referred to *Sus scrofa* assembly Build 10.2.

[2] Proportion of the genetic variance explained by the marker on the total genetic variance of the trait.

[3] Posterior probability of the marker inclusion in the model iterations.

[4] BF_i stands for the Bayes Factor of the marker.

Docosapentaenoic acid (C22:5Ω3), docosahexaenoic acid (C22:6Ω3), polyunsaturated fatty acids omega 3 (PUFAΩ3) and short chain fatty acids (SCFA) are not reported in the table because no marker showed relevant association with these fatty acids (Bayes Factors were less than or equal to 3.2).

Supplementary Table S3. List of the 1Mb windows associated with pigs backfat fatty acids composition explaining at least 0.5% of the trait genetic variance. In the table are reported also the extended regions, composed of at least two contiguous windows associated with the same trait, and whose explained genetic variances were summed.

Traits	SSC	Region[1] (Mb)	SNPs in the region	Genetic Variance[2]	$\widehat{P}_1$[3]	Extended region[4]	
						Region (Mb)	Genetic Variance[5]
Palmitic acid (C16)	8	119-120	15	9.2	0.99	119-122	10.19
	8	9-10	26	0.68	0.58	8-10	1.15
	7	52-53	31	0.59	0.57		
	10	10-11	26	0.57	0.59		
	1	279-280	28	0.57	0.59		
	18	3-4	15	0.57	0.50		
Palmitoleic acid (C16:1, cis-9)	8	119-120	15	1.27	0.37	119-122	3.01
	2	107-108	24	0.47	0.23	107-109	0.79
Stearic acid (C18)	7	53-54	31	1.01	0.65		
	8	33-34	22	0.84	0.52		
	1	77-78	22	0.79	0.43	75-80	2
	14	117-118	25	0.68	0.54	116-118	1.18
	7	5-6	31	0.57	0.51		
Oleic acid (C18:1, cis-9)	8	119-120	15	1.78	0.74		

	9	28-29	32	0.6	0.64		
	18	41-42	22	0.57	0.52		
	14	122-123	23	0.45	0.51	121-124	1.23
Linoleic acid (C18:2, cis-9, 12)	5	104-105	21	0.97	0.59		
	10	55-56	15	0.52	0.45		
	7	57-58	28	0.39	0.52	56-58	0.75
Linolenic acid (C18:3Ω3)	17	17-18	10	1.2	0.52		
	9	26-27	30	0.78	0.59		
	5	65-66	25	0.67	0.48		
	3	32-33	14	0.61	0.37		
	11	7-8	23	0.6	0.45		
	1	18-19	20	0.57	0.37		
Arachidic acid (C20)	5	72-73	19	0.9	0.38	70-75	2.57
	14	138-139	31	0.48	0.34	138-142	1.13
Arachidonic acid (C20:4Ω6)	7	53-54	31	0.72	0.48		
	9	22-23	27	0.65	0.42		
	16	36-37	25	0.56	0.39		
Docosahexaenoic acid (C22:6Ω3)	14	50-51	29	0.4	0.16	47-52	1.65
Saturated Fatty Acids (SFA)	7	53-54	31	0.97	0.68	51-54	1.58
	1	77-78	22	0.74	0.50	75-80	2.04
Monounsaturated Fatty Acids (MUFA)	9	28-29	32	0.72	0.70		

	14	10-11	28	0.51	0.57		
	7	56-57	38	0.51	0.62		
	14	122-123	23	0.5	0.53	121-124	1.26
Polyunsaturated Fatty Acids (PUFA)	5	104-105	21	0.95	0.59		
	10	55-56	15	0.53	0.45		
Unsaturated Fatty Acids (UFA)	7	53-54	31	0.91	0.68	51-54	1.42
	1	77-78	22	0.73	0.49	75-80	1.98
	8	9-10	26	0.62	0.60		
	7	5-6	31	0.6	0.59		
	1	277-278	12	0.59	0.39	277-280	0.87
	8	141-142	31	0.51	0.60		
	14	117-118	25	0.49	0.56	116-118	0.82
Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	5	104-105	21	0.99	0.61		
	10	55-56	15	0.5	0.44		
Polyunsaturated Fatty Acids omega 3 (PUFAΩ3)	14	50-51	29	0.43	0.22	48-52	1.64
Short Chain Fatty Acids (SCFA)	14	50-52	29	0.57	0.19	46-52	2.15
	13	173-174	13	0.34	0.08	167-174	1.21
Medium Chain Fatty Acids (MCFA)	8	119-120	15	70.24	1.00		
Long Chain Fatty Acids (LCFA)	8	119-120	15	70.38	1.00		

Backfat thickness	7	15-16	25	1.38	0.63	10-12;15-16	2.02
	7	122-123	15	0.91	0.60		
	16	0-1	18	0.69	0.50	0-3	0.92
	11	65-66	17	0.53	0.39		

[1] Windows positions referred to *Sus scrofa* assembly Build 10.2, expressed in Mb.

[2] Proportion of the collective genetic variance explained by the markers in the window on the total genetic variance of the trait.

[3] Posterior probability of inclusion in the model iterations for the markers in the window.

[4] Extended region comprising contiguous windows explaining at least 0.5% of the genetic variance.

[5] The proportion of genetic variance explained by the markers placed in the neighbouring windows of the extended regions.

Docosapentaenoic acid (C22:5Ω3) is not reported in the table because no marker window explained at least 0.5% of the genetic variance.

Supplementary Table S4. List of the pathways in which are involved the genes located in the regions relevantly associated with the traits.

Trait	Pathway	P-value	Adjusted P-value	Z-score	Combined score	Genes
Palmitic acid (C16)	Acetylcholine Binding And Downstream Events	4.52E-05	2.51E-03	-2.32	13.88	*CHRNA3;CHRNB; CHRNA5*
	Regulation of cholesterol biosynthesis by SREBP (SREBF)	1.82E-03	6.73E-02	-1.90	5.14	*SEC24B; INSIG1;ELOVL6*
	Synthesis of very long-chain fatty acyl-CoAs	4.14E-03	1.25E-01	-2.05	4.25	*ACSBG1; ELOVL6*
Palmitoleic acid (C16:1, cis-9)	Metabolism of lipids and lipoproteins	2.08E-02	2.04E-01	-2.18	3.47	*SEC24B; ELOVL6; PLA2G12A; ETNPPL*
	Regulation of cholesterol biosynthesis by SREBP (SREBF)	3.94E-03	2.04E-01	-1.97	3.13	*SEC24B; ELOVL6*
Stearic acid (C18)	Biosynthesis of amino acids	2.01E-03	2.03E-01	-1.77	2.83	*GOT1; PGAM1; ALDH18A1; IDH3A*
	Arginine and proline metabolism	5.77E-03	2.52E-01	-1.86	2.56	*GOT1; ALDH18A1; HOGA1*
	B cell receptor signaling pathway	1.55E-02	2.61E-01	-1.83	2.46	*CHUK; BLNK; PIK3AP1*
	PPAR signaling pathway	1.34E-02	2.61E-01	-1.59	2.14	*SCD; ACSBG1; SORBS1*
	Fatty acid metabolism	4.68E-02	4.30E-01	-1.41	1.19	*SCD; ACSBG1*

Oleic acid (C18:1, cis-9)	Biosynthesis of unsaturated fatty acids	8.26E-03	3.61E-01	-1.59	1.62	*SCD; ELOVL6*
	Fatty acid elongation	9.60E-03	3.61E-01	-1.41	1.44	*ELOVL3; ELOVL6*
Linoleic acid (C18:2, cis-9, 12)	Myogenesis	6.22E-04	3.89E-02	-2.41	7.83	*MYF6; MYF5*
	CDO in myogenesis	6.22E-04	3.89E-02	-2.40	7.78	*MYF6; MYF5*
	Neurotransmitter Release Cycle	1.82E-03	7.56E-02	-1.97	5.08	*GAD2; LIN7A*
	Developmental Biology	6.53E-02	3.64E-01	-2.21	2.23	*APBB1IP; MYF6; MYF5*
Linolenic acid (C18:3Ω3)	Cytokine Signaling in Immune system	1.32E-02	3.39E-01	-2.38	2.57	*GRIN2A; CIITA; SOCS1; TNFRSF17; TAB2; NUP43*
	Amino acid synthesis and interconversion (transamination)	3.35E-03	3.39E-01	-2.26	2.45	*FOLH1B; NAALAD2*
	Cellular responses to stress	2.99E-02	3.39E-01	-2.26	2.45	*PHC1; HSPH1; NOX4; NUP43*
	Metabolism of amino acids and derivatives	2.23E-02	3.39E-01	-2.23	2.41	*IYD; FOLH1B; NAALAD2; TYR*
	Immune System	1.95E-02	3.39E-01	-2.20	2.38	*CIITA; GRIN2A; SOCS1; KLRB1; TNFRSF17; TAB2; NUP43; CLEC2D; ULBP1; KLRG1*
Arachidic acid (C20)	Digestion of dietary lipid	6.59E-08	8.37E-06	-1.94	22.68	*PNLIPRP1; PNLIPRP2;*

						PNLIPRP3; PNLIP
	Lipid digestion, mobilization, and transport	1.25E-05	7.94E-04	-1.82	12.98	*PNLIPRP1; PNLIPRP2; PNLIPRP3; PNLIP*
	NCAM1 interactions	2.68E-03	1.14E-01	-2.03	4.41	*GFRA1; CACNA1C*
	Metabolism of lipids and lipoproteins	3.91E-02	3.23E-01	-2.13	2.41	*PNLIPRP1; PNLIPRP2; PNLIP; PNLIPRP3*
	Glycerolipid metabolism	5.50E-06	3.25E-04	-1.97	15.81	*PNLIPRP1; PNLIPRP2; PNLIPRP3; PNLIP*
	Fat digestion and absorption	7.96E-05	2.35E-03	-1.82	11.01	*PNLIPRP1; PNLIPRP2; PNLIP*
	Pancreatic secretion	8.83E-04	1.74E-02	-1.67	6.78	*PNLIPRP1; PNLIPRP2; PNLIP*
Arachidonic acid (C20:4Ω6)	Cholinergic synapse	3.67E-02	2.46E-01	-1.98	2.77	*CHRNB4; CHRNA3*
	Neuroactive ligand-receptor interaction	6.02E-03	2.35E-01	-1.89	2.74	*CHRNB4; CHRNA3; CHRNA5; GZMA*
	Carbon metabolism	3.79E-02	2.46E-01	-1.58	2.22	*ME3; IDH3A*
	Fatty acid degradation	1.15E-01	3.13E-01	-1.33	1.55	*ACSBG1*
	Arachidonic acid metabolism	1.57E-01	3.13E-01	-1.29	1.50	*GPX8*
Docosahexaenoic acid (C22:6Ω3)	Phospholipid metabolism	1.28E-04	1.24E-02	-1.95	8.58	*MTMR3; INPP5J; PLA2G3;PISD*
	Metabolism of lipids and	2.08E-02	1.39E-01	-2.13	4.21	*MTMR3;INPP5J;*

	lipoproteins					*PLA2G3; PISD*
Saturated Fatty Acids (SFA)	Neuroactive ligand-receptor interaction	5.59E-04	1.73E-02	-1.89	7.68	*CHRNB4; CHRNA3; CHRNA5; GRIK2; MCHR2*
	Fatty acid degradation	1.09E-01	2.82E-01	-1.38	1.75	*ACSBG1*
	Fatty acid biosynthesis	3.53E-02	2.59E-01	-1.19	1.61	*ACSBG1*
Monounsaturated Fatty Acids (MUFA)	Dectin-1 mediated noncanonical NF-kB signaling	1.97E-02	2.88E-01	-2.24	2.79	*BTRC; NFKB2*
	Organelle biogenesis and maintenance	2.55E-02	2.88E-01	-2.15	2.68	*ACTR1A; GBF1; PPRC1; MRPL43*
	Fatty acid and ketone body metabolism	1.74E-01	4.70E-01	-1.37	1.04	*ELOVL3; MED17*
	Linoleic acid (LA) metabolism	3.06E-02	2.88E-01	-0.82	1.02	*ELOVL3*
	Metabolism of lipids and lipoproteins	1.91E-01	4.81E-01	-1.33	0.97	*STARD5; ELOVL3; MED17; CYP17A1*
Polyunsaturated Fatty Acids (PUFA) and Polyunsaturated Fatty Acids omega 6 (PUFAΩ6)	Myogenesis	2.40E-04	1.01E-02	-2.41	11.09	*MYF6; MYF5*
	CDO in myogenesis	2.40E-04	1.01E-02	-2.40	11.02	*MYF6; MYF5*
	Insulin receptor signalling cascade	2.04E-01	3.13E-01	-2.14	2.49	*APBB1IP*
	Gastrin-CREB signalling pathway via PKC and MAPK	2.93E-01	3.41E-01	-1.93	2.07	*APBB1IP*
Unsaturated Fatty Acids (UFA)	Acetylcholine Binding And	9.12E-05	5.32E-03	-2.32	12.14	*CHRNA3; CHRNB4;*

	Downstream Events					*CHRNA5*
	Presynaptic nicotinic acetylcholine receptors	6.15E-05	5.32E-03	-2.27	11.89	*CHRNA3; CHRNB4; CHRNA5*
	Fatty Acyl-CoA Biosynthesis	2.05E-01	5.97E-01	-0.90	0.46	*ACSBG1*
	Synthesis of very long-chain fatty acyl-CoAs	1.11E-01	5.97E-01	-0.84	0.43	*ACSBG1*
Polyunsaturated Fatty Acids omega 3 (PUFAΩ3)	Phospholipid metabolism	7.49E-06	7.27E-04	-1.95	14.13	*MTMR3; PITPNB; INPP5J; PLA2G3; PISD*
	Glycerophospholipid biosynthesis	7.76E-04	1.88E-02	-1.87	7.42	*PITPNB; PLA2G3; PISD*
	Metabolism of lipids and lipoproteins	4.76E-03	6.60E-02	-2.18	5.93	*MTMR3; PITPNB; INPP5J; PLA2G3; PISD*
	Metabolism of lipids and lipoproteins	4.76E-03	6.60E-02	-2.18	5.93	*MTMR3;PITPNB;INPP5J ;PLA2G3;PISD*
Short Chain Fatty Acids (SCFA)	Phospholipid metabolism	7.93E-05	9.67E-03	-1.95	9.07	*MTMR3; PITPNB; INPP5J; PLA2G3; PISD*
	Glycerophospholipid biosynthesis	2.95E-03	7.19E-02	-1.86	4.89	*PITPNB; PLA2G3; PISD*
	Metabolism of lipids and lipoproteins	3.33E-02	2.64E-01	-2.13	2.83	*MTMR3; PITPNB; INPP5J; PLA2G3; PISD*
	Metabolism of carbohydrates	5.51E-01	6.35E-01	-1.41	0.64	*SLC5A1*
Backfat thickness	MicroRNA (miRNA) biogenesis	2.38E-02	1.75E-01	-2.23	3.90	*DICER1*
	Hemostasis	1.49E-02	1.75E-01	-2.12	3.71	*SERPINA1; ITPK1;*

	SERPINA5

In grey are reported the pathways obtained from Reactome, the remaining were obtained from KEGG (in white).

In this table are not reported the Medium Chain Fatty Acids and the Long Chain Fatty Acids categories as their amounts in backfat showed in particular to be associated with three markers already indicated in previous studies to be in linkage disequilibrium with mutations in the promoter region of the ELOVL Fatty Acid Elongase 6.

Docosapentaenoic acid (C22:5Ω3) is not reported in the table because no marker window explained at least 0.5% of the genetic variance.

Chapter 2: Transcriptional profiling of subcutaneous adipose tissue in Italian Large White pigs divergent for backfat thickness.

P. Zambonelli [1]*, E. Gaffo[2]*, M. Zappaterra[1], S. Bortoluzzi[2], R. Davoli[1]

[1]Department of Agricultural and-Food Sciences (DISTAL), Bologna University, Via Fratelli Rosselli 107, 42123 Reggio Emilia, Italy.

[2]Department of Molecular Medicine, University of Padova, Via Gabelli 63, 35121 Padova, Italy

*These authors contributed equally to this work

Published in

Animal Genetics

Zambonelli et al. *Anim Genet.* 2016 Jun;47(3):306-23.

doi: 10.1111/age.12413.

Summary

Fat deposition is a widely studied trait in pigs for the implications with animal growth efficiency, technological and nutritional characteristics of meat products, but the global framework of the biological and molecular processes regulating fat deposition in pigs is still incomplete. This paper describes the backfat tissue transcription profile in Italian Large White pigs and reports genes differentially expressed between fat and lean animals according to RNA-seq data. The backfat transcription profile was characterized by the expression of 23,483 genes of which 54.1% were represented by known genes. Of 63,418 expressed transcripts, about 80% were non previously annotated isoforms. By comparing the expression level of fat *vs*. lean pigs we detected 86 robust differentially expressed transcripts, 72 more expressed (e.g. *ACP5, BCL2A1, CCR1, CD163*, *CD1A, EGR2*, *ENPP1*, *GPNMB*, *INHBB, LYZ, MSR1, OLR1, PIK3AP1, PLIN2, SPP1*, *SLC11A1, STC1*) and 14 less expressed (e.g. *ADSSL1, CDO1, DNAJB1, HSPA1A, HSPA1B, HSPA2, HSPB8, IGFBP5, OLFML3*) in fat pigs. The main functional categories enriched in differentially expressed genes were immune system process, response to stimulus, cell activation, and skeletal system development, for the overexpressed, unfolded protein binding and stress response, for the under-expressed genes, which include five heat shock proteins. Adipose tissue alterations and impaired stress response are linked to inflammation and, in turn, to adipose tissue secretory activity similarly to what is observed in human obesity. Our results open the opportunity to identify biomarkers of carcass fat traits to improve pig production chain and to identify genetic factors that regulate the observed differential expression.

Introduction

Backfat deposition and fat traits are among the most important characters studied in pigs, due to their strong relation with human nutrition of pig products and for the technological characteristics of high quality Protected Designation of Origin (PDO) dry-cured hams. The amount of fat laid on the external part of the pig body (subcutaneous fat or backfat) is of extreme importance for growth performances, as the lesser is the deposed fat, the better the growth performances. Regarding technological aspects related to the dry-cured high

quality products and meat industry, an adequate layer of fat is required for the seasoning process of PDO products, like dry cured hams (Bosi and Russo, 2004; Čandek-Potokar and Škrlep, 2012).

During the last decade, pig transcriptomic data have been obtained initially by expressed sequence tag sequencing (Mikawa et al., 2004; Uenishi et al., 2004; Chen et al., 2006; Gorodkin et al., 2007; Uenishi et al., 2007) and microarrays (Hornshøj et al., 2007; Ferraz et al., 2008; Moon et al., 2009), which allowed the comparison of gene expression level in several pig tissues. More recently, the RNA-seq approach was used to compare the transcription profile of different pig fat tissues or different pig breeds (Chen et al., 2011; Li et al., 2012; Corominas et al., 2013; Jiang et al., 2013; Zhou et al., 2013; Sodhi et al., 2014; Toedebush et al., 2014; Wang et al. 2014). The differentially expressed genes (DEG) reported in these studies are useful to investigate the metabolic pathways activated by or associated with an increased fat deposition in pig body. However, the large amount of data produced and the results reported in literature are often hardly comparable because of differences in the studied breeds; heterogeneous animals' ages; and fat deposition stages. Moreover, these researches identified several new genes and transcripts not reported in swine or other species. To date, the global framework of the biological processes regulating backfat deposition in pigs is still incomplete, and literature is poor of studies carried out on a homogeneous sample of individuals of the same breed reared on the same environmental conditions.

The objective of this research was to investigate the transcription profile of Italian Large White (ILW) pig backfat tissue and to compare the transcriptome of animals reared in the same herd and farming conditions and showing high (FAT) and low (LEAN) backfat thickness. Moreover a first functional characterization of DEGs has been obtained to provide new insights on genes, pathways and processes influencing the divergent aptitude of subcutaneous adipose tissue deposition in ILW pigs.

Materials and methods

Samples collection and RNA extraction

We sampled twenty individuals from a purebred population of 949 ILW sib-tested pigs provided by the Italian National Association of Pig Breeders (Associazione Nazionale

Allevatori Suini, ANAS, http://www.anas.it. Accessed 22 June 2015). All animals used in this study were kept according to Italian and European law for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering. The animals were reared on the ANAS Sib-Test genetic station from about 30 kg live weight to at least 155 kg live weight. For the genetic evaluation of a boar, full sib triplets (two females and one castrated male) were farmed on the genetic station to be performance tested. The formula and amount of the ration were the same for all. It was based mainly on cereals and soybean, given in excess calculated using the "*quasi ad libitum*" rule (a ration sufficiently abundant that the 60% of pigs was able to ingest the full supplied food). At the end of tests, animals were transported to a commercial abattoir located at about 25 km far from the test station according to the Council Rule (EC) No 1/2005 on the protection of animals during transport and related operations and amending Directives 64/432/EEC and 93/119/EC and Regulation (EC) No 1255/97. At slaughterhouse the pigs were electrical stunned and bled in a lying position in agreement with the Council Regulation (EC) No 1099/2009 on the protection of animals at the time of killing. All slaughter procedures were monitored by the Veterinary team appointed by the Italian Ministry of Health. Backfat samples were collected after slaughter, from 949 ILW pigs slaughtered at an average hot carcass weight of 118.97 kg (±0.29 SEM) and at an average age of eight months during the years 2011 and 2012 in 27 different slaughtering days. The collected samples were immediately frozen in liquid nitrogen and stored at -80°C in a deep freezer until RNA extraction. For the RNA-seq analysis we selected the animals according to the estimated breeding value (EBV) for backfat thickness (BFT) calculated by ANAS as described by Russo et al. (2000; 2008). EBVs were determined through a BLUP multiple-trait animal model procedure (Henderson and Quaas, 1976) using the BFT, measured in mm, recorded post mortem in correspondence of the *gluteus medius* muscle. The model included fixed effects of batch in test, sex, age at beginning of test, age of sow, weight at slaughter, age at slaughter, and inbreeding coefficient as well as the random effects of litter, individual permanent environment, and animal. Pigs' genetic merit for the BFT trait was calculated taking into account the additive relationship matrix. EBVs were expressed as differences from the genetic mean value for the considered trait in the year 1993. Backfat thickness genetic index may present negative values because the value of the trait is referred to the fixed genetic base defined by ANAS as mean values of the pigs born in 1993 and considered

as “zero”, so the more negative values indicate lower values of BFT. The animals were selected to compose two groups of 10 pigs showing extreme and divergent characteristics for the BFT EBV with respect to the larger population of the 949 pigs (Table 1). The twenty animals considered for RNA-seq analysis were slaughtered in 12 dates, with 5 dates common to both groups. The animals were selected also according to their pedigree, in order to avoid the presence of full sibs in the considered groups. From now on the two groups will be referred as FAT and LEAN samples.

RNA extraction, library preparation, sequencing

Total RNA was extracted with Trizol (Invitrogen) according to the manufacturer’s instruction. RNA extracted samples were quantified using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies) and the quality of the RNA was assayed using an Agilent 2100 BioAnalyzer (Agilent Technologies). The RNA libraries were prepared from total RNA using the TruSeq RNA sample preparation kit (Illumina) and version 3 of the reagents, following the manufacturer’s suggested protocol. The libraries were tagged and couples of libraries were run on a single lane of an Illumina HiSeq2000. Reads are 100 nt paired-end represented in FASTQ format.

Architecture of the bioinformatics pipeline

A computational pipeline to process the sequencing data for gene/transcript expression estimation and to perform differential expression analysis between the two sample groups was developed. The pipeline components to achieve expression estimates were assembled using Scons software (http://www.scons.org/. Accessed 22 June 2015), which allows the parallelization and automation of the pipeline tasks. The pipeline and its following steps are detailed in the next paragraphs.

RNA-seq data pre-processing and mapping to swine genome

Exploratory analyses on the raw reads quality were carried out using the FastQC v0.10.1 software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 22 June 2015), which generates an HTML report for each sample read set. Read fragments with

quality Phred score lower than 30 were trimmed using the DynamicTrim script of the SolexaQA v2.1 (Cox et al., 2010). The FASTX-Toolkit v0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/. Accessed 22 June 2015) was used for trimming result report. A custom Python script using the HTSeq package (Anders et al., 2015) filtered out the trimmed reads shorter than 50 nucleotides. To maintain a consistent paired-end read set, discarded read mates were also filtered out, despite their length and quality. Each sample paired-end clean read set was mapped to the swine genome (Sscrofa10.2.70) by Tophat v2.0.8 (Kim et al., 2013) using default parameters with transcriptome inference from Ensembl annotation (Tophat2 used Bowtie v2.1.0.0; Langmead and Salzberg, 2012) and SAMtools v0.1.19()Li et al., 2009).

Gene/transcript expression evaluation and transcript reconstruction

Gene annotation for the reference genome was retrieved from Ensembl (BioMart) using the biomaRt R package (Durinck et al., 2009). Read alignments were processed by Cufflinks v2.1.1 (Roberts et al., 2011a; Roberts et al., 2011b; Trapnell et al., 2010) to identify and discover expressed genes and transcripts, and to quantify their expression. Expression data were indicated as Fragments Per Kilobase of transcript per Million mapped reads (FPKM). Cufflinks was applied to each sample alignment; then, we merged the transcript predictions in a non-redundant reference using the Cuffmerge tool from the Cufflinks package. To reduce artefacts deriving from the transcript prediction and normalisation strategies, only predicted transcripts at least 200 nt long and with minimal expression of 100 (Cufflinks normalised) reads in at least one of the two groups were considered for transcriptome reconstruction and for the following analyses.

Gene and transcript differential expression assessment

The samples were inspected by principal component analysis to examine their similarities. The read counts of each gene in the 20 considered samples were transformed with the variance stabilizing transformation function provided by the DESeq2 package (Anders and Huber, 2010) and used to compute the principal components.

The genes identified by Cufflinks were assessed for differential expression (DE) between the LEAN and FAT groups, by means of two strategies, namely Cuffdiff2 (v2.1.1 from the Cufflinks package; Trapnell et al., 2012) and DESeq2 v1.2.1 (Anders and Huber, 2010). Instead, transcript DE was assayed only with Cuffdiff2. To represent gene expression, the two methods use similar statistical approaches based on generalized linear model (GLM) of the negative binomial family. Cuffdiff2 extends the model using a beta negative binomial distribution to handle uncertainty of multi-mapped reads. On the contrary, DESeq2 considers only uniquely mapped reads (counted by means of the htseq-count script of the HTSeq package (Anders et al., 2015), but facilitate the specification in the statistical model of additional factors effecting the fit of the GLM. In this study, the statistical model included sex effect as a potential conditioning factor. Gene and transcript DE test computed *P* values were corrected according to the Benjamini-Hochberg procedure. Differentially expressed genes and transcripts were considered statistically significant according to false discovery rate less than or equal to 0.05.

Transcript characterisation

Using custom scripts including BEDTools v2.17.0 software (Quinlan and Hall, 2010), we retrieved the nucleotide sequences of the transcripts extracting from the *Sus scrofa* genome the stretches of nucleotides according to the annotation generated by the RNA-seq analysis tools. Transcripts were identified or characterised by sequence similarity using BLASTN and BLAST2 from the NCBI BLASTN suite (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome. Accessed 22 June 2015) using Megablast algorithm (Morgulis et al., 2008). To assign a gene name, the sequences IDs obtained with this comparison were used to query the NCBI Gene and the UniGene databases (http://www.ncbi.nlm.nih.gov/unigene/. Accessed 22 June 2015). We used two strategies for transcript annotation. DE transcripts and genes were annotated by similarity using nr/nt nucleotide collection. The threshold considered for the identification of our transcripts was identity ≥80% in at least 70% of the sequence length of a transcript present in the database. Transcripts from new genes were characterized using a comparative genomics approach. We compared the new transcripts from intergenic regions with known human transcripts (RefSeq Release 72) by aligning with

BLASTN (NCBI BLAST 2.2.29+). For each transcript the best hit was considered, and then alignments with E-value greater than 10e-6, identity less than 60%, and length less than 100 nucleotides were discarded.

Prediction of coding/non-coding potential

The transcript coding potential was predicted by CPC (Coding Potential Calculator; Kong et al., 2007). CPC is a support vector machine-based classifier of transcript protein-coding potential grounding on six features of sequence. Three features assess the extent and quality of the predicted transcript ORF: the Framefinder software identifies the longest ORF in the three forward and in the three reverse frames, then the coverage and the integrity of the predicted ORF are evaluated. Another three features derive from results of BLASTX search against UniProt Reference Clusters. All the features contribute together to a final score, and to the classification of transcripts as coding or non-coding. Only transcripts not including uncalled bases were considered for CPC analysis.

Validation by quantitative real time-PCR

The validation of selected RNA-seq results was performed using a quantitative real time-PCR (qRT-PCR) approach using 18 out of the 20 samples used for the RNA-seq analysis. Two samples, one in the FAT group and one in the LEAN group, were not considered because the total RNA extracted was used completely for the RNA-seq analysis. QRT-PCR validation was carried out using Rotor-Gene TM 6000 (Qiagen - Corbett Research). After DNase treatment (TURBO DNA-free^{TM}, Ambion, Applied Biosystems), 1 µg of total RNA was reverse transcribed using the iScript cDNA Synthesis kit (BIORAD) according to the manufacturers' instructions.

The samples were first used to analyze four candidate normalizing genes beta-2-microglobulin (*B2M*), polymerase (RNA) II (DNA directed) polypeptide A, 220kDa (*POLR2A*), hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*). The primer pairs and the PCR conditions used are reported in Supplementary Table 1. The expression levels of these four genes were evaluated using NormFinder and *B2M* and *HPRT1*, the two most stably expressed normalizing genes, were utilized as reference genes. For each gene selected for

validation, we designed an external primer pair to obtain the amplicon for the standard curve construction and an internal primer pair for the qRT-PCR on Rotor Gene 6000 (Table S1). Standard curves for each gene were generated from 10-12 serial dilutions (from 10^9 to 25 molecules/µl) of the PCR amplicons obtained with the external primer pairs and containing the internal primers used in the qRT-PCR analysis. Amplifications were performed in a total volume of 10 µl containing using 5 µl of the SYBR® Premix Ex Taq™ (Takara Bio Inc.), 0.5 µl of each primer and about 100 ng of cDNA. The used Premix Ex Taq™ is optimized for a two-step cycling, and the amplification conditions for the tested genes are reported in Table S1. The PCR efficiency was calculated as E=10 exp(-1/slope), with a range between -2.7 and -4.3, indicating a good PCR efficiency result. All the PCR products were checked on a polyacrylamide gel and the specificity of the amplification was checked by a final melting curve analysis.

Threshold cycles obtained for the samples were converted by Rotor Gene 6000 to mRNA molecules/µl using for each gene the relative standard curve (Bustin and Nolan, 2004). Moreover, the average mRNA molecules/µl for each sample was normalized dividing the mRNA molecules of a gene /µl by the geometric average of *B2M* and *HPRT1* mRNA molecules/µl in the given sample, as suggested by Bustin and Nolan, 2004 and Vandesompele et al., 2002. Differences on the expression level calculated for FAT and LEAN samples were tested by two-tailed Student's t test. Statistical analyses were performed with SAS version 9.3 (SAS 9.3 Help and Documentation, Cary, NC. SAS Institute Inc.) and nominal *P* value ≤0.05 was considered as significance threshold.

Functional characterization

Functional annotation, classification and clustering of selected gene sets were carried out by DAVID Tools 6.7 (Huang et al., 2009) using Biological Processes, Molecular Function gene ontology categories and KEGG pathways. A threshold for significance of P<0.01 and P<0.05 after Benjamini correction was considered for the selection of the functional categories respectively in the characterization of most expressed transcripts and for the selection of the functional categories of DEG.

Results

Samples

In this study we applied RNA-seq by Illumina technology to the study of gene expression in backfat tissue of 20 ILW pigs. We considered a large group of 949 sampled animals, with EBV for BFT ranging from -10.64 mm to 7.28 mm, with mean value and standard deviation (SD) -1.96 mm and 3.01, respectively. We selected, from the whole collected population, two groups of 10 unrelated pigs (FAT and LEAN) with extremely divergent EBVs for BFT, with 1:1 sex ratio within each group. The mean values of each of the two selected groups of pigs are outside the range -7.98 mm / 4.06 mm defined by the mean value of the 949 samples ±2 standard deviations. Specifically, FAT and LEAN animals were associated to average BFT values of +5.22 mm (±1.30 SD) and -8.63 mm (±1.40 SD) as indicated in Table 1.

Sequencing, reads pre-processing and mapping

Pairs of samples were run together, after barcoding, on a single lane of an Illumina HiSeq 2000 apparatus, obtaining a total of 3,917,123,414 raw reads for the 20 considered samples, with an average of 195,856,171 raw reads per sample (Table S2; GEO accession GSE68007). After trimming and length filtering the clean reads per sample were on the average 113,934,264 (58.04%) and were used for read-to-genome mapping (Figure S1A). Reads that align on a single genome locus (uniquely mapped reads) were on the average the 91.07% of the mapped reads (Table S2). The 72.42% of the uniquely mapped reads (72,219,306.45 on the average aligned to annotated exons, the 19.15% mapped on intergenic regions and the 8.43% mapped on introns of annotated genes. The deep sequencing allowed the identification of genes expressed at low level and relatively rare alternatively spliced transcripts. We observed splicing events in the 21.19% of the reads on the average, providing useful information for the reconstruction of alternative transcript isoforms (Figure S1B).

Transcripts and genes expressed in backfat samples

The deep sequencing analysis of backfat transcripts performed on two groups of pigs divergent for fat deposition in this tissue allowed the detection of 63,418 transcripts. Many of them have not yet been annotated in the porcine genome, thus providing new consistent

resources for pig genome annotation and studies of adipose tissue biology. We identified the expression of genes on all porcine autosomes, sex chromosomes and mitochondrial genome. Chromosome 1 has the largest number of expressed genes (8.23%), followed by chromosomes 6 (7.84%) and 2 (7.25%). Furthermore, a non-negligible part (12.48%) of the expressed genes is located in genomic scaffolds (Figure S1C), as about the 7.5% of the genome has no assigned location yet, as described in Ensembl annotation of pig genome (database version 78 at the time of the analysis ; http://www.ensembl.org/Sus_scrofa/Location/Genome. Accessed 22 June 2015). In term of genes, we identified 23,483 expressed pig genes: 12,707 known and 10,776 putative new genes.

Transcripts were split in different classes according to their matching with the genome annotations (Figure 1A, Table S3). Transcripts matching exactly the reference annotation are indicated as “known” transcripts; annotated transcripts’ new isoforms or overlapping with annotated transcript are indicated as “novel isoforms; and all other transcripts, such as those expressed from extragenic regions, are referred as “new” transcripts and might represent putative new genes. The majority of expressed transcripts are novel isoforms (35,030; the 55.2%) or known transcripts (12,969, representing the 20.5%) that are prevalently annotated as protein coding (12,883; 99.3%). The expressed new transcripts are 15,419 (24.3%).

Transcript lengths range from 200 to 50,610 nt, with median and average values of 3,224 and 3,979. Average size exceeds the 2 kb pig mean transcript size that can be estimated according to Ensembl pig coding transcript annotation. We observed that the novel isoforms reconstructed are longer than “known” pig transcripts (Figure 1B).

Sequences longer than 5 kb compose the 25% of the expressed transcripts. Noteworthy, we detected two transcripts overlapping *ZBTB16* gene and two new transcripts from chromosome 16 that are longer than 40 kb.

Considering transcripts expression, we observed that new transcripts are less expressed in fat tissue than known transcripts (Figure 1C). Nevertheless, all the three transcript categories span a considerably large range of expression values.

The majority of the expressed genes (12,138; 52%) present only one transcript isoform expressed in fat tissue (Figure 1D); the 27.0% and the 18.3% of the genes present two and three expressed isoforms, respectively, whereas the remaining 12.7% of the genes are

associated each one to 4 to 31 different isoforms. We identified 31 isoforms for the gene *MAP4K4*, for which a complex expression pattern is reported in humans: Ensembl release 79 lists 20 *MAP4K4* transcripts generated by at least 3 different promoters, by complex alternative splicing and by polyadenylation patterns, whereas five protein isoforms are reported in UniProt release 2015_3.

Looking at isoform types, Figure 1E shows that many genes expressing only one transcript (first bar from the left) in fat tissue are putative new genes (green portion). Interestingly, some genes expressing only one transcript in fat tissue are represented only by a novel isoform (first bar, red shadows). The proportion of novel isoforms (red portion) increases along with the numbers of expressed transcripts per gene. Moreover, the transcripts classes showing exonic overlap compared to a reference transcript are found in genes with a varying number of transcripts and are particularly abundant in genes with up to three isoforms. The remaining transcript classes are very rare.

Interesting new isoforms derived from known genes regard Perilipin 2 (*PLIN2;* alias *ADFP*, adipofilin), an important gene for fat metabolism in pigs (Davoli et al., 2010; Gandolfi et al., 2011) whose expression in humans correlates positively with cytosolic triacylglicerol levels (Conte et al., 2013). Only one transcript is currently annotated in Ensembl for pig *PLIN2* (ENSSSCT00000005701), whereas according to our results, *PLIN2* expressed four different isoforms. The most expressed *PLIN2* transcript (expressed two times more in FAT than in LEAN pigs) is a non-annotated isoform (TCONS_00002441 in Table 2; 2441DE in Figure S3) characterized by the skipping of the fourth exon. The same transcript has also a shorter 3' sequence with respect to the canonical *PLIN2/ADFP* form, probably due to the use of an alternative polyadenylation site. Importantly, the skipping of the 83 nt long exon four introduces downstream a shift in the reading frame and a premature stop codon. Thus, this transcript encodes a truncated protein (only 80 aa) corresponding to the N-terminal region and of the Perilipin domain of the *PLIN2* protein annotated isoform (463 aa). The other two new transcripts differ from the annotated isoform, one for the skipping of exon 2, and the other for a longer first exon, probably due to alternative TSS usage by different promoters. The four expressed isoforms are also heterogeneous in the length of the 3' UTR region.

Coding and non-coding transcripts from new genes

We obtained a characterization of intergenic transcripts from new genes first both by similarity, comparing them against human transcripts, and by predicting their coding potential. New pig transcripts with an assigned human best hit were 10,020 (65%), expressed by 7,099 genes (66%), and corresponding to 4633 human Refseq sequences (3,882 unique gene symbols; Table S4).

We considered 12,702 intergenic transcripts for protein coding potential analysis. For each transcript, the coding potential of both the forward and the reverse complement sequence were evaluated. According to CPC results, we classified the 35.8% (4,551) of transcripts as coding, and the 64.2% (8,151) as non-coding. As done by Zhou et al., (2014), we considered as proper non-coding only those transcripts classified as non-coding and having a CPC score lower than -1 for both the forward and the reverse sequence. A portion of the non-coding transcripts (37.5%) resulted with CPC score < -1 for both the forward and the reverse complement sequences. We refer to these transcripts as “reliable non-coding” class, which represented 24% (3,056) of the intergenic transcripts (Figure 2A). We observed that intergenic coding transcripts are on average longer than intergenic non-coding transcripts (4,149 and 3,083 nt, respectively), and that the reliable non-coding fraction has a even shorter average length (2,571 nt; Figure 2B and Table S5). Reportedly, non-coding transcripts tend to be shorter and to have fewer exons than coding transcripts in mammalian genomes (Iyer et al., 2015).

Coding transcripts have an average expression in fat tissue higher than the non-coding transcripts (5.32 and 2.28 FPKM respectively, and 3.23 FPKM for the reliable non-coding group; Figure 2C). One reliable non-coding transcript is ranked within the 100 most expressed transcripts detected in backfat tissue; 15 reliable non-coding transcripts are within the 1,000 most expressed transcripts; and 98 are within the 10% most expressed transcripts (Table S6). In agreement with previous results showing that coding transcripts tend to present higher expression than non-coding ones (Cabili et al., 2011; Iyer et al., 2015), we observe that intergenic transcripts ranking in the 10% most expressed in backfat tissue are enriched in the coding category (55%) and particularly if compared with the proportion of the coding category within the set of intergenic transcripts (35.8%; Figure 2D, green portions).

Function of most expressed transcripts

A global view of the transcription profile of porcine backfat tissue was obtained by averaging the FPKM values of all 20 analysed samples. The 1411 most expressed transcripts, accounting together for 75% of expression, were chosen to extract the most expressed genes (Table S6).
Among these genes, 59 are indicated as reliable non-coding (CPC score <1) and 66 showing a positive CPC score are indicated as putative coding.
According to DAVID functional annotation and clustering, we characterized the biological processes (Table S7) associated to the most expressed genes. Ribosomal activity, oxidative phosphorylation, protein metabolic processes, intracellular protein transport, regulation of translation initiation, fatty acid metabolism, response to oxidative stress resulted to be the biological processes more represented in subcutaneous adipose tissue of the analysed samples.

Gene/transcript differential expression

Unsupervised analysis of gene expression profiles was carried out to inspect similarities among the samples. Principal component analysis revealed a clear separation of the LEAN and FAT samples according to the first two most informative components (Figure S4 A), which, notably, do not separate the samples by sex (Figure S4 B).
Average gene expression values for FAT and LEAN groups were 32.46 and 33.63 FPKM). In both groups, few highly expressed genes contribute to the majority of the cumulative expression. For instance, roughly 25% of expressed genes (5,908 and 5,728 in FAT and LEAN, respectively) constitute 95% of the total detected expression (Figure S2). As expected, transcript expression distribution is similar to the gene expression distribution being positively skewed, with mean and median corresponding to 11.84 and 0.64 FPKM, respectively. Transcripts average expression values are lower than genes expression values since the latter was computed as the sum of transcripts expression of each gene.
To identify a set of robust DEG and DET the transcription profiles of FAT and LEAN samples were compared with the integration of two methods applied at gene and at transcript levels. Cuffdiff2 identified 414 DEGs between FAT and LEAN groups, corresponding to 1,187 transcripts: 266 DEGs are more expressed and 148 DEGs are less expressed in FAT samples. Fold changes in base two logarithmic scale of DEGs range from 0.46 to 8.95 for the higher

expressed genes, and from -6.19 to -0.47 for the less expressed ones (Table S8). DESeq2 identified 586 DEGs (185 in common with the DEGs identified by Cuffdiff2) corresponding to 1,504 transcripts: 358 genes are up-regulated and 228 genes are less expressed in FAT samples. DEGs base two logarithmic scale transformed fold changes (Log_2 FC) range from -1.13 to -0.20 for the less expressed genes and from 0.21 to 1.18 for the higher expressed genes (Table S9). Cuffdiff2 differential expression analysis at the transcript-level identified 154 DE transcripts (corresponding to 153 genes): 48 were less expressed and 106 transcripts were more expressed in FAT samples, with Log_2 FC ranging from -3.44 to -0.54 and from 0.64 to 3.66, respectively (Table S10). On the whole, 818 genes were DE, or associated to at least one DE transcript, according to at least one method, were detected (Figure 3A).
The overlapping of the different lists of DEGs and the list of DE transcripts (DET) evidenced a group of 86 DET that are identified by all the approaches, from now on referred as "common DET" (cDET). These DET belongs to 78 DEG, from now on referred as "common DEG" (cDEG) since five genes are represented by more than one isoform (Table 2).
The cDET present the same fold change sign of the corresponding cDEG (Figure 3B): 72 DET were more expressed in FAT (max Cuffdiff2 gene-level Log_2 FC 2.55 for *DSC2* gene) and 14 DET were less expressed in FAT (minimum Log_2 FC -3.44 for an intergenic gene located in GL894890.1 scaffold). Among the 86 cDET, 44 are known transcripts, 16 are novel isoforms and 26 come from intergenic regions.
cDEG are found in all chromosomes except for chromosomes 16 and Y, with up to 11 DE genes in chromosome 4 and 19 DE genes in scaffolds (Figure 3C). The most expressed (average FPKM greater than 100) known cDEG, reported in decreasing FPKM order, are *DNAJB1*, *CTSH*, *CTGF*, *C1QC*, *SPP1* and *CDO1*.

Coding and con-coding intergenic DET

We considered the 41 novel isoforms or new transcript cDET for CPC analysis. In 14 of these transcripts both the forward and reverse sequence is probably non-coding, according to integrated ORF analyses and to similarity searches, and to CPC score thresholds used before. Five cDET with CPC score <-1 were scored as "reliable non-coding". Of the remaining transcripts, nine presented low coding potential both in the forward and in the reverse

complement sequence but with CPC score ranging from -1 to 0 ("non-coding"), and 27 were classified as coding transcripts (Table S11).

qRT-PCR confirmation of DE for selected genes

To validate the results obtained by RNA-seq, eleven cDEG were chosen according to the absolute value of the Log_2 FC between FAT and LEAN pigs, or for their functional role and involvement in relevant pathways. As reported in Figure 4, the DE of all selected genes has been validated, with high correlation between the fold changes obtained by RNA-seq and by qRT-PCR data.

DE transcript characterisation

We characterized the cDEG in terms of their functional role in adipose tissue. Using DAVID Bioinformatics Resources we first identified the functional categories, enriched in genes differentially regulated between FAT and LEAN groups.
The Biological Process categories enriched in higher expressed DEG are response to stimulus, immune system process and cell activation, skeletal system development (Table 3). DAVID clustering of the few lower expressed genes detected (*ADSSL1, CDO1, DNAJB1, HSPA1A, HSPA1B, HSPA2, HSPB8, IGFBP5, OLFML3*) allowed to identify the functional categories unfolded protein binding and stress response represented by five heat shock protein genes that are involved in protein stabilization after cellular stress.
Apart from the Gene Ontology-based functional characterization of the whole subsets of higher- and lower-expressed genes we considered cDEG function and involvement in specific pathways, according to literature and knowledge bases.
Several more expressed genes in FAT animals (*ACP5, BCL2A1, CD1A, EGR2*, *ENPP1*, *GPNMB*, *INHBB, LYZ, MSR1, OLR1, PIK3AP1, PLIN2, SPP1*, *STC1*) are characterized by a metabolic function mainly related to adipocyte growth regulation, while others (*CCR1, CD163*, *SLC11A1*) are known to be involved in immune defence of the organism.

Discussion

Transcriptome data highlight the adipose tissue complexity

The deep sequencing analysis of pig backfat transcriptome performed allowed finding thousands of genes and transcripts expressed. In the present study, we applied stringent cleaning and filtering procedures of the sequencing data and, on average, 90 million reads per sample were mapped, obtaining a higher sequencing depth compared to previous studies (Chen et al., 2011; Jiang et al., 2013; Sodhi et al., 2014; Wang et al., 2014).
The adipose tissue is not only metabolically and transcriptionally active, but has been recognized as an important endocrine organ (Kershaw et al., 2004; Trayhurn et al., 2005). Adipocytes are a dynamic and highly regulated population of cells (Rosen and MacDougald, 2006; Moreno-Navarrete and Fernández-Real, 2012). Our results agree with these data supporting the characterization of the adipocytes as highly specialised endocrine cells that can play key roles in various physiological processes. The multifunctionality and the complexity of the tissue is witnessed also by the high number of transcripts (more than sixty thousands) found in the present study, including many new transcripts from previously non-annotated loci in porcine genome. The majority of the reconstructed sequences are novel isoforms of already known genes that express more than two different transcripts each. Similar patterns observed in human cells (Djebali et al., 2012) and the high quality of the sequenced reads used in our analysis support the idea that this is more attributable to an incomplete annotation of the transcript isoforms expressed in pig backfat, than to transcript reconstruction artefacts. The different isoforms derived from the same locus arisen from our analysis and observed for almost half of the expressed genes, may contribute to improve the knowledge of the porcine transcriptome, and to refine the current swine genome annotation. The new *PLIN2* isoforms reported above are an interesting example, especially if compared to the human genome where at least eight *PLIN2* transcript isoforms are annotated and only four of them are coding. Remarkably, three human *PLIN2* isoforms encode N-terminal truncated amino acid chains that are similar to the truncated isoform we reconstructed in our study, and whose function has not yet been elucidated. Furthermore, Russell et al. (2008) identified in a *PLIN2* deficient mouse cell line the expression of a *PLIN2* C-terminal truncated protein that may partially replace the function of the full-length protein. Additional studies are needed to understand if and how the short transcript we found differentially expressed could change the gene functions compared to the wild type long protein.

Functional characterisation of the adipose tissue expression profile

The profile of the subcutaneous adipose tissue transcriptome in pigs was delineated and the functional analysis of the genes expressed in backfat tissue was performed to know their metabolic role and to connect them to specific competences of the tissue. We didn't find particular differences between the functional categories of the genes expressed in the backfat tissue of FAT and LEAN pigs. More in details among the most expressed genes in the fat tissue, many are involved in metabolic pathways and biological processes related to protein metabolism, oxidoreductase activity for ATP production, regulation of lipid synthesis and degradation.

Genes differentially expressed between LEAN and FAT animals converge and connect to specific functions

The detection of DE genes and transcripts has been obtained by a stringent procedure grounding on integration of different methods for expression estimation and differential expression testing, as done in a recent study (Ropka-Molik et al., 2014) focused to muscle tissue gene expression in pigs of different breeds. In the present study, which compares pigs of the same breed and reared under standard conditions, we detected significant gene expression variations. The sensitivity of our approach was supported by the successful validation of all the eleven DEG assayed.

We analyzed the biological functions of genes differentially expressed between FAT and LEAN animals (Figure 5). It is interesting to note that the main differences were found for functional categories of genes related Inflammation and immunity that resulted more expressed in FAT pigs. The genes less expressed in FAT animals include some heat shock protein genes. The biological functions of DEGs show a stronger activation in adipose tissue of FAT pigs of genes for important processes involved in hypertrophy and adipogenesis, such as differentiation and maturation. Supposedly, these biological processes could be altered in adipose tissue of FAT pigs due to dysregulated adipose metabolism and endocrinology similarly to what was hypothesized in humans (Sethi, 2010). On the whole, there is a consistent difference concerning the biological functions characterizing the most expressed

genes on backfat tissue and those of the genes differentially expressed between FAT and LEAN pigs.

Some genes higher expressed in FAT animals could modulate backfat physiological processes

Specific DEGs more expressed in FAT pigs participate to biochemical pathways related to and involved in adipocytes metabolism and adipose tissue physiology. Ectonucleotide pyrophosphatase/phosphodiesterase 1 (*ENPP1*) encodes a catalytic enzyme involved in adipocyte maturation (Liang et al., 2007). Pan et al. (2011) showed that the over-expression of *ENPP1* in a human cell line resulted in adipocyte insulin resistance and demonstrated an association with fatty liver, hyperlipidemia, and dysglycemia. Accordingly, the study of Chandalia et al. (2012) underlined an increased *ENPP1* expression in adipose tissue associated with defective adipocyte maturation leading to pathogenesis of insulin resistance and its associated complications for glucose and lipid metabolism in absence of obesity. In addition, Meyre et al. (2005) reported the presence of three *ENPP1* SNPs in human gene associated with adult obesity and increased risk of glucose intolerance and type 2 diabetes. Furthermore, also the genes acid phosphatase 5, tartrate resistant (*ACP5*) and lysozyme (*LYZ*) that in this research have higher transcriptional level in FAT pigs have been reported to be involved in excessive backfat deposition in pigs and in the development of atherosclerosis (Padilla et al., 2013).

In the present research, some genes overexpressed in the adipose tissue of FAT pigs, namely *STC1, EGR2,* and *INHBB*, are related to adipocyte differentiation and adipocyte maturation. *STC1* (Stanniocalcin 1) has been reported in literature to be up-regulated during adipogenesis and to modulate steroidogenesis. Serlachius and Andersson (2004) related *STC1* up-regulation to the set of survival genes in adipocyte differentiation, which is also associated to overexpression of the anti-apoptotic proteins *BCL2* reported to be involved in inflammation pathway. *EGR2* (early growth response 2) is a direct target of *mir-224-5p*, a negative regulator of adipocyte differentiation that is down regulated during the early process of mouse adipocyte differentiation, and the expression of *EGR2* is increased (Peng et al., 2013). The *INHBB* (Inhibin beta B) gene coding for the activin B subunit is part of the inhibins/activins family of proteins with cytokine and hormone activity. In human and mice,

INHBB has been associated to the physiological and metabolic modifications during adipogenesis when it is highly expressed and is the predominant activin in human adipose tissue (Hoggard et al., 2009). *INHBB* is member of TGF-protein superfamily of secreted growth factors involved in many biological responses including regulation of apoptosis; proliferation and differentiation of human adipocytes; tissue remodeling; and inflammatory immune response (Dani C., 2013). It can be hypothesized that in FAT pigs the pro-adipogenic *INHBB* gene expression increases as it is involved in the differentiation of preadipocytes into mature adipocyte, and that *INHBB* is involved in many physiological processes and including the control of food intake and to energy metabolism through the regulation of hypothalamic and pituitary hormone secretions. Another gene overexpressed in FAT pigs related to feeding and pituitary secretions is *GPNMB* (glycoprotein transmembrane NMB). GPNMB is one of the receptors activated by bombesin-like endogenous peptide ligands, such as gastrin-releasing peptide (*GRP*), neuromedin B (*NMB*) and neuromedin C (*GRP18-27*). These receptors are involved in the regulation of many biological functions including thermoregulation, feeding, pituitary, gastric and pancreatic secretion. The NMB/NMB-R pathway is involved in the regulation of a wide variety of behaviours, such as spontaneous activity, feeding, and anxiety-related behaviour (Yamada et al., 2002).
The *OLR1* (Oxidized low density lipoprotein (lectin-like) receptor 1) gene resulted more expressed in FAT pigs compared to LEAN animals. This gene codes for a LDL receptor that belongs to the C-type lectin superfamily, one of many target genes, including perilipins, of the PPAR signalling, which is involved specifically in lipid metabolism and fatty acids transport. In this way, *OLR1* is a receptor that mediates the recognition, internalization and degradation of oxidatively modified low-density lipoprotein by vascular endothelial cells. *OLR1* removes oxidised low-density lipoproteins from the circulation, as part of lipid metabolism pathways (Mehta et al., 2002).

Genes involved in immunity and inflammation are more expressed in FAT animals

Some other genes overexpressed in FAT pigs are related to immunity. Inflammatory links between human obesity and metabolic diseases are well known mechanisms based on the recruitment of immune cells into adipose tissue (Kabir et al., 2014). The development of a pre-inflammatory condition in presence of dysregulated excessive adipogenesis is

associated with adipose macrophage infiltration and activation. From our study, we can hypothesize a similar process in backfat tissue of FAT pigs where we identified the over expression of the gene macrophage scavenger receptor 1 (*MSR1*), a membrane glycoprotein that in humans is involved in the pathologic deposition of cholesterol in arterial walls during atherogenesis (Haasken et al., 2013). Additionally, the overexpression of secreted phosphoprotein 1 (*SPP1*) in FAT pigs can suggest the hypothesis that this gene is acting as a proinflammatory cytokine that promotes monocyte chemotaxis and cell motility and might link, in pigs like in mice, fat accumulation to the development of insulin resistance by sustaining inflammation and the accumulation of macrophages in adipose tissue (Nomiyana et al. 2007). Interestingly, a porcine *SPP1* gene polymorphism was associated to backfat thickness in the Landrace × Jeju (Korea) Black pig F2 population (Han et al., 2012). *SPP1* might play a key role in the pathway that leads to type I immunity enhancing interferon-gamma and interleukin-12 production and suppressing interleukin-10 (Ashkar et al., 2000). Therefore, these data allow hypothesizing *SPP1* as a gene associated, in pigs like to in human, to the link between obesity, adipose tissue inflammation, and insulin resistance. In addition, phosphoinositide-3-kinase adaptor protein 1 (*PIK3AP1*), higher expressed in FAT pigs, is a positive regulator of phosphatidylinositol 3-kinase (*PI3K*) signalling. *PI3K* signalling pathway has a key role in the insulin-dependent regulation of adipocyte metabolism (glucose and lipid metabolism). Besides, *PI3K* participate in obesity-associated inflammatory cell recruitment (neutrophils and macrophages), as well as in the CNS-dependent neurohumoral regulation of food intake/energy expenditure (McCurdy and Clemm, 2013; Beretta et al., 2015).

Other genes found in the present research and related to inflammatory condition of the adipose tissue in FAT pigs are particularly interesting to mention. CD163, member of the scavenger receptor cysteine-rich superfamily (Guo et al., 2014; Smith et al., 2014); solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1 (*SLC11A1*), a gene involved in the resistance to *Salmonella* infection (Kommadath et al., 2014) as well as the chemokine (C-C motif) receptor 1 (*CCR1*), that was previously found overexpressed in obese pigs (Kogelman et al., 2014); BCL2-related protein A1 (*BCL2A1*), a gene found to be overexpressed in pigs with an high obesity index and that is related to immunity, inflammatory pathway, and osteoclast differentiation (Kogelman et al., 2014); CD1a molecule (*CD1A,* indicated as *PCD1A* on the cited paper), a surface antigen involved in

immunity was found to be overexpressed in obese pigs by Kogelman et al. (2014). The same Authors highlighted a strong connection between fat deposition on the body (obesity), immunity and bone development. They also indicated that *CCR1* gene is a strong candidate regulator of immune response as it is a receptor of pro-inflammatory chemokines in adipose tissue playing a pivotal role in obesity-associated diseases (Kabir et al. 2014; Lumeng and Saltiel, 2011).

Heat shock response, protein folding and repair are impaired in FAT animals

Considering the 14 genes less expressed in FAT animals, direct relationships with lipid metabolism are not apparent. However, the "unfolded protein binding" function is enriched among these genes, which include five functionally linked heat shock proteins (*DNAJB1, HSPA1A, HSPA1B, HSPA2* and *HSPB8*). Heat shock proteins are involved in stabilization of existing proteins against aggregation, mediating the folding of newly translated proteins in the cytosol and in organelles, and also in the ubiquitin-proteasome pathway. *DNAJB1*, a member of the Hsp40 family, is a molecular chaperon involved in protein folding and protein complex assembly. *DNAJB1,* a member of the Hsp40 family, promotes protein folding and prevent misfolded protein aggregation, as *HSPB8*, a member of the Hsp20 family, does (Vicario et al., 2014). *DNAJB1* also stimulates the ATPase activity of protein of the Hsp70 family to which other genes less expressed in FAT pigs (*HSPA1A*, *HSPA1B*, and *HSPA2)* belong, indicating a possible functional link between these four genes. Our results suggest a general impairment of the protein folding and repair in the fattest animals, in accordance to previous observations of studies carried out on human obesity. Obesity is a pathological human condition in which a chronically positive energy balance induces in adipocytes, the cells in charge to store the excess of energy in fat depots, a persistent stress activating in turn defence processes as autophagy or apoptosis.
As reviewed by Newsholme and de Bittencourt (2014), if the heat shock response, a key component of the physiological response to resolve inflammation, is hampered in adipose tissue, the adipocyte metabolic stress triggers fat cell senescence with reduction of the heat shock proteins activity. In this condition, the advance of inflammasome mediated secretory activity from adipose to other tissues promotes cellular senescence in many other cells of the organism, aggravating obesity-dependent chronic inflammation. This mechanism could

have been activated also in the FAT pigs of our experiment (Figure 5) due to a genetic aptitude of the fattest animals toward a higher fat deposition and adiposity similar to obesity. Indeed, a decrease in the synthesis of the mRNAs of the heat shock proteins and an increase of the expression of many genes related to an inflammatory status and to immune response is a characteristic of the fattest pigs. Increase of the expression of *INHBB* and *SPP1* denotes for instance the augmented production of cytokines and the higher expression of *ENPP1* and *PIK3AP1* may indicate a status of insulin resistance, one of the typical signals connected with obesity.

Pig backfat deposition and impaired stress response may activate inflammation

Our results agree with recent studies showing that several immune system and anti-inflammatory processes are activated and play a critical role in the response to fat accumulation in porcine backfat tissue (Sodhi et al., 2014) and in visceral fat tissue (Toedebusch et al., 2014; Wang et al., 2014). Wang et al. (2014) and Zhou et al. (2013) used three female Landrace pigs to identify DEG between subcutaneous, visceral and intramuscular fat indicating that visceral and intramuscular adipose tissues were mainly associated with inflammatory features of the tissue and immune response. Our data suggest that also in backfat a predominant role of immunity processes is related to an increased adipose tissue deposition.

The results obtained seem to sustain the hypothesis that the high fat accumulation in adipose tissue of pigs can determine the development of an inflammatory process producing a cascade of defence and adaptive reactions in the tissue, such as activation of immune system and mesenchymal cells differentiation in adipocytes.

A deeper knowledge of the metabolic processes involved in fat deposition can be very important to develop the use the pig as model species to study obesity and related disorders for humans because of similar anatomy and physiology (Spurlock and Gabler, 2008; Litten-Brown et al., 2010; Varga et al., 2010) and considering the above described similarities between pigs and humans.

In order to fully elucidate the complex gene network regulating backfat deposition on pigs, it will be important to extend the basic knowledge by further coding and non-coding transcriptome characterization. Additional information would probably come from studying

interactions between the differentially expressed long RNAs identified in the present paper and the regulatory microRNAs expressed in porcine adipose tissue identified on some of the same animals (Gaffo et al., 2014).

The results of the present work unlock the opportunity that some of the identified differentially expressed genes might be used as biomarkers (Ibáñez-Escriche et al., 2014) to improve carcass fat traits in to look for SNPs regulating their expression to be included in selection schemes to make more sustainable the pig production chain.

Acknowledgements

We thank Dr. Maurizio Gallo from Associazione Nazionale Allevatori Suini (ANAS) and Dr. Luca Buttazzoni (CREA-PCM) for providing the porcine backfat tissue samples, BMR genomics for performing the RNA-Seq. This work was supported by Progetto "AGER - Agroalimentare e ricerca": Advanced research in genomics and processing technologies for the Italian heavy pig production – Hepiget (Grant N. 2011- 0279).

References

Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11:R106.

Anders S, Pyl PT & Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.

Ashkar S, Weber GF, Panoutsakopoulou V, Sanchirico ME, Jansson M, Zawaideh S, Rittling SR, Denhardt DT, Glimcher MJ & Cantor H (2000) Eta-1 (osteopontin): an early component of type-1 (cell-mediated) immunity. *Science* 287:860–864.

Beretta M, Bauer M & Hirsch E (2015) PI3K signaling in the pathogenesis of obesity: The cause and the cure. *Adv Biol Regul* 58:1-15.

Bosi P & Russo V (2004) The production of the heavy pig for high quality processed products. *Ital J Anim Sci* 3:309–321.

Bustin SA & Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 15:155–166.

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A & Rinn JL (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 25:1915–1927.

Čandek-Potokar M & Škrlep M (2012) Factors in pig production that impact the quality of dry-cured ham: a review. *Animal* 6:327–338.

Chandalia M, Davila H, Pan W, Szuszkiewicz M, Tuvdendorj D, Livingston EH & Abate N (2012) Adipose tissue dysfunction in humans: a potential role for the transmembrane protein ENPP1. *J Clin Endocrinol Metab* 97:4663–4672.

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

Chen CH, Lin EC, Cheng WTK, Sun HS, Mersmann HJ & Ding ST (2006) Abundantly expressed genes in pig adipose tissue: an expressed sequence tag approach. *J Anim Sci* 84:2673–2683.

Conte M et al. (2013) Increased Plin2 expression in human skeletal muscle is associated with sarcopenia and muscle weakness. *PLOS ONE* 8:e73709.

Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Estellé J, Castelló A, Alves E, Pena RN, Ballester M & Folch JM (2013) Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics* 14:843.

Cox MP, Peterson DA & Biggs PJ (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11:485.

Dani C (2013) Activins in adipogenesis and obesity. *Int J Obes (Lond)* 37:163–166.

Davoli R, Gandolfi G, Braglia S, Comella M, Zambonelli P, Buttazzoni L & Russo V (2010) New SNP of the porcine Perilipin 2 (PLIN2) gene, association with carcass traits and expression analysis in skeletal muscle. *Mol Biol Rep* 38:1575–1583.

Durinck S, Spellman PT, Birney E & Huber W (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 4:1184–1191.

Djebali S et al. (2012) Landscape of transcription in human cells. *Nature* 489:101–108.

Ferraz ALJ, Ojeda A, López-Béjar M, Fernandes LT, Castelló A, Folch JM & Pérez-Enciso M (2008) Transcriptome architecture across tissues in the pig. *BMC Genomics* 9:173.

Gaffo E, Zambonelli P, Bisognin A, Bortoluzzi S & Davoli R (2014) miRNome of Italian Large White pig subcutaneous fat tissue: new miRNAs, isomiRs and moRNAs. *Anim Genet* 45:685–698.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V & Davoli R (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631–637.

Gorodkin J et al. (2007) Porcine transcriptome analysis based on 97 non-normalized cDNA libraries and assembly of 1,021,891 expressed sequence tags. *Genome Biology* 8:R45.

Guo L, Niu J, Yu H, Gu W, Li R, Luo X, Huang M, Tian Z, Feng L & Wang Y (2014) Modulation of CD163 expression by metalloprotease ADAM17 regulates porcine reproductive and respiratory syndrome virus entry. *J Virol* 88:10448–10458.

Haasken S, Auger JL, Taylor JJ, Hobday PM, Goudy BD, Titcombe PJ, Mueller DL & Binstadt BA (2013) Macrophage scavenger receptor 1 (Msr1, SR-A) influences B cell autoimmunity by regulating soluble autoantigen concentration. *J Immunol* 191:1055–1062.

Han SH, Shin KY, Lee SS, Ko MS, Oh HS & Cho IC (2012) Porcine SPP1 gene polymorphism association with phenotypic traits in the Landrace × Jeju (Korea) Black pig F2 population. *Mol Biol Rep* 39:7705–7709.

Henderson CR & Quaas RL (1976) Multiple Trait Evaluation Using Relatives' Records. *J Anim Sci* 43:1188–1197.

Hoggard N, Cruickshank M, Moar KM, Barrett P, Bashir S & Miller JDB (2009) Inhibin betaB expression in murine adipose tissue and its regulation by leptin, insulin and dexamethasone. *J Mol Endocrinol* 43:171–177.

Hornshøj H, Conley LN, Hedegaard J, Sørensen P, Panitz F & Bendixen C (2007) Microarray Expression Profiles of 20.000 Genes across 23 Healthy Porcine Tissues. *PLOS ONE* 2:e1203.

Huang DW, Sherman BT & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.

Ibáñez-Escriche N, Forni S, Noguera JL & Varona L (2014) Genomic information in pig breeding: Science meets industry needs. *Livest Sci* 166:94–100.

Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, Poliakov A, Cao X, Dhanasekaran SM, Wu YM, Robinson DR, Beer DG, Feng FY, Iyer HK & Chinnaiyan AM (2015) The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* 47:199–208.

Jiang S, Wei H, Song T, Yang Y, Peng J & Jiang S (2013) Transcriptome Comparison between Porcine Subcutaneous and Intramuscular Stromal Vascular Cells during Adipogenic Differentiation. *PLOS ONE* 8:e77094.

Kabir SM, Lee ES & Son DS (2014) Chemokine network during adipogenesis in 3T3-L1 cells: Differential response between growth and proinflammatory factor in preadipocytes vs. adipocytes. *Adipocyte* 3:97–106.

Kershaw EE & Flier JS (2004) Adipose tissue as an endocrine organ. *J Clin Endocrinol Metab* 89:2548–2556.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R & Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.

Kogelman LJA, Cirera S, Zhernakova DV, Fredholm M, Franke L & Kadarmideen HN (2014) Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. *BMC Med Genomics* 7:57.

Kommadath A, Bao H, Arantes AS, Plastow GS, Tuggle CK, Bearson SMD, Guan LL & Stothard P (2014) Gene co-expression network analysis identifies porcine genes associated with variation in Salmonella shedding. *BMC Genomics* 15:452.

Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L & Gao G (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucl Acids Res* 35:W345–349.

Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G & Durbin R (2009) 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Li XJ, Yang H, Li GX, Zhang GH, Cheng J, Guan H & Yang GS (2012) Transcriptome profile analysis of porcine adipose tissue by high-throughput sequencing. *Anim Genet* 43:144–152.

Liang J, Fu M, Ciociola E, Chandalia M & Abate N (2007) Role of ENPP1 on Adipocyte Maturation. *PLOS ONE* 2:e882.

Litten-Brown JC, Corson AM & Clarke L (2010) Porcine models for the metabolic syndrome, digestive and bone disorders: a general overview. Animal 4, 899–920.

Lumeng CN & Saltiel AR (2011) Inflammatory links between obesity and metabolic disease. *J Clin Invest* 121:2111–2117.

McCurdy CE & Klemm DJ (2013) Adipose tissue insulin sensitivity and macrophage recruitment: Does PI3K pick the pathway? *Adipocyte* 2:135–142.

Mehta JL & Li D (2002) Identification, regulation and function of a novel lectin-like oxidized low-density lipoprotein receptor. *J Am Coll Cardiol* 39:1429–1435.

Meyre D et al. (2005) Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nat Genet* 37:863–867.

Mikawa A, Suzuki H, Suzuki K, Toki D, Uenishi H, Awata T & Hamasima N (2004) Characterization of 298 ESTs from porcine back fat tissue and their assignment to the SSRH radiation hybrid map. *Mamm Genome* 15:315–322.

Moon JK, Kim KS, Kim JJ, Choi BH, Cho BW, Kim TH & Lee CK (2009) Differentially expressed transcripts in adipose tissue between Korean native pig and Yorkshire breeds. *Anim Genet* 40:115–118.

Moreno-Navarrete JM & Fernández-Real JM (2012) Adipocyte Differentiation. In *Adipose Tissue Biology.* [Symonds ME editor], pp. 17-38. Springer, New York

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R & Schäffer AA (2008) Database indexing for production MegaBLAST searches. *Bioinformatics* 24:1757–1764.

Newsholme P & de Bittencourt PIH (2014) The fat cell senescence hypothesis: a mechanism responsible for abrogating the resolution of inflammation in chronic disease. *Curr Opin Clin Nutr Metab Care* 17, 295–305.

Padilla J, Jenkins NT, Lee S, Zhang H, Cui J, Zuidema MY, Zhang C, Hill MA, Perfield JW, Ibdah JA, Booth FW, Davis JW, Laughlin MH & Rector RS (2013) Vascular transcriptional alterations produced by juvenile obesity in Ossabaw swine. *Physiol Genomics* 45:434–446.

Pan W, Ciociola E, Saraf M, Tumurbaatar B, Tuvdendorj D, Prasad S, Chandalia M & Abate N (2011) Metabolic consequences of ENPP1 overexpression in adipose tissue. Am. J. Physiol. *Endocrinol Metab* 301:E901–911.

Peng Y, Xiang H, Chen C, Zheng R, Chai J, Peng J & Jiang S (2013) MiR-224 impairs adipocyte early differentiation and regulates fatty acid metabolism. *Int J Biochem Cell Biol* 45:1585–1593.

Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.

Roberts A, Pimentel H, Trapnell C & Pachter L (2011a) Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* 27:2325–2329.

Roberts A, Trapnell C, Donaghey J, Rinn JL & Pachter L (2011b) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12:R22.

Ropka-Molik K, Zukowski K, Eckert R, Gurgul A, Piórkowska K & Oczkowicz M (2014) Comprehensive analysis of the whole transcriptomes from two different pig breeds using RNA-Seq method. *Anim Genet* 45:674–684.

Rosen ED & MacDougald OA (2006) Adipocyte differentiation from the inside out. *Nat Rev Mol Cell Biol* 7:885–896.

Russell TD, Palmer CA, Orlicky DJ, Bales ES, Chang BHJ, Chan L & McManaman JL (2008) Mammary glands of adipophilin-null mice produce an amino-terminally truncated form of adipophilin that mediates milk lipid droplet formation and secretion. *J Lipid Res* 49:206–216.

Russo V, Buttazzoni L, Baiocco C, Davoli R, Nanni Costa L, Schivazappa C & Virgili R (2000) Heritability of muscular cathepsin B activity in Italian Large White pigs. *J Anim Breed Genet* 117:37–42.

Russo V, Fontanesi L, Scotti E, Beretti F, Davoli R, Nanni Costa L, Virgili R & Buttazzoni L (2008) Single nucleotide polymorphisms in several porcine cathepsin genes are associated with growth, carcass, and production traits in Italian Large White pigs. *J Anim Sci* 86:3300–3314.

Serlachius M & Andersson LC (2004) Upregulated expression of stanniocalcin-1 during adipogenesis. Exp. Cell Res. 296, 256–264.Sethi, J.K. (2010) Activatin' human adipose progenitors in obesity. *Diabetes* 59:2354–2357.

Sethi JK (2010) Activatin' Human Adipose Progenitors in Obesity. *Diabetes* 59:2354–2357.

Smith SH, Wilson AD, Van Ettinger I, MacIntyre N, Archibald AL & Ait-Ali T (2014) Down-regulation of mechanisms involved in cell transport and maintenance of mucosal integrity in pigs infected with Lawsonia intracellularis. *Vet Res* 45:55.

Sodhi SS, Song KD, Ghosh M, Sharma N, Lee SJ, Kim JH, Kim N, Mongre RK, Adhikari P, Kim JY, Hong SP, Oh SJ & Jeong DK (2014) Comparative transcriptomic analysis by RNA-seq to discern differential expression of genes in liver and muscle tissues of adult Berkshire and Jeju Native Pig. *Gene* 546:233–242.

Spurlock ME & Gabler NK (2008) The development of porcine models of obesity and the metabolic syndrome. *J Nutr* 138:397–402.

Toedebusch RG, Roberts MD, Wells KD, Company JM, Kanosky KM, Padilla J, Jenkins NT, Perfield JW, Ibdah JA, Booth FW & Rector RS (2014) Unique transcriptomic signature of omental adipose tissue in Ossabaw swine: a model of childhood obesity. *Physiol Genomics* 46:362–375.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL & Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7:562–578.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ & Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515.

Trayhurn P (2005) Endocrine and signalling role of adipose tissue: new perspectives on fat. *Acta Physiol Scand* 184:285–293.

Uenishi H, Eguchi T, Suzuki K, Sawazaki T, Toki D, Shinkai H, Okumura N, Hamasima N & Awata T (2004) PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries. *Nucl Acids Res* 32:D484–D488.

Uenishi H, Eguchi-Ogawa T, Shinkai H, Okumura N, Suzuki K, Toki D, Hamasima N & Awata T (2007) PEDE (Pig EST Data Explorer) has been expanded into Pig Expression Data Explorer, including 10 147 porcine full-length cDNA sequences. *Nucl Acids Res* 35:D650–D653.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034.

Varga O, Harangi M, Olsson IS & Hansen AK (2010) Contribution of animal models to the understanding of the metabolic syndrome: a systematic overview. *Obes Rev* 11:792–807.

Vicario M, Skaper SD, Negro A (2014) The small heat shock protein HspB8: role in nervous system physiology and pathology. *CNS Neurol Disord Drug Targets* 13:885–895.

Wang RN, Green J, Wang Z, Deng Y, Qiao M, Peabody M, Zhang Q, Ye J, Yan Z, Denduluri S, Idowu O, Li M, Shen C, Hu A, Haydon RC, Kang R, Mok J, Lee MJ, Luu HL & Shi LL (2014) Bone Morphogenetic Protein (BMP) signaling in development and human diseases. *Genes Dis* 1:87–105.

Yamada K, Santo-Yamada Y & Wada K (2002) Restraint stress impaired maternal behavior in female mice lacking the neuromedin B receptor (NMB-R) gene. *Neurosci Lett* 330:163–166.

Zhou C, Zhang J, Ma J, Jiang A, Tang G, Mai M, Zhu L, Bai L, Li M & Li X (2013) Gene expression profiling reveals distinct features of various porcine adipose tissues. *Lipids Health Dis* 12:75.

Zhou ZY, Li AM, Adeola AC, Liu YH, Irwin DM, Xie HB & Zhang YP (2014) Genome-Wide Identification of Long Intergenic Noncoding RNA Genes and Their Potential Association with Domestication in Pigs. *Genome Biol Evol* 6:1387–1392.

Table 1 - Genetic indexes and phenotypes for BFT and hot carcass weight of the pigs selected for the transcriptome analysis.

Group	Sample ID	Sex	Day of slaughter	Slaughter weight (kg) (*)	BFT phenotype (mm)	BFT EBV		
							Mean	SD
FAT	477	M	6	120	43	7.36	5.22	1.3
	476	F	6	119	37	7.17		
	474	M	2	113	38	6.03		
	482	F	9		42	5.75		
	478	F	7	118	33	5.05		
	516	F	3	115	36	4.88		
	479	M	8		41	4.76		
	483	F	10	119	38	4.41		
	489	M	18	108	35	3.54		
	484	M	15	128	35	3.27		
LEAN	490	M	19	113	24	-6.46	-8.63	1.4
	473	F	2	132	23	-7.54		
	487	M	18	110	23	-7.61		
	517	M	4	117	20	-7.71		
	485	F	17	126	20	-7.82		
	475	M	5	119	20	-8.03		
	481	M	9		22	-9.91		
	486	F	17	123	19	-10.27		
	488	F	18	128	19	-10.37		
	480	F	9		16	-10.59		

EBV: estimated breeding value

BFT: backfat thickness.

(*) slaughter weight: the hot carcass slaughter weight is reported. For four animals the weight is not available due to a problem of the automatic recording system at the slaughterhouse.

Table 2. List of the DE genes and transcripts.

Cufflinks transcript ID	Cufflinks gene ID	Gene locus	Gene symbol	Cuffdiff2 gene log2 FC FAT *vs*. LEAN	Transcript group	Coding potential
TCONS_00102010	XLOC_040987	JH118612.1:113132-140205	*DSC2*	2.55	Known	-
TCONS_00061823	XLOC_023331	4:78928264-78930654	-	2.46	New	NON CODING
TCONS_00033774	XLOC_013001	15:140797584-140847461	*NYAP2*	2.38	New	CODING
TCONS_00061359	XLOC_023211	4:35670339-35685878	*DCSTAMP*	2.23	Novel isoform	CODING
TCONS_00095554	XLOC_036823	GL893451.1:11131-27485	*CRLF2*	2.21	Known	-
TCONS_00093244	XLOC_035190	9:50996895-51001264	-	2.17	New	NON CODING
TCONS_00087029	XLOC_032796	8:140307937-140315415	*SPP1*	2.09	Known	-
TCONS_00003007	XLOC_000806	1:283547172-283552108	-	2.07	New	CODING
TCONS_00095549	XLOC_036822	GL893451.1:7060-10625	-	2.03	New	NON CODING
TCONS_00067029	XLOC_025404	5:36179189-36186325	*LYZ*	2.03	Known	-
TCONS_00042581	XLOC_016514	18:6731368-6733669	*GIMAP2*	1.98	Known	-
TCONS_00061600	XLOC_023265	4:55660234-55715444	*ATP6V0D2*	1.96	Novel isoform	CODING
TCONS_00039556	XLOC_015432	17:53815353-53827092	*MMP9*	1.92	Known	-
TCONS_00039900	XLOC_015518	17:4110395-4192029	*MSR1*	1.92	Known	-
TCONS_00061643	XLOC_023283	4:62172539-62226917	*STMN2*	1.85	Known	-
TCONS_00034645	XLOC_013236	15:62409564-62414328	-	1.84	New	RELIABLE NON CODING
TCONS_00091509	XLOC_034399	9:63158999-63198155	*ST14*	1.79	Novel isoform	CODING
TCONS_00098750	XLOC_038994	GL895411.1:0-1073	*INHBB*	1.65	New	CODING
TCONS_00022322	XLOC_008474	13:32323641-32330286	*CCR1*	1.63	Known	-
TCONS_00044383	XLOC_017319	2:11807281-11850646	*MPEG1*	1.63	Known	-
TCONS_00075056	XLOC_028007	6:70039585-70099223	*PADI2*	1.6	Known	-
TCONS_00095875	XLOC_037025	GL893645.1:0-307	-	1.57	New	RELIABLE NON CODING
TCONS_00084869	XLOC_032187	8:71288921-71302169	*AMBN*	1.56	Known	-
TCONS_00033691	XLOC_012975	15:133452328-133456736	*SLC11A1*	1.56	Known	-

TCONS_00089513	XLOC_033895	9:90266412-90348498	*SCIN*	1.55	Known	-
TCONS_00042660	XLOC_016535	18:8306789-8313120	-	1.52	New	CODING
TCONS_00059834	XLOC_022860	4:99905518-99915176	*CD1A*	1.52	Novel isoform	CODING
TCONS_00059837	XLOC_022860	4:99905518-99915176	*CD1A*	1.52	Known	-
TCONS_00093519	XLOC_035465	9:101443296-101443885	*GPNMB*	1.46	New	NON CODING
TCONS_00098157	XLOC_038614	GL894967.1:126-517	*GPNMB*	1.42	New	CODING
TCONS_00018804	XLOC_007247	12:23439824-23441829	-	1.4	New	CODING
TCONS_00103084	XLOC_041497	X:37303173-37393818	*CYBB*	1.38	Known	-
TCONS_00065337	XLOC_024931	5:52504178-52625145	*BCAT1*	1.37	Novel isoform	CODING
TCONS_00098113	XLOC_038589	GL894923.1:47-563	*GPNMB*	1.36	New	CODING
TCONS_00002441	XLOC_000664	1:227333991-227356844	*PLIN2*	1.32	Novel isoform	CODING
TCONS_00044392	XLOC_017322	2:12191483-12243400	*LPXN*	1.31	Known	-
TCONS_00084565	XLOC_032101	8:33970571-33982450	*UCHL1*	1.27	Novel isoform	CODING
TCONS_00067389	XLOC_025495	5:64579162-64590512	*OLR1*	1.26	Known	-
TCONS_00059747	XLOC_022835	4:97720982-97736619	*CD48*	1.25	Known	-
TCONS_00028769	XLOC_011055	14:143745489-143752509	*GMFG*	1.23	Known	-
TCONS_00029056	XLOC_011139	14:8804077-8816800	*STC1*	1.23	Novel isoform	CODING
TCONS_00098643	XLOC_038938	GL895339.1:13269-61205	*COTL1*	1.15	Known	-
TCONS_00100592	XLOC_040068	GL896326.1:1999-3913	*ACP5*	1.13	Known	-
TCONS_00096837	XLOC_037668	GL894123.1:0-400	*CD163*	1.13	New	CODING
TCONS_00097297	XLOC_037990	GL894401.1:0-471	*CD163*	1.13	New	CODING
TCONS_00005002	XLOC_001331	1:125897935-125953413	*AQP9*	1.09	Known	-
TCONS_00096863	XLOC_037686	GL894145.1:0-401	*CD163*	1.09	New	CODING
TCONS_00071337	XLOC_027094	6:74616232-74621248	*C1QC*	1.08	Known	-
TCONS_00012469	XLOC_005058	11:21534980-21685851	*LCP1*	1.07	Novel isoform	CODING
TCONS_00079920	XLOC_030238	7:94900207-94906867	*AKAP5, LOC100153460*	1.06	Novel isoform	CODING

TCONS_00041537	XLOC_016257	18:6613761-6621027	*GIMAP4*	1.06	Known	-
TCONS_00097908	XLOC_038444	GL894747.1:3047-10617	*HMOX1*	1.06	Novel isoform	CODING
TCONS_00030401	XLOC_011444	14:71516962-71521335	*EGR2*	1.05	Known	-
TCONS_00030878	XLOC_011579	14:117265093-117349965	*BLNK*	1.04	Known	-
TCONS_00056578	XLOC_021190	3:77408776-77439119	*PLEK*	1.04	Known	-
TCONS_00071335	XLOC_027093	6:74609911-74612993	*C1QA*	1.02	Known	-
TCONS_00081915	XLOC_030757	7:54395230-54406136	*BCL2A1*	1.01	Known	-
TCONS_00041554	XLOC_016261	18:6872940-6875292	*GIMAP1*	1	Known	-
TCONS_00085005	XLOC_032236	8:79743274-79751980	*SFRP2*	0.99	Known	-
TCONS_00098919	XLOC_039115	GL895590.1:0-1327	*GPNMB*	0.91	New	NON CODING
TCONS_00068526	XLOC_026077	5:52625315-52630242	*BCAT1*	0.89	New	RELIABLE NON CODING
TCONS_00062055	XLOC_023401	4:97099149-97103132	*FCER1G*	0.87	Known	-
TCONS_00009719	XLOC_003695	10:48841010-48961015	*MRC1*	0.86	Novel isoform	CODING
TCONS_00030894	XLOC_011584	14:117670639-117938624	*PIK3AP1*	0.85	Known	-
TCONS_00017526	XLOC_006800	12:36561025-36604089	*CLTC*	0.8	Novel isoform	CODING
TCONS_00062959	XLOC_023614	4:119674090-119703427	*CD53*	0.78	Known	-
TCONS_00081898	XLOC_030753	7:53623061-53644262	*CTSH*	0.78	Known	-
TCONS_00060570	XLOC_023035	4:119013307-119039899	*ADORA3*	0.74	Known	-
TCONS_00052401	XLOC_020144	3:11035819-11055510	*LAT2*	0.71	Known	-
TCONS_00004118	XLOC_001095	1:35133812-35137388	*CTGF*	0.68	Known	-
TCONS_00045043	XLOC_017499	2:59214054-59218018	*IFI30*	0.65	Known	-
TCONS_00004124	XLOC_001096	1:35240242-35281384	*ENPP1*	0.62	Known	-
TCONS_00062884	XLOC_023592	4:116704501-116707235	*OLFML3*	-0.54	Known	-
TCONS_00035484	XLOC_013426	15:131680309-131684630	*IGFBP5*	-0.65	Known	-
TCONS_00101718	XLOC_040809	JH118426.1:306724-312138	-	-0.77	New	RELIABLE NON CODING
TCONS_00063805	XLOC_024145	4:77261119-77264781	-	-0.77	New	NON CODING
TCONS_00050164	XLOC_018733	2:124815021-124828122	*CDO1*	-0.9	Novel isoform	CODING

TCONS_00101559	XLOC_040715	GL896532.1:212-2567	*ADSSL1*	-1.02	New	NON CODING
TCONS_00079927	XLOC_030240	7:94987617-94990126	*HSPA2*	-1.1	Known	-
TCONS_00083805	XLOC_031620	7:66542203-66555641	-	-1.18	New	CODING
TCONS_00041725	XLOC_016313	18:15292592-15295178	-	-1.61	New	CODING
TCONS_00048853	XLOC_018425	2:65175406-65180520	*DNAJB1*	-1.66	Novel isoform	CODING
TCONS_00029533	XLOC_011248	14:35688332-35701411	*HSPB8*	-1.81	Known	-
TCONS_00094194	XLOC_036009	GL892492.1:0-3540	*HSPA1B*	-2.32	New	NON CODING
TCONS_00101505	XLOC_040677	GL896522.1:9039-10877	*HSPA1A*	-2.57	New	RELIABLE NON CODING
TCONS_00098059	XLOC_038555	GL894890.1:5-696	*HSP70*	-3.44	New	NON CODING

Table 3. David functional annotation clustering obtained considering the significant Biological Processes GO terms (Benjamini adjusted *P* values <0.05) of genes more expressed in FAT than in LEAN animals.

Annotation Cluster 1	Enrichment Score: 7.0	
Term	Count	Genes
GO:0006954~inflammatory response	12	*C1QA, SLC11A1, CYBB, ADORA3, OLR1, HMOX1, CCR1, LYZ, C1QC, BLNK, CD163, SPP1*
GO:0006952~defense response	15	*ADORA3, OLR1, CCR1, LYZ, COTL1, C1QC, CD163, INHBB, CD48, C1QA, SLC11A1, CYBB, HMOX1, SPP1, BLNK*
GO:0009611~response to wounding	14	*ADORA3, PLEK, OLR1, CCR1, LYZ, C1QC, CD163, C1QA, SLC11A1, CYBB, CTGF, HMOX1, SPP1, BLNK*
GO:0009605~response to external stimulus	17	*ADORA3, PLEK, OLR1, CCR1, LYZ, C1QC, CD163, INHBB, C1QA, SLC11A1, CYBB, CTGF, SFRP2, HMOX1, STC1, SPP1, BLNK*
GO:0050896~response to stimulus	29	*ADORA3, AQP9, ENPP1, CCR1, UCHL1, ACP5, C1QC, CD48, SLC11A1, PLIN2, CTGF, HMOX1, FCER1G, BLNK, SPP1, EGR2, OLR1, PLEK, LYZ, CD1A, COTL1, CD163, INHBB, C1QA, CYBB, LAT2, SFRP2, STC1, LCP1*
GO:0006950~response to stress	19	*ADORA3, AQP9, PLEK, OLR1, CCR1, UCHL1, LYZ, COTL1, C1QC, CD163, INHBB, CD48, C1QA, SLC11A1, CYBB, CTGF, HMOX1, SPP1, BLNK*
Annotation Cluster 2	Enrichment Score: 2.7	
Term	Count	Genes
GO:0001775~cell activation	7	*CD48, SLC11A1, LAT2, PLEK, LCP1, BLNK, GIMAP1*
GO:0002274~myeloid leukocyte activation	4	*CD48, SLC11A1, LAT2, GIMAP1*
GO:0046649~lymphocyte activation	6	*CD48, SLC11A1, LAT2, LCP1, BLNK, GIMAP1*
Annotation Cluster 3	Enrichment Score: 2.4	
Term	Count	Genes
GO:0048583~regulation of response to stimulus	10	*C1QA, SLC11A1, LAT2, PLEK, ENPP1, HMOX1, FCER1G, C1QC, SPP1, GIMAP1*
GO:0050776~regulation of immune response	7	*C1QA, SLC11A1, LAT2, HMOX1, FCER1G, C1QC, GIMAP1*
GO:0050778~positive regulation of immune response	6	*C1QA, SLC11A1, LAT2, FCER1G, C1QC, GIMAP1*
GO:0002443~leukocyte mediated immunity	5	*C1QA, SLC11A1, LAT2, FCER1G, C1QC*
GO:0002682~regulation of immune system process	8	*C1QA, SLC11A1, LAT2, HMOX1, SCIN, FCER1G, C1QC, GIMAP1*
Annotation Cluster 4	Enrichment Score: 2.0	
Term	Count	Genes
GO:0060348~bone development	6	*AMBN, CTGF, ACP5, STC1, GPNMB, SPP1*
GO:0031214~biomineral formation	4	*AMBN, ENPP1, GPNMB, SPP1*
GO:0001503~ossification	5	*AMBN, CTGF, STC1, GPNMB, SPP1*

GO:0001501~skeletal system development	7	*AMBN, CTGF, MMP9, ACP5, STC1, GPNMB, SPP1*
Annotation Cluster 5	Enrichment Score: 1.6	
Term	Count	Genes
GO:0001775~cell activation	7	*CD48, SLC11A1, LAT2, PLEK, LCP1, BLNK, GIMAP1*

Figures

Figure 1. Transcripts and isoforms classification.

(A) Expressed transcript were classified, according to current gene annotations, into 8 types, reported with different colors (see legend) and grouped in three categories: K (known) collects transcripts found in reference annotation (yellow); I (isoform) collects alternative forms of transcripts (red shades); N collects new transcripts from not-annotated loci (green shades). The pie chart shows the number of transcripts detected, for each type, and their mutual proportions. Three transcript types of the N group have few elements (43 intronic; 5 possible polymerase run-on fragments; 3 transcript intron overlap a reference intron on the opposite strand) and are barely visible in the chart. (B) Transcript length distributions in the three categories. (C) Transcript expression level distribution for the three categories. (D) Number of genes (vertical axis) with their number of transcript isoforms detected (horizontal axis). Genes with only one transcript isoforms detected are the most frequent; however, genes with up to 31 different isoforms were detected. (E) The proportion of each transcript type for the transcript isoforms grouped as in (D). Genes with only one isoform (first bar) are mainly intergenic genes (green part). For genes having more than one isoform expressed, the proportion of novel isoforms detected increases along with the number of different isoforms for a gene (red part).

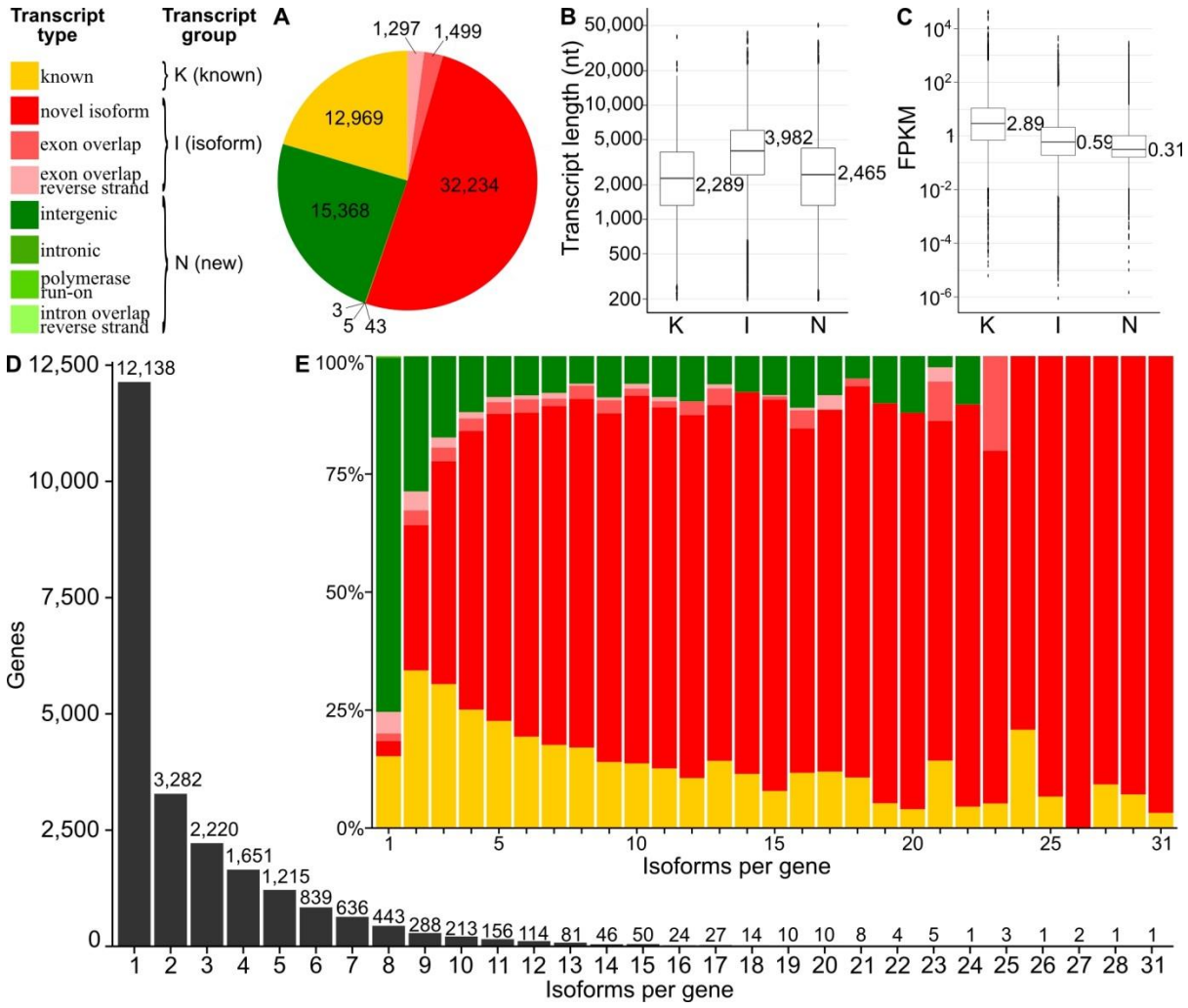

Transcript type
Transcript group
known
K (known)
novel isoform
I (isoform)
exon overlap
exon overlap reverse strand
intergenic
intronic
polymerase run-on
intron overlap reverse strand
N (new)
A
1,297
1,499
12,969
32,234
15,368
3
5
43
B
Transcript length (nt)
50,000
20,000
10,000
5,000
2,000
1,000
500
200
2,289
3,982
2,465
K
I
N
C
FPKM
10^4
10^2
1
10^{-2}
10^{-4}
10^{-6}
2.89
0.59
0.31
K
I
N
D
Genes
12,500
10,000
7,500
5,000
2,500
0
12,138
3,282
2,220
1,651
1,215
839
636
443
288
213
156
114
81
46
50
24
27
14
10
10
8
4
5
1
3
1
2
1
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 31
Isoforms per gene
E
100%
75%
50%
25%
0%
1 5 10 15 20 25 31
Isoforms per gene

Figure 2. Coding potential of new intergenic transcripts.

According to CPC scores, calculated both for the forward and for the reverse complement sequence, the intergenic transcripts were classified as “coding”, “non-coding” and “reliable non-coding”. (A) The pie chart shows numbers and proportions of intergenic transcripts falling in each category and provides the color code for the figure panels. (B) and (C) show respectively the distribution of lengths and of expression levels of intergenic transcripts, binned in the three categories. (D) Percentages of transcripts per category are compared, considering all the intergenic transcripts and the subset of the intergenic transcripts ranked within the 10% most expressed transcripts considering the whole transcriptome.

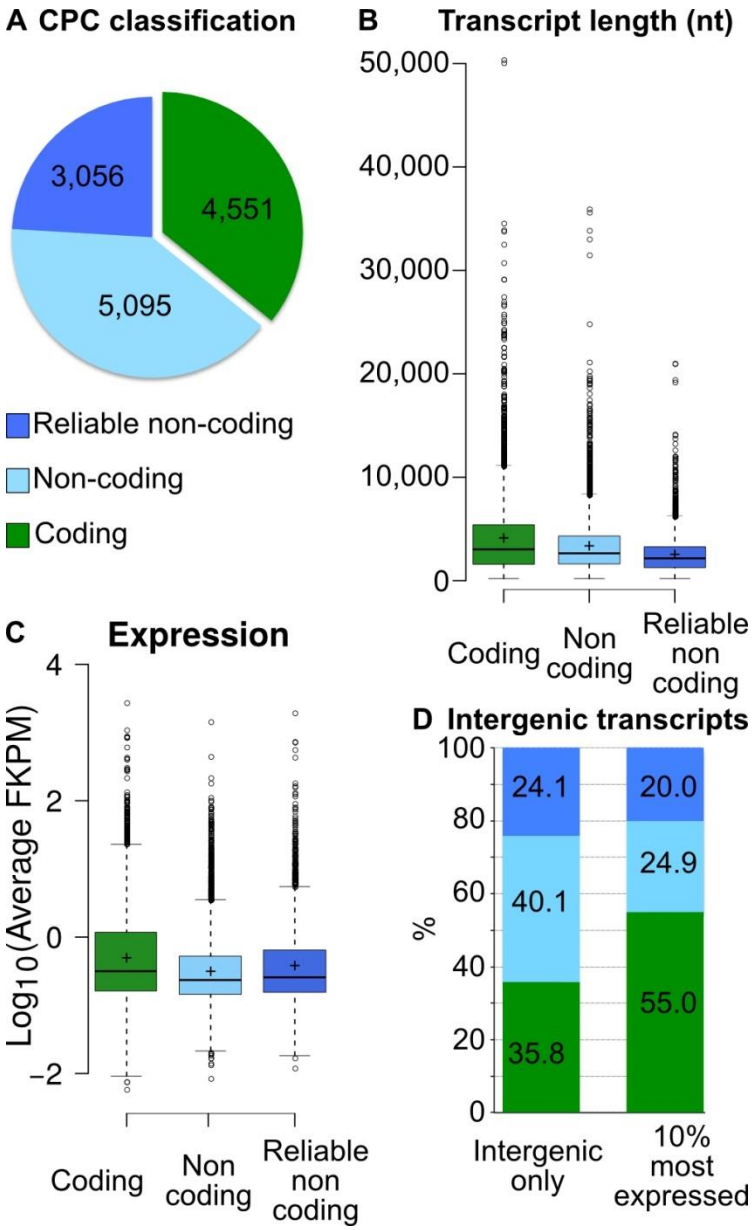
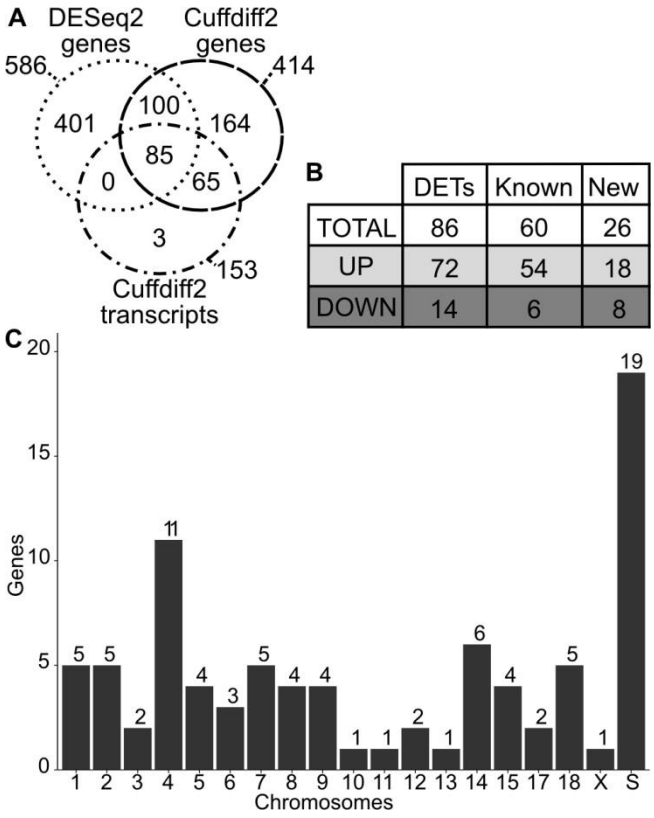


Figure 3. Differentially expressed genes and transcripts identified.

(A) Intersection of genes resulting differentially expressed (DE) according to DESeq2 and Cuffdiff2 analysis, and genes with at least one transcript resulting DE according to the transcript-level Cuffdiff2 analysis. We focused on the transcripts belonging to the 85 *loci* commonly identified by all the methods. (B) Proportions of the new and known DETs resulting higher- and lower-expressed in FAT *vs*. LEAN samples. (C) Number of DE genes mapping to chromosomes or to genome scaffolds (S).



	DETs	Known	New
TOTAL	86	60	26
UP	72	54	18
DOWN	14	6	8

Figure 4. qRT-PCR validation of eleven genes differentially expressed according to RNA-seq data.

(A)Log_2 FC values obtained from RNA-seq, according to Cuffdiff2 estimates, (black bars) and from qRT-PCR data (grey bars), for the eleven tested genes; (B) scatterplot showing the good correlation between the Log_2 FC values calculated with the two experimental methods.

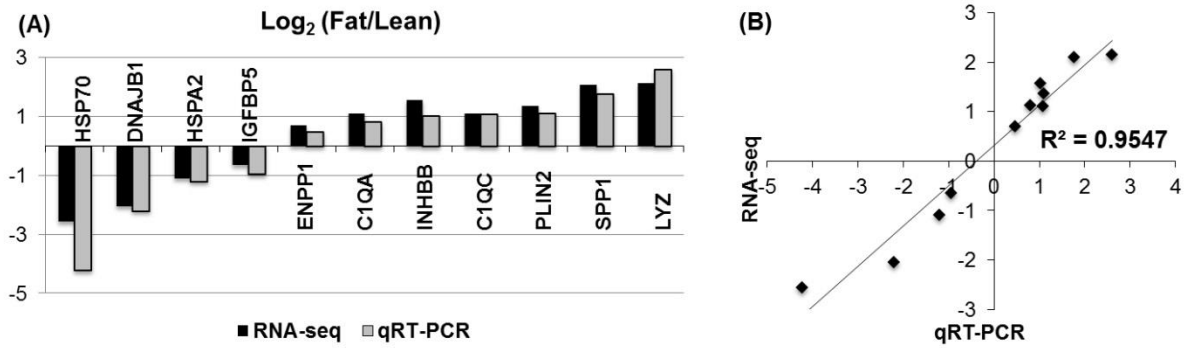
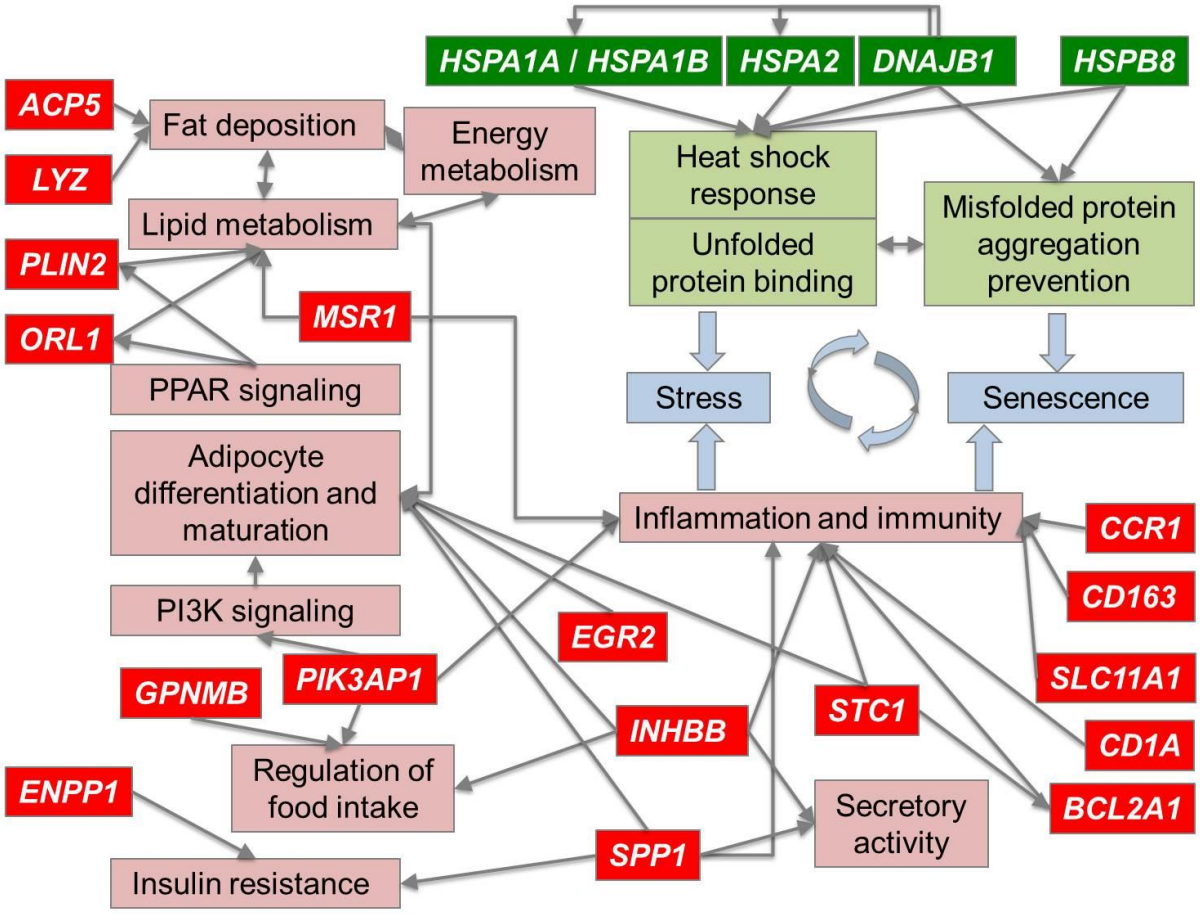


Figure 5. Genes differentially expressed between FAT and LEAN animals impact on specific and connected biological processes.

Genes differentially expressed in FAT *vs*. LEAN pigs converge to specific functions that are more activated or impaired in FAT pigs. Genes and functions upregulated and downregulated in FAT pigs are shown in red and green shades, respectively. Several genes more expressed in FAT pigs are linked to fat deposition and lipid metabolism, to adipocyte differentiation and maturation or to signaling pathways regulating them; FAT pigs show as well increased expression of genes involved in inflammation and immunity and increased expression of genes involved in the control of complex behavior, also by inflammation-mediated secretory activity of adipocytes. Metabolic alterations induce chronic stress in the adipose tissue. FAT pigs shows under-expression of several genes involved in stress response by unfolded protein binding and misfolded protein aggregation prevention. The impairment of these functions might in turn augment inflammation and the consequent secretory activity and possibly induce senescence.



Supporting information

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

Supplementary Tables are included in the file: TranscriptomeILW_SupplementaryTables.xlsx

Table S1 - Primers and PCR condition used for the validation.

EXT: primer pairs used for the amplification of a larger PCR product

INT: primer pairs used for the creation of the standard curve and for the qRT-PCR analysis

Table S2 - Number of reads for each sample

For each sample is indicated the total raw reads sequenced, total clean reads after the trimming and length filters and total reads mapped to the reference genome. Reported values refer to reads as they were single end (total clean paired reads are half the value in the table). Respective percentages are shown in the last three columns.

Table S3 - Types of transcripts expressed in backfat tissue, according to the considered genome annotations.

Transcripts, associated to eight Cufflinks class codes (see http://cufflinks.cbcb.umd.edu/manual.html#class_codes), were classified into three major informative groups.

Table S4 – Intergenic transcript annotations.

Table S5 - Transcript coding potential predicted by Coding Potential Calculator (CPC) for intergenic transcripts.

Reliable noncoding: CPC score <-1

Noncoding: CPC score -1=> / <=0

Coding: CPC score >0

Table S6 - Most expressed transcripts (top 75%) detected in porcine backfat.

Table S7 - David functional annotation clustering of the most expressed genes.

The 10 most relevant clusters are reported

Table S8 - List of differentially expressed genes detected by Cuffdiff2.

Table S9 - List of differentially expressed genes detected by DeSeq2.

Table S10 - List of differentially expressed transcripts detected by Cuffdiff2.

Table S11 - Transcript coding potential predicted by Coding Potential Calculator (CPC) for the differentially expressed transcripts.

Reliable noncoding: CPC score <-1

Noncoding: CPC score -1=> / <=0

Coding: CPC score >0

Figure S1. Read processing and alignment results.

(A) The boxplots show the distribution of the reads considered in different steps and filters of the computational analysis pipeline, in the 20 considered samples. From left to right we show the number of raw reads sequenced, of clean reads resulted from the filtering steps, of reads successfully mapped to the reference genome, and of reads with unique alignment in the genome. (B) From the left, the bars show the average amounts, in the 20 considered samples, of reads spliced, aligned to an exon, to an intron, to intergenic regions (according to the *Sus scrofa* 10.2 genome annotation), or spanning exon-intron borders. Different colors indicate the proportion of read aligning to chromosomes (blue), genome scaffolds (red) or mitochondrial genome (yellow). (C) Number of expressed genes detected in different chromosomes, in mitochondrial genome (Mt) or in genome scaffolds (S).

Figure S2. Gene expression distribution in FAT and LEAN groups.

Cumulative gene expression is shown for the two groups. The figure represents the number of genes (horizontal axis) required to reach different percentages (vertical axis) of the overall gene expression. The inner panel focus on the cumulative expression curves for 50% and 75% of the expression.

Figure S3.Alignment of the four detected isoforms of *PLIN2* gene (red box) with the porcine and vertebrates transcripts present in Ensembl.

Figure S4. Principal component analysis (PCA) based on gene expression profiles.
The figure presents sample separation according to the two principal components, explaining most of the gene expression variation in the data. Samples are represented by dots, with green and orange colours indicating LEAN and FAT samples, respectively in Panel A) and red and blue indicating females and castrated males in Panel B). The PCA shows a clear separation of LEAN and FAT samples, with no separation of samples by sex.

Chapter 3: Comparison of expression levels of fourteen genes involved in the lipid and energy metabolism in two pig breeds.

M. Zappaterra[1], S. Braglia[1], M. Bigi[1], P. Zambonelli[1], R. Davoli[1]

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

Published in

Livestock Science

Zappaterra et al. *Livest Sci.* 2015 Nov;181:156-162.

http://dx.doi.org/10.1016/j.livsci.2015.09.007

Summary

Fat content, fatty acid composition and lean cut weight are important parameters which influence meat and carcass quality in pigs. Up to now, the genes involved in the regulation of the lipid and energy metabolism in porcine skeletal muscle and fat tissue are still relatively unknown. The aim of this study was to investigate the expression levels of fourteen genes (*ACACA*, *ACLY*, *CES3*, *ENO3*, *FASN*, *INSIG2*, *LMNA*, *MTTP*, *ACVR1C*, *NAMPT*, *PLIN1*, *PLIN2*, *PLTP* and *SORT1*) mapped on different chromosomes (1, 4, 7, 8, 9, 12, 15 and 17) which were chosen for their involvement in lipid or energy metabolism in porcine muscle and backfat tissue. Tissue samples from Italian Large White and Italian Duroc pig breeds were collected at the slaughterhouse and frozen in liquid nitrogen. After extraction, the mRNA was quantified by quantitative real time polymerase chain reaction (RT-PCR) and the transcription levels of the genes analysed were compared between breeds for each tissue. In the backfat tissue, differences were found for the *ACACA*, *ACLY*, and *FASN* genes whose highest gene expression levels were found in Italian Large White pigs. In addition, a correlation analysis was carried out between the transcription levels of the genes considered in each tissue and breed. Co-expression relationships still relatively unknown were identified, suggesting new associations between genes which in some cases differed between the two breeds. These results suggest differences between Italian Large White and Italian Duroc pig breeds determined at the genome level affecting carcass quality and fat traits.

Introduction

The levels of leanness and fatness in pigs are the main determinants affecting meat and carcass quality traits. Porcine meat marbling, backfat thickness and lean cut amount are all essential elements for the quality of fresh meat and seasoned products. Several studies have been carried out on different pig breeds in order to identify the causative genes and to investigate the genetic basis of fat deposition. However, nowadays, the genes responsible for the phenotypic variation and for the differences among breeds in intramuscular fat content and backfat thickness are still relatively unknown (Puig-Oliveras et al., 2014). The analysis in different tissues of the expression levels of a set of potential candidate genes for fat deposition or energy metabolism and the identification of their co-expression relationships represent a useful step in identifying the genes involved in the lipid metabolism and the search for markers to be used in selection. Moreover, the knowledge of the correlation between the mRNA levels of this set of genes and some important carcass traits may help both to shed light on the metabolic pathways affecting pig performance and carcass quality, and to define the interactions among candidate genes for fat traits.

The 14 genes were chosen from previous studies cited in the literature which were carried out in pig or other mammalian species (Table 1 and Supplementary Table S1). These genes have a direct or indirect involvement in fat deposition and are, in some cases, involved in human metabolic diseases. The majority of the genes selected, such as *PLIN2, SORT1, PLIN1, MTTP, ACACA, ACLY, ENO3, FASN, ACVR1C, PLTP* and *CES3* are well known for their role in energy and lipid metabolism (Supplementary Table S1), while *LMNA, INSIG2* and *NAMPT* are not directly involved in lipid metabolism but have already been considered in other species for a putative role in adipogenesis or lipid metabolism. *LMNA,* whose main role is related to nuclear stability and chromatin structure, also has a putative role in fat deposition (Verstraeten et al., 2011; Lopez-Mejia et al., 2014) and for this reason was included among the genes selected. Furthermore, *INSIG2* was reported by Malzahn et al. (2014) as a gene associated with the body mass index in humans, even if this association is still under debate. In addition*, NAMPT* was also considered. This gene is mainly known to be involved in the rate-limiting step in nicotinamide adenine dinucleotide (NAD) biosynthesis. Čepica et al. (2010) found, in different pig populations, that *NAMPT* is located in a chromosomal region near a Quantitative Trait Locus (QTL) for backfat thickness, growth, body composition, feed intake, fatty acid composition and loin pH_2. The literature regarding these genes mainly

focuses on several mammalian species, but very little or nothing at all is known regarding the importance of these genes in pig lipid metabolism.

The aim of the present research was to compare between Italian Duroc (IDU) and Italian Large White (ILW) breeds the expression levels in backfat and skeletal muscle tissues of a set of potential candidate genes involved in lipid metabolism and/or adipogenesis.

The identification of gene co-expressions divergent between the two breeds may help to shed light on the gene interactions related to the different phenotypic characteristics of the IDU and ILW pigs concerning lean and fat deposition aptitude. In addition, the study can contribute to improving knowledge regarding the expression profile in the pig muscle and fat tissue of the genes considered.

Materials and Methods

Animals and traits

Italian Large White and Italian Duroc pig breeds were considered in this study. The samples were obtained from pure-breed sib-tested pigs included in the Italian sib-test genetic evaluation scheme carried out by Italian National Association of Pig Breeders (*Associazione Nazionale Allevatori Suini*, ANAS; www.anas.it). All the animals used in this study were kept according to Italian and European laws for pig production, and all the procedures described were in compliance with Italian and European Union regulations for animal care and slaughtering. The animals were reared on the ANAS Sib-Test genetic station from about 30 kg live weight to at least 155 kg live weight. The pigs were fed with a *quasi ad libitum* nutrition level, which means that about 60% of pigs were able to ingest the entire supplied ration. At the end of the test, the animals were transported to a commercial abattoir located about 25 km from the test station in accordance with Council Rule (EC) No. 1/2005 regarding the protection of animals during transport and related operations and, amending Directives 64/432/EEC and 93/119/EC and Regulation (EC) No. 1255/97. At the slaughterhouse, the pigs were electrically stunned and bled in a supine position in agreement with Council Regulation (EC) No. 1099/2009 regarding the protection of animals at the time of slaughter. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health. For each breed, samples of backfat and skeletal muscle tissue (obtained from *semimembranosus* muscle) were collected, immediately frozen in liquid nitrogen and

stored at -80°C until RNA isolation. At the slaughterhouse, it was possible to sample muscle tissue from 11 IDU (M_{IDU}) and 20 ILW pigs (M_{ILW}), and backfat tissue from 12 IDU (B_{IDU}) and 26 ILW pigs (B_{ILW}). From this point on, these abbreviations will be used to indicate the four groups of pigs studied. The gene expression study for each tissue was carried out on the two groups of pigs in order to compare the gene transcription levels between breeds. In Table 2, the number of males and females for each group is indicated. Moreover, for all the pigs, ANAS provided the Estimated Breeding Values (EBVs) and the backfat thickness phenotypic trait, expressed in mm and measured at the slaughterhouse. For each group of samples, the mean values for the traits considered and the EBVs are reported in Table 2. The EBVs provided are Average Daily Gain (ADG, calculated from 30 to 155 kg of live weight with *quasi ad libitum* feeding level, expressed in grams), Backfat Thickness (BFT, recorded *post mortem* at the level of the *gluteus medius* muscle, expressed in mm), Lean Cuts (LC, the sum of neck and loin weight, expressed in kg), Feed Conversion Ratio (FCR, obtained from feed intake recorded daily and body weight measured bimonthly, expressed in units) and Ham Weight (HW, expressed in kg). The EBVs were calculated by ANAS according to the statistical model reported by Russo et al. (2000), using a BLUP-multiple trait animal model (Henderson and Quaas, 1976). The models considered were different for each trait and included fixed effects of batch, sex and age at the beginning of the test, age of the pig, weight at slaughter, age at slaughter, inbreeding coefficient, the random effects of litter and the individual permanent environmental and animal effects. Furthermore, ANAS considered the additive relationship matrix to calculate the genetic merit of the pigs for the traits analysed; the EBVs are always expressed as the difference from the genetic basis evaluated as “zero” calculated on pigs born in 1993.

RNA source, total RNA extraction and cDNA preparation

Total RNA was extracted from *semimembranosus* muscle and backfat tissue of ILW and IDU pigs using TRIZOL reagent (Invitrogen Corporation, Carlsbad, California), as described in Davoli et al. (2011). The quality and integrity of the RNAs were both checked by a reading on the ND-1000 Spectrophotometer (NanoDrop Technologies, Willmington, DE, USA), and visualisation on 1% agarose gel. The RNA samples were treated with DNaseI (Invitrogen Corporation) and, for each sample, 1 µg of total RNA was retrotranscribed to cDNA, according to the manufacturer’s instructions, using the Improm-II ™ Reverse Transcription

System and Oligo-dT primers (Promega Corporation, Madison, Wi, USA). The intronic primer sequence of *AGPAT1* (*1-acylglycerol-3-phosphate O-acyltransferase 1*) gene was utilised in the PCR analysis to verify the contamination of possible residual DNA, amplifying the retrotranscribed products.

Quantification of genes expression

The expression levels of the genes studied in backfat and skeletal muscle tissues in IDU and ILW pigs were analysed using the relative quantitative Real-Time polymerase chain reaction (qRT-PCR) standard curve method (Pfaffl, 2004) and the data from each sample were normalised against the two different normalising genes for each tissue, according to Vandesompele et al. (2002) and GENORM (http://medgen.ugent.be/~jvdesomp/genorm/). The qRT-PCR analyses were carried out with the Light Cycler 1.0 System (Roche Diagnostics, Mannheim, Germany) using SYBR® Premix Ex Taq™ (TAKARA Bio INC, Olsu, Shiga, Japan), 10 pmol of each primer and 2 µl of cDNA template diluted 1:10, for a total volume of 10 µl. The Light Cycler protocol was optimised using specific annealing temperatures for each primer couple (Supplementary Table S2). For the normalising genes and the 14 genes studied, a specific standard curve was obtained, amplifying 12 serial dilutions (from 10^9 to 50 molecules/µl) of a known-concentration cDNA amplicon, obtained with a PCR using the external primer pairs. The PCR efficiency was calculated as $E = 10 \exp(-1/\text{slope})$, with a range between -2.7 and -4.3, indicating good PCR efficiency results. All the PCR products were checked on a polyacrylamide gel, and detection of the melting temperature allowed checking the specificity of the amplicon. The coefficient of variation (CV = Standard Deviation of the Crossing Points/Average of the Crossing Points) of the replicated analysis for each sample (three in two different qRT PCRs) was accepted for CV < 0.5. The data obtained were normalised using different pools of normalising genes for the two tissues and, for that reason, it was not possible to compare the normalised transcription levels obtained in the backfat tissue with those obtained for muscle samples. To normalise the result of the expression data in backfat and muscle tissues, the genes *beta-2-microglobulin* (*B2M*), *polymerase (RNA) II (DNA directed) polypeptide A, 220 kDa* (*POLR2A*), *Peptidylprolil isomerase A (cyclophilin A)* (*PPIA*) and *Thymosin beta 4, X-linked* (*TMSB4X*) reported in different papers as normalising genes (Hsiao et al., 2001; Wang et al., 2012) were tested. The *B2M* and *POLR2A* genes were utilised to normalise skeletal muscle genes expression while

PPIA and *TMSB4X* were used for the mRNA quantification in backfat tissue. For each tissue, the normalised gene transcription levels were compared between breeds.
In muscle and backfat tissues, only the expression levels for 12 genes out of the 14 initially chosen were measured due to the undetectable transcription level of *MTTP* and *ACVR1C* in the tissues considered.

Statistical analysis

R software was used for the statistical analyses, in particular, the Stats package (R Core Team, 2015). Within each tissue, the expression levels of the 12 detectable genes were compared between breeds using the Student's t-test, considering $P < 0.05$ as a significance threshold. Moreover, Spearman's correlations between gene expressions and the EBVs were calculated for the IDU groups (M_{IDU} and B_{IDU}) and the ILW groups (M_{ILW} and B_{ILW}). In addition, correlations between the transcription levels of the genes analysed were determined in order to find the co-expressed genes in each tissue for each breed.

Results

Expression study

The expression study results are reported in Figure 1 for muscle tissue and in Figure 2 for backfat. In muscle tissue, the only significant comparison between M_{IDU} and M_{ILW} pigs was observed for the *ACLY* gene (Figure 1), with a t test *P* value of 0.0178. The *ENO3*, *NAMPT* and *LMNA* genes were those most expressed in muscle tissue of both breeds.
When observing the backfat tissue results (Figure 2), the higher *PLIN1* and *CES3* expression levels clearly stand out with respect to the other genes tested. Furthermore, in backfat tissue, the *FASN*, *ACLY* and *ACACA* mRNA levels differed significantly between the two breeds ($P = 0.0273$, $P = 0.0024$, $P = 0.0039$, respectively) and were notably higher in ILW than in IDU samples.

Correlation analysis

Muscle gene expression correlations. The gene expression data obtained in muscle tissue for the 11 M_{IDU} and the 20 M_{ILW} pigs were used to carry out a Spearman's correlation analysis between transcription levels and EBVs (Table 3). The M_{ILW} group never showed any

significant correlation with the performance EBVs in muscle tissue. On the contrary, in the IDU sample, the transcription levels of the *PLIN1*, *ENO3*, *PLIN2*, *PLTP* and *LMNA* genes were correlated with the traits considered. In particular, *PLTP* showed significant correlation coefficients with all the EBVs considered except LC; its expression was positively correlated with BFT (P = 0.021) and FCR (P = 0.003), and negatively correlated with ADG (P = 0.013) and HW (P = 0.007). In the IDU sample, HW was negatively affected by an increase in both *PLTP* and *PLIN1* transcription levels; moreover, a worsening of feed conversion to body mass (expressed with FCR EBV) was observed. The significant correlation values between muscle gene expressions are reported in Table 4. In muscle tissue, a larger number of significant correlations between the gene expressions were found in ILW pigs than in IDU pigs; in particular, M_{IDU} did not show any significant correlation of the transcription levels of *PLIN1*, *PLIN2*, *CES3* and *ACACA* with the other genes. The *FASN* and *ACLY* genes did not show any significant correlations in either breed and, for that reason, the data are not reported in Table 4. In both breeds, all the correlations obtained between gene expressions in the skeletal muscle tissue were positive, and the most interesting ones will be commented on in the Discussion section.

Backfat gene expression correlations. Spearman's correlations between backfat gene expression levels and performance traits are reported in Table 5. The IDU pigs showed a higher number of correlations between gene expressions and EBVs in comparison to the ILW pigs which, in contrast, presented fewer and less significant correlation coefficients. The *CES3, SORT1, PLIN1*, *LMNA, FASN* and *ACACA* genes were significantly correlated with the EBVs and, in particular, *CES3* showed the highest number of correlations and the most significant *P* value for FCR (P = 0.005) in IDU pigs. In IDU pigs, BFT EBV was correlated with the expressions of the *CES3, PLIN1*, *LMNA, FASN* and *ACACA* genes and, in particular, the highest significant correlations for BFT were found for *FASN* and *ACACA* expression (Table 5). In the backfat tissue of both breeds, when observing Spearman's correlations between the transcription levels, some putative gene interactions and pathways could be hypothesised. Some interesting differences between breeds are visible. In particular, for the *ENO3*, *PLIN2 and ACLY* genes, only the ILW group presented significant correlations between the expressions of these three *loci* and the other genes studied. Furthermore, the B_{ILW} pigs showed the most significant correlations (P values < 0.0001) for *PLIN1 vs. NAMPT*, *PLIN1 vs. LMNA*, *SORT1 vs. LMNA*, *SORT1 vs. INSIG2*, and between *FASN* and *ACLY vs. ACACA*

correlation. The only negative correlation found in our samples was between *INSIG2* and *ACLY* in B_{ILW} pigs (Table 6).

Discussion

The results of the research indicate that there are differences between the two breeds for the correlations of gene expressions (co-expressions) in backfat and muscle tissues for the majority of the considered loci. The IDU and the ILW breeds have different characteristics regarding muscle and fat deposition; ILW pigs have a better feed conversion efficiency than IDU pigs (http://www.anas.it/circolari/201500000.PDF, page 22), however, maintaining the correct balance between muscle and backfat deposition. On the other hand, the IDU breed tends to a higher intramuscular fat deposition as compared to the ILW breed. This different aptitude for lean and fat tissue deposition between the two breeds is also visible in the results obtained. Indeed, the IDU and the ILW pigs showed some significant discrepancies concerning both the gene expression levels in the two tissues analysed and the gene co-expressions within the tissues. Moreover, significant correlations were found between gene transcription and the EBVs. In backfat tissue, the genes involved in lipid synthesis (*ACACA*, *ACLY* and *FASN*) showed different expression levels between the breeds, with the highest transcription values identified in ILW pigs (Figure 2). The *ACLY* and *ACACA* genes code for the enzymes ATP citrate lyase and Acetyl-CoA carboxylase, respectively; the first is the primary enzyme responsible for the synthesis of cytosolic Acetyl-CoA while the second is a complex multifunctional enzyme system which catalyses the carboxylation of acetyl-CoA to malonyl-CoA, the rate-limiting step in fatty acid synthesis (Davoli et al., 2014). Furthermore, a polymorphism of the *Fatty acid synthase* gene has already been reported in association with backfat deposition in ILW pigs (Braglia et al., 2014). In B_{ILW} pigs, the *FASN* and *ACLY* genes presented a high number of co-expressions with the other genes studied which were, for the most part, missing in B_{IDU} pigs (Table 6). The *ACLY, ACACA* and *FASN* genes have a key role in lipogenesis, and the higher transcription level observed in backfat for the ILW samples can be considered consistent with the greater backfat deposition in this breed as compared to the IDU samples (Tables 2 and 6). On the other hand, the IDU samples showed a higher expression of *ACLY* and *FASN genes* in muscle tissue (Figure 1), in agreement with the predisposition of the IDU breed towards a more marked intramuscular fat deposition than the ILW breed. In line with these data, Cánovas et al. (2010) reported a more intense

transcriptional activity of the genes involved in fat metabolism for IDU pigs with extreme intramuscular fat deposition in the muscle tissue. Based on the results obtained, it is likely that the ILW and IDU pig breeds present different co-expression pathways for the *ACLY*, *ACACA* and *FASN* genes induced by the different lipogenic capability in the backfat and muscle tissue characterising the two breeds.

In addition, in Table 3, the M_{IDU} pigs presented significant correlations between the EBVs and the expression levels of the genes related to intracellular lipid droplet storage (*PLIN1*, *PLIN2*), lipoprotein metabolism (*PLTP*), glycolysis (*ENO3*) and adipose tissue differentiation (*LMNA*). On the contrary, for the same genes, the M_{ILW} pigs did not show significant values, likely highlighting the less pronounced aptitude towards intramuscular fat deposition of the ILW pigs as compared to the Duroc pigs. Verstraeten et al. (2011) and Boguslavsky et al. (2006) have proven the importance of *LMNA* gene for fat tissue deposition and adipocyte differentiation in mouse-cultured cells. Although evidence links *LMNA* to fat deposition in mice, this gene remains poorly explored in swine.

Perilipin family genes are the main actors intervening on the modulation of intracellular lipid droplet hydrolysis. Even if the Perilipin 2 protein in the liver seems to promote lipid storage (Sun et al., 2012), in oxidative muscles, the same protein coats the cellular lipid droplets to promote hydrolysis in energy-requiring processes (Gandolfi et al., 2011; Sztalryd and Kimmel, 2014). In M_{ILW} pigs, *PLIN2* presented a relevant number of significant co-expressions with the genes studied while M_{IDU} pigs did not present any significant co-expression (Table 4). This result seems to point out a more pronounced activity of this gene in the ILW samples than in the IDU samples which promotes the hydrolysis of the intracellular lipids stored in the muscle.

The same results were obtained for *CES3*, a gene which catalyses a protein having carboxylic ester hydrolase activity (Shanghani et al., 2009). Similar to the data obtained for *PLIN2* expression, the presented results indicate that only in ILW muscle *CES3* showed significant co-expressions with the majority of the genes considered. This result could allow hypothesising that this gene is involved in some processes in the ILW muscle more than in the same tissue of the IDU samples.

Moreover, it is worth noting that *SORT1* expression was positively correlated to lean cut deposition in B_{IDU} pigs and negatively related to ADG in B_{ILW} pigs (Table 5). Sortilin 1 functions are still not completely clear and, in the literature, Huang et al. (2013)

hypothesised the involvement of sortilin 1 in the formation of insulin responsive vesicles on perinuclear membranes and a possible engagement of SORT1 in the increase in insulin sensitivity in differentiating adipocytes. The two breeds considered have a different lipogenic activity for the tissues analysed as the ILW breed show a higher tendency to fat deposition in backfat tissue while the IDU pigs have a strong aptitude for muscle fat deposition as reported in the literature (Cánovas et al., 2010). It can be suggested that the differences observed in backfat tissue between the IDU pigs and the ILW pigs according to the *SORT1* correlations with the EBVs may be associated to the different adipogenic disposition of the two breeds.

In summary, differences between ILW and IDU concerning the transcription rate and the gene co-expressions were observed, dissimilarities which could be partially related to the different phenotypic characteristics of the two breeds. These phenotypic and molecular differences could be due to the distinct genetic origin of the Duroc breed as compared to Large White pigs, hypothesising that selection in European domestic pig breeds led to genetic differentiation in chromosomal regions harbouring genes involved in many metabolic pathways (Amaral et al*.,* 2011).

Moreover, the results suggest still unexplored roles and pathways linking some genes, such as the strong correlation intervening between *SORT1* and *LMNA* (Tables 4 and 6), between *INSIG2* and *ACLY* and between *INSIG2* and *SORT1* expressions (Table 6). However, additional studies will be required to prove the latter results.

In our study, the correlations found in muscle (Tables 4 and 5) between *PLTP* gene expression and the EBVs can only be partially explained considering that the effect of this gene on pig fat deposition has been poorly investigated to date. *PLTP* codes for a phospholipid transfer protein involved in the cholesterol metabolism and this gene is currently considered for its influence on coronary diseases, diabetes and obesity (Qin et al., 2014). The present results offer some new insights, providing additional information on pigs regarding *PLTP* co-expression pathways which may be useful in increasing knowledge of this gene.

Conclusions

The different aptitudes for lean and fat deposition characterising the Italian Duroc and Italian Large White breeds could also be partially identified at a transcriptional level for the

genes studied, suggesting that, for each breed, distinct and specific gene effects on carcass quality and fat traits could exist. Additional functional or regulative roles for the genes considered can be hypothesised, but much remains to be tested in additional studies. However, the results obtained may help in offering new insights on some genes which are still relatively unknown in pigs.

Aknowledgements

We thank Dr. Luca Buttazzoni (CREA-PCM) and Dr. Maurizio Gallo (Associazione Nazionale Allevatori Suini, ANAS) for providing the porcine tissue samples. This work was supported by Progetto "AGER - Agroalimentare e ricerca": Advanced research in genomics and processing technologies for the Italian heavy pig production – Hepiget (Grant N. 2011- 0279).

References

Amaral AJ, Ferretti L, Megens HJ, Crooijmans RPMA, Nie H, Ramos-Onsins SE, Perez-Enciso M & Shook LB (2011) Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLOS ONE* 6:e14782.

Boden G, Salehi S, Cheung P, Homko C, Song W, Loveland-Jones C & Jayarajan S (2013) Comparison of In Vivo Effects of Insulin on SREBP-1c Activation and INSIG-1/2 in Rat Liver and Human and Rat Adipose Tissue. *Obesity* 21:1208–214.

Boguslavsky RL, Stewart CL & Worman HJ (2006) Nuclear lamin A inhibits adipocyte differentiation: implications for Dunnigan-type familial partial lipodystrophy. *Hum Mol Genet* 15:653–663.

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. *Livest Sci* 162:15-22.

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

Čepica S, Bartenschlager H, Ovilo C, Zrůstová J, Masopust M, Fernández A, López A, Knoll A, Rohrer GA, Snelling WM & Geldermann H (2010) Porcine NAMPT gene: search for polymorphism, mapping and association studies. *Anim Genet* 41:646-651.

Davoli R, Braglia S, Russo V, Varona L & te Pas MFW (2011) Expression profiling of functional genes in prenatal skeletal muscle tissue in Duroc and Pietrain pigs. *J Anim Breed Genet* 128:15-27.

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. *Livest Sci* 165:1-7.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V & Davoli R (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631-637.

Henderson CR & Quaas RL (1976) Multiple trait evaluation using relatives' records. *J Anim Sci* 43:1188-1197.

Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, Dillon W, Lee KF, Clark KE, Haverty P, Weng Z, Mutter GL, Frosh MP, MacDonald ME, Milford EL, Crum CP, Bueno R, Pratt RE, Maha-devappa M, Warrington JA, Stephanopoulos G & Gullans SR (2001) A compendium of gene expression in normal human tissues. *Physiol Genomics* 7:97-104.

Huang G, Buckler-Pena D, Nauta T, Singh M, Asmar A, Shi J, Kim JY & Kandror KV (2013) Insulin responsiveness of glucose transporter 4 in 3T3-L1 cells depends on the presence of sortilin. *Mol Biol Cell* 24:3115-3122.

Lopez-Mejia IC, de Toledo M, Chavey C, Lapasset L, Cavelier P, Lopez-Herrera C, Chebli K, Fort P, Beranger G, Fajas L, Amri EZ, Casas F & Tazi J (2014) Antagonistic functions of LMNA isoforms in energy expenditure and lifespan. *EMBO Rep* 15:529-539.

Malzahn D, Müller-Nurasyid M, Heid IM, Wichmann HE, KORA study group & Bickeböller H (2014) Controversial association results for INSIG2 on body mass index may be explained by interactions with age and with MC4R. *Eur J Hum Genet* 22:1217-1224.

Pfaffl MW (2004) Quantification strategies in real-time PCR. In A–Z of quantitative PCR [Bustin SA editor] International University Line (IUL), La Jolla, CA, USA.

Puig-Oliveras A, Ballester M, Corominas J, Revilla M, Estellé J, Fernández AI, Ramayo-Caldas Y & Folch JM (2014) A co-association network analysis of the genetic determination of pig conformation, growth and fatness. *PLOS ONE* 9:e114862.

Qin S, Song G & Yu Y (2014) Phospholipid transfer protein in diabetes, metabolic syndrome and obesity. *Cardiovasc Hematol Disord Drug Targets* 14:149-153.

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Russo V, Buttazzoni L, Baiocco C, Davoli R, Nanni Costa L, Schivazappa C & Virgili R (2000) Heritability of muscular cathepsin B activity in Italian Large White pigs. *J Anim Breed Genet* 117:37-42.

Sanghani SP, Sanghani PC, Schiel MA & Bosron WF (2009) Human carboxylesterases: an update on CES1, CES2 and CES3. *Protein Pept Lett* 16:1207-1214.

Sun Z, Miller RA, Patel RT, Chen J, Dhir R, Wang H, Zhang D, Graham MJ, Unterman TG, Shulman GI, Sztalryd C, Bennett MJ, Ahima RS, Birnbaum MJ & Lazar MA (2012) Hepatic Hdac3 promotes gluconeogenesis by repressing lipid synthesis and sequestration. *Nat Med* 18:934-942.

Sztalryd C & Kimmel AR (2014) Perilipins: lipid droplet coat proteins adapted for tissue-specific energy storage and utilization, and lipid cytoprotection. *Biochimie* 96:96-101.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F (2002) Accurate normalization of real- time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:research0034.

Verstraeten VLRM, Renes J, Ramaekers FCS, Kamps M, Kuijpers HJ, Verheyen F, Wabitsch M, Steijlen PM, van Steensel MAM & Broers JLV (2011) Reorganization of the nuclear lamina and cytoskeleton in adipogenesis. *Histochem Cell Biol* 135:251-261.

Wang Q, Ishikawa T, Michiue T, Zhu BL, Guan DW & Maeda H (2012) Stability of endogenous reference genes in post mortem human brains for normalization of quantitative real-time PCR data: comprehensive evaluation using geNorm, NormFinder, and BestKeeper. *Int J Legal Med* 126:943-952.

Table 1. List of the 14 genes considered for the present expression study in pig muscle and backfat tissues sorted by chromosome location.

GENE	Coded Protein	Chromosome (SSC)[1]
PLIN2	Perilipin 2	SSC1
LMNA	Lamin A/C	SSC4
SORT1	Sortilin1	SSC4
INSIG2	Insulin induced gene 2	SSC7
PLIN1	Perilipin 1	SSC7
MTTP	Microsomal triglyceride transfer protein	SSC8
NAMPT	Nicotinamide phosphoribosyltransferase	SSC9
ACACA	Acetyl-CoA carboxylase alpha	SSC12
ACLY	ATP citrate lyase	SSC12
ENO3	Enolase 3 (Beta, muscle)	SSC12
FASN	Fatty acid synthase	SSC12
ACVR1C	Activin A receptor type-1C	SSC15
PLTP	Phospholipid transfer protein	SSC17
CES3	Carboxylesterase 3	Unplaced

[1] SSC is the acronym used for *Sus scrofa* chromosome.

The position of the studied genes is indicated in NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene) accessed on 11 September 2015.

Table 2. The considered groups of pigs and their composition in the top half of the table and in the bottom half the means ± the standard deviations of the traits and Estimated Breeding Values (EBVs) for the groups.

	Skeletal muscle samples		Backfat samples	
	M_{IDU}	M_{ILW}	B_{IDU}	B_{ILW}
Castrated males	6	8	6	11
Sows	5	12	6	15
Total	11	20	12	26
Traits/EBVs				
Backfat thickness (mm)[1]	23.80 ± 6.63	27.75 ± 8.83	24.25 ± 6.40	29.42 ± 7.68
ADG (g)[2]	48.00 ± 48.79	26.35 ± 48.55	45.00 ± 49.89	38.46 ± 27.49
BFT (mm)[3]	-3.26 ± 3.81	-2.29 ± 7.34	-3.55 ± 3.94	-2.00 ± 6.83
LC (kg)[4]	3.32 ± 4.45	1.91 ± 3.34	3.34 ± 4.45	2.27 ± 2.73
HW (kg)[5]	1.06 ± 0.91	0.54 ± 0.99	1.16 ± 0.97	0.67 ± 0.70
FCR[6]	-0.21 ± 0.15	-0.13 ± 0.27	-0.21 ± 0.15	-0.18 ± 0.16

IDU is used to identify the groups of Italian Duroc breed pigs.

ILW is used to identify the groups of Italian Large White breed pigs.

[1] backfat thickness phenotype in millimetres, phenotype measured at 8 cm off the midline of the carcass at the level placed between the third and fourth last ribs.

[2] Estimated Breeding Values obtained with a BLUP-multiple trait animal model for Average Daily Gain (ADG) calculated from 30 to 155 kg of live weight with *quasi ad libitum* feeding level, expressed in grams.

[3] Estimated Breeding Values obtained with a BLUP-multiple trait animal model for Backfat Thickness (BFT) recorded *post mortem* at the level of *gluteus medius* muscle, expressed in mm.

[4] Estimated Breeding Values obtained with a BLUP-multiple trait animal model for Lean Cuts (LC), the sum of neck and loin weights, expressed in kg.

[5] Estimated Breeding Values obtained with a BLUP-multiple trait animal model for Ham Weight (HW), the weight of thighs expressed in kg.

[6] Estimated Breeding Values calculated with a BLUP-multiple trait animal model for Feed Conversion Ratio (FCR) obtained from feed intake daily recorded and body weight bimonthly measured, expressed in units.

Table 3. Spearman's correlation coefficients and relative *P* values between the Estimated Breeding Values (EBVs) and the expression levels in pig muscle tissue of the Italian Duroc breed pigs group (M_{IDU}).

		Estimated Breeding Values				
		ADG[1]	BFT[2]	LC[3]	HW[4]	FCR[5]
PLIN1	M_{IDU}	-	-	-	-0.654*	0.690*
ENO3	M_{IDU}	-	-	-	-	0.636*
PLIN2	M_{IDU}	-	-	-0.645*	-	-
PLTP	M_{IDU}	-0.745*	0.709*	-	-0.781**	0.818**
LMNA	M_{IDU}	-	0.601*	-	-	0.627*

The correlation is expressed with the r coefficient and * for $P < 0.05$, ** for $P < 0.01$.

- is used for not significant values. In the table are reported only genes with significant correlations, expressed with the r coefficient and between brackets the relative *P* value. ILW pigs are not included in the table, as they did not showed significant corrrelations.

[1] EBV of Average Daily Gain calculated from 30 to 155 kg of live weight with *quasi ad libitum* feeding level, expressed in grams.

[2] EBV of Backfat Thickness recorded *post mortem* at the level of *gluteus medius* muscle, expressed in mm.

[3] EBV of Lean Cuts, the sum of neck and loin weights, expressed in kg.

[4] EBV of Ham Weight, expressed in kg.

[5] EBV of Feed Conversion Ratio obtained from feed intake daily recorded and body weight bimonthly measured, expressed in units.

Table 4. Spearman's correlation coefficients between the expression levels of the 12 detectable genes in pig muscle tissue of the Italian Duroc pigs (M_{IDU}) and Italian Large White pigs (M_{ILW}).

		ENO3	*PLIN1*	*PLIN2*	*NAMPT*	*SORT1*	*CES3*	*PLTP*	*LMNA*	*INSIG2*	*ACACA*
ENO3	M_{IDU}	1	-	-	0.927***	0.939***	-	0.794***	0.903***	0.903***	-
	M_{ILW}	1	-	0.767***	0.744***	0.829***	-	0.707***	0.810***	0.773***	0.875***
PLIN1	M_{IDU}	-	1	-	-	-	-	-	-	-	-
	M_{ILW}	-	1	0.516*	0.505*	0.575**	0.680**	-	0.619**	0.553*	0.643**
PLIN2	M_{IDU}	-	-	1	-	-	-	-	-	-	-
	M_{ILW}	0.767***	0.516*	1	0.892***	0.889***	-	0.705***	0.851***	0.901***	0.870***
NAMPT	M_{IDU}	0.927***	-	-	1	0.900***	-	-	-	0.909***	-
	M_{ILW}	0.744***	0.505*	0.892***	1	0.789***	0.516*	0.668*	0.833***	0.974***	0.761***
SORT1	M_{IDU}	0.939***			0.900***	1	-	-	0.773**	0.836**	-
	M_{ILW}	0.829***	0.575**	0.889***	0.789***	1	-	0.807***	0.872***	0.809***	0.851***
CES3	M_{IDU}	-	-	-	-	-	1	-	-	-	-
	M_{ILW}	-	0.680**	-	0.516*	-	1	-	-	0.554*	0.531*
PLTP	M_{IDU}	0.794**	-	-	-	-	-	1	0.854**	-	-
	M_{ILW}	0.707**	-	0.705***	0.668**	0.807***	-	1	0.807***	0.676**	0.659**
LMNA	M_{IDU}	0.903***	-	-	0.727*	0.773**	-	0.854**	1	0.773**	-
	M_{ILW}	0.810***	0.619**	0.851***	0.833***	0.872***	-	0.807***	1	0.815***	0.810***
INSIG2	M_{IDU}	0.903***	-	-	0.909***	0.836**	-	-	0.773**	1	-
	M_{ILW}	0.773***	0.553*	0.901***	0.974***	0.809***	0.554*	0.676**	0.815***	1	0.819***
ACACA	M_{IDU}	-	-	-	-	-	-	-	-	-	1
	M_{ILW}	0.875***	0.643**	0.870***	0.761***	0.851***	0.531*	0.659**	0.810***	0.819***	1

The correlation is expressed with the r coefficient and * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

- is used for not significant *P* values.

Table 5. Spearman's correlation coefficients and relative *P* values between the Estimated Breeding Values and the expression levels in pig backfat tissue of the Italian Duroc breed pigs group (B_{IDU}) and the Italian Large White group (B_{ILW}).

		Estimated Breeding Values				
		ADG[1]	BFT[2]	LC[3]	HW[4]	FCR[5]
CES3	B_{IDU}	-0.671*	0.697*	-	-0.622*	0.748**
	B_{ILW}	-	0.478*	-	-	-
SORT1	B_{IDU}	-	-	0.700*	-	-
	B_{ILW}	-0.395*	-	-	-	-
PLIN1	B_{IDU}	-	0.672*	-	-	-
	B_{ILW}	-	-	-	-	-
LMNA	B_{IDU}	-	0.648*	-	-	0.615*
	B_{ILW}	-	-	-	-	-
FASN	B_{IDU}	-	0.701*	-	-	-
	B_{ILW}	-	-	-	-	-
ACACA	B_{IDU}	-	0.752**	-	-	-
	B_{ILW}	-0.452*	-	-	-	0.489*

The correlation is expressed with the r coefficient and * for $P < 0.05$, ** for $P < 0.01$.

- is used for not significant P values.

[1] EBV of Average Daily Gain calculated from 30 to 155 kg of live weight with *quasi ad libitum* feeding level, expressed in grams.

[2] EBV of Backfat Thickness recorded *post mortem* at the level of *gluteus medius* muscle, expressed in mm.

[3] EBV of Lean Cuts, the sum of neck and loin weights, expressed in kg.

[4] EBV of Ham Weight, expressed in kg.

[5] EBV of Feed Conversion Ratio obtained from feed intake daily recorded and body weight bimonthly measured, expressed in units.

Table 6. Spearman's correlation coefficients between the expression levels of the 12 detectable genes in pig backfat tissue of the Italian Duroc pigs (B_{IDU}) and Italian Large White pigs (B_{ILW}).

		ENO3	*PLIN1*	*PLIN2*	*NAMPT*	*SORT1*	*CES3*	*PLTP*	*LMNA*	*INSIG2*	*FASN*	*ACLY*	*ACACA*
ENO3	B_{IDU}	1	-	-	-	-	-	-	-	-	-	-	-
	B_{ILW}	1	-	-	-	-	-	-	-	-	-	-	0.952***
PLIN1	B_{IDU}	-	1	-	-	-	0.783**	-	0.769**	-	0.804**	-	0.836**
	B_{ILW}	-	1	0.668***	0.842***	0.509**	0.550**	0.528*	0.738***	-	0.535**	-	0.669**
PLIN2	B_{IDU}	-	-	1	-	-	-	-	-	-	-	-	-
	B_{ILW}	-	0.668***	1	0.637***	-	-	-	-	-	0.570**	-	0.681**
NAMPT	B_{IDU}	-	-	-	1	-	-	-	-	-	-	-	0.818**
	B_{ILW}	-	0.842***	0.637***	1	-	0.652***	0.543*	0.585**	-	0.524**	-	0.678**
SORT1	B_{IDU}	-	-	-	-	1	-	-	-	0.773**	-	-	-
	B_{ILW}	-	0.509**	-	-	1	-	0.588**	0.827***	0.810***	-	-	-
CES3	B_{IDU}	-	0.783**	-	-	-	1	-	0.790**	-	-	-	-
	B_{ILW}	-	0.550**	-	0.652***	-	1	-	-	-	-	-	0.687***
PLTP	B_{IDU}	-	-	-	-	-	-	1	-	-	-	-	0.806**
	B_{ILW}	-	0.528*	-	0.543*	0.588**	-	1	0.606**	-	-	-	-
LMNA	B_{IDU}	-	0.769**	-	-	-	0.790**	-	1	-	-	-	-
	B_{ILW}	-	0.737***	-	0.585**	0.827***	-	0.606**	1	0.611**	-	-	-
INSIG2	B_{IDU}	-	-	-	-	0.773**	-	-	-	1	-	-	-
	B_{ILW}	-	-	-	-	0.810***	-	-	0.611**	1	-	-0.474*	-
FASN	B_{IDU}	-	0.804**	-	-	-	-	-	-	-	1	-	0.873***
	B_{ILW}	-	0.535**	0.570**	0.524**	-	-	-	-	-	1	0.573**	0.785***

ACLY	B_{IDU}	-	-	-	-	-	-	-	-	-	-	1	-
	B_{ILW}	-	-	-	-	-	-	-	-	-0.474*	0.573**	1	0.762***
ACACA	B_{IDU}	-	0.836**	-	0.818**	-	-	0.806**	-	-	0.873***	-	1
	B_{ILW}	0.952***	0.669**	0.681**	0.678**	-	0.687***	-	-	-	0.785***	0.762***	1

The correlation is expressed with the r coefficient and * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

- is used for not significant *P* values.

Figures

Figure 1. The studied gene expression levels in pig muscle tissue for the Italian Duroc (M_{IDU}) and the Italian Large White (M_{ILW}) samples.

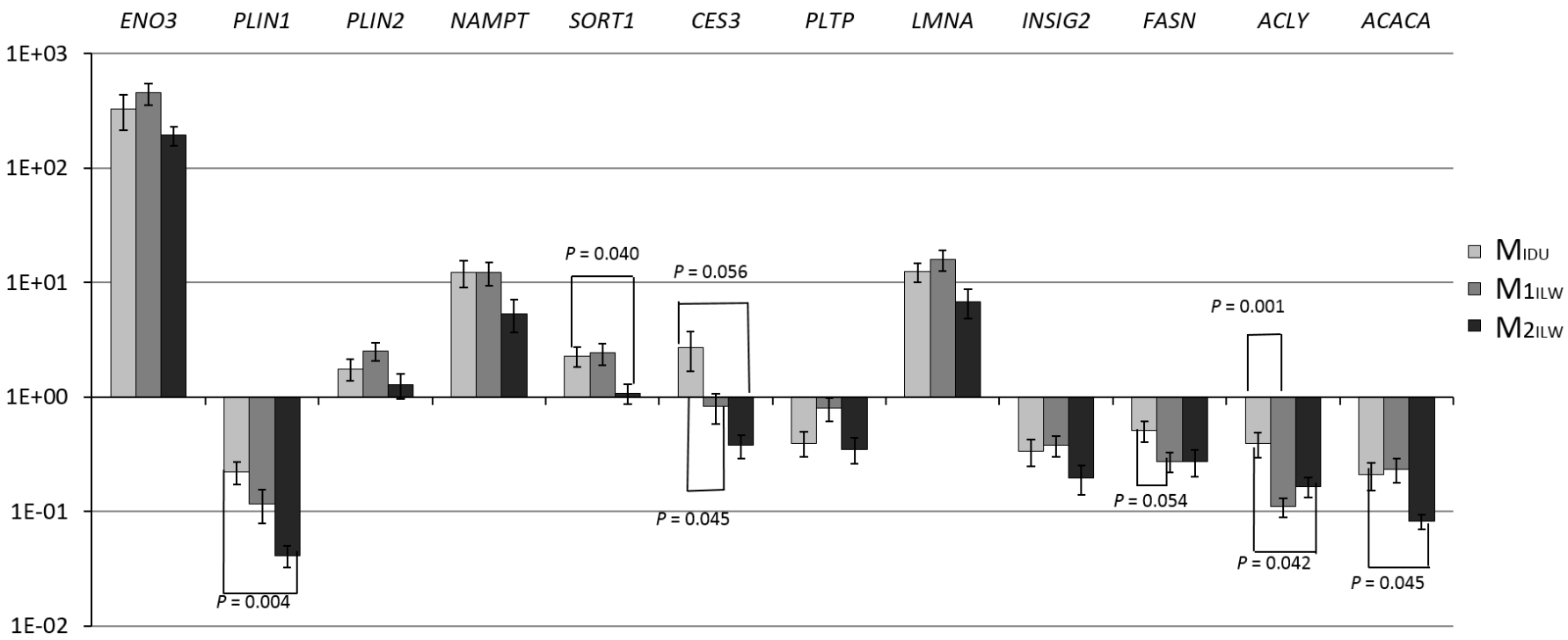
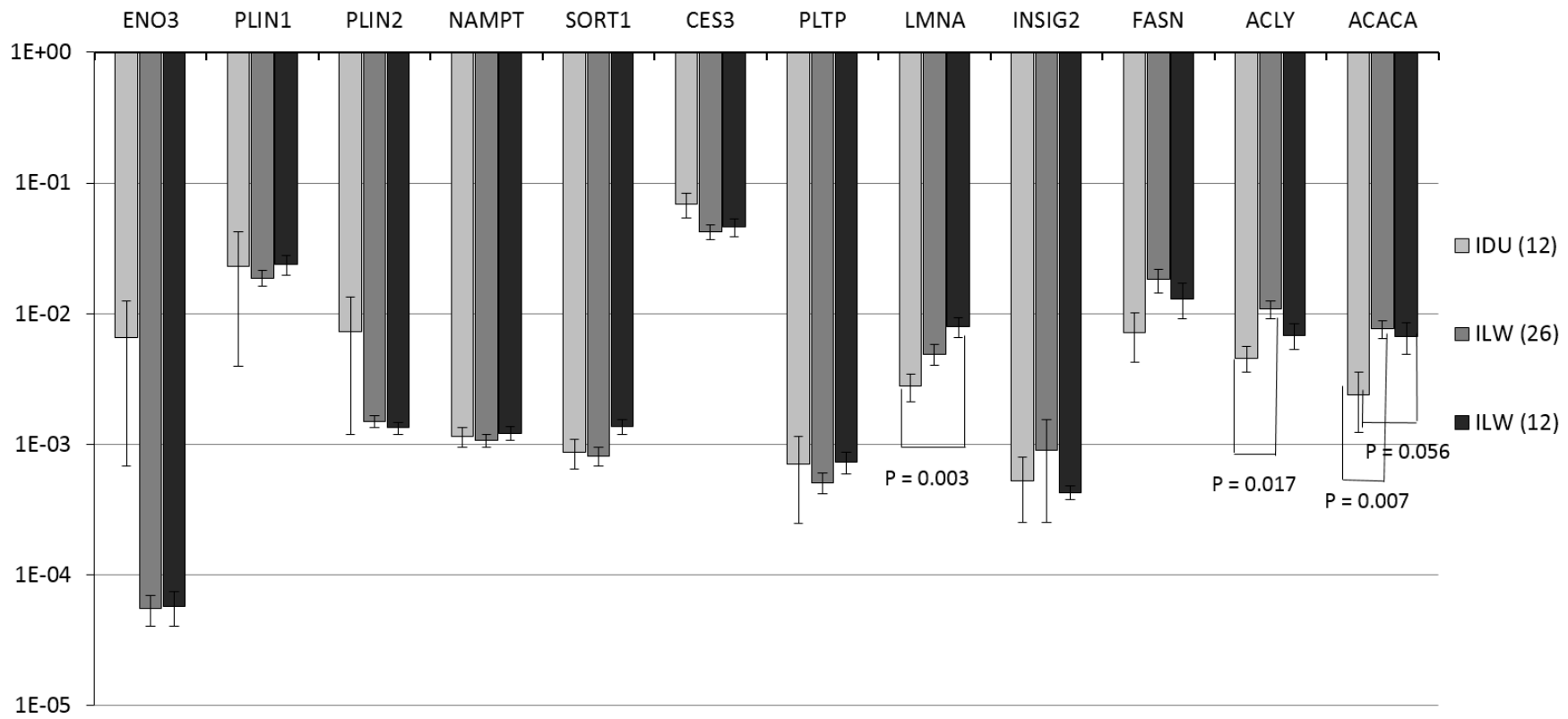


Figure 2. The studied gene expression levels in pig backfat tissue for the Italian Duroc (B_{IDU}) and the Italian Large White (B_{ILW}) groups.



Supplementary Table S1. The list of the 14 genes considered for the present expression study in pig, their chromosome mapping obtained from NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene) accessed on 11 September 2015, and the role of the coded protein.

GENE	Coded Protein	Chromosome[1]	Role	Literature
PLIN2	Perilipin 2	SSC1	*PLIN2* codes for a structural protein involved in development of adipose tissue.	Gandolfi *et al.,* 2011
LMNA	Lamin A/C	SSC4	*LMNA* codes for a protein involved in nuclear stability and chromatin structure. It showed an increased expression level during adipogenesis.	Lopez-Mejia *et al*., 2014
SORT1	Sortilin1	SSC4	*SORT1* codes for a protein receptor that also participates in lipoprotein metabolism. Its overexpression in adipocyte and neurons increases the insulin-stimulated glucose uptake.	Coutinho *et al*., 2001
INSIG2	Insulin induced gene 2	SSC7	*Insulin Induced Gene 2* mediates feedback control of cholesterol synthesis and body mass index.	Malzahn *et al*., 2014
PLIN1	Perilipin 1	SSC7	*PLIN1* codes for a lipid droplet coat protein that modulates lipid droplets hydrolysis	Nowacka-Woszuk *et al*., 2008 Vykoukalová *et al.,* 2009
MTTP	Microsomal triglyceride transfer protein	SSC8	*MTTP* codes for a protein that catalyses the transport of triglyceride, cholesteryl ester and phospholipid.	Estellé *et al*., 2009
NAMPT	Nicotinamide phosphoribosyltransferase	SSC9	*NAMPT* codes for a protein involved in a rate-limiting reaction in mammalian NAD biosynthesis. This gene shows different expression levels in adipose tissue and regulates glucose-stimulated insulin secretion.	Čepica *et al*., 2010
ACACA	Acetyl-CoA carboxylase alpha	SSC12	*ACACA* codes for a protein catalysing the biogenesis of long-chain fatty acids.	Muñoz *et al.,* 2012
ACLY	ATP citrate lyase	SSC12	*ACLY* codes for a protein catalysing the synthesis of cytosolic acetyl-CoA and lipid synthesis. It may be involved in the biosynthesis of acetylcholine.	Davoli *et al*., 2014

ENO3	Enolase 3 (Beta, muscle)	SSC12	*ENO3* codes for a protein promoting the interconversion from diphosphoglycerate to phosphoenolpyruvate in glycolysis.	Jian *et al*,.2008
FASN	Fatty acid synthase	SSC12	*FASN* codes for a multifunctional protein that catalyses formation of long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH. It is a putative candidate gene for body fat development.	Braglia *et al.,* 2014
ACVR1C	Activin A receptor type-1C	SSC15	*ACVR1C* codes for a multifunctional receptor in adipose tissue and its role is also correlated to carbohydrate and lipid metabolism. *ACVR1C* gene expression is elevated during adipocyte differentiation.	Murakami *et al.,* 2013
PLTP	Phospholipid transfer protein	SSC17	*PLTP* codes for a transporting phospholipids protein in plasma, which has multifaceted functions in lipoprotein metabolism and in cholesterol metabolism.	Chang *et al.,* 2011
CES3	Carboxylesterase 3	Unplaced	*CES3* codes for an endoplasmic reticulum carboxylesterase that may be involved in cholesterol and ester metabolism. This gene is highly expressed in adipose tissue.	Chang *et al*., 2014

[1] SSC is the acronym used for *Sus scrofa* chromosome.

References for Supplementary Table S1

Čepica S, Bartenschlager H, Ovilo C, Zrůstová J, Masopust M, Fernández A, López A, Knoll A, Rohrer GA, Snelling WM & Geldermann H (2010) Porcine NAMPT gene: search for polymorphism, mapping and association studies. *Anim Genet* 41:646-651.

Chang J, Oikawa S, Iwahashi H, Kitagawa E, Takeuchi I, Yuda M, Aoki C, Yamada Y, Ichihara G, Kato M & Ichihara S (2014) Expression of proteins associated with adipocyte lipolysis was significantly changed in the adipose tissues of the obese spontaneously hypertensive/NDmcr-cp rat. *Diabetol Metab Syndr* 6:8.

Coutinho MF, Bourbon M, Prata MJ & Alves S (2013) Sortilin and the risk of cardiovascular disease. *Rev Port Cardiol* 32:793-799.

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. *Livest Sci* 165:1-7.

Estellé J, Fernández AI, Pérez-Enciso M, Fernández A, Rodríguez C, Sánchez A, Noguera JL Folch JM (2009) A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. *Anim Genet* 40:813-820

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V & Davoli R (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631-637.

Jian W, Donghai Z, Changyan D, Xiaoxiong W, Liangqi L & Yuanzhu X (2008) Characterization of porcine ENO3: genomic and cDNA structure, polymorphism and expression. *Genet Sel Evol* 40:563-579.

Lopez-Mejia IC, de Toledo M, Chavey C, Lapasset L, Cavelier P, Lopez-Herrera C, Chebli K, Fort P, Beranger G, Fajas L, Amri EZ, Casas F & Tazi J (2014) Antagonistic functions of LMNA isoforms in energy expenditure and lifespan. *EMBO Rep* 15:529-539.

Malzahn D, Müller-Nurasyid M, Heid IM, Wichmann HE, KORA study group, Bickeböller H (2014) Controversial association results for INSIG2 on body mass index may be explained by interactions with age and with MC4R. *Eur J Hum Genet* 22:1217-1224.

Muñoz M, Alves E, Corominas J, Folch JM, Casellas J, Noguera JL, Silió L & Fernández A (2012) Survey of SSC12 regions affecting fatty acid composition of intramuscular fat using high-density SNP data. *Front Genet* 2:101.

Murakami M, Shirai M, Ooishi R, Tsuburaya A, Asai K, Hashimoto O, Ogawa K, Nishino Y & Funaba M (2013) Expression of Activin Receptor-like Kinase 7 in Adipose Tissues. *Biochem Genet* 51:202-210.

Nowacka-Woszuk J, Szczerbal I, Fijak-Nowak H & Switonski M (2008) Chromosomal localization of 13 candidate genes for human obesity in the pig genome. *J Appl Genet* 49:373-377.

Vykoukalová Z, Knoll A & Čepica S (2009) Porcine perilipin (PLIN) gene: structure, polymorphism and association study in large white pigs. *Cz J Anim Sci* 54:359-364.

Supplementary Table S2. Primer used for the expression study in pig, in muscle and backfat tissues.

Primer name	Primer sequence (5'-3')	TM (°C)	Size (bp)
Primers used to check DNA contamination in mRNA samples			
AGPAT1			
AGPAT1 FOR	5'- AGGACGCAACGTCGAGAACA -3'	61°C	110
AGPAT1 REV	5'- GTGAGGGAGGGAAGTGGTGAG -3'		
Primers used for expression quantification			
PLIN2			
PLIN2-E-FOR	5'- GGGCAAAAGATGCTATGACG -3'	65°C	244
PLIN2-E-REV	5'- TCAGTGAGAGGGAGGTACTGG -3'		
PLIN2-I-FOR	5'- ATCACTGAGGTGGTGGACAAG -3'	63°C	112
PLIN2-I-REV	5'- GCTGCATCATCCGACTTCC -3'		
SORT1			
SORT1-E-FOR	5'- ACCAAGGAGGAAGGTGGAAA -3'	59°C	340
SORT1-E-REV	5'- TTGATAGGATGGCTGCTGTG -3'		
SORT1-I-FOR	5'- ACGCTTCCTACAGCATCTCC -3'	60°C	179
SORT1-I-REV	5'- GGGTCCTTCCAGCATCTTC -3'		
LMNA			
LMNA-E-FOR	5'- GTGACCATGATTGAGGACGA -3'	63°C	313
LMNA-E-REV	5'- AGTTGCCCAGGAGGTAGGAG -3'		
LMNA-I-FOR	5'- GAGGATGAGGATGGAGATGA -3'	66°C	189
LMNA-I-REV	5'- GGAGGAGCCAGAGGAGATG -3'		
CES3			
CES3-E-FOR	5'- TTCTGGTGTTGTCTCCCTTG -3'	67°C	380
CES3-E-REV	5'- GCCAAACTCTTGCTTGTTGA -3'		
CES3-I-FOR	5'- CACCTCGGCTGTCTTTGTTC -3'	60°C	176
CES3-I-REV	5'- GAATCTCTTCAGGCATCTTGG -3'		
PLIN1			
PLIN1-E-FOR	5'- AGACAAAGTCCTCGGTGCTG -3'	60°C	477
PLIN1-E-REV	5'- TGGTCATCGTGATCCTCCTC -3'		
PLIN1-I-FOR	5'- TCCAGCCAAGGAAGAGTCAG -3	63°C	125
PLIN1-I-REV	5'- GGAAGGTGTGTTGAGAGATGG -3'		

MTTP			
MTTP-E-FOR	5’- TGACCTTCATTCAGCACCTC -3’	56°C	399
MTTP-E-REV	5’- TTTCAAGTCCACCCAGGATT -3’		
MTTP-I-FOR	5’- TGTGGATGCTGTCACCTCT -3’	66°C	194
MTTP-I-REV	5’- TATCGTTGCTCCCAAAGGAA -3’		
NAMPT			
NAMPT-E-FOR	5‘- TGGGCATCTTCCAATAGAGG -3’	57°C	410
NAMPT-E-REV	5’- CTGCTGCTGGCACTGAATAG -3’		
NAMPT-I-FOR	5’- CCTGGTATCCAATCACAGTGG -3’	63°C	158
NAMPT-I-REV	5’- CCCTATGCCAGCAGTCTCTT -3’		
ENO3			
ENO3-E-FOR	5’- GTCATCAAGGGCAAATACGG -3’	61°C	334
ENO3-E-REV	5’- AGGTCCAGGTCTTCCAGTCA -3’		
ENO3-I-FOR	5’- CTACCGCAACGGGAAGTATG -3’	63°C	162
ENO3-I-REV	5’- TCAATGGAGACCACAGGATAG -3’		
FASN			
FASN-E-FOR	5’- CCAGCATCACCATAGACACG -3’	59°C	317
FASN-E-REV	5’- CTCCTTGGAACCGTCTGTGT -3’		
FASN-I-FOR	5’- AACGTCCTGCTGAAGCCTAA -3’	62°C	181
FASN-I-REV	5’- CATTGAGGATGGTGGCGTAT -3’		
ACACA			
ACACA-E-FOR	5’- GCGAGCAACATCACATCAGT -3’	57°C	496
ACACA-E-REV	5’- GCAAATGGGAGGCAATAAGA -3’		
ACACA-I-FOR	5’- TACCTGCGAGTGGAGACACA -3’	60°C	127
ACACA-I-REV	5’- TGGTGACTTGAGCGTGAGAG -3’		
ACLY			
ACLY-E-FOR	5’- CCCAGACATGAGAGTGCAGA -3’	61°C	359
ACLY-E-REV	5’- ATCCCAAGGGTGACGATACA -3’		
ACLY-I-FOR	5’- CCTTATCCTGAATGTAGACGGTTT -3’	63°C	151
ACLY-I-REV	5’- AATGAAGCCCATACTCCTTCC -3’		
INSIG2			
INSIG2-E-FOR	5’- GGAGAGACAAAGTCACCTGGA -3’	63°C	329
INSIG2-E-REV	5’- GCTACGCACCGCATTACACT -3’		

INSIG2-I-FOR	5’- TCTTTCCACCGGATGTGATT -3’	63°C	164
INSIG2-I-REV	5’- GCACCGCATTACACTGGAC -3’		
ACVR1C			
ACVR1C-E-FOR	5’- CAGCAAGATTACCCCTTCACA -3	60°C	340
ACVR1C -E-REV	5’- GGCACTCCATATCGTCCTTG -3’		
ACVR1C -I-FOR	5’- AGAGTAATTTTGGAAGCCCACA -3’	60°C	168
ACVR1C -I-REV	5’- TGTAAAGGTTCTCCCTCAG -3’		
PLTP			
PLTP-E-FOR	5’- CTGGAGCAAGAGCTGGAGAC -3’	62°C	315
PLTP-E-REV	5’- GCAGGAGACGTTGGACACTT -3’		
PLTP-I-FOR	5’- AGCCAGAGCAGGAGCTGA -3’	66°C	139
PLTP-I-REV	5’- AGAGCCGTGTGGATGGAA -3’		
Normalising genes			
PPIA			
PPIA-E-FOR	5’- TAACCCCACCGTCTTCTTC -3’	64°C	317
PPIA-E-REV	5’- GTTTGCCATCCAACCACTC -3’		
PPIA-I-FOR	5’- AAAACTTCCGTGCTCTGAG -3’	63°C	125
PPIA-I-REV	5’- GCCACCAGTGCCATTAT -3’		
TMSB4X			
TMSB4X-E-FOR	5’- ACTGCGTAGACCGGATTCCT -3’	62°C	339
TMSB4X-E-REV	5’- TTCGCCATTCTTTGATGTGA -3’		
TMSB4X-I-FOR	5’- GCCAGCTTGCTTCTCTTGTTC -3’	63°C	120
TMSB4X-I-REV	5’- GACAAACCCGATATGGCTG -3’		
B2M			
B2M-E-FOR	5’- CACTTTTCACACCGCTCCA -3’	57°C	218
B2M-E-REV	5’- TCTCCCCGTTTTTCAGCA -3’		
B2M-I-FOR	5’- TCGGGCTGCTCTCACTGT -3’	56°C	135
B2M-I-REV	5’- TCTGGGGCGGATGGAAC -3’		
POLR2A			
POLR2A-E-FOR	5’- CACCCACAGCACCCATCC -3’	61°C	573
POLR2A-E-REV	5’- CCCTCCACATTCTGCTG -3’		
POLR2A-I-FOR	5’- TCGCCTCTTCTATTCCAA -3’	60°C	165
POLR2A-I-REV	5’- GCCTTCTCGATGACCTC -3’		

For all the considered genes in the table below are reported the used primer, the annealing temperature expressed in °C and the size in base pairs (bp).

Each primer name is composed by the gene name, “I” for internal primer or “E” for external primer, “FOR” for the left primer and “REV” for the right primer.

[1]TM (°C) Annealing temperature

Chapter 4: Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs.

S. Gol[1], R. Ros-Freixedes[1], P. Zambonelli[2], P. Zambonelli[2], M. Tor[1], R.N. Pena[1], S. Braglia[2], M. Zappaterra[2], J. Estany[1] and R. Davoli[2]

[1]Departament de Producció Animal-Agrotecnio Center, Universitat de Lleida, 25198 Lleida, Catalonia, Spain

[2]Department of Agricultural and Food Science (DISTAL), University of Bologna, via Fratelli Rosselli, 107, 42123 Reggio Emilia, Italy

Published in

Journal of Animal Breeding and Genetics

Gol et al. *J Anim Breed Genet.* 2016 Feb;133(1):24-30.

doi: 10.1111/jbg.12159.

Summary

The perilipins (PLIN) belong to a family of structural proteins that play a role regulating intracellular lipid storage and mobilization. Here, *PLIN1* and *PLIN2* have been evaluated as candidate genes for growth, carcass, and meat quality traits in pigs. A sample of 607 Duroc pigs were genotyped for two single nucleotide polymorphisms, one in intron 2 of the *PLIN1* gene (*JN860199:g.173G>A)* and the other at the 3' untranslated region of the *PLIN2* gene (*GU461317:g.98G>A*). Using a Bayesian approach we have been able to find evidence of additive, dominant, and epistatic associations of the *PLIN1* and *PLIN2* polymorphisms with early growth rate and carcass length. However, the major effects were produced by the dominant allele A at the *PLIN2* polymorphism, which also affected the carcass lean weight. Thus, pigs carrying an additional copy of the allele A at the *g.98G>A PLIN2* polymorphism had a probability of at least 98% of producing carcasses with heavier lean weight (+0.41 kg) and ham weight (+0.10 kg). The results obtained indicate that the *PLIN2* polymorphism can be a useful marker for lean growth. In particular, it may help to reduce the undesired negative correlated response in lean weight to selection for increased intramuscular fat content, a common scenario in some Duroc lines involved in the production of high quality pork products.

Introduction

Growth rate and carcass lean content are crucial characteristics for the economic viability of pork production. Selection emphasizing lean content has led to reduce some pork quality attributes, including the intramuscular fat (IMF) content. The use of molecular markers may be useful to improve the genetic progress in traits that are difficult and expensive to measure (Dekkers, 2004), but also to break down unfavorable genetic correlations between antagonistic traits, such as those between lean growth rate or carcass lean content and IMF content (Ros-Freixedes et al., 2012; Ros-Freixedes et al., 2013). In this scenario, performing association studies with candidate genes related to proteins affecting fat metabolism is of particular interest. The perilipins (PLIN) belong to a family of structural proteins that coat intracellular lipids into cytosolic droplets (Kimmel et al., 2010), where they regulate intracellular lipid storage and mobilization by fine-tuning the activity of lipases (Bickel et al., 2009). The composition of PLIN changes as lipid droplets enlarge and mature. Perilipin 2 (PLIN2) is the most prominent PLIN protein in most adult cell types and in immature adipocytes. In contrast, the large central mature lipid droplets of mature adipocytes are largely coated by perilipin 1 (PLIN1). Recently, PLIN1 and PLIN2 have been shown to co-localize in the skeletal muscle of pigs (Gandolfi et al., 2011).

Mutations in the *PLIN* genes have been associated to body fat mass in mice (Saha et al., 2004) and humans (Qi et al., 2004; Corella et al., 2005; Ruiz et al., 2011). So far only two reports in pigs have investigated the association of *PLIN1* and *PLIN2* polymorphisms with a limited number of production traits. In the first report, two synonymous single nucleotide polymorphisms (SNP) in exons 3 and 6 of *PLIN1* showed suggestive associations with average daily gain (ADG) and backfat thickness in Large White pigs (Vykoukalová et al., 2009). In a second study, a 3' untranslated region (UTR) SNP at the *PLIN2* gene (*GU461317:g.98G>A*) was found to be associated to lean growth and content but not to visible intermuscular fat (Davoli et al., 2011). The aim of the present study was to further investigate the contribution of *PLIN1* and *PLIN2* genes to a wider range of performance, carcass, and meat quality traits in pigs and, in particular, to confirm whether *PLIN1* and *PLIN2* genotype variants exert a differential effect on lean growth and IMF content.

Materials and methods

Animals, traits and sample collection

A total of 607 Duroc barrows from 88 sires and 348 dams were used for the analyses. Pigs were randomly sampled in seven batches from the same commercial line and performance-tested from 75 d to 210 d of age under commercial conditions (Ros-Freixedes et al., 2012). During the test period they had *ad libitum* access to commercial diets. A complete description of the line and of the procedures followed for testing and sample collection is given in Ros-Freixedes et al. (2012). The traits recorded included live body weight (BW), backfat thickness, and loin thickness at 120, 180, and 205 d. Backfat and loin thickness was ultrasonically measured at 5 cm off the midline at the position of the last rib (Piglog 105, Herlev, Denmark). After slaughter at 210 days, the carcass weight and length, the carcass backfat and loin thickness, and the ham weight were measured. Carcass backfat and loin thickness at 6 cm off the midline between the third and fourth last ribs, together with the carcass lean percentage, were estimated using an on-line ultrasound automatic scanner (AutoFOM, SFK-Technology, Herlev, Denmark). After chilling for about 24 h at 2ºC, the pH was measured in the *longissimus dorsi* and in the *semimembranosus* muscles. Samples of at least 50 g of *gluteus medius* muscle and *longissimus dorsi* were taken, immediately vacuum packaged, and stored in deep freeze until required for IMF content and fatty acid determination (Bosch et al., 2009).

Single nucleotide polymorphism genotyping

Genomic DNA was isolated from freeze-dried muscle samples using standard protocols (Sambrook et al., 1989). To search for sequence variation in the pig *PLIN1* gene, the genomic, cDNA, and EST sequences available in the GenBank (http://www.ncbi.nlm.nih.gov/Genbank) and in the Ensembl databases (http://www.ensembl.org) were compared for an *in silico* variability analysis. Italian heavy pigs were used to validate the in silico-identified SNPs.

Seven primer pairs (Supplementary Table S1) were designed using Primer3 v.0.4.0 software (http://frodo.wi.mit.edu/primer3/) to amplify seven porcine *PLIN1* gene

fragments. The PCR products were sequenced on both strands using the BigDye Terminator v3.1 Cycle Sequencing kit (Life Technologies, Grand Island, NY, USA) in a ABI PRISM 3100-Avant Genetic Analyzer (Life Technologies). The sequences obtained were compared by multiple alignments, performed with MEGA software v4.0 (www.megasoftware.net/).

The *JN860199:g.173G>A PLIN1* SNP polymorphism, which was selected for subsequent analyses, was genotyped by PCR-restriction fragment length polymorphism assay. PCR products obtained with the "P2" primer set (Supplementary Table S1) were digested with *Hin*1II (Fermentas, Vilnius, Lithuania) and the resulting products were resolved on polyacrylamide gels. For *PLIN2*, the *GU461317:g.98G>A* SNP*,* in the 3' UTR region of the gene, was genotyped by High Resolution Melting PCR in a Rotor-Gene™ 6000 (Corbett Research, Mortlake, New South Wales, Australia) following the protocol described in Davoli et al. (2011). The linkage disequilibrium between SNP was estimated as r^2 using the Haploview software (Barrett, 2009).

Association analysis

The additive, dominant, and epistatic effects of the *PLIN* genotypes were estimated independently for each trait using a Bayesian setting, in line with the methodology described in Ros-Freixedes et al. (2012). A two-generation pedigree was used for the analyses. In matrix notation, the model used for the *i*th trait was $y_i = X_i b_i + Z_i a_i + e_i$, where y_i is the vector of observations for trait *i*; b_i, a_i, and e_i are the vectors of systematic, polygenic, and residual effects, respectively; and X_i and Z_i the known incidence matrices that relate b_i and a_i with y_i, respectively. The systematic effects were the batch (7 levels), the age at test as a covariate, and orthogonal coefficients for additive (a), dominance deviation (d) and first-order epistatic effects (aa: additive × additive; ad: additive × dominance; da: dominance × additive; and dd: dominance × dominance) for *PLIN1* and *PLIN2* SNPs. Pigs in a given batch were contemporaneous pigs tested at the same unit and slaughtered in the same abattoir. The litter effect was not included because, on average, there were less than 2 piglets per litter. The orthogonal coefficients for the genetic effects were calculated using the algorithm proposed by Alvarez-Castro & Carlborg (2007).

The models were solved using Gibbs sampling with the TM software (Legarra et al., 2008). The traits were assumed to be conditionally normally distributed as $[y_i | b_{i,} a_{i,} I\sigma^2_{ei}] \sim N(Xb_i + Za_{i,} I\sigma^2_{ei})$, where σ^2_{ei} is the residual variance and I the appropriate identity matrix. The animal effects conditionally on the additive genetic variance σ^2_{ai} were assumed multivariate normally distributed with mean zero and variance $A\sigma^2_{ai}$, where A was the numerator relationship matrix. The matrix A was calculated using 1043 animals in the pedigree. Flat priors were used for b_i while the variance components were set to the values obtained by Ros-Freixedes et al. (2013) with data and pedigree from 1996 onwards. Statistical inferences were derived from the samples of the marginal posterior distribution using a unique chain of 500,000 iterations, where the first 100,000 were discarded and one sample out of 100 iterations retained. The additive, dominance, and epistatic effects were assessed by calculating both the probability of each of these components being greater or lower than zero and their highest posterior density interval at 95% of probability (HPD95). Statistics of marginal posterior distributions and the convergence diagnostics were obtained using the BOA package (Smith, 2005). Convergence was tested using the *Z*-criterion of Geweke (Geweke, 1992) and visual inspection of convergence plots.

Results and discussion

Polymorphisms and sequence variation of PLIN genes

The *in silico* analysis on publicly available genomic, EST, and cDNA sequences revealed ten SNPs (detected at least twice) within the coding sequence of *PLIN1*, located in the exons 1, 2, 5, and 8 (data not shown) and five SNP in intronic regions. Seven genomic regions containing these putative SNP were subjected to direct sequencing in 20 animals from three Italian heavy pig breeds. A total of 2,437 bp of the pig *PLIN1* gene were screened, which covered 1,126 bp of the coding sequence, the complete 183-bp 5' UTR, and 1,128 bp of intronic regions and part of the promoter and 3' downstream genomic region, according to the annotation of the Ensembl entry [ENSSSCG00000001844]. The sequencing covered the positions of the putative SNP detected *in silico*, with the exception of the SNPs on exon 8, which were not analyzed

due to the unsuccessful amplification of this region. Four SNP (two intronic and two exonic) were detected by sequencing in the Italian heavy pig breeds (Table 1). The two intronic SNPs were novel and the sequences were reported to GenBank [JN860199; SNP *g.173G>A* and *g.3484C>G*]. The two exonic SNPs, which were detected in our *in silico* analysis, were both synonymous and had been reported before (GenBank: AM931171; SNP *g.4119A>G* and *g.7966T>C*; Vykoukalová et al., 2009). The four SNP were in complete linkage disequilibrium in the initial panel of 20 pigs. The intronic *JN860199g.173G>A* SNP was selected for subsequent analyses because a restriction enzyme was available to analyze this mutation.

To assess the association of these mutations with productive parameters, the *PLIN1 JN860199:g.173G>A* and *PLIN2 GU461317:g.98G>A* SNPs were genotyped in a population of 607 Duroc pigs, which had data available on performance, fattening and meat quality traits (Ros-Freixedes et al. 2012).

The allele frequencies and the distribution genotypes for *PLIN1* and *PLIN2* SNPs are reported in Table 2. In both SNPs the alleles were segregating at intermediate frequencies, with the allele G being the less frequent in *JN860199:g.173G>A* (minor allele frequencies of 0.38) and alleles G and A showing identical gene frequency for *GU461317:g.98G>A*. As expected, since *PLIN1* and *PLIN2* are lying in different chromosomes, the two SNPs were in linkage equilibrium (r^2= 0.04).

Effect of PLIN genotypes

The additive, dominant, and epistatic effects of *PLIN1 g.173G>A* and *PLIN2 g.98G>A* SNPs associated to BW and growth rate at different ages during the fattening period are given in Table 3. The substitution of A for G in *PLIN1* showed some evidence of a negative additive effect on BW (-0.66 kg at 120 d and -0.68 kg at 180 d, with a probability of 6% and 10% of being greater than zero, respectively), but a strong evidence of a positive additive effect in *PLIN2*, with values of +0.95 kg, +1.19 kg, and +1.08 kg at 120 d, 180 d and 205 d, respectively, with an associated probability of being greater than zero superior to 95% in the three ages. The substitution effect of A for G for BW was similar at 120 d, 180 d, and 205 d, thereby indicating that the beneficial effect of allele A on BW was due to increased growth at early stages. In agreement with this, the effect of allele A at *PLIN2* for ADG was evident up to 120 d

(+7.26 g/d, with a probability of being positive of 98%) but not thereafter, both from 120 to 180 d (+4.15 g/d) and from 180 to 205 d (-0.42 g/d). Consequently, the variance associated to the additive effects of *PLIN2 g.98G>A* SNP (Falconer and Mackay, 1996) is able to capture a greater proportion of the total additive variance of BW (Ros-Freixedes et al., 2013) at 120 d (1.49%) than at 205 d (1.12%). A negative dominant effect for BW at 120 and 180 days in *PLIN1* (-1.04 kg and -1.56 kg, respectively, with associated probabilities of being negative of 95% and 97%) and a positive dominant effect for BW at 180 days in *PLIN2* (+1.17 kg, with associated probability of being positive of 94%) were observed. No clear evidence of epistasis between *PLIN1* and *PLIN2* SNP was observed for BW and ADG, with the exception of an additive × additive effect for BW at 120 d (-0.88 kg, with associated probability of being positive of 6%) and for ADG up to 120 d (-7.94 g/d, with associated probability of being positive of 4%).

The additive, dominant, and epistatic effects of *PLIN1 g.173G>A* and *PLIN2 g.98G>A* SNPs associated to backfat and loin thickness at 120 d, 180 d and 205 d of age are given in Table 4. The *PLIN1 g.173G>A* SNP did not show a clear pattern of association with fatness traits, but results on the *PLIN2 g.98G>A* SNP indicated that allele A is positively associated to backfat thickness at early ages (+0.17 mm and +0.19 mm, at 120 d and at 180 d, respectively, with a probability of being positive of 91% and 98%) and negatively to backfat thickness at 205 d (-0.22 mm, with a probability of being positive of 10%). The effect of the *PLIN2 g.98G>A* SNP on backfat thickness followed a similar pattern as for ADG, with the positive effect of allele A at 120 d vanishing at later ages.

In agreement with these results, no strong evidence of association of *PLIN1* and *PLIN2* SNPs with carcass backfat thickness, and carcass loin thickness was observed (Table 5). However, allele G at *PLIN1* and allele A at *PLIN2* had some beneficial effects on carcass traits. Thus, pigs carrying an additional copy of allele G at *PLIN1* and allele A at *PLIN2* had longer carcasses (+0.62 cm and +0.43 cm, with a probability of being positive greater than 96% and 99%, respectively) and, more interestingly, those carrying allele A at *PLIN2* showed a higher carcass lean weight (+0.41 kg, with a probability of being positive of 99.9%). This latter effect should be interpreted as a result of a moderate but favorable change in both carcass weight (+0.58 kg, with a probability of being

positive of 86%), mostly as a consequence of increased growth rate at early ages, and carcass lean percentage (+0.23%, with a probability of being positive of 80%). As a result, the *PLIN2 g.98G>A* SNP reached to explain 0.59% of the additive variance of lean weight. Moreover, a positive effect of allele A at *PLIN2* on ham weight was also detected (0.10 kg, with a probability of being positive of 94%).

No evidence was found indicating that meat quality traits (pH and IMF) were additive by *PLIN1* and *PLIN2* SNP, although some minor changes were observed for IMF fatty acid composition (Table 6). In particular, allele A at *PLIN1* decreased PUFA (-0.20%) and increased MUFA (-0.20%) while allele A at *PLIN2* decreased SFA (-0.24%). Evidence supporting the existence of dominant and epistatic effects associated to carcass and meat quality traits was mostly circumscribed to traits where the additive effects were more evident (carcass length and carcass lean weight), thereby suggesting that the mode of action of *PLIN1* and *PLIN2* on the traits that they are influencing is subjected to complex regulations. As for BW and ADG, the dominant effect associated to lean weight was negative in *PLIN1* (-0.19 kg, with a probability of 2% of being positive) but positive in *PLIN2* (0.41 kg, with 99.9% probability of being positive). These dominant values were around two-fold higher than their respective additives, a result which supports for an underdominant *PLIN1* and overdominant *PLIN2* gene action for lean weight. To check for overparameterization, the additive and dominance effects were estimated ignoring the epistatic effects. The estimates obtained (results not shown), although slightly higher, were in line with those reported previously including epistatis, thereby confirming the favourable effects of allele G at *PLIN1* and allele A in *PLIN2*.

Our findings are consistent with the results in Vykoukalová et al. (2009), who found suggestive associations of the two exonic *PLIN1* SNP with ADG in Large White pigs, and, particularly, with those in Davoli et al. (2011), who reported a favourable effect of allele A at *PLIN2* on ADG, feed conversion ratio, lean cuts, and ham weight estimated breeding values in Italian Duroc. The five members of the *PLIN* family have been studied in depth in humans and model animals. Most reports have focussed on *PLIN1*, the main perilipin protein in mature adipocytes, particularly in relation to BW and obesity-related phenotypes (Smith and Ordovas, 2012), but results do not show a consistent trend across them. It must be taken into account that, depending on the energy state of the organism, *PLIN1* either limits lipase access to stored triglycerides

(in the fed state) or facilitates hormonally stimulated lipolysis (in the fasted state). This dual activity is illustrated by the fact that both *PLIN1*-null and *PLIN1*-overexpressing mice are protected from diet-induced obesity (Saha et al, 2004). In our pig population, mutations in the *PLIN1* did not correlate with growth or fat deposition traits. This indicates that genes other than *PLIN1* are the main players of fat deposition in pig, or that other mutations outside the transcribed sequence, for instance in the regulatory 5' or 3' regions, might have a more relevant effect over the expression of the gene. In contrast, only few reports in humans and mice have focused on *PLIN2* gene. Our results indicate that allele A at the *PLIN2 g.98G>A* SNP has beneficial effects on early growth, lean growth and prime retail cuts. In agreement with this, the genomic position of *PLIN2* on chromosome 1 co-localises with quantitative trait loci for ADG (Liu et al., 2007), BW at birth (Guo et al., 2008), and daily feed intake (Kim et al., 2000) (Supplementary Table S2). Of the five PLIN proteins, PLIN2 and 3 are by far the most prominent in human skeletal muscle (Gjelstad et al., 2012), with PLIN2 accounting for >60% of total perilipin content. It has been shown that PLIN2 is also the main perilipin in pig muscle (Gandolfi et al., 2012). Therefore, it is not surprising that *PLIN2* is related to growth and lean weight, as perilipins regulate not the deposition of fat *per se*, but more importantly, the accessibility of lipases to the stored fats in response to the energy demands of the cells.

Our results indicate that *PLIN2 g.98G>A* SNP can be a useful marker for lean growth, which is a relevant trait for the pig industry in general, very interested in fast- growing lean animals. Although results are encouraging for Duroc, further association studies are needed to confirm whether this polymorphism similarly affects other pig breeds. However, it is in this breed where it can be of particular interest. Duroc lines are the most used in premium quality markets, where pigs are raised to heavy weights and IMF becomes a key trait. In such scenario it is very convenient to find selection criteria addressed to reduce the undesired negatively correlated response on BW to selection for IMF.

Acknowledgements

This work was partially funded by national projects AGL2009-09779 and AGL2012-33529 (Ministry of Education and Innovation, Spain) and AGER-HEPIGET grant N. 2011-0279 (Italy).

References

Alvarez-Castro JM & Carlborg O (2007) A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. *Genetics* 176:1151–67.

Barret JC (2009) Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc* 2009(10):pdb.ip71.

Beeckmann P, Schröffel J, Moser G, Bartenschlager H, Reiner G & Geldermann H (2003) Linkage and QTL mapping for Sus scrofa chromosome 1. *J. Anim Breed Genet* 120:1-10.

Bickel PE, Tansey JT & Welte MA (2009) PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. *Biochim Biophys Acta* 1791:419-440.

Bosch L, Tor M Reixach J & Estany J (2009) Estimating intramuscular fat content and fatty acid composition in live and post-mortem samples in pigs. *Meat Sci* 82:432-437.

Corella D, Sorli QL, Godoy JV, Portoles D & Coltell O (2005) Obese subjects carrying the 11482G>A polymorphism at the perilipin locus are resistant to weight loss after dietary energy restriction. *J Clin Endocrinol Metab* 90:5121-5126.

Davoli R, Gandolfi G, Braglia S, Comella M, Zambonelli P & Buttazzoni L (2011) New SNP of the porcine perilipin 2 (PLIN2) gene, association with carcass traits and expression analysis in skeletal muscle. *Mol Biol Rep* 38:1575-1583.

Falconer S & Mackay FC (1996) *Introduction to Quantitave Genetics*. Pearson Prentice Hall. Harlow, England.

Dekkers JC (2004) Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *J Anim Sci* 82:E313-E328.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A & Russo V (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of european cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631-637.

Geldermann H, Čepica S, Stratil A, Bartenschlager H, Preuss S (2010) Genome-wide mapping of quantitative trait loci for fatness, fat cell characteristics and fat metabolism in three porcine F2 crosses. *Genet Sel Evol* 28:42-31.

Geweke J (1992) Evaluating the accuracy of sampling-based approached to the calculation of posterior moments (with discussion). In *Bayesian statistics.* [Bernardo JM, Berger JO, Dawid AP, Smith AFM editors] Oxford Univ. Press, Oxford, UK. 69–193.

Gjelstad I, Haugen MF, Gulseth H, Norheim LF, Jans A & Bakke SS (2012) Expression of perilipins in human skeletal muscle in vitro and in vivo in relation to diet, exercise and energy balance. *Arch Physiol Biochem* 118:22-30.

Guo YM, Lee GJ, Archibald AL & Haley CS (2008) Quantitative trait loci for production traits in pigs: A combined analysis of two Meishan x Large White populations. *Anim Genet* 39:486-495.

Harmegnies N, Davin F, De Smet F, Buys N, Georges M & Coppieters W (2006) Results of a whole-genome quantitative trait locus scan for growth, carcass composition and meat quality in a porcine four-way cross. *Anim Genet* 37:543-53

Kim KS, Larsen N, Short T, Plastow G, Rothschild MF (2000) A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm. Genome* 11:131-135.

Kimmel AR, Brasaemle DL, McAndrews-Hill M, Sztalryd C & Londos C (2010) Adoption of PERILIPIN as a unifying nomenclature for the mammalian PAT-family of intracellular lipid storage droplet proteins. *J Lipid Res* 51:468-471.

Legarra A, Varona L & Maturana EL (2008) TM Threshold Model. Accessed Set.20, 2012. http://snp.toulouse.inra.fr/alegarra/manualtm.pdf

Liu G, Jennen DG, Tholen E, Juengst H, Kleinwachter T, Holker M (2007) A genome scan reveals QTL for growth, fatness, leanness and meat quality in a duroc-pietrain resource population. *Anim Genet* 38:241-252.

Onteru SK, Gorbach DM, Young JM, Garrick DJ, Dekkers JCM, Rothschild MF (2013) Whole Genome Association Studies of Residual Feed Intake and Related Traits in the Pig. *PLOS ONE* 8:1756-1773.

Qi L, Shen H, Larson I, Schaefer EJ, Greenberg AS & Tregouet DA (2004) Gender-specific association of a perilipin gene haplotype with obesity risk in a white population. *Obes Res* 12:1758-1765.

Mohrmann M, Roehe R, Knap PW, Looft H, Plastow GS & E Kalm (2006) Quantitative trait loci associated with AutoFOM grading characteristics, carcass cuts and chemical body composition during growth of Sus scrofa. *Anim Genet* 37:435-43.

Ros-Freixedes R, Reixach J, Bosch L, Tor M & Estany J (2013) Response to selection for decreased backfat thickness at restrained intramuscular fat content in Duroc pigs. *J Anim Sci* 91:3514-21.

Ros-Freixedes R, Reixach J, Bosch L, Tor M & Estany J (2012) Expected genetic response for oleic acid content in pork. *J Anim Sci* 90:4230-4238.

Rückert C & Bennewitz J (2010) Joint QTL analysis of three connected F2-crosses in pigs. *Genet Sel Evol* 42:40-48.

Ruiz JR, Larrarte E, Margareto J, Ares R, Alkorta P & Labayen I (2011) Preliminary findings on the role of PLIN1 polymorphisms on body composition and energy metabolism response to energy restriction in obese women. *Br J Nutr* 106:486-490.

Saha PK, Kojima H, Martinez-Botas J, Sunehag AL & Chan AL (2004) Metabolic adaptations in the absence of perilipin: Increased beta-oxidation and decreased hepatic glucose production associated with peripheral insulin resistance but normal glucose tolerance in perilipin-null mice. *J Biol Chem* 279:35150-35158.

Sambrook J, Fritschi EF & Maniatis T (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York.

Smith BJ (2005) Bayesian Output Analysis Program (BOA). Version 1.1.5. The University of Iowa. Acceded Set. 20, 2012. http://www.public-health.uiowa.edu/boa.

Smith CE & Ordovás JM (2012) Update on perilipin polymorphisms and obesity. *Nutr Rev* 70:611-621.

Tansey JT, Huml AM, Vogt R, Davis KE, Jones JM & Fraser KA (2003). Functional studies on native and mutated forms of perilipins. A role in protein kinase A-mediated lipolysis of triacylglycerols. *J Biol Chem* 278:8401-8406.

Tao X, Jihong Y, Li G, Bin F, Yi Z & Xiaodong C (2008) Cloning, chromosome mapping and expression pattern of porcine PLIN and M6PRBP1 genes. *Genet Sel Evol* 40:215-226.

Thomsen H, Lee HK, Rothschild MF, Malek M & Dekkers JC (2004) Characterization of quantitative trait loci for growth and meat quality in a cross between commercial breeds of swine. *J Anim Sci* 82:2213-2228.

Vykoukalová Z, Knoll A & Čepica S (2009) Porcine perilipin (PLIN) gene: structure, polymorphism and association study in large white pigs. *Cz J Anim Sci* 54:359-364.

Table 1. Single nucleotide polymorphisms (SNP) detected by sequencing the porcine *PLIN1* gene in Italian heavy pigs.

SNP[1]	Gene position [2]	Gene location	Amino acid change
JN860199 g.173G>A	691	Intron 1	-
JN860199 g.3484C>G	4,004	Intron 1	-
AM931171g.4119A>G	4,119	Exon 2	Synonymous[3]
AM931171g.7966T>C	7,966	Exon 5	Synonymous[3]

[1] GenBank accession number is indicated.

[2] Position from the start codon as referred to the entry [Ensembl:ENSSSCG00000001844; assembly Sscrofa10.2: chromosome 7; 601266014:60139897:-1].

[3] These SNPs are also reported by Vykoukalová et al., 2009

Table 2. Number of pigs (N), frequency of the allele G (f (G)), and number of pigs per *PLIN1* and *PLIN2* genotypes by batch.

		PLIN1 (*JN860199:g.173G>A*)				*PLIN2* (*GU461317:g.98G>A*)			
	N	f(G)	GG	AG	AA	f(G)	GG	AG	AA
Batch 1	108	0.51	36	38	34	0.49	23	60	25
Batch 2	102	0.51	31	42	29	0.37	16	44	42
Batch 3	66	0.35	13	20	33	0.50	15	36	15
Batch 4	69	0.33	6	34	29	0.43	16	27	26
Batch 5	84	0.26	6	32	46	0.60	31	39	14
Batch 6	95	0.31	8	42	45	0.61	37	42	16
Batch 7	83	0.32	8	37	38	0.48	19	42	22
Total	607	0.38	108	245	254	0.50	157	290	160

Table 3. Mean (standard deviation) and additive, dominant, and epistatic effects of *PLIN1 JN860199:g.173G>A* and *PLIN2 GU461317:g.98G>A* polymorphisms associated to live body weight and growth rate at different ages

		Additive (a) and dominant (d) effects[1]								Epistatic effects[1]							
		PLIN1,g.173G>A				*PLIN2, g.98G>A*											
Trait	Mean (SD)	a_1	P(>0)	d_1	P(>0)	a_2	P(>0)	d_2	P(>0)	a_1a_1	P(>0)	a_1d_2	P(>0)	d_1a_2	P(>0)	d_1d_2	P(>0)
Body weight, kg																	
120 d	61.28 (12.13)	-0.66	**0.06**	-1.04	**0.05**	0.95	**0.99**	0.77	0.89	-0.88	**0.06**	0.47	0.71	-0.51	0.29	1.35	0.86
180 d	107.32 (11.01)	-0.68	**0.10**	-1.56	**0.03**	1.19	**0.98**	1.17	**0.94**	-0.78	0.14	0.64	0.73	0.13	0.55	0.59	0.65
205 d	122.15 (11.33)	-0.42	0.27	-0.51	0.29	1.08	**0.96**	1.03	0.87	-1.01	0.12	0.19	0.56	0.46	0.63	0.18	0.55
Daily gain, g/d																	
0-120 d	500.77 (80.94)	-4.76	**0.09**	-6.93	**0.09**	7.26	**0.98**	5.51	0.86	-7.94	**0.04**	4.70	0.76	-4.59	0.27	12.04	0.88
120-180 d	766.88 (112.88)	-1.95	0.38	-6.83	0.29	4.15	0.74	4.37	0.69	2.26	0.60	1.10	0.54	15.38	0.87	-10.22	0.30
180-205 d	596.23 (193.43)	5.72	0.70	22.65	**0.94**	-0.42	0.48	-9.57	0.48	-8.23	0.28	-3.27	0.41	20.03	0.82	-22.91	0.24

[1] The numbers 1 and 2 refers to *PLIN1* and *PLIN2*, respectively, with the additive effects expressed as A-G; P (>0): Posterior probability of a value being positive. In bold, probabilities above 0.90 or below 0.10.

Table 4. Mean (standard deviation) and additive, dominant, and epistatic effects of *PLIN1 JN860199:g.173G>A* and *PLIN2 U461317:g.98G>A* polymorphisms associated to backfat and loin thickness at different ages.

		Additive (a) and dominant (d) effects[1]								Epistatic effects[1]							
		PLIN1,*g.173G>A*				*PLIN2*, *g.98G>A*											
Trait	Mean (SD)	a_1	P(>0)	d_1	P(>0)	a_2	P(>0)	d_2	P(>0)	a_1a_1	P(>0)	a_1d_2	P(>0)	d_1a_2	P(>0)	d_1d_2	P(>0)
Backfat thickness, mm																	
120 d	11.05 (2.72)	-0.07	0.29	-0.18	0.17	0.17	**0.91**	-0.07	0.33	-0.23	**0.07**	0.03	0.55	-0.14	0.29	0.59	**0.95**
180 d	17.76 (3.74)	-0.06	0.27	-0.15	0.14	0.19	**0.98**	-0.10	0.31	-0.76	0.16	0.54	0.69	0.15	0.56	0.79	0.68
205 d	20.66 (4.15)	0.01	0.52	-0.24	0.16	-0.22	**0.10**	-0.03	0.46	-0.41	**0.03**	0.06	0.58	0.12	0.63	0.05	0.54
Loin thickness, mm																	
120 d	40.38 (3.25)	0.33	**0.92**	-0.40	0.15	-0.42	**0.04**	-0.59	**0.04**	0.07	0.59	-0.23	0.31	-0.91	**0.04**	0.31	0.66
180 d	45.04 (3.97)	0.26	0.85	-0.56	0.20	-0.05	0.41	-0.63	**0.03**	0.23	0.75	1.51	**0.93**	0.49	0.82	-0.42	0.28
205 d	48.57 (4.49)	0.00	0.51	0.11	0.61	0.02	0.52	-0.08	0.42	-0.46	0.09	-0.33	0.25	-0.47	0.19	0.31	0.65

[1] The numbers 1 and 2 refers to *PLIN1* and *PLIN2*, respectively, with the additive effects expressed as A-G; P (>0): Posterior probability of a value being positive. In bold, probabilities above 0.90 or below 0.10.

Table 5. Mean (standard deviation) and additive, dominant, and epistatic effects of *PLIN1 JN860199:g.173G>A* and *PLIN2 U461317:g.98G>A* polymorphisms associated to carcass traits.

		Additive (a) and dominant (d) effects[1]								Epistatic effects[1]							
		PLIN1,g.173G>A				*PLIN2*, *g.98G>A*											
Trait	Mean (SD)	a_1	P(>0)	d_1	P(>0)	a_2	P(>0)	d_2	P(>0)	a_1a_1	P(>0)	a_1d_2	P(>0)	d_1a_2	P(>0)	d_1d_2	P(>0)
Carcass weight, kg	93.69 (9.28)	-0.20	0.36	0.41	0.70	0.58	0.86	-0.95	0.11	1.09	**0.94**	0.19	0.57	-0.07	0.47	-0.50	0.38
Carcass backfat, mm	22.59 (3.68)	-0.09	0.33	0.02	0.52	-0.15	0.24	0.10	0.65	0.32	0.88	0.41	0.85	0.19	0.69	-0.21	0.36
Carcass loin, mm	45.25 (7.23)	0.23	0.69	-0.19	0.39	0.28	0.73	-0.52	0.22	0.58	0.83	0.69	0.78	-0.74	0.22	-0.70	0.31
Carcass lean, %	43.77 (4.96)	0.08	0.62	-0.01	0.50	0.23	0.80	-0.47	0.11	-0.17	0.32	-0.20	0.36	-0.14	0.41	0.20	0.59
Carcass length, cm	86.58 (2.96)	-0.62	**0.04**	0.81	**>0.99**	0.42	**0.99**	-0.82	**<0.01**	0.92	**0.98**	-0.22	0.24	-0.45	0.11	-0.14	0.39
Lean weight, kg	40.73 (5.29)	0.07	0.85	0.19	**0.98**	0.41	**>0.99**	-0.72	**<0.01**	0.30	**>0.99**	-0.11	0.20	-0.37	**<0.01**	-0.06	0.38
Ham weight, kg	12.09 (1.16)	0.00	0.51	-0.04	0.34	0.10	**0.94**	-0.05	0.28	0.09	0.86	0.20	**0.95**	-0.04	0.39	-0.10	0.28

[1] The numbers 1 and 2 refers to *PLIN1* and *PLIN2*, respectively, with the additive effects expressed as A-G; P (>0): Posterior probability of a value being positive. In bold, probabilities above 0.90 or below 0.10.

Table 6. Mean (standard deviation) and additive, dominant, and epistatic effects for *PLIN1 JN860199:g.173G>A* and *PLIN2 U461317:g.98G>A* polymorphisms associated to meat quality traits

		Additive (a) and dominant (d) effects[2]								Epistatic effects[2]							
		PLIN1,*g.173G>A*				*PLIN2*, *g.98G>A*											
Trait[1]	Mean (SD)	a_1	P(>0)	d_1	P(>0)	a_2	P(>0)	d_2	P(>0)	a_1a_1	P(>0)	a_1d_2	P(>0)	d_1a_2	P(>0)	d_1d_2	P(>0)
pH24 LM	5.71 (0.25)	0.00	0.58	0.01	0.61	-0.01	0.23	0.02	0.86	-0.01	0.24	0.03	**0.90**	0.00	0.47	-0.03	0.20
pH24 SM	5.72 (0.25)	0.01	0.79	0.00	0.52	0.00	0.43	0.03	**0.92**	-0.02	0.12	0.00	0.57	0.01	0.61	-0.03	0.22
IMF, %	4.50 (1.66)	0.10	0.85	-0.07	0.32	0.04	0.67	0.06	0.67	-0.16	0.11	0.05	0.59	0.11	0.70	0.18	0.73
SFA, %	34.99 (3.68)	0.01	0.53	0.01	0.53	-0.24	**0.04**	0.07	0.66	-0.15	0.19	-0.22	0.19	-0.08	0.40	-0.08	0.41
MUFA, %	50.54 (3.11)	0.20	**0.94**	-0.05	0.40	0.30	0.99	-0.17	0.17	0.04	0.59	-0.15	0.29	-0.06	0.42	0.74	**0.98**
PUFA, %	14.47 (2.75)	-0.20	**0.06**	0.04	0.59	-0.06	0.32	0.10	0.73	0.12	0.77	0.40	**0.95**	0.15	0.71	-0.60	**0.05**
pH24 LM	5.71 (0.25)	0.00	0.58	0.01	0.61	-0.01	0.23	0.02	0.86	-0.01	0.24	0.03	**0.90**	0.00	0.47	-0.03	0.20

[1] IMF: intramuscular fat; SFA: saturated fatty acids (C14:0+C16:0+C18:0); MUFA: monounsaturated fatty acids (16:1+C18:1+C20:1); PUFA: polyunsaturated fatty acids (C18:2+C18:3+C20:2+C20:4) in muscle *gluteus medius*

[2] The numbers 1 and 2 refers to *PLIN1* and *PLIN2*, respectively, with the additive effects expressed as A-G; P (>0): Posterior probability of a value being positive. In bold, probabilities above 0.90 or below 0.10.

Supplementary information

Supplementary Table S1. Primers used for single nucleotide polymorphism discovery in *PLIN1* gene.

Primer	Sequence (5'-3')	Gene regions	Product size (bp)	Ta [1]
P1	F GTCAAATAACCATAGCAACCAAC R ATTCCCAGAAGACCCTAACC	partial promoter; exon 1; partial intron 1	253	61
P2	F AGGGAACTGATGGTGAGAGG R TCCGCAAGAAGGAGTGAGG	partial intron 1; exon 2, partial intron 2	306	60
P3	F AGAGCCAAGGTTGTGACCAG R CAGGCAGTGAACGAGCAAG	partial intron 2; exon 3, partial intron 3	415	61
P4	F ATCTGCACGCCTGACTCC R TGGTGGCCTCTTGGTAATTC	partial intron 4; exon 5; partial intron 5	375	60
P5	F CGGGATGACCACTTTCTAACC R GCTCAGGGCAGACACTCAC	partial intron 5; exon 6	289	60
P6	F AGGTGCTGTGAAGTCAGTGG R TGTTCCAGGGTGAGGTGAAG	partial intron 6; exon 7; partial intron 7	368	61
P7	F GGATAGTGAGGAGGGGAAGG R CAGGAGACTGGGGAAGGAG	partial intron 7; exon 8; 3'downstream genomic region	431	63

[1] Annealing temperature

Supplementary Table S2. Quantitative trait loci (QTL) co-localizing with the porcine *PLIN2* mapping position[1].

QTL trait	QTL (cM)	Reference
PLIN2 (SSC1q2.3-2.7; 227.3 Mb on SSC assembly 10.2)		
Abdominal fat	107.6	Geldermann et al. (2010)
Adipocyte diameter	94.3-122.6	Geldermann et al. (2003)
Average daily gain	3.0-140.5	Liu et al. (2007)
Average daily gain	42.36-134.76	Onteru et al. (2013)
Average daily gain	49.4-79.4	Ruckert and Bennewitz (2010)
Average daily gain	73-140.5	Harmegnies et al. 2006
Average daily gain	100.8-118.5	Mohrmann et al. (2006)
Average daily gain	127.1-140.5	Evans et al. (2004)
Backfat thickness	80.0-110.5	Liu et al. (2007)
Body weight at birth	16.4-132	Guo et al (2008)
Daily feed intake	78.7-79.4	Kim et al. (2000)
Ham weight	94.3-122.6	Geldermann et al. (2003)
Lean meat percentage	94.3-122.6	Geldermann et al. (2003)
pH48 hours post mortem (loin)	102.9-119.5	Thomsen et al. (2004)

[1] Source: animal genome gbrowse (http://www.animalgenome.org/cgi-bin/gbrowse/pig/), accessed on 22-11-2014

Chapter 5: Investigation of Perilipin 5 gene and its role for porcine meat quality.

M. Zappaterra [1], M. Mazzoni [2], P. Zambonelli [1], R. Davoli [1]

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

[2] *Department of Veterinary Medical Sciences (DIMEVET), Bologna University, Via Tolara di Sopra, 50, 40064 Ozzano dell'Emilia (BO), Italy.*

Submitted to

Animal

Summary

The improvement of pig meat quality by means of selection represents one of the main goals of pork producers, processors and researchers. Although many efforts have been done to obtain better meat characteristics following consumers preference, the growing request of animal products with healthy fatty acid composition is raising the necessity of a deeper knowledge of the main players controlling fatty acids storage in muscle and backfat. Perilipin (PLIN) 5, and more in general the whole Perilipin family, seems to play a crucial role in the regulation of lipids deposition in muscle and carcass fat depots. The Perilipin genes can be considered important candidates for the genetic selection in livestock addressed towards the qualitative improvement of the products. This gene family codes for proteins coating intracellular lipid droplets surface and despite the number of researches focused on PLINs in human and mouse is increasing, the knowledge of these genes in pig is still poor and quite incomplete. This is the first study on PLIN5 in pigs: the present research is aimed at investigating in Italian Large White and Italian Duroc pigs the PLIN5 associations with meat and carcass quality traits. PLIN5 protein was localised through immunofluorescence, resulting to be widely expressed inside *Semimembranosus* muscle (SM) myofibers. Then, the SNP rs327694326 (NC_010444.3:g.74806942 C>T) was identified in both considered pig breeds by sequencing the 3'UTR and downstream region of *PLIN5* gene. Using High Resolution Melt analysis 512 Italian Large White and 300 Italian Duroc pigs were genotyped for this SNP. The two breeds showed different allelic frequencies, with a very low frequency for C allele in Duroc pigs. In Large White pigs this SNP resulted to be associated with backfat contents of docosatetraenoic fatty acid and with the ratio between the amounts of C18 tri-unsaturated fatty acids and C18 di-unsaturated fatty acids. Furthermore, since PLINs are known to regulate lipases activity, we tested if the studied rs327694326 SNP on *PLIN5* downstream region was associated to differences in *Hormone sensitive lipase* (*LIPE*) gene expression levels. In SM of Italian Large White pigs, C allele was associated with significantly lower *LIPE* mRNA levels than T allele, suggesting that variations in *PLIN5* sequence may be linked to a still poorly known molecular process regulating intracellular LIPE activity.

Implications

This is the first study in pigs investigating *PLIN5* gene, and the results of the association study suggest that variations in its sequence may be related to a still unclear regulative function exerted by *PLIN5* on *LIPE* gene activity. This study provides results that, if confirmed, could improve at least in part the knowledge of one of the regulative pathways linking *PLIN5* and lipases activity, indicating that SNPs in *PLIN5* may be useful markers for selection breeding programs addressing the improvement of meat quality.

Introduction

For many years pig selection breeding programs focused mainly on the increase of carcass muscle deposition, determining a consistent reduction in intramuscular fat and backfat deposition and a worsening in pork meat quality (Wood, 1990; Wood et al., 2008). During the last decade, the improvement of meat quality has been an objective of pork industry, and several researches have demonstrated that a selection for increased intramuscular fat and meat quality is possible (Schwab et al., 2009; Jeong et al., 2015). On the other hand, meat quality traits have been found to be correlated with intramuscular and backfat fatty acid composition (Wood et al., 2008), suggesting that selection for meat and carcass traits may also determine changes in pork fatty acid composition. This issue, along with the increasing interest towards healthier pig products, has raised the need to have a better knowledge of the key molecular regulators at the basis of body lipid storage and utilisation. The Perilipin family (Perilipin 1 to 5) is one of the most promising gene families likely involved in controlling lipids deposition in muscle and carcass. *Perilipin 5* (*PLIN5*) gene, and more in general the whole Perilipin family, has been investigated mainly in human and mouse, and recent studies have found in mice and rats that Plin5 protein probably regulates intracellular fatty acid fluxes and oxidation (Wang et al., 2011; Laurens et al., 2016). Different Authors (Brasaemle, 2007; Laurens et al., 2016) reported that this regulation may be mediated by coordinating the access of lipases to lipid droplets surface. In particular, *PLIN5* was indicated as an essential player interacting with, among others, Hormone sensitive lipase (LIPE). *LIPE* gene maps in a QTL region linked to sensory quality in porcine meat (Pena et al., 2013), and its protein expression was found associated with intramuscular fat content in porcine *Semimembranosus* muscle (Zappaterra et al., 2016). The literature

suggests that LIPE may play an important role in pig meat and carcass traits, and the investigation of the physiological and molecular patterns regulating its activity can provide new tools for an improved genetic selection in livestock species. Despite these promising results, it is still unclear how PLIN5 interacts with lipases and *PLIN5* gene has not yet been analysed in pigs. This research is aimed at investigating in Italian Large White and Italian Duroc pigs: i) PLIN5 protein localisation in *Semimembranosus* muscle (SM), ii) associations between *PLIN5* SNP and carcass traits, fat deposition and backfat fatty acid composition, iii) *PLIN5* gene expression levels in SM samples of the two pig breeds. Furthermore, considering the results obtained for LIPE in a previous study (Zappaterra et al., 2016), we decided to test if the identified *PLIN5* gene variability could be associated with variations in *PLIN5* and *LIPE* transcription levels.

Material and methods

Utilised population

The samples used in the present study were chosen among 949 Italian Large White (ILW) pigs and 484 Italian Duroc (IDU) pigs, two breeds considered in this research for their importance in selection schemes addressed to obtain high quality seasoned products. The sampled animals are pure breed gilts and castrated males included in the Italian sib test genetic evaluation scheme performed by ANAS (Associazione Nazionale Allevatori Suini, ANAS; www.anas.it), reared in the same environmental conditions at the genetic test station, fed the same diet *quasi ad libitum* (60% of the pigs was able to ingest the entire supplied ration). At about 150 kg of live weight the animals were transported to a commercial abattoir located at about 25 km from the test station in accordance with Council Rule (EC) No. 1/2005 regarding the protection of animals during transport and related operations. At the slaughterhouse, the pigs were electrically stunned and bled in a supine position in agreement with Council Regulation (EC) No. 1099/2009 regarding the protection of animals at the time of slaughter. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health. After slaughtering, for each individual two samples of backfat and skeletal muscle tissue (obtained from SM) were collected: one aliquot of each tissue was stored at -20°C and the leftover aliquots were frozen in liquid nitrogen and stored at -80°C. For the immunohistochemical analysis, 10

samples of SM, 5 taken from ILW and 5 from IDU were included in Tissue-Tek® O.C.T. Compound (Sakura Finetek Europe, Zoeterwoude, The Netherlands).

Measured phenotypes

Carcass traits. ANAS and the slaughterhouse technicians provided us with the measure of the main carcass phenotypic traits: the carcass weight expressed in kg (carcass weight), the percentage of lean mass over the total carcass weight measured with Fat-O-Meter (lean %), the backfat thickness measured with Fat-O-Meter at 8 cm off the midline of the carcass between the third and fourth last ribs. Furthermore, the pH measured one hour after slaughter (pH_1) and at 24 hours post-mortem (pH_{24}), the objective colour measurements of International Commission on Illumination (CIE, 1976) L*, a* and b* were measured on *Semimembranosus* muscle.

Determination of backfat fatty acid composition in ILW pigs. ILW backfat tissue samples stored at -20°C were used to obtain backfat fatty acid (FA) composition. FA composition has been detected by direct trans-esterification, following the protocol reported by Murrieta et al. (2003). For each sample, 50 mg of frozen backfat was used for the total lipid extraction and then, in each tube, 0.5 mg of C19:0 methyl ester in hexane was added as internal standard. Gas chromatography was performed on GC- 2010 Plus High-end Gas Chromatograph (Shimadzu Corporation, Tokyo, Japan), using SPTM- 2560 Capillary GC Column (Sigma- Aldrich, Merck, Darmstadt, Germany). Backfat FA composition was expressed as the ratio between each individual FA or FA family and the total backfat FAs.

Immunohistochemical analysis for the detection of PLIN5 antigens

Immunofluorescence reaction. Semimembranosus muscle samples of 5 ILW and 5 IDU individuals were randomly chosen among the 949 ILW and 484 IDU pigs and were analysed (Table 1). Immediately after collection, muscle samples were cut into 1 × 1 cm pieces, parallel to the muscle fiber direction, rapidly frozen in liquid nitrogen-cooled isopentane and embedded in Tissue-Tek®. Serial cross-sections (10 µm thick) were cut on a cryostat microtome at −20°C and mounted on poly-l-lysine coated glass slides (Sigma-Aldrich, St. Louis, MO, USA). Immunofluorescence was performed using the procedure described previously by Gandolfi et al. (2011). The sections were fixed for 10 min in 4%

paraformaldehyde in phosphate buffer saline (PBS, 0.1 M, pH 7.2), rinsed in PBS and incubated in 5% normal donkey serum (NDS) for 30 min at room temperature (RT) to reduce the non-specific binding of the secondary antibodies. The sections were then incubated at 4°C in a humid chamber for 24 h in a mixture of 3% NDS and the primary antibodies *rabbit anti-PLIN5* (Cat. NB110-60509, Novus Biologicals, Littleton, Colorado, USA) diluted 1:2000. The next day, after washing in PBS, the sections were incubated for 1 hour at room temperature in a PBS mixture containing 3% of NDS and the secondary antibody donkey anti-rabbit IgG Alexa Fluor 488-conjugated antibody (Invitrogen, Thermo Fisher Scientific Inc., Waltham, MA, USA) diluted 1:1300. After washing in PBS, the sections were mounted with buffered glycerol, pH 8.6. The specificity of the secondary antibodies was tested by performing the staining in the absence of primary antibodies as control. The specificity of primary antibodies was tested by Western blot.

Morphometrical analysis. The sections were examined on a Zeiss Axioplan microscope equipped with the appropriate filter cubes to discriminate different fluorochromes. The images were recorded with a Polaroid DMC digital camera (Polaroid, Cambridge, Mass., U.S.A.) and the DMCV 2 software images were further processed using Corel Photo Paint and Corel Draw software programs (Corel, Milan, Italy). Morphometrical analyses were carried out considering for each sample the cross-sectional area (CSA, measured in μm^2) of 100 myofibers with a 20x objective lens using KS 300 image analysis software (Kontron Elektronic, Munich, Germany). For each sample the total CSA defined by PLIN5 negative myofibers was measured by outlining the profiles on the monitor screen using a computer mouse. The relative percentage of PLIN5 immunoreactive (PLIN5-IR) myofibers was calculated considering the total area of the 100 myofibers measured for each pig from which the total area of the PLIN5 negative myofibers was deducted.

Statistical analysis. Data are reported as mean and SD. Differences were tested by Wilcoxon test, using Stats package on R environment version 3.3.2 (R Core Team, 2016). Nominal *P* value ≤ 0.05 was considered as the threshold for significance.

Validation of PLIN5 antibody specificity through Western blot. Western blot was performed to assure the specificity of the primary antibodies used for PLIN5 immunostaining. About 300

mg of SM sample stored at −80 °C was homogenized on ice in 900 µl of T-PER Tissue Protein Extraction (Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA) following the protocol reported in Zappaterra et al. (2016). About 20–60 µg of the total protein was separated in a 12% SDS-polyacrylamide gel and blotted into a nitrocellulose membrane (GE Healthcare/Amersham, Uppsala, Sweden). After blocking in 5% non-fat milk-Tris-buffered saline (TBS), membranes were incubated overnight with 1:4000 dilution of *Rabbit anti-PLIN5*. After washing, membranes were incubated for 1 h with 1:1000 dilution of horseradish peroxidase-conjugated anti-rabbit IgG (Santa Cruz Biotechnology, Dallas, Texas, USA). The membrane blots were developed with diaminobenzidin (DAB, Santa Cruz Biotechnology, Dallas, Texas, USA). The observed molecular weight was then compared using UniProt database (The UniProt Consortium, 2017) with PLIN5 weights in other species.

PLIN5 SNP detection and association study

SNP detection. Porcine *PLIN5* gene sequence obtained from Ensembl database (last accession on 13th of March 2017; http://www.ensembl.org/index.html) was compared with sequences in GenBank database (https://www.ncbi.nlm.nih.gov/genbank/). The *PLIN5* 3'UTR and downstream gene regions of fifteen ILW samples randomly chosen were sequenced using ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). A primer pair was designed for the sequencing using Primer3web software version 4.0.0 (http://primer3.ut.ee/; Supplementary Table S1). PCR were performed in a total volume of 20 µl containing 4 µl of 5x standard buffer, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 160 µM dNTP, 1 U of GoTaq® G2 DNA Polymerase (Promega Corporation, Madison, USA) and 20–50 ng of template DNA. Cycling conditions were: initial denaturation at 95°C for 5 min, 35 cycles consisting of one step at 95°C for 30 s, the annealing temperature of 64°C for 30 s, and the third step at 72°C for 30 s, followed by a final extension step of 72°C for 5 min. The PCR products were sequenced on both strands using the BigDye v. 3.1 Cycle Sequencing kit (Applied Biosystems, Thermo Fisher Scientific Inc., Waltham, MA, USA) and before sequencing the obtained products were purified using ethanol/sodium acetate precipitation. The fifteen ILW samples were then sequenced on the four-capillary system ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems). The nucleotide sequences obtained from sequencing and obtained and those retrieved from databases were compared by multiple alignments performed with MEGA software version 6.06 (Tamura et al., 2013).

SNP genotyping. 512 ILW samples out of the total 949 individuals and 300 IDU pigs out of the 484 were genotyped for the SNP identified through sequencing. Samples were genotyped using High Resolution Melt (HRM) analysis on Rotor Gene 6000 (Corbett Life Science, Concorde, New South Wales). Primers were designed using Primer3web software version 4.0.0 (Supplementary Table 1), and the amplification procedure was performed in 35 cycles constituted of a denaturation step at 95°C for 20 s, the annealing step at 63°C for 15 s and the extension at 72°C for 10 s. A HRM analysis was performed after amplification by increasing the temperature of 0.1°C each 1 s, starting from 65°C until 90°C. The fluorescence was read at each temperature increase.

Statistical analysis. Allelic frequencies of the identified SNP were calculated in ILW and IDU samples. Deviations from Hardy-Weinberg equilibrium were evaluated by a $\chi 2$ test. The association study between the SNP and the measured traits was performed using TM, a Bayesian statistics based software (Legarra et al., 2011). Different models were used depending on the considered traits: for backfat fatty acid composition traits, the associations were analysed using backfat thickness, animal's age (days of life) and carcass weight as covariates, slaughter batch as random effect (27 slaughter batches), and the effects of sex (castrated males or female) and animal (using a pedigree with 1,724 individuals) were taken into account. For the association study between the SNP and the phenotypic measures, the same model was performed, considering among the covariates the interaction between carcass weight and animal's age instead of backfat thickness. For each trait, the estimated means and differences between genotypes were assessed on Rabbit program (Blasco, 2012) using the same model performed in TM.

Gene expression study

Sampling. For the comparison of *PLIN5* gene expression between breeds, 30 ILW and 30 IDU were chosen avoiding as much as possible full and half sibs and balancing the two groups for sex (Table 1). Then, with the aim of deepening the understanding of the results obtained from the association study, both *PLIN5* and *LIPE* gene expressions were tested on samples displaying different genotypes for the studied SNP. Taking into account the higher frequency

of T allele in the studied populations, 30 TT, 20 CT and 19 CC samples were considered among ILW pigs; for IDU samples 13 TT, 18 CT and 3 CC were studied.

Total RNA extraction and preparation. Semimembranosus muscle samples were withdrawn at the slaughterhouse and immediately frozen at −80 °C. Total RNA was then extracted using TRIZOL reagent (Invitrogen Corporation, Carlsbad, California), the quality and integrity of the RNAs were both checked and RNA was retrotranscribed to cDNA as described in Davoli et al. (2011). Primers were designed using Primer3web software version 4.0.0 (Supplementary Table 1). The quantitative Real-Time PCR (qRT-PCR) standard curve method was used to analyse the genes expression. qRT-PCR was performed on Rotor Gene™ 6000 (Corbett Life Science, Concorde, New South Wales) as described in Zappaterra et al. (2016), and *beta-2-microglobulin* (*B2M*), *tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta* (*YWHAZ*) and *hypoxanthine phosphoribosyltransferase 1* (*HPRT1*) have been chosen as normalising genes. The protocols used for the amplification reactions and standard curves creation are the same described in Zappaterra et al. (2016).

Statistical analysis. Differences in gene expression levels were tested by two-tailed Student's t test. Gene expressions were compared: i) between breeds, to test *PLIN5* gene expression variations; ii) within breed between different genotypes in order to find possible associations of the analysed SNP with *PLIN5* and *LIPE* expression levels. Additionally, with the purpose of identifying the strength of the relationship between *PLIN5* and *LIPE* transcriptions, a Pearson's correlation analysis was performed between the mRNA levels of these two genes using cor.test of stats package in R environment version 3.3.2 (R Core Team, 2016). For this analysis were considered the samples analysed for the comparisons of *PLIN5* and *LIPE* gene expressions between the SNP different genotypes (69 ILW and 34 IDU individuals). Nominal *P* value ≤ 0.05 was considered as the threshold for significance.

Results

PLIN5 protein localisation in Semimembranosus muscle

This is the first study that investigates PLIN5 protein localisation in porcine SM. The specificity of the antibody was confirmed by Western-blot, as a unique band was identified

(Figure 1A). PLIN5 molecular weight was found to be approximately 50 kDa, a weight that is similar to that of PLIN5 in mouse (Q8BVZ1 in UniProt database) and in human (Q00G26 in UniProt database). All porcine SM samples showed to be positive to PLIN5 antigens. PLIN5 protein was widely expressed in SM: almost all the myofibers were PLIN5-IR, with isolated (Figure 1B) or grouped in pairs (Figure 1C) PLIN5 negative myofibers.PLIN5-IR myofibers showed a broad fluorescence pattern comprising the whole sarcoplasma, and a strong intensity staining was found in particular at the sarcolemma and perimysial collagen levels (Figure 1B and C).

This high prevalence of PLIN5 positive myofibers is also visible in Table 2, where the results of the morphometrical analysis for both ILW and IDU samples are reported. The CSA of PLIN5 staining myofibers, the prevalence of PLIN5-IR myofibers expressed as percentage and the average sizes of PLIN5 positive and negative myofibers were calculated in both ILW and IDU samples (Table 2). No significant difference for the morphometrical measures between the two breeds was detected: all the samples showed PLIN5 staining CSA percentages above 87%, with peaks of the 100% of immunoreactivity in some sections.

SNP genotyping and allele frequencies

The results of the sequencing performed on ILW DNA samples for *PLIN5* 3'UTR and downstream region showed a unique SNP lying in *PLIN5* gene downstream region. Comparing the identified variation with those reported in Ensembl database we noticed that the same SNP was already reported among the nucleotide variants associated with *PLIN5* and *PLIN4* genes (variant ID: rs327694326, HGVS name NC_010444.3:g.74806942T>C). The SNP rs327694326 deviated from Hardy-Weinberg equilibrium in ILW samples ($P < 0.0001$) with the C allele showing a lower frequency than the T allele in ILW samples (0.14 for the C allele against 0.86 observed for the T allele; Table 3). In IDU samples the C allele was quite infrequent, displaying an allele frequency of 0.04 and an expected number of CC animals that was too low to test Hardy-Weinberg equilibrium in this population (expected number of CC animals < 5). Due to the very low frequency of the C allele in IDU samples, the subsequent association study was conducted exclusively in ILW samples.

Association study with backfat fatty acid composition and carcass traits

In ILW individuals, several associations between the rs327694326 SNP and backfat fatty acid composition were identified (Table 4). A dominant effect was found for trans-9, cis-12 octadecenoic acid and for the ratio between the amounts of C18 tri-unsaturated fatty acids and C18 di-unsaturated fatty acids, while the SNP showed an additive effect on docosatetraenoic acid. On the whole, CT genotype showed a dominant effect on both CC and TT for trans-9, cis-12 octadecenoic and for the ratio between C18 tri-unsaturated and C18 di-unsaturated fatty acids. For docosatetraenoic fatty acid TT homozygotes showed the highest estimated means (Table 4). Oleic, cis-vaccenic, linolenic, eicosadienoic, docosahexaenoic, monounsaturated fatty acids, long chain and omega 3 fatty acids were reported in Table 4 as the observed P(>0) were near to the considered thresholds of 0.90 and 0.10. Consistent associations were also found with carcass traits (Table 5): in this case too, the CT genotype presented a dominant effect on the majority of the considered traits: lean %, backfat thickness, drip loss, pH_{24}, *CIE* L* and *CIE* b*. However, an additive effect was observed for CIE a* meat colour parameter.

Gene expression study

Initially, a comparison between *PLIN5* gene expression in ILW and IDU samples was performed as the primary aim of the present study was to investigate this gene in the two pig breeds. The observed results highlighted a higher transcription level of this Perilipin in ILW samples than in IDU ($P = 0.005$; Figure 2A). Furthermore, based on the results of the association study, we also tested if rs327694326 was associated with changes in *PLIN5* and *LIPE* gene expressions. These two genes were quantified and compared across samples displaying different genotypes for the studied SNP. ILW individuals showing TT genotype presented higher *LIPE* mRNA levels than CT ($P = 0.02$) and CC pigs ($P < 0.0001$) (Figure 2B). The homozygotes for the T allele showed also increased levels of *PLIN5* expression, with a trending towards significance for the comparison between TT and CC ILW samples ($P = 0.06$). Conversely, in IDU samples the highest *PLIN5* and *LIPE* transcriptional levels were found in CT pigs ($P = 0.01$ for *PLIN5* gene expressions in CT vs. TT, and $P = 0.08$ for *LIPE* levels in CT vs. TT; Figure 2C). Anyway, due to the low frequency of the C allele in IDU, the expression levels detected in the 3 CC subjects were not compared with the mRNA levels detected in CT and TT IDU samples. Additionally, a correlation analysis between *PLIN5* and *LIPE* expression levels was performed and the results are reported in Table 6. The two breeds presented

significant correlations between *PLIN5* and *LIPE* expressions, and IDU samples showed a higher correlation compared to ILW (r = 0.597 in IDU and r = 0.371 in ILW pigs).

Discussion

The wide expression of PLIN5 protein detected in SM myofibers of both ILW and IDU pigs is consistent with the literature, which reports PLIN5 to be highly expressed in oxidative tissues, such as type I skeletal muscle fibers, cardiac muscle and liver (Dalen et al., 2007; MacPherson et al., 2012). Ruusunen and Puolanne (2004) and Lefaucheur et al. (2010) observed that, despite being classified as a white skeletal muscle, porcine SM is mainly composed of type IIA myofibers, which confer to this muscle a higher oxidative capacity and a red appearance when compared to glycolytic (type IIB) muscle fibers. Hence, the prevalence of PLIN5-IR myofibers detected in the present research is in accordance with the SM oxidative capacity reported by Ruusunen and Puolanne (2004) and Lefaucheur et al. (2010). Furthermore, the broad fluorescence pattern comprising the whole sarcoplasma reported here confirms the evidences already noticed in other animal species, where PLIN5 protein was found to be localised both on intracellular lipid droplets surface (Yamaguchi et al., 2006), on mitochondria (Bosma et al., 2012) and also as a free form in cytosol (Wolins et al., 2006). The two considered pig breeds showed no difference in the number and area of PLIN5-IR myofibers, but displayed distinct PLIN5 gene expression levels, with a higher transcription level in ILW samples. The difference detected for *PLIN5* expression levels between ILW and IDU may be related to the different adipogenic potential of the two breeds and to the lower capacity of Large White breed to store intramuscular fat respect to Duroc (Wood et al., 2004; Jung et al., 2015). The tight relation between PLIN5 and lipases was evidenced by the correlation linking *PLIN5* and *LIPE* transcription levels, identified in both breeds with a higher correlation coefficient in IDU pigs. Current literature lacks dedicated studies about *PLIN5* and *LIPE* transcription patterns in livestock species. However, the *PLIN5* and *LIPE* co-expression found in the present research may be consistent with the observations of Wang et al. (2009), who noticed through fluorescence microscopy that cultured cells displaying lipid droplets with higher rates of PLIN5 coated proteins showed also the largest amount of LIPE localised on their surfaces. Though the co-expression between *PLIN5* and *LIPE* identified in the present work seems to contradict the generally-accepted view that PLIN5 limits lipolysis (MacPherson and Peters, 2015; Wang et al., 2015),

it must not be forgotten that the information about PLIN5 mode of actions are still contradictory: while some studies identify this protein as a player limiting lipid droplets hydrolysis (Laurens et al., 2016), others report that its overexpression bears to increased fat oxidation (Bosma et al., 2012). The controversy around PLIN5 role in skeletal muscle can partly be solved by the fact that PLIN5 protein amount and gene expression are not the only ways by which lipases activity is controlled, but more steps and more processes might be needed to activate lipolysis. Bosma et al. (2012) and Pollak et al. (2015) reported that the phosphorylation of PLIN5 protein particular sites, such as the PAT-1 domain, and the phosphorylation of LIPE are among the hypothesised regulative mechanisms for the control of lipase activity. The results of the present study suggest that there could also be a molecular control affecting the interaction of PLIN5 with lipases. Indeed, in both breeds the different genotypes for the SNP rs327694326 showed changes in *PLIN5* and *LIPE* trancription levels (in IDU due to the low frequency of the CC genotype it was possible to compare only the expression levels for the CT and TT samples). Furthermore, in ILW samples the same SNP was associated with carcass and backfat fatty acid composition. Except for CIE *a* and* backfat content of docosatetraenoic acid for which we detected an additive effect, in most of the cases the ILW individuals displaying the CT genotype presented a dominant effect on the associated traits. A dominant effect was also observed by Gol et al. (2016) for two SNPs identified on the sequence of two other members of the Perilipins family: *PLIN1* and *PLIN2*. The two mutations (for *PLIN1* JN860199:g.173G>A and for *PLIN2* GU461317:g.98G>A) were studied in pigs belonging to a Spanish Duroc line and showed dominant effects on body weight at different ages, on average daily gain and on lean weight (Gol et al., 2016). In the present study, the dominant effect of CT genotypes is particularly enhanced for drip loss and pH at 24 hours post-mortem. The same dominant effect observed for the heterozygote in the present study was also identified for *PLIN1* JN860199:g.173G>A and *PLIN2* GU461317:g.98G>A SNPs in an Italian Duroc pig population: for the considered carcass traits the GA individuals showed estimated means diverging substantially from both homozygotes (unpublished data).

In the present investigation, the SNP rs327694326 was associated with both backfat thickness and fatty acid composition. This result is consistent with literature, where changes in backfat thickness are known to be directly correlated with backfat fatty acid composition, in particular with oleic acid content (Wood et al., 2008).

The complete understanding of the effects associated to rs327694326 remains to be elucidated and further studies are needed to deepen *PLIN5* knowledge in pig. Anyway, the changes in *PLIN5* and *LIPE* expression levels detected between samples with different genotypes may suggest that the studied SNP may be a marker of other causal mutations in *PLIN5* sequence affecting *PLIN5* and *LIPE* gene expression. On the whole *PLIN5* seems to affect pig meat quality, and the results indicate that SNPs in *PLIN5* sequence may be useful markers for selection breeding programs addressed towards improving pork quality.

Acknowledgements

We thank Dr. Maurizio Gallo from Associazione Nazionale Allevatori Suini (ANAS) and Dr. Luca Buttazzoni from Consiglio per la Ricerca in agricoltura e l'analisi dell'economia agraria, Centro di ricerca per la produzione della carni e il miglioramento genetico (CREA-PCM) for providing the samples and the information about the studied individuals. This work was supported by AGER – Hepiget project (Grant N. 2011- 0279) and by PRIN 2015 funds. Moreover, we want to thank Professor Paolo Trevisi for his help during Western blot analysis.

References

Blasco A (2012) The program "Rabbit". Retrieved on 13th January 2017 from http://www.dcam.upv.es/dcia/ablasco/Programas/THE%20PROGRAM%20Rabbit.pdf

Bosma M, Minnaard R, Sparks LM, Schaart G, Losen M, de Baets MH, Duimel H, Kersten S, Bickel PE, Schrauwen P & Hesselink MKC (2012) The lipid droplet coat protein perilipin 5 also localizes to muscle mitochondria. *Histochem Cell Biol* 137:205-216.

Brasaemle DL (2007) Thematic review series: Adipocyte Biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *J Lipid Res* 48:2547-2559.

CIE (International Commission on Illumination) 1976. International commission on illumination, colorimetry: official recommendations of the international commission on illumination. Publication CIE no. 15 (E-1.3.1). Bureau Central de la CIE, Paris.

Dalen KT, Dahl T, Holter E, Arntsen B, Londos C, Sztalryd C & Nebb HI (2007) LSDP5 is a PAT protein specifically expressed in fatty acid oxidizing tissues. *Biochim Biophys Acta* 1771:210-227.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V & Davoli R (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. *Meat Sci* 88:631-637.

Gol S, Ros-Freixedes R, Zambonelli P, Tor M, Pena RN, Braglia S, Zappaterra M, Estany J & Davoli R (2016) Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs. *J Anim Breed Genet* 133:24-30.

Jeong H, Song KD, Seo M, Caetano-Anollés K, Kim J, Kwak W, Oh J, Kim E, Jeong DK, Cho S, Kim H & Lee HK (2015) Exploring evidence of positive selection reveals genetic basis of meat quality traits in Berkshire pigs through whole genome sequencing. *BMC Genetics* 16.

Jung JH, Shim KS, Na CS & Choe HS (2015) Studies on Intramuscular Fat Percentage in Live Swine Using Real-time Ultrasound to Determine Pork Quality. *Asian-Australas J Anim Sci* 28:318-322.

Laurens C, Bourlier V, Mairal A, Louche K, Badin PM, Mouisel E, Montagner A, Marette A, Tremblay A, Weisnagel JS, Guillou H, Langin D, Joanisse DR & Moro C (2016) Perilipin 5 fine-tunes lipid oxidation to metabolic demand and protects against lipotoxicity in skeletal muscle. *Sci Rep* 6:38310.

Lefaucheur L, Lebret B, Ecolan P, Louveau I, Damon M, Prunier A, Billon Y, Sellier P & Gilbert H 2011. Muscle characteristics and meat quality traits are affected by divergent selection on residual feed intake in pigs. *J Anim Sci* 89:996-1010.

Legarra A, Varona L & de Maturana EL (2011) TM. Retrieved on 26th February 2017 from http://genoweb.toulouse.inra.fr/~alegarra/tm_folder/manualtm.pdf.

MacPherson REK & Peters SJ (2015) Piecing together the puzzle of perilipin proteins and skeletal muscle lipolysis. *Appl Physiol Nutr Metab* 40:641-651.

MacPherson REK, Herbst EAF, Reynolds EJ, Vandenboom R, Roy BD & Peters SJ (2012) Subcellular localization of skeletal muscle lipid droplets and PLIN family proteins OXPAT and ADRP at rest and following contraction in rat soleus muscle. *Am J Physiol Regul Integr Comp Physiol* 302:R29-R36.

Murrieta CM, Hess BW & Rule DC (2003) Comparison of acidic and alkaline catalysts for preparation of fatty acid methyl esters from ovine muscle with emphasis on conjugated linoleic acid. *Meat Sci* 65:523-529.

Pollak NM, Jaeger D, Kolleritsch S, Zimmermann R, Zechner R, Lass A & Haemmerle G (2015) The Interplay of Protein Kinase A and Perilipin 5 Regulates Cardiac Lipolysis. *J Biol Chem* 290:1295-1306.

Pena RN, Gallardo D, Guardia MD, Reixach J, Arnau J, Amills M & Quintanilla R (2013) Appearance, flavor, and texture attributes of pig dry-cured hams have a complex polygenic genomic architecture. *J Anim Sci* 91:1051-1058.

R Core Team (2016) R: A language and environment for statistical computing. Retrieved on 10th of February 2017 from http://www.R-project.org/.

Ruusunen M & Puolanne E (2004) Histochemical properties of fibre types in muscles of wild and domestic pigs and the effect of growth rate on muscle fibre properties. *Meat Sci* 67:533-539.

Tamura K, Stecher G, Peterson D, Filipski A & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725-2729.

The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucl Acids Re*s 45:D158-D169.

Wang C, Zhao Y, Gao X, Li L, Yuan Y, Liu F, Zhang L, Wu J, Hu P, Zhang X, Gu Y, Xu Y, Wang Z, Li Z, Zhang H & Ye J (2015) Perilipin 5 improves hepatic lipotoxicity by inhibiting lipolysis. *Hepatology* 61:870-882.

Wang H, Hu L, Dalen K, Dorward H, Marcinkiewicz A, Russell D, Gong D, Londos C, Yamaguchi T, Holm C, Rizzo MA, Brasaemle D & Sztalryd C (2009) Activation of Hormone-sensitive

Lipase Requires Two Steps, Protein Phosphorylation and Binding to the PAT-1 Domain of Lipid Droplet Coat Proteins. *J Biol Chem* 284:32116-32125.

Wang H, Sreenivasan U, Hu H, Saladino A, Polster BM, Lund LM, Gong D, Stanley WC & Sztalryd C (2011) Perilipin 5, a lipid droplet-associated protein, provides physical and metabolic linkage to mitochondria. *J Lipid Res* 52:2159-2168.

Wolins NE, Brasaemle DL & Bickel PE (2006) A proposed model of fat packaging by exchangeable lipid droplet proteins. *FEBS Letters* 580:5484–5491.

Wood JD (1990) Consequences for meat quality of reducing carcass fatness. In Reducing fat in meat animals [JD Wood and AV Fisher editors], pp. 344-397. Elsevier Applied Science, London, UK.

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

Wood JD, Nute GR, Richardson RI, Whittington FM, Southwood O, Plastow G, Mansbridge R, da Costa N & Chang KC (2004) Effects of breed, diet and muscle on fat deposition and eating quality in pigs. *Meat Sci* 67:651-667.

Yamaguchi T, Matsushita S, Motojima K, Hirose F & Osumi T 2006. MLDP, a novel PAT family protein localized to lipid droplets and enriched in the heart, is regulated by peroxisome proliferator-activated receptor alpha. *J Biol Chem* 281:14232-14240.

Zappaterra M, Deserti M, Mazza R, Braglia S, Zambonelli P & Davoli R (2016) A gene and protein expression study on four porcine genes related to intramuscular fat deposition. *Meat Sci* 121:27-32.

Table 1. The considered Italian Large White and Italian Duroc samples utilised for the detection of PLIN5 antigens through immunohistochemical analysis, for the genotyping of rs327694326 SNP and for the comparison between breeds of PLIN5 gene expressions.

	Immunohistochemical analysis		Genotyping		Gene expression study	
	ILW[1]	IDU[2]	ILW[1]	IDU[2]	ILW[1]	IDU[2]
Total	5	5	512	300	30	30
Sows	3	3	339	203	20	10
Barrows	2	2	173	97	18	12
Days of slaughter	2	2	27	19	16	10
Litters	5	5	270	148	28	25

[1] Italian Large White pigs.

[2] Italian Duroc pigs.

Table 2. Immunostaining results for Perilipin 5 (PLIN5) protein and average fiber size in pig Semimembranosus muscle in Italian Large White and Italian Duroc samples. Samples belonging to the two breeds have been compared using Wilcoxon test, and the relative P-values were reported.

	Immunostaining results				
	ILW[1]		IDU[2]		
	Mean	SD	Mean	SD	*P*-value
PLIN5 staining total CSA[3] (μm^2)	715,261.66	100,166.11	708,011.97	84,280.84	0.931
PLIN5 staining CSA percentage (%)	92.94	5.13	94.96	1.92	0.931
Average PLIN5-positive myofiber size (μm^2)	7,673.47	776.34	7,459.79	909.47	0.931
Average PLIN5-negative myofiber size (μm^2)	6,677.67	2,967.15	6,156.80	2,607.23	0.792

[1] Italian Large White pigs.

[2] Italian Duroc pigs.

[3] CSA stands for cross-sectional area.

Table 3. Genotyping results and rs327694326 allele frequencies in Italian Large White and Italian Duroc samples.

		TT		CT		CC		Allele frequencies		
	N	N	%	N	%	N	%	T	C	HWE[1]
ILW[2]	512	395	0.77	92	0.18	25	0.05	0.86	0.14	P<0.0001
IDU[3]	300	278	0.93	19	0.06	3	0.01	0.96	0.04	-

For Italian Duroc pigs the C allele frequency was too low to calculate the deviance from Hardy-Weinberg equilibrium

[1] Hardy-Weinberg equilibrium.

[2] Italian Large White pigs.

[3] Italian Duroc pigs.

Table 4. Additive and dominant effects of rs327694326 SNP on backfat fatty acid composition, with the means and differences of the estimated marginal posterior distribution for the genotypes in Italian Large White pigs.

Backfat fatty acids[1]	Additive (a) and dominant (d) effects				Estimated means			Estimated differences between genotypes					
								TT-CT		CT-CC		TT-CC	
	a	P(>0)	d	P(>0)	TT	CT	CC	Mean	P(>0)	Mean	P(>0)	Mean	P(>0)
C18:1, cis-9[2]	0.001	0.64	0.002	0.85	38.41	38.66	38.39	-0.25	**0.07**	0.27	0.79	0.02	0.52
C18:1, cis-11[3]	-0.00002	0.48	0.0003	0.83	2.18	2.22	2.20	-0.03	**0.10**	0.03	0.69	-0.008	0.44
C18:2, trans-9, cis-12[4]	-0.000004	0.30	-0.00002	**0.05**	0.021	0.019	0.022	0.002	**0.98**	-0.002	**0.09**	-0.0004	0.38
C18:3, Ω3[5]	-0.00014	0.20	-0.00029	0.12	0.77	0.76	0.80	0.012	0.74	-0.04	0.15	-0.028	0.20
C18:3/C18:2[6]	-0.00001	0.14	-0.00002	**0.10**	0.049	0.048	0.051	0.0006	0.67	-0.003	**0.10**	-0.002	0.14
C20:1, cis-11[7]	-0.00002	0.47	0.00036	0.89	0.80	0.83	0.81	-0.040	**0.04**	0.028	0.73	-0.012	0.38
C20:2, Ω6[8]	0.0001	0.71	0.00015	0.88	0.77	0.79	0.77	-0.014	**0.08**	0.013	0.75	-0.001	0.48
C22:4, Ω6[9]	0.00002	**0.95**	0.00002	0.85	0.093	0.092	0.088	0.0004	0.58	0.0037	0.87	0.0041	**0.91**
C22:6, Ω3[10]	-0.00001	0.30	-0.00001	0.15	0.013	0.012	0.011	0.0007	0.80	0.0012	0.77	0.0019	**0.90**
LCFA[11]	0.0010	0.80	0.0018	0.86	73.26	73.41	73.15	-0.15	0.13	0.26	0.84	0.11	0.68
MUFA[12]	0.0005	0.61	0.003	0.89	43.61	43.94	43.61	-0.33	**0.05**	0.33	0.81	-0.001	0.50
PUFA Ω3[13]	-0.00014	0.22	-0.00031	0.11	0.84	0.82	0.87	0.014	0.77	-0.043	0.14	-0.028	0.20

[1] All the fatty acids or fatty acid categories are expressed as percentage on the total fatty acids amount.

[2] Oleic acid; [3] Cis-vaccenic acid; [4] Trans-9, cis-12 octadecenoic acid; [5] Linolenic acid; [6] the ratio between C18 tri-unsaturated and C18 di-unsaturated fatty acids; [7] Eicosenoic acid; [8] Eicosadienoic acid; [9] Docosatetraenoic acid; [10] Docosahexaenoic acid; [11] Long chain fatty acids; [12] Monounsaturated fatty acids; [13] Polyunsaturated fatty acids omega3.

P (>0): Posterior probability of a value being positive. In bold, probabilities above (or equal to) 0.90 or below 0.10.

Table 5. Additive and dominant effects of rs327694326 SNP on FCR and carcass phenotypes, with the means and differences of the estimated marginal posterior distribution for the genotypes in Italian Large White pigs.

								Estimated differences between genotypes					
	Additive (a) and dominant (d) effects				Estimated means			TT-CT		CT-CC		TT-CC	
Traits	a	P(>0)	d	P(>0)	TT	CT	CC	Mean	P(>0)	Mean	P(>0)	Mean	P(>0)
Lean[1] (%)	-0.002	0.20	-0.005	**0.07**	48.50	48.09	48.97	0.41	0.91	-0.89	**0.06**	-0.48	0.18
Backfat thickness[2]	0.003	0.71	0.009	**0.90**	28.29	29.02	27.67	-0.73	**0.10**	1.35	**0.90**	0.62	0.74
Drip loss[3]	0.008	0.81	0.04	**>0.99**	52.59	56.14	51.04	-3.55	**<0.01**	5.09	**0.99**	1.54	0.79
pH_{24}[4]	0.0002	0.77	0.001	**>0.99**	5.96	6.05	5.92	-0.09	**0.01**	0.14	**0.98**	0.04	0.77
CIE *L**	-0.005	0.11	-0.009	**0.07**	33.18	32.86	34.27	0.31	0.73	-1.41	**0.07**	-1.09	**0.10**
CIE *a**	0.003	**0.91**	0.003	0.86	9.17	9.22	8.69	-0.05	0.42	0.53	0.88	0.48	0.88
CIE *b**	-0.002	0.09	-0.004	**0.03**	3.64	3.45	4.01	0.19	0.88	-0.56	**0.03**	-0.37	**0.08**

[1] The percentage of lean mass measured with Fat-O-Meter.

[2] Backfat thickness measured with Fat-O-Meter at 8 cm off the midline of the carcass between the third and fourth last ribs.

[3] Drip loss measured with Filter Paper Press method

[4] pH measured at 24 hours post-mortem.

CIE stands for International Commission on Illumination.

P (>0): Posterior probability of a value being positive. In bold, probabilities above (or equal to) 0.90 or below 0.10.

Table 6. The Pearson's correlation results between PLIN5 and LIPE expressions in Italian Large White and Italian Duroc samples.

	PLIN5 gene expression		
	n[1]	r[2]	*P* value
LIPE gene expression			
ILW[3]	69	0.371	0.003
IDU[4]	34	0.597	0.0002

[1] The number of the considered samples

[2] the Pearson's correlation coefficient.

[3] Italian Large White pigs.

[4] Italian Duroc pigs.

Figures

Figure 1. Western blot results confirming the specificity of primary antibody rabbit anti-PLIN5 (A). Cross-sections from *Semimembranosus* muscle in Italian Duroc (B) and Italian Large White (C) pigs stained by PLIN5. The images show that the majority of muscle fibers are PLIN5 immunoreactive, while some isolated or coupled fibers appear unlabeled (asterisks).

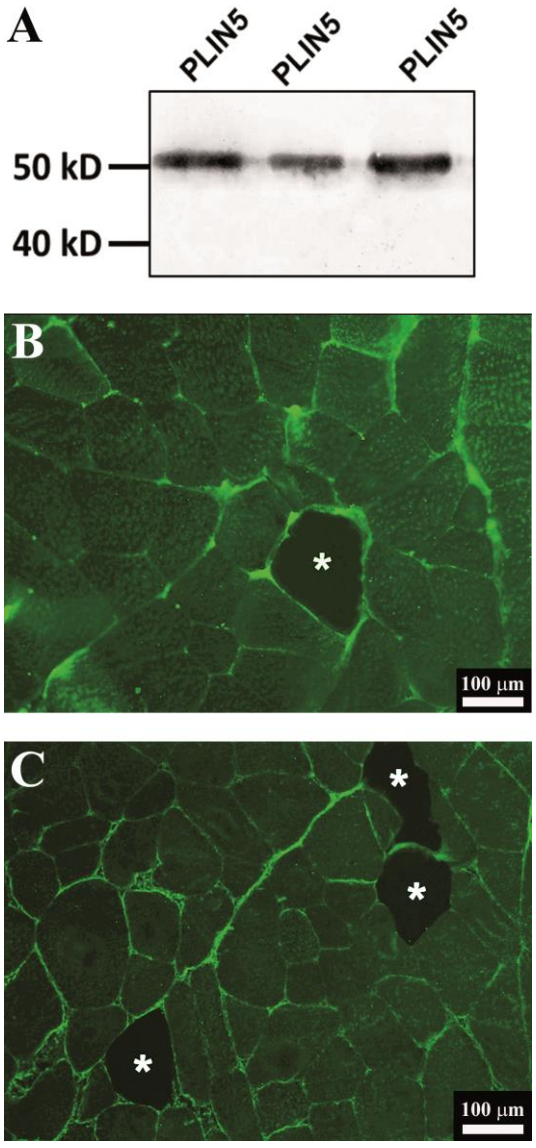
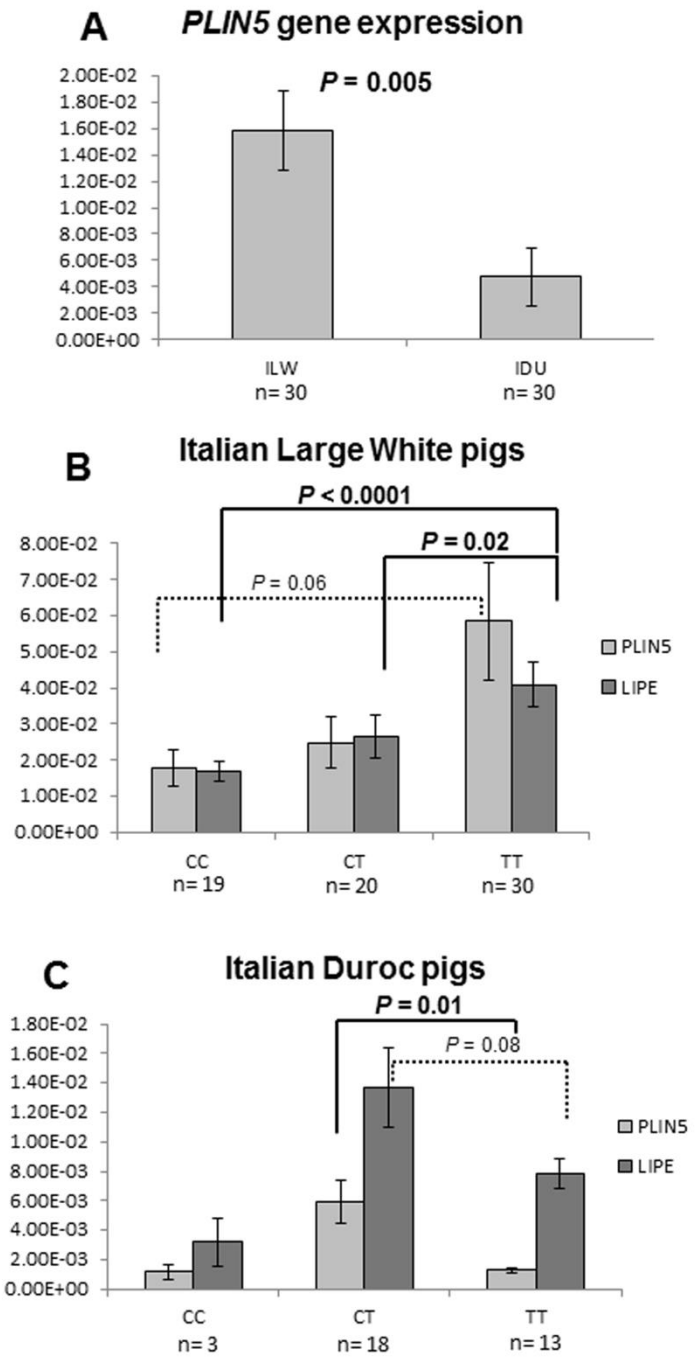


Figure 2. Gene expression results and Student's t test *P* values of the comparisons between *PLIN5* levels in the two breeds (A) and between *PLIN5* and *LIPE* expressions for the different genotypes of rs327694326 SNP in Italian Large White (B) and Italian Duroc pigs (C).



Supplementary Table S1. Primers used for the sequencing, the genotyping and the gene expression study.

Primer name (sequence ID)	Primer sequence (5'-3')	TM (°C)	Size (bp)
Sequencing (ENSSSCG00000013513.2)			
PLIN5 -FOR	5'- TTGGGGCTCTGAAAAGTGAG -3'		
PLIN5 -REV	5'- GCAGTTGTGGTTCAGATCCTG -3'	64°C	600
Genotyping(ENSSSCG00000013513.2)			
PLIN5 -FOR	5'- CCAGGCTAGGGGTGGAATC -3'		
PLIN5 -REV	5'- AAGAATCCAGCATCGCCATG -3'	63°C	119
PLIN5 (ENSSSCG00000013513.2)			
PLIN5 -E-FOR	5'- GCGGTCTCCGATGCTTATAG -3'		
PLIN5 -E-REV	5'- CCCTGTTGTCTCCTCTGCTC -3'	64°C	476
PLIN5 -I-FOR	5'- GTGGAGCTCAAACGATCCAT -3'		
PLIN5 -I-REV	5'- TCAGTCATGGGCAGGAAGT -3'	67°C	89
LIPE (AY686758.1)			
LIPE -E-FOR	5'- CCGAGACGAGATTAGCACCA -3'		
LIPE -E-REV	5'- CCTAGCGAACATGACCGAGT -3'	66°C	247
LIPE -I-FOR	5'- AAGTCTACAGTGTGAGGGCC -3'		
LIPE -I-REV	5'- CGATGGGAGCTGAGTAGAGG -3'	70°C	96
Normalising genes			
HPRT1 (AK346023.1)			
HPRT1-E-FOR	5'- GCCCCAGCGTCGTGATTA -3'		
HPRT1-E-REV	5'- AGAGGGCTACGATGTGATGG -3'	64°C	183
HPRT1-I-FOR	5'- CCCAGCGTCGTGATTAGTGA -3'		
HPRT1-I-REV	5'- CCTTTTCCAAATCCTCGGCA -3'	66°C	88
YWHAZ (AK344707.1)			
YWHAZ-E-FOR	5'- TGGAGCACTTACAAGGCGTA -3'		
YWHAZ -E-REV	5'- ACCGTTTCTGCCCTTATCCA -3'	64°C	168
YWHAZ -I-FOR	5'- AAGGCGTAGTGGAAGTGGAT -3'		
YWHAZ -I-REV	5'- GCTGTAGTCAAAGGTGTGCA -3'	66°C	98
B2M (AK239552.1)			
B2M-E-FOR	5'- AAACGGGGAGAAGATGAACG -3'		
B2M-E-REV	5'- ACATCTACCTGCTCAGACAGT -3'	63°C	377

B2M-I-FOR	5’- CCTTCTGGTCCACACTGAGT -3’		
B2M-I-REV	5’- TCCCACTTAACTATCTTGGGCT -3’	66°C	99

For all the considered genes in the table are reported the used primer, the annealing temperature expressed in °C and the size in base pairs (bp).

Each primer name is composed by the gene name, “I” for internal primer or “E” for external primer, “FOR” for the left primer and “REV” for the right primer.

[1]TM (°C) Annealing temperature.

Chapter 6: A gene and protein expression study on four porcine genes related to intramuscular fat deposition.

M. Zappaterra[1], M. Deserti[2 3], R. Mazza[4], S. Braglia[1], P. Zambonelli[1], R. Davoli[1].

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

[2] Center for Applied Biomedical Research (C.R.B.A.), S. Orsola-Malpighi University Hospital, University of Bologna, Via Massarenti 9,40138, Bologna, Italy.

[3] Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Via Massarenti 9,40138, Bologna, Italy.

[4] Center for Applied Biomedical Research and Endocrinology Unit, Department of Medical and Surgical Sciences, University of Bologna, Via Massarenti 9,40138, Bologna, Italy.

Published in

Meat Science

Zappaterra et al. *Meat Sci.* 2016 Nov;121:27-32.

doi: 10.1016/j.meatsci.2016.05.007

Summary

Intramuscular fat (IMF) content has a prominent role in meat quality, affecting sensory attributes such as flavour and texture. In the present research, we studied in samples of porcine *Semimembranosus* muscle four genes related to lipid metabolism and whose gene expressions have been associated to IMF deposition: *FASN*, *SCD*, *LIPE* and *LPL*. We analysed both mRNA and protein expressions in two groups of Italian Large White pigs divergent for *Semimembranosus* IMF deposition, with the aim of comparing the levels of four genes and enzymes between the two groups and identifying possible coexpression links. The obtained results suggest a prominent role of LIPE enzyme in IMF hydrolysis, as the samples with low IMF deposition show a significantly higher amount of this lipase. Finally, a poorly known correlation was found between LIPE and FASN enzymes only in female individuals. These results provide new information for the understanding of IMF deposition.

Introduction

During the last decades, pig selection has aimed at satisfying the request of the pork industry mainly focused on the increase of muscle deposition and carcass lean cut amount, resulted in a reduction in fat storage and adipogenesis (Wood, 1990). This selective pressure has also led to a progressive lowering in the total lipid content of muscle, in particular in some breeds selected for their predisposition to lean mass deposition (Wood and Warriss, 1992). In Large White breed the selection lowered noticeably the marbling percentage, passing from an average of 2-4% of intramuscular fat (IMF) in Large White pigs bred in 1960's (Wood, 1990) to less than 1% in *Longissimus* muscle of the modern Large White pigs (Wood et al., 2008). IMF is composed of lipid droplets stored within myofibers cytoplasm and adipocytes located between the fiber *fasciculi*, consequently IMF amount is strongly related to the number of intramuscular adipocytes (Zheng and Mei, 2009). It is known that sensory attributes such as flavour and juiciness are influenced by IMF content (Wood, 1990; Fernandez et al., 1999).

The number of Genome Wide Association studies aimed at identifying SNPs and genes affecting IMF deposition and composition in different pig breeds (Ma et al., 2013; Muñoz et al., 2013; Nonneman et al., 2013; Kim et al., 2015; Davoli et al., 2016) is increasing, however little is known about the relative enzyme quantity of the putative genes involved in IMF

deposition in pig muscle tissue. Therefore, in the present study we considered four genes involved in lipid metabolism and whose mRNA levels in literature are reported to be linked to IMF deposition in different pig breeds (Zhao et al., 2009; Wang et al., 2012). Two of the selected enzymes are involved in synthesis and desaturation of fatty acids (*fatty acid synthase, FASN*; *stearoyl-CoA desaturase (delta 9 desaturase)*, *SCD*) and the two remaining enzymes are involved in fatty acids catabolism (*hormone sensitive lipase*, *LIPE*; *lipoprotein lipase*, *LPL*). *FASN* gene is located on Sus scrofa chromosome 12 (Muñoz et al., 2003), and the coded protein plays an essential role in long-chain fatty acid synthesis, starting from acetyl CoA and using malonyl-CoA as a 2 carbon donor and NADPH as reducing equivalent (Wakil, 1989; Menendez et al., 2009). In pigs both *SCD* and *LPL* genes are localised on chromosome 14 (Gu et al., 1992; Ren et al., 2003). SCD catalyses the desaturation of palmitoyl-CoA and stearoyl-CoA at the position Δ9 producing de novo palmitoleoyl-CoA and oleoyl-CoA, while LPL has the dual function of hydrolysing the circulating chylomicron triglycerides to diglycerides and of ligand/bridging factor for receptor mediated lipoprotein uptake. On the other hand, LIPE hydrolyses the triglycerides stored in muscle to diglycerides, then to monoglycerides and at the end to free fatty acids. *LIPE* gene has been assigned to porcine chromosome 6, and its position coincides with a Quantitative Trait Locus (QTL) region linked to sensory quality in porcine meat (Pena et al., 2013). In the present study, we analysed both the protein quantifications and the transcription profiles of these four genes in the *Semimembranosus* muscle of two groups of Italian Large White (ILW) pigs divergent for IMF deposition, with the aims: i) of testing whether the mRNA and enzyme levels of FASN, LIPE, SCD and LPL differed between two groups of pigs divergent for IMF (LOW IMF group *vs*. HIGH IMF group), ii) of identifying common trends in the expression levels of the four studied genes and proteins, suggesting coexpression links.

The focus both on gene and protein levels of the analysed enzymes could be useful to highlight a possible involvement of FASN, LIPE, SCD and LPL proteins on IMF deposition in pig meat.

Materials and methods

Protein expression study

Sampling. For the present study, a set of 155 pigs was selected among a population of 950 Italian Large White (ILW) pigs. As reported in Davoli et al. (2016), using the Soxhlet extraction method the whole sample of 950 ILW pigs has been characterised for IMF content, reported as percentage (grams of IMF on 100 grams of pig *Semimembranosus* muscle). The 155 individuals used for the quantitation of FASN, SCD, LIPE and LPL enzymes in *Semimembranosus* muscle have been selected among the 950 ILW pigs based on their extreme and divergent IMF phenotype (Table 1) and divided in two groups differing for IMF content (LOW IMF group and HIGH IMF group). The LOW IMF group consists of 77 pigs presenting IMF values lower than the average IMF level of the 950 pigs (total population) minus one standard deviation unit. Furthermore, the HIGH IMF group consists of individuals with IMF contents higher than the average IMF value of the total population plus 3.5 standard deviation units. The two IMF divergent groups have been chosen avoiding as much as possible full and half sibs, in order to prevent the family effect on protein and gene quantitation results. Additionally, since sex and batch may influence protein and gene expressions the two groups were balanced for these two factors. The pigs are pure breed animals included in the Italian sib test genetic evaluation scheme performed by ANAS (Associazione Nazionale Allevatori Suini, ANAS; www.anas.it), reared in the same environmental conditions at the genetic test station with a *quasi ad libitum* feeding level (60% of the pigs was able to ingest the entire supplied ration). The sib test program calculates the estimated genetic value of each candidate boar testing three of its full sibs, two females and one castrated male. For this reason, the considered sample of 155 individuals is composed of two thirds of sows and one third of castrated males (Table 1), with the purpose of maintaining the proportions of the total population and of using a representative sample of the 950 pigs population. At the end of the test, the animals were transported to a commercial abattoir located about 25 km from the test station in accordance with Council Rule (EC) No. 1/2005 regarding the protection of animals during transport and related operations and, amending Directives 64/432/EEC and 93/119/EC and Regulation (EC) No. 1255/97. At the slaughterhouse, the pigs were electrically stunned and bled in a supine position in agreement with Council Regulation (EC) No. 1099/2009 regarding the protection of animals at the time of slaughter. All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health. Moreover, for all the pigs, ANAS provided us with the Estimated Breeding Values (EBVs): Average Daily Gain (ADG, calculated from 30 to 155 kg of live weight with *quasi ad libitum* feeding level, expressed in grams), Backfat Thickness (BFT, recorded *post mortem* at the level of *Gluteus medius* muscle, expressed in mm), Lean Cuts (LC, the sum of neck and loin weight, expressed in kg), Feed Conversion Ratio (FCR, obtained from feed intake recorded daily and body weight measured bimonthly, expressed in units),

and Ham Weight (HW, expressed in kg). The listed EBVs have been calculated by ANAS according to the statistical model reported by Russo et al. (2000). Moreover, the slaughterhouse technicians provided us with the phenotypic measures of the carcass weight expressed in kg (carcass weight), the percentage of lean cuts over the total carcass weight (%lean), and the backfat and loin thickness measured at 8 cm off the midline of the carcass at the level placed between the third and fourth last ribs measured with Fat-O-Meter (FOM).

Protein extraction and total protein quantitation. At the slaughterhouse, *Semimembranosus* muscle tissue was quickly frozen in Liquid Nitrogen (LN2) and then stored at -80°C for subsequent total protein extraction. Approximately 300 mg of tissue sample were homogenized on ice in 900 µl of T-PER Tissue Protein Extraction (Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA), to which 1x protease and phosphatase inhibitors (PhosSTOP™, Roche, Hoffmann-La Roche, Basel, Switzerland) were added. Tissue lysates were then centrifuged at 15,000 x g for 20 min at 4°C to remove lipids and insoluble debris. For LPL, before the quantitation of the total proteins and the quantification of LPL enzyme, the extracted proteins were concentrated through a filtering step, using Amicon® Ultra-4 filters (Merck Millipore, Darmstadt, Germany). This additional step was needed to obtain LPL quantitation within the sensitivity range of the commercial ELISA kit used for the quantitation of this protein.

The total extracted proteins were quantified using the BCA reagent (Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA) and the optical density of each sample was determined using a microplate optical reader.

Quantification of intramuscular FASN, LIPE, SCD and LPL enzymes. FASN, LIPE, SCD and LPL were quantitatively detected using ELISA kits produced by Cusabio (Cusabio Biotech, Wuhan, China), according to manufacturer's instructions. Briefly, 100 µl of blank, standards or samples were added to well, covered, incubated for 2 hours at 37°C and the liquid of each well was removed without washing. Then 100 µl of biotin antibody working solution was added to each well, incubated for 1 hour at 37°C and then washed three times with wash buffer (200 µl). At each well, 100 µl of HRP avidin working solution were added, the well covered, incubated for 1 hour at 37°C and washed three times with wash buffer (200 µl). Finally, TMB substrate (90 µl) was then added to each well, covered, dark incubated for 30

minutes at 37°C and stopped with 50 µl of stop solution. The optical density of each well was determined within 5 minutes using a microplate reader at 450 nm.

For all the samples, the relative quantification of each enzyme was calculated, through the ratio between the enzyme absolute quantitation and the total extracted protein amount.

Gene expression study

Sampling. The gene expression study was carried out on a smaller sample, composed of 47 ILW pigs chosen among the overall group of 155 animals. These selected pigs, which were extreme and divergent for IMF phenotype (Table 1), were the progeny of 36 boars and 44 sows. Both Subset 1 (for the low IMF pigs) and Subset 2 (for the high IMF individuals) have been chosen avoiding as much as possible full and half sibs.

RNA source, total RNA extraction and cDNA preparation. Semimembranosus muscle samples were withdrawn at the slaughterhouse and immediately frozen at -80°C. Total RNA was then extracted using TRIZOL reagent (Invitrogen Corporation, Carlsbad, California), the quality and integrity of the RNAs were both checked and RNA was retrotranscribed to cDNA as described in Davoli et al. (2011).

Gene expression quantitation. The quantitative Real-Time PCR (qRT-PCR) standard curve method (Pfaffl, 2004) was used to analyse the four genes expression. qRT-PCR was performed on Rotor Gene™ 6000 (Corbett Life Science, Concorde, New South Wales) using 5 µl of SYBR® Premix Ex Taq™ (TAKARA Bio INC, Olsu, Shiga, Japan), 5 pmol of each primer, 2 µl of cDNA template diluted 1:10 and make up to the total volume of 10 µl with water. Rotor Gene™ 6000 protocol was optimised using specific annealing temperatures for each primer couple (Supplementary Table S1). Six genes were tested to be used as normalising genes (polymerase (RNA) II (DNA directed) polypeptide A, 220kDa, POLR2A; beta-2-microglobulin, B2M; tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta, YWHAZ; hypoxanthine phosphoribosyltransferase 1, HPRT1; TATA box binding protein, TBP; peptidylprolyl isomerase A (cyclophilin A), PPIA) and using NormFinder (Andersen et al., 2004) were identified the three most stable ones to be used as normalising genes (*B2M*, *YWHAZ*, and *HPRT1*). For the normalising and the 4 studied genes a specific standard curve was obtained, amplifying 11 serial dilutions (from 10^9 to 50 molecules/µl) of a known concentration sample of cDNA amplicon, obtained with a PCR with the external primer pairs (Supplementary Table S1). The PCR efficiency was calculated

as E = 10 exp (-1/slope), with a range between -2.7 and -4.3, indicating a good PCR efficiency results. All the PCR products were checked on a polyacrylamide gel and washed from primer dimers with the QIAquick® PCR Purification Kit (QIAGEN, Venlo, Netherlands). The absence of unspecific amplicons during qRT-PCR on Rotor Gene™ 6000 was tested using the melt step after the cycling. The variation coefficient (CV = Standard Deviation of the Crossing Points/Average of the Crossing Points) of the replicated analysis for each samples (three in two different qRT PCRs) was accepted for CV < 0.2.

Statistical analysis. All the statistical analyses were carried out using SAS software, version 9.4 (SAS Inst. Inc., Cary, NC). To decide what statistical approach to use, we performed a general linear model (GLM) where we tested the fixed effects of sex, batch and group (LOW and HIGH IMF groups) on the studied gene and protein levels. As sex and batch did not affect the gene and protein levels, we decided to take into account only the group effect, comparing the protein and gene levels between the divergent groups of pigs. PROC NPAR1WAY command of SAS (the Wilcoxon test) was used to compare the protein levels between the groups of pigs divergent for IMF deposition (as protein values were not normally distributed), and the PROC TTEST of SAS (Student's t test) was utilised to compare the gene expression data between groups (as gene expression values were normally distributed). The PROC CORR SAS command was considered to calculate the correlations between the different gene and protein expressions.

Results

The main results are reported in Figure 1, Figure 2 and in Table 2, where for the two IMF divergent groups the protein and gene expression quantifications are showed. In particular we investigated both the mRNA and protein levels of *FASN*, *LIPE*, *SCD* and *LPL* genes, which have been already described in literature to be likely involved in IMF deposition. However, to date, the association of these genes with IMF deposition was reported mainly on the basis of their mRNA levels while in the present research both the mRNA and protein amounts were considered. For the quantification of these four enzymes in the two groups of ILW pigs extreme for IMF, we used commercially available ELISA kits. Owing to some problems in SCD and LPL protein quantitation and in order to avoid uncertain data, we decided to present and compare only the most reliable quantification values, maintaining however the

divergence between the two groups for IMF value and balancing the samples for sex (Supplementary Table S2).

The means ± the standard errors of the quantified enzymes in the LOW IMF and HIGH IMF groups are represented in Figure 1. It is worth noting that LIPE is the only enzyme showing a higher amount in samples with LOW IMF deposition ($P<0.05$) (Figure 1 C), while SCD protein presents a trend ($P<0.1$) and on average a higher enzyme expression in the HIGH IMF group (Figure 1 A). The transcription levels obtained for the four considered genes are reported in Figure 2: the two groups (LOW and HIGH IMF) show differences in the mRNA levels for *FASN*, *LPL* and *SCD* genes. In Table 2 are indicated the means ± the standard deviations of the gene and enzyme expressions in the LOW and HIGH IMF groups, in the LOW and HIGH IMF sows and castrated males, and in the two sexes. In the same table are reported the significant comparisons between groups. In order to identify links between gene transcription levels and enzyme expressions, the correlations between the mRNA and the protein levels were performed, and the results reported in Table 3. The lack of correlation between the mRNA level of each gene and the expression of its coded protein clearly stands out. Furthermore, we calculated the correlations between each one of the analysed mRNA and protein expressions and all the other available quantitation data, in order to look for the presence of some possible coexpressions (Table 3). This approach revealed the presence of two significant correlations between the mRNA of *LIPE* gene and the amount of LPL and SCD proteins. Furthermore, *FASN* gene resulted highly correlated to *SCD* and *LPL* gene expressions, while for the enzymes a single positive correlation between LIPE and FASN was identified (Table 3). In particular, this positive and high correlation between LIPE and FASN enzymes showed to be strongly linked to sex, as only sows presented a highly significant correlation coefficient, not observed in castrated males (Table 4). Furthermore, in order to find some possible links between gene and protein expressions and carcass traits, a correlation analysis considering gene and enzymes expression levels towards EBVs and phenotypic measures of carcass traits was carried out (Supplementary Table S3). The results in Supplementary Table S3 show low correlations between the gene and protein expression data and the EBVs or phenotypic measures, but all the associations we found are consistent with the roles of the studied genes and enzymes. For example, FASN is negatively correlated to ADG and HW EBVs, and to carcass weight measure, in agreement with the lipogenic role of this enzyme. Moreover, IMF content resulted to be positively associated with *FASN*, *SCD*

and *LPL* gene expression levels, and among the tested enzymes only LIPE showed a negative correlation with IMF deposition (Supplementary Table S3).

Discussion

The primary aim of the study was to investigate the expression levels of FASN, LIPE, SCD and LPL enzymes between two groups of ILW pigs divergent for IMF deposition, in order to look for the involvement of these proteins in porcine IMF deposition. The results suggest a prominent role of LIPE enzyme in IMF hydrolysis: indeed, low IMF pigs have a significantly higher amount of LIPE enzyme as compared to the high IMF group (Figure 1 C and Table 2). This result agrees with the role of LIPE, which hydrolyses to free fatty acids the triglycerides stored in adipose tissue. According to this result, we also found a negative correlation between IMF deposition and LIPE enzyme amount (Supplementary Table S3). The different LIPE relative quantitation between high and low IMF individuals is also maintained between IMF divergent pigs of the same sex, indicating the importance of LIPE activity in IMF deposition for both castrated male and female individuals (Table 2). To date, in different animal species several authors have found associations between *LIPE* gene or enzyme expressions and the carcass traits (Zhao et al., 2010 in pigs; Ying et al., 2013 in mice; Zhang et al., 2014 in yaks). In particular, Zhao et al. (2009) found a divergent expression of *LIPE* gene and protein between 12 Wujin pigs with high IMF deposition and 12 Landrace pigs with low IMF deposition. Due to the different diet compositions and the use of two samples coming from distinct pig breeds, the diverging *LIPE* mRNA levels found by the authors between the two breeds may be affected by the different characteristics of the diets. However, on the basis of the results of the present study, it is possible to suggest a diet independent role of LIPE on IMF deposition, as LIPE enzyme showed different expressions between IMF divergent pigs, although they were reared in the same environmental conditions.

On the other hand, in pigs with a more pronounced IMF deposition a higher level of SCD protein was found, but the difference between the SCD enzyme expressions in LOW IMF and HIGH IMF groups showed only a trend towards significance ($P < 0.1$) (Figure 1). This result suggests that SCD enzyme could have an important role in IMF metabolism and is in agreement with the data reported by Wu et al. (2013), who detected an overexpression of *SCD* gene in high IMF Jinhua pigs as compared to Landrace pigs.

In addition, we tested also the linear correlation between the genes and the relative protein expressions (Table 3), finding that no one of the enzyme amounts significantly covaried with the relative gene transcription level. This result indicates that post-transcriptional regulation for these genes may play an essential role in the modulation of their enzyme synthesis and activity. However, despite this result, it is worth noting that *FASN*, *LPL* and *SCD* maintained on average a higher expression of both mRNA and protein levels in high IMF pigs as compared to the low IMF group (Figure 1 and 2) suggesting the existence of a conserved trend between the transcription and translation levels of these three genes.

Furthermore, the strong correlation found between LIPE and FASN (Table 3) was further investigated through an additional correlation analysis performed in castrated males and in sows (Table 4). Based on the results obtained, in muscle tissue we can suppose the activation of distinct biological pathways controlling IMF adipogenesis between female and male individuals. Mukherjee et al. (2014) found that in rat white adipose tissue the quantity of LIPE enzyme varies together with the expression of Caveolin1 (CAV1): in particular, CAV1 showed to be stimulated by estrogen (the primary female sex hormone) and suppressed by androgen (typically the male sex hormone). On the other side, CAV1 is coexpressed with FASN and interacts with this enzyme in human prostate cancer (Di Vizio et al., 2008). CAV1 is an important target in sex hormone dependent regulation of several metabolic pathways, particularly in cancer and diabetes (Mukherjee et al., 2014). Considering the results reported in literature, it is possible to hypothesise that the strong correlation we found in sows between LIPE and FASN proteins may be the effect of an estrogen activated pathway, possibly linking these two enzymes through the mediation of CAV1. Anyway, further studies are needed to prove in pigs this hypothesis.

In Table 3, also the correlation between *FASN* and *SCD* gene expressions stands out. The expressions of these two genes have already been found to covary in *Longissimus dorsi* muscle between pigs with divergent muscle lipid deposition (Wang et al., 2015). Nevertheless this coexpression was not detected at the protein level (Table 3).

Moreover, the correlations found between the EBVs or carcass traits and the gene and protein expression levels agree with the known roles of the studied genes: FASN and SCD are both enzymes involved in the synthesis and desaturation of fatty acids, and in fact showed

negative correlations with the EBV of Average Daily Gain and more generally to lean mass deposition, causing a worsening in Feed Conversion Ratio EBV (Supplementary Table 3).

Conclusions

The study of FASN, LIPE, LPL and SCD gene and protein expression levels allowed obtaining a more comprehensive view of their involvement in pig intramuscular fat deposition. The results suggest for LIPE and SCD enzyme a role in muscle fat deposition and indicate that these two genes may be involved in similar regulatory pathways and respond to similar transcription cues. Anyway, future studies are needed to better understand the role of *LIPE* gene and its coded protein in pork quality. Finally, in female individuals a poorly known correlation between LIPE and FASN enzymes was found, suggesting the need of dedicated studies aimed at identifying and elucidating the pathways involved.

Aknowledgements

This work was supported by the MiPAAF Innovagen project.

References

Andersen CL, Ledet-Jensen J & Ørntoft T (2004) Normalization of real-time quantitative RT-PCR data: a model based variance estimation approach to identify genes suited for normalization - applied to bladder- and colon-cancer data-sets. *Canc Res 64:5245–5250.*

Davoli R, Braglia S, Russo V, Varona L & te Pas MFW (2011) Expression profiling of functional genes in prenatal skeletal muscle tissue in Duroc and Pietrain pigs. *J Anim Breed Genet 128*:15-27.

Davoli R, Luise D, Mingazzini V, Zambonelli P, Braglia S, Serra M & Russo V (2016) Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *J Anim Breed Genet* 133:277-282.

Di Vizio D, Adam RM, Kim J, Kim R, Sotgia F, Williams T, Demichelis F, Solomon KR, Loda M, Rubin MA, Lisanti MP & Freeman MR (2008) Caveolin-1 interacts with a lipid raft-associated population of fatty acid synthase. *Cell Cycle* 7:2257–2267.

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

Gu F, Harbitz I, Chowdhary BP, Davies W & Gustavsson I (1992) Mapping of the porcine lipoprotein lipase (LPL) gene to chromosome 14q12----q14 by in situ hybridization. *Cytogenet Cell Genet* 59:63-64.

Kim ES, Ros-Freixedes R, Pena RN, Baas TJ, Estany J & Rothschild MF (2015) Identification of signatures of selection for intramuscular fat and backfat thickness in two Duroc populations. *J Anim Sci* 93:3292-3302.

Ma J, Yang J, Zhou L, Zhang Z, Ma H, Xie X, Zhang F, Xiong X, Cui L, Yang H, Liu X, Duan Y, Xiao S, Ai H, Ren J & Huang L (2013) Genome-Wide Association Study of Meat Quality Traits in a White Duroc×Erhualian F2 Intercross and Chinese Sutai Pigs. *PLOS ONE* 8:e64047.

Menendez JA, Vazquez-Martin A, Ortega FJ & Fernandez-Real JM (2009) Fatty Acid Synthase: Association with Insulin Resistance, Type 2 Diabetes, and Cancer. *Clin Chem* 55:425–438.

Mukherjee R, Kim SW, Choi MS, Yun JW (2014) Sex-Dependent Expression of Caveolin 1 in Response to Sex Steroid Hormones Is Closely Associated with Development of Obesity in Rats. *PLOS ONE* 9:e90918.

Muñoz G, Óvilo C, Noguera JL, Sánchez A, Rodríguez C & Silió L (2003) Assignment of the fatty acid synthase (FASN) gene to pig chromosome 12 by physical and linkage mapping. *Anim Genet* 34:234–235.

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

Nonneman DJ, Shackelford SD, King DA, Wheeler TL, Wiedmann RT, Snelling WM & Rohrer G A (2013) Genome-wide association of meat quality traits and tenderness in swine. *J Anim Sci* 91:4043–4050.

Pena RN, Gallardo D, Guardia MD, Reixach J, Arnau J, Amills M, Quintanilla R (2013) Appearance, flavor, and texture attributes of pig dry-cured hams have a complex polygenic genomic architecture. *J Anim Sci* 91:1051–1058.

Pfaffl MW (2004) Quantification strategies in real-time PCR. In *A–Z of quantitative PCR* [Bustin SA ed.], pp. 87-112, International University Line (IUL), La Jolla.

Ren J, Knorr C, Habermann F, Fries R, Huang LS & Brenig B (2003) Assignment of the porcine stearoyl-CoA desaturase (SCD) gene to SSC14q27 by fluorescence in situ hybridization and by hybrid panel mapping. *Anim Genet* 34:471–473.

Russo V, Buttazzoni L, Baiocco C, Davoli R, Nanni Costa L, Schivazappa C & Virgili R (2000) Heritability of muscular cathepsin B activity in Italian Large White pigs. *J Anim Breed Genet* 117:37-42.

Wakil SJ (1989) Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* 28:4523-4530.

Wang W, Xue W, Jin B, Zhang X, Ma F & Xu X (2012) Candidate gene expression affects intramuscular fat content and fatty acid composition in pigs. *J Appl Genet* 54:113–118.

Wang Z, Li Q, Chamba Y, Zhang B, Shang P, Zhang H & Wu C (2015) Identification of Genes Related to Growth and Lipid Deposition from Transcriptome Profiles of Pig Muscle Tissue. *PLOS ONE* 10:e0141138.

Wood JD (1990) Consequences for meat quality of reducing carcass fatness. In *Reducing fat in meat animals* [Wood JD, Fisher AV editors], pp. 344–397, Elsevier Applied Science, London.

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

Wood JD, Warriss PD (1992) The influence of the manipulation of carcass composition on meat quality. In *The Control of Fat and Lean Deposition* [Boorman KN, Buttery PJ, Lindsay DB editors], pp. 331-353, Butterworth Heinemann, Oxford.

Wu T, Zhang Z, Yuan Z, Lo LJ, Chen J, Wang Y & Peng J (2013) Distinctive Genes Determine Different Intramuscular Fat and Muscle Fiber Ratios of the longissimus dorsi Muscles in Jinhua and Landrace Pigs. *PLOS ONE* 8:e53181.

Ying HZ, Zang JN, Deng LL, Wang ZY & Yu C (2013) Pentamethylquercetin reduces fat deposition via Sirt1-mediated pathways in male obese mice induced by a high fat diet. *Food Chem Toxicol* 62:463–469.

Zhang HB, Wang MS, Wang ZS, Zhou AM, Zhang XM, Dong XW & Peng QH (2014) Supplementation Dietary Zinc Levels on Growth Performance, Carcass Traits, and Intramuscular Fat Deposition in Weaned Piglets. *Biol Trace Elem Res* 161:69–77.

Zhao SM, Ren LJ, Chen L, Zhang X, Cheng ML, Li WZ, Zhang YY & Gao SZ (2009) Differential Expression of Lipid Metabolism Related Genes in Porcine Muscle Tissue Leading to Different Intramuscular Fat Deposition. *Lipids* 44:1029–1037.

Zhao S, Wang J, Song X, Zhang X, Ge C & Gao S (2010) Impact of dietary protein on lipid metabolism-related gene expression in porcine adipose tissue. *Nutr Metab (Lond)* 7:6.

Zheng GS & Mei ZS (2009) Physiology, affecting factors and strategies for control of Pig meat intramuscular Fat. *Recent Pat Food Nutr Agric* 1:59–74.

Table 1. Composition and intramuscular fat (IMF) mean value of the two divergent and extreme groups of pigs considered for the enzymes quantification (top half of the table) and for mRNA expression study (bottom half of the table).

	PROTEIN QUANTITATION STUDY		
	LOW IMF Group (IMF <1.50%)	HIGH IMF Group (IMF > 3.50%)	Total
Castrated males	21	28	49
Sows	56	50	106
Total	77	78	155
IMF means (%)± st. dev.	0.838±0.131	4.663 ± 1.030	
	GENE EXPRESSION STUDY		
	Subset 1 (IMF <1.50%)	Subset 2 (IMF > 3.50%)	Total
Castrated males	11	10	21
Sows	14	12	26
Total	25	22	47
IMF means(%) ± st. dev.	0.738±0.093	5.718±1.130	

Table 2. The means and the standard deviations of the considered groups of pigs for the studied genes and enzymes.

	FASN		LIPE		SCD		LPL	
	Gene expression	Enzyme (ng/mg Prot)	Gene expression	Enzyme (u/mg Prot)	Gene expression	Enzyme (ng/mg Prot)	Gene expression	Enzyme (ng/mg Prot)
LOW IMF group	0.03±0.02[a]	0.045±0.068	0.04±0.02	0.123±0.168[g]	0.008±0.007	0.047±0.036	0.07±0.03[f]	35.28±73.28
HIGH IMF group	0.07±0.08[a]	0.048±0.085	0.04±0.02	0.069±0.101[g]	0.014±0.014	0.066±0.051	0.19±0.06[f]	61.51±123.34
LOW IMF sows	0.02±0.03[b]	0.051±0.076	0.05±0.03	0.126±0.159[d]	0.01±0.01	0.048±0.038	0.06±0.02[e]	42.65±84.30
LOW IMF castrated males	0.04±0.06	0.032±0.043	0.03±0.02	0.12±0.19[c]	0.01±0.01	0.046±0.032	0.08±0.03	26.76±59.16
HIGH IMF sows	0.07±0.09[b]	0.046±0.087	0.05±0.02	0.078±0.12[d]	0.01±0.01	0.062±0.061	0.21±0.07[e]	42.08±87.017
HIGH IMF castrated males	0.08±0.05	0.050±0.085	0.04±0.02	0.055±0.063[c]	0.02±0.02	0.069±0.037	0.16±0.06	82.69±152.99
Sows	0.04±0.07	0.049±0.081	0.05±0.02	0.10±0.14	0.01±0.01	0.054±0.049	0.14±0.09	42.35±84.77
Castrated males	0.06±0.06	0.043±0.071	0.04±0.02	0.081±0.13	0.01±0.01	0.059±0.037	0.12±0.06	56.77±121.08

Samples are grouped by intramuscular (IMF) deposition and/or by sex. Gene expression data were compared between groups using Student's t test, and enzyme amounts were compared between groups using Wilcoxon test.

Superscript letters indicate the significant differences; groups with the same superscript letter are statistically different. The *P* values of the comparisons are: *P*<0.05 (a, b, c, d); *P*<0.01 (e, f); *P*=0.001 (g).

Table 3. Significant correlations between gene and protein relative quantitation.

	FASN gene expression	FASN enzyme (ng/mg Prot)	*SCD* gene expression	SCD enzyme (ng/mg Prot)	*LPL* gene expression	LPL enzyme (ng/mg Prot)	*LIPE* gene expression	LIPE enzyme (u/mg Prot)
FASN gene expression	1	-	r=0.808 *P*<.0001	-	r=0.560 *P*<.0001	-	-	-
FASN enzyme (ng/mg Prot)	-	1	-	-	-	-	-	r=0.727 *P*<.0001
SCD gene expression	r=0.808 *P*<.0001	-	1	-	-	-	-	-
SCD enzyme (ng/mg Prot)	-	-	-	1	-	-	r=0.4709 *P*<.01	-
LPL gene expression	r=0.560 *P*<.0001	-	-	-	1	-	-	-
LPL enzyme (ng/mg Prot)	-	-	-	-	-	1	r=0.527 *P*=.001	-
LIPE gene expression	-	-	-	r=0.4709 *P*<.01	-	r=0.527 *P*=.001	1	-
LIPE enzyme (u/mg Prot)	-	r=0.727 *P*<.0001	-	-	-	-	-	1

- is reported for not significant values.

Table 4. Correlation analysis between the relative quantitation of LIPE and FASN enzymes in groups of samples divergent for IMF deposition and sex.

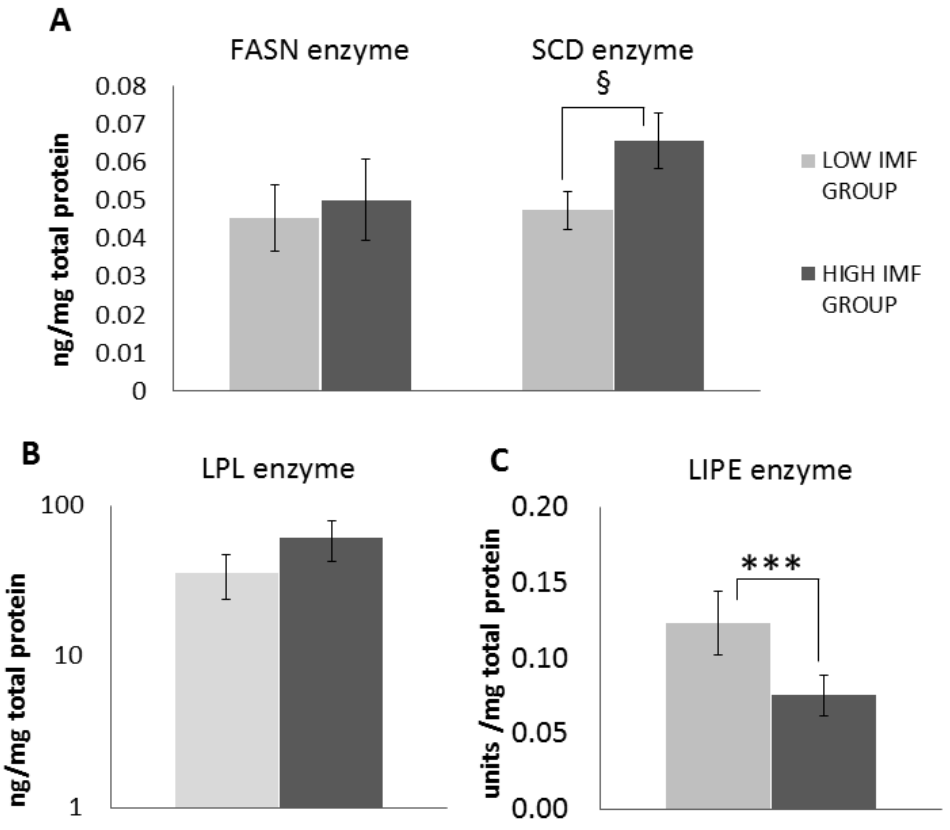
	LOW IMF group	HIGH IMF group	Sows	Castrated males
FASN enzyme (ng/mg Prot) with LIPE enzyme (u/mg Prot)	r=0.741 *P*<.0001 N=56	r=0.833 *P*<.0001 N=62	r=0.911 *P*<.0001 N=76	r=0.263 *P*=0.096 N=40

N stands for the number of samples considered for each considered groups of pigs.

Figures

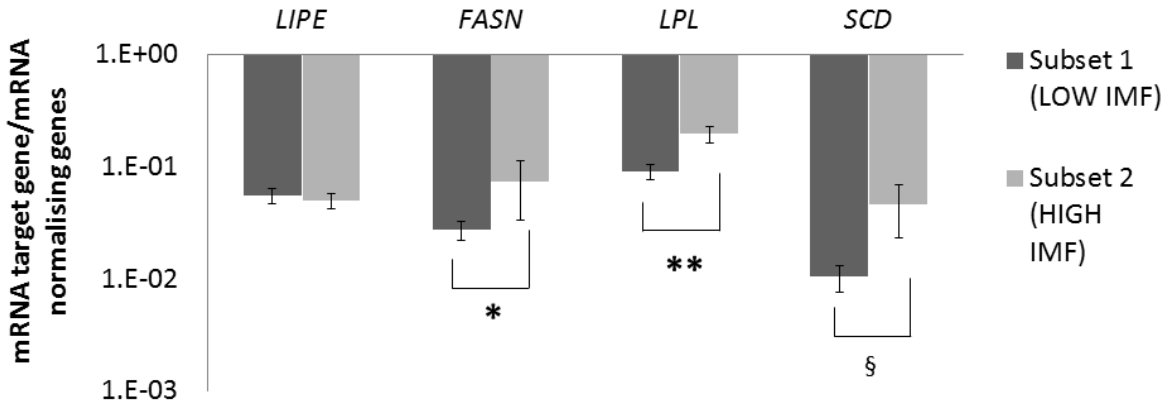
Figure 1. Relative quantitation of the studied enzymes in the LOW and HIGH IMF groups. Mean values are shown and error bars represent the standard errors.

Section A: relative quantitation of FASN and SCD enzymes in the low IMF and high IMF groups; section B: relative quantitation of LPL enzyme in the low IMF and high IMF groups; section C: relative quantitation of LPL enzyme in the low IMF and high IMF groups.



Significant differences between the two divergent groups are expressed with *** for $P<0.001$. [§] is used to indicate differences with P value<0.10.

Figure 2. Relative quantitation of the gene expressions in Subset 1 (LOW IMF) and Subset 2 (HIGH IMF). Mean values are shown and error bars represent the standard errors.



Significant differences between the two divergent groups are expressed with * for $P<0.05$ and ** for $P<0.01$. [§] is used to indicate differences with P value<0.1.

Supplementary Table S1. Primer used for the expression study in pig *Semimembranosus* muscle tissue.

Primer name	Primer sequence (5'-3')	TM (°C)	Size (bp)
FASN			
FASN-E-FOR	5'- CCAGCATCACCATAGACACG -3'	59°C	317
FASN-E-REV	5'- CTCCTTGGAACCGTCTGTGT -3'		
FASN-I-FOR	5'- ATGCCGAAGGGACCGGCTAT -3'	67°C	93
FASN-I-REV	5'- CATTGAGGATGGTGGCGTAT -3'		
LIPE			
LIPE -E-FOR	5'- CCGAGACGAGATTAGCACCA -3'	66°C	247
LIPE -E-REV	5'- CCTAGCGAACATGACCGAGT -3'		
LIPE -I-FOR	5'- AAGTCTACAGTGTGAGGGCC -3'	70°C	96
LIPE -I-REV	5'- CGATGGGAGCTGAGTAGAGG -3'		
LPL			
LPL -E-FOR	5'- GCCCTGTAACTTCTACCCCA -3'	66°C	350
LPL -E-REV	5'- CCTCTTGTATAGGGCAGCCA -3'		
LPL -I-FOR	5'- CTGCTCCTAGTGGCTCTGAG -3'	67°C	85
LPL -I-REV	5'- CTCCTGAAATTCTGTCGGCG -3'		
SCD			
SCD -E-FOR	5'- TCGCCACCTTTCTTCGTTAC -3'	66°C	266
SCD -E-REV	5'- CTTCCGGTCATAAGCCAGAC -3'		
SCD -I-FOR	5'- CCGGGAGAATATCCTGGTTT -3'	66°C	56
SCD -I-REV	5'- GGTAGTTGTGGAAGCCCTCA -3'		
	Normalising genes		
HPRT1			
HPRT1-E-FOR	5'- GCCCCAGCGTCGTGATTA -3'	64°C	183
HPRT1-E-REV	5'- AGAGGGCTACGATGTGATGG -3'		
HPRT1-I-FOR	5'- CCCAGCGTCGTGATTAGTGA -3'	66°C	88
HPRT1-I-REV	5'- CCTTTTCCAAATCCTCGGCA -3'		
YWHAZ			
YWHAZ-E-FOR	5'- TGGAGCACTTACAAGGCGTA -3'	64°C	168
YWHAZ -E-REV	5'- ACCGTTTCTGCCCTTATCCA -3'		
YWHAZ -I-FOR	5'- AAGGCGTAGTGGAAGTGGAT -3'	66°C	98
YWHAZ -I-REV	5'- GCTGTAGTCAAAGGTGTGCA -3'		
B2M			
B2M-E-FOR	5'- AAACGGGGAGAAGATGAACG -3'	63°C	377
B2M-E-REV	5'- ACATCTACCTGCTCAGACAGT -3'		
B2M-I-FOR	5'- CCTTCTGGTCCACACTGAGT -3'	66°C	99
B2M-I-REV	5'- TCCCACTTAACTATCTTGGGCT -3'		
POLR2A			
POLR2A-E-FOR	5'- CACCCACAGCACCCATCC -3'	61°C	573
POLR2A-E-REV	5'- CCCTCCACATTCTGCTG -3'		
POLR2A-I-FOR	5'- GGGACTCCATTGCTGATTCT -3'	66°C	92
POLR2A-I-REV	5'- GCCTTCTCGATGACCTC -3'		

For all the considered genes in the table are reported the used primer, the annealing temperature expressed in °C and the size in base pairs (bp).

Each primer name is composed by the gene name, “I” for internal primer or “E” for external primer, “FOR” for the left primer and “REV” for the right primer.

[1]TM (°C) Annealing temperature.

Supplementary Table S2. The samples used for the statistical analyses of the enzymes quantitation.

	FASN enzyme		
	LOW IMF Group (IMF <1.50%)	HIGH IMF Group (IMF > 3.50%)	Total
Castrated males	19	26	45
Sows	44	39	83
Total	63	65	128
	LIPE enzyme		
	LOW IMF Group (IMF <1.50%)	HIGH IMF Group (IMF > 3.50%)	Total
Castrated males	18	25	43
Sows	46	43	89
Total	64	68	132
	SCD enzyme		
	LOW IMF Group (IMF <1.50%)	HIGH IMF Group (IMF > 3.50%)	Total
Castrated males	18	23	41
Sows	33	24	57
Total	51	47	98
	LPL enzyme		
	LOW IMF Group (IMF <1.50%)	HIGH IMF Group (IMF > 3.50%)	Total
Castrated males	19	22	41
Sows	22	24	46
Total	41	46	87

Supplementary Table S3. Correlations between gene and protein expressions and the Estimated Breeding Values (EBVs) or carcass traits.

		FASN		SCD		LPL		LIPE	
		Expression[1]	Enzyme (ng/mg Prot)	Expression[1]	Enzyme (ng/mg Prot)	Expression[1]	Enzyme (ng/mg Prot)	Expression[1]	Enzyme (u/mg Prot)
EBVs									
	ADG	-	r=-0.189 *P*=0.032	r=0.295 *P*=0.046	r=-0.243 *P*=0.016	-	-	-	-
	BFT	-	-	-	-	-	-	-	-
	LC	-	-	-	-	-	-	-	-
	HW	-	r=-0.226 *P*=0.010	-	-	-	-	-	r=-0.179 *P*=0.040
	FCR	-	-	-	r=0.262 *P*=0.009	-	-	r=0.247 *P*=0.094	-
Traits									
	IMF	r=0.330 *P*=0.023	-	r=0.313 *P*=0.034	r=0.169 *P*=0.096	r=0.310 *P*=0.034	-	-	r=-0.202 *P*=0.020
	Carcass weight	-	r=-0.255 *P*=0.005	-	-	-	-	-	r=-0.226 *P*=0.011
	%lean	-	-	-	-	-	-	r=-0.318 *P*=0.046	-
	Backfat	-	-	-	-	-	-	r=0.310 *P*=0.052	-
	Loin	-	-	-	-	-	-	r=-0.299 *P*=0.061	-

[1] The gene expressions are reported as relative values obtained as ratios between the molecules/µl of the target gene mRNA and the molecules/µl of the normalising genes mRNA. All the EBVs and traits reported in the above table are explained in Materials and Methods section.

Chapter 7: Metabolomics characterization of colostrum in three sow breeds and its influences on piglets' survival and litter growth rates.

G. Picone[1], M. Zappaterra[1], D. Luise[1], A. Trimigno, F. Capozzi, V. Motta, R. Davoli, P. Bosi, P. Trevisi

Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum - University of Bologna, Viale Fanin 46, Bologna, 40127, Italy;

[1] Contributed equally to the research.

Submitted to

Journal of Animal Science

Summary

Colostrum differs from milk and is the first secretion of mammary gland produced in late pregnancy and during the hours immediately preceding and succeeding parturition. This secretion represents an essential vehicle of passive immunity, prebiotic compounds and growth factors involved in intestinal development. Its composition has been investigated mainly in human and cow, but very little is known about pig colostrum metabolome and how it varies between pig breeds and different farrowing parity. Thus, the aim of the present research is to provide new information about pig colostrum composition and the associations between some metabolites, the breed of the sow and the survival and growth rates of their litters. Colostrum samples were gathered from 58 parturitions of sows belonging to three different breeds chosen for their importance in the Italian heavy pig production: 31 Large White, 15 Landrace and 12 Duroc respectively. Farrowing was not induced, and colostrum samples were collected after the first birth and before the last piglet was born, defatted and centrifuged in a 10kDa cut-off membrane. The eluted was analysed using ^{1}H-NMR spectroscopy. The Principal Components Analysis (PCA) was assessed on the obtained spectra. In addition, using a Stepwise Regression and a Linear Regression analyses the metabolites named after the signals assignment were tested for their associations with piglets' performances. 25 metabolites were identified, comprehending monosaccharides, disaccharides (such as lactose), organic acids (lactate, citrate, acetate and formate), peptides (such as creatine) and others compounds, including nucleotides. PCA results evidence a clustering due to breed and season effects. Lactose was the main compound determining the assignment of the samples into different clusters according to the sow breed. The amount of dimethylamine identified in colostrum was associated with the piglets' mortality at three days after birth (P = 0.004). This is the first study characterizing swine colostrum metabolome using the ^{1}H-NMR spectroscopy technology and the results obtained will contribute to improve the knowledge on colostrum deeper composition and variability. Furthermore, this work may help understanding the compounds that influence piglets' survival and growth in addition to the best- known immunoglobulin colostrum fraction.

Introduction

The pre-weaning litter environment has been proven to affect the pigs development and performances during later life (Vallet et al., 2016) and in particular colostrum intake, coupled with birth weight, was found to influence piglets' growth and mortality (Devillers et al., 2011; Ferrari et al., 2014; Decaluwé et al., 2014b). Colostrum provides new-borns with energy and passive immunity (Noblet et al., 1997; Rooke and Bland, 2002): in particular, most of the literature concerns the effects of the different immunoglobulins on piglets' health and survival capacities (Vallet et al., 2013, 2015; Ogawa et al., 2016). Studies assessed on human and bovine colostrum suggested important roles in new-borns' health also for other bioactive molecules, such as nucleotides, oligosaccharides, organic acids and peptides (Gopal and Gill, 2000; Schlimme et al., 2000; Korhonen, 2013; He et al., 2014), but little is still known about the presence of these metabolites in sows' colostrum and their association with piglets' performances. Furthermore, to date little or no information about pig breed influence on colostrum composition is available and most of the knowledge about metabolites composition of swine colostrum was produced on samples gathered after farrowing induction, fact that may alter colostrum composition (Foisnet et al., 2011). In this study, 58 colostrum samples were collected during a natural parturition with the aims i) to analyse through a NMR-based metabolomics approach the colostrum compounds with a maximum 10 kDa molecular weight in three pig breeds, ii) to evaluate which factors mostly affect the colostrum composition, iii) to test the associations between the identified metabolites, the sow maternal attitude, and the piglets' survival and growth rates.

Materials and Methods

The procedures complied with Italian law pertaining to experimental animals and were approved by the Ethic-Scientific Committee for Experiments on Animals of the University of Bologna, Italy.

Animals and sampling

Fifty-eight colostrum samples were collected from 58 different farrowing of pure breed sows: 12 Duroc (D), 15 Landrace (L) and 31 Large White (LW). All sows were raised in the same farm and were not treated with antibiotics and medical products during gestation. Farrowing was not induced, and the colostrum sampling was carried out during natural parturition, after the birth of the first piglet and before the parturition of the last, across all teats. All samples were immediately frozen at -20 °C and then stored at -80 °C until the samples preparation for NMR analysis.

For each sow, parity and data related with the reproductive performances were recorded: the number of piglets alive and the litter body weight (LBW) were recorded at birth and at day 3, cleansed from the weight of the piglets dead. The litter weight gain (BWG) was then calculated for the period from birth to day 3. Furthermore, the number of weaners per litter was recorded as well as the occurrence of diarrhoea during suckling (1= presence of diarrhoea events from piglets' birth until weaning, 0= absence of diarrhoea event).

Colostrum preparation for ^{1}H-NMR analysis

Colostrum was de-frozen, carefully mixed by inversion, and 15 ml of each colostrum sample were diluted 1:1 with pure water. To each diluted sample, 0.02% of sodium azide was added, to inhibit bacterial growth during the sample preparation. Then the sample was defatted through a centrifugation at 4 °C for 30 minutes at 1500 x g. The aqueous phase was transferred to a clean falcon avoiding the outer layer of fat, and centrifuged again; this procedure was repeated three times. 5 ml of the obtained aqueous phase were then transferred in Amicon Ultra 10 kDa membrane centrifugal filters (Merck Millipore, Merck KGaA, Darmstadt, Germany) and filtered by centrifugation at room temperature for 90 minutes at 5500 x g. This step was needed to eliminate immunoglobulins and other proteins with high molecular weight. The eluted sample was then weighted and lyophilized.

^{1}H-NMR measurements

At the time of sample processing, for each gram of lyophilized sample 250 µl distilled water was added. Eighty µl of thawed sample was centrifuged at 14,000 × g for 5 min (Scilogex D3024 High Speed Micro-Centrifuge, Rocky Hill, CT, USA) and then added to 720 µl of distilled water and 100 µl of a D_2O solution of 3-(trimethylsilyl)-propioniate-2,2,3,3-d_4 (TMSP) (Cambridge Isotope Laboratories Inc, Tewksbury, MA, USA) with a final

concentration of 6.25 mM. ^{1}H-NMR spectra were recorded at 298 K with an AVANCE spectrometer (Bruker BioSpin, Karlsruhe, Germany) operating at a frequency of 600.13 MHz, equipped with an autosampler with 60 holders. The HOD residual signal was suppressed by applying the NOESYGPPR1D sequence (a standard pulse sequence included in the Bruker library) incorporating the first increment of the NOESY pulse sequence and a spoil gradient. The HOD residual signal was suppressed by applying the first increment of the NOESY pulse sequence and a spoil gradient. Each spectrum was acquired using 32 K data points over a 7211.54 Hz spectral width (12 ppm) and adding 256 transients. A recycle delay of 5 s and a 90° pulse of 11.4 μs were set up. Acquisition time (2.27 s) and recycle delay were adjusted to be 5 times longer than the longitudinal relaxation time of the protons under investigation, which has been not longer than 1.4 s. The data were Fourier transformed and phase and baseline corrections were automatically performed using TopSpin version 3.0 (Bruker BioSpin, Karlsruhe, Germany). Signals were assigned through a combination of literature assignments and by the use of a multimedia library included in Chenomx NMR suite 8.2 professional software (Chenomx, Edmonton, Alberta, Canada).

Data analysis

The collected data were aggregated to create homogenous classes. The farrowing parity (from 1 to 3 were classified as 1, parity order > 4 were considered as 2), the parturition season (from 1 to 4). The seasons were assigned as follows: 1= parturition between the 1st of December and the 28th of February; 2= between the 1st of March and the 31st of May; 3= between the 1st of June and the 31st of August; 4= between the 1st of September and the 30th of November. Among the studied animals, 6 sows gave birth during season 1, 19 during season 2, 21 during season 3 and 12 during season 4.

Statistical analyses on spectra data were performed using R computational language (ver. 3.1.2) and MATLAB (ver R2014b, MathWorks Inc.). Each NMR spectrum was processed by means of scripts developed in-house as follows: spectra baseline was adjusted by employing the signals identification algorithm named “baseline.peakDetection” from R (version 3.1.2) package “Baseline” (https://cran.r project.org/web/packages/baseline/index.html). Chemical shift referencing was performed by setting the TMSP signal to 0.00 ppm. The following spectral regions were removed prior to data analysis: the regions including only

noise (the spectrum edges between 11.00 and 8.65 ppm and between 0.15 and -1.00 ppm), the NMR signal which is strongly affected by the residual solvent signals (water, between 4.90 and 4.50 ppm) and the glycerol's signals from 3.82 and 3.76 ppm, from 3.69 and 3.63 ppm and from 3.60 and 3.54 ppm**.** Spectra were then normalized by means of probabilistic quotient normalization method (PQN) (Dieterle *et al.*, 2006) and binned. The first operation is aimed at removing possible dilution effects. The second one avoids the effect of signals misalignments among different spectra due to variations in chemical shift of signals belonging to some titratable acids. The binning operation is performed by subdividing the spectra into 369 bins, each integrating 120 data points (0.0219 ppm each). In order to focus on the real information contained in the spectra, bins that an average higher value than noise were selected. In this way, a total of 201 bins were kept.

The spectra obtained were then analysed through an unsupervised multivariate approach using Principal Component Analysis (PCA). The PCA was conducted on the 201 bins matrix to identify the outlier samples, and test the existence variables contributing to samples clustering. The multivariate models were calculated and the results were visualized on both scores and loadings' plot. This multivariate analysis is the predominant linear dimensionality reduction technique used when dealing with scientific dataset and it is defined as an unsupervised method as it does not use class labels for discriminating between groups. After reducing the dimensionality of the dataset, it produces new linear combinations of the originals variables which can be plotted in a score plot (Bailey et al., 2003). In order to determine the spectral regions encompassing most of the discriminative information, bins with a loading value greater than 1% of the overall standard deviation of all loading values were selected. The identified metabolites included in the significant bins emerged from the loadings' plot and additional metabolites relevant for their biological function were selected and grouped in a new dataset named C-dataset. The C-dataset was used to conduct an analysis of variance (ANOVA) with the aim to confirm if the amounts of the identified compounds were influenced by the effects of breed and farrowing season identified with the PCA and parity order. The model utilized for this analysis was:

$$y = \beta 0 + \beta p * b + \beta p * s + \beta p * o + E$$

Where:

β0 was the intercept;

βp was the corresponding regression coefficient;

y was the amount of each identified metabolite;

b was the sow breed (LW; D; LA);

s was the farrowing season (1; 2; 3; 4);

o was the parity order (1; 2);

E was the error.

This first part was conducted to test if sows breed influences colostrum profile, and if in addition to breed there are other "environmental" variables affecting colostrum quality (in this case the farrowing season and the parity order were tested).
Then, a stepwise regression analysis was used to select, among the metabolites included in the C-dataset and sows' reproductive performances, the variables that had to be included with the breed, the farrowing season and the parity order in the final GLM model for the identification of the metabolites related to piglets' performances. This statistical analysis involves starting with no variables in the model and adding gradually each metabolite and sow reproductive parameter (the litter weight and the number of alive piglets at birth) to evaluate which one of the colostrum identified compounds and sows' reproductive abilities most influenced the piglets' survival and growth. The results obtained from the stepwise regression analysis were then confirmed through General Regression Analysis (GLM), considering as *y* variables the BWG, the number of weaned piglets, the number of dead piglets from birth until day 3 or the number of piglets dead from day 3 to weaning; as independent variables were considered the sows breed, the farrowing season, the parity order of the sow and the significant factors identified through the stepwise regression analysis. The utilized GLM model was:

$$y = \beta 0 + \beta p * b + \beta p * s + \beta p * o + \beta p * p + \beta p * q + \beta p * r + \ldots + E$$

Where:

y was BWG, the number of weaned piglets, the number of piglets dead from the birth day until day 3 or the number of dead from day 3 to weaning;

β0 was the intercept;

βp was the corresponding regression coefficient;

b was the sow breed (LW; D; LA);

s was the farrowing season (1; 2; 3; 4);

o was the parity order (1; 2);

p; *q*; *r* were the significant metabolites identified through the stepwise regression analysis;

E was the error.

Finally, all the variables that did not show an effect on the dependent variables were removed from the model and only the significant effects were maintained.
The *prcomp* function of R environment was used to perform the PCA analysis on bins matrix (R Core Team, 2015). The ANOVA analysis, the stepwise regression analysis and the regression model were carried out on SAS software using PROC REG and PROC GLM respectively (SAS® 9.4, SAS Inst. Inc., Cary, NC).

Results

Dataset description

In Table 1 the complete dataset is detailed Duroc sows had on average a lower number of piglets at birth (8.92 ± 2.28) respect to Landrace (12.60 ± 1.72) and Large White (11.90 ± 2.26), while the newborns of Landrace and Large White breeds presented a lower weight at birth (1.38 ± 0.15 kg and 1.43 ± 0.16 kg, respectively) compared to Duroc piglets (on average 1.59 ± 0.23 kg).

Colostrum spectra

In Figure 1 a NMR molecular profile of sow colostrum is represented. The ^{1}H spectrum is mainly dominated by the carbohydrate signals overlapping in the midfield region between

3.49 and 4.49 ppm (Figure 1B). Those belong to lactose and nucleotides sugars such as UDP-glucose and UDP-galactose and nucleotide as UMP. In this area, also signals from creatine and its products arise (3.04-3.05 ppm). Amino acids mainly fall in the upfield region, between 0.99 and 3.49 ppm, together with signals from organic acids (Figure 1A). In this part of the spectrum fall also signals from threonine (1.33 ppm) and alanine (1.49 ppm), lactic acid (1.33 ppm), acetic acid (1.92 ppm), succinic acid (2.41 ppm) and citric acid (2.54 and 2.67 ppm). Finally, in the downfield region (Figure 1C) signals of different phenolic compounds can be observed, but in this case, only formic acid was assigned (8.4 ppm), together with signals from the nucleotide sugars UDP-glucose and UDP-galactose (5.5-6 ppm, 7.9-8 ppm) and UMP (8.1 ppm, 5.98-5.99 ppm, 4.42 ppm) as listed in table 2. The 25 compounds have been identified through a combination of literature assignments (Wu et al., 2016) and by the use of a multimedia library included in Chenomx NMR suite 8.2 professional software (Chenomx, Edmonton, Alberta, Canada).

Factors affecting colostrum composition

After alignment, normalization and binning, the dataset contained a total of 58 colostrum spectra characterized by 201 bins and PCA was applied on it to investigate differences on the metabolome between groups. For sow's parity order, in the total colostrum spectra no PCA clustering was identified (data not showed). While Figure 2A and B shows that samples clustered on PC1-PC2 due to the effects of the sow breeds (Figure 2A) and on PC2-PC3 due to the farrowing seasons (Figure 2B). The PC1 explained the 81% of the total variance and separated the colostrum spectra of D and LW, while PC2 (10% of the variance) discriminated the L colostrum composition from the ones of LW and D sows. The PC2-PC3 plot highlighted the season effect, in particular along PC2 the differences in the colostrum spectra due to seasons 1 and 4 (winter-autumn) against season 2 and 3 (spring-summer). The weighting of each variable (bin) is represented by the loadings plot in Figure 2C and 2D in which are displayed the loadings from PC1 and PC2 respectively as a bar plot, where each bar corresponds to a single spectral variable (bin). The main bins accounting for the spectral differentiation and their relative chemical shift were listed in the Supplementary Table S1 (SS1). As emerging from SS1 table, most of the signals included in these discriminant bins

were assigned to the corresponding metabolites. The C-dataset, which was used for the following statistical analyses, resulted to be composed of 25 metabolites, listed in Table 2. The parity, breed and season effects on colostrum composition were then confirmed through the ANOVA analysis on the identified metabolites described in the C-dataset, and the results are reported in Table 3. For sow's parity order, not significant data are obtained. Parity order showed only trends nearing statistical significance for succinate ($P = 0.097$), creatine phosphate ($P = 0.091$), creatinine ($P = 0.061$) and UDP-glucose ($P = 0.061$) (data not show). Breed and season resulted to be the major factors affecting the assigned compounds. Indeed, the amounts of alanine ($P = 0.004$; $P = 0.004$), citrate ($P < 0.0001$; $P = 0.006$), succinate ($P < 0.0001$; $P = 0.024$), dimethylamine ($P = 0.030$; $P = 0.0001$), creatine ($P < 0.0005$; $P < 0.0005$), creatine phosphate ($P = 0.003$; $P < 0.0001$), cis-aconitate ($P = 0.030$; $P < 0.0001$), taurine ($P = 0.002$; $P = 0.001$), glycolate ($P < 0.0001$; $P = 0.001$) and UMP ($P = 0.001$; $P = 0.009$) were affected by both breed and the season of the farrowing respectively. Season affects significantly ($P < 0.0001$) the amount of acetate, creatinine and formate, where a higher level was registered for all the metabolites during the cold months. On the other hand, the amounts of o-acetylcholine ($P < 0.0001$), sn-glycerophosphocholine ($P = 0.036$), UDP-n-acetylglucosamine ($P = 0.001$), lactose ($P < 0.0001$), myo-inositol ($P = 0.001$) and UDP-glucose ($P < 0.0001$) were affected only by sow breed. In particular, the colostrum of L samples showed upper signals for UDP-Glucosio, UDP-galactosio and sn-glycerophosphocholine compared to the other two breeds, while LW colostrum was characterized by a major quantity of lactose, taurine, myo-inositol and glycolate.

Factors affecting litter performances

The stepwise regression analysis revealed that, in addition to the influence of sows' reproductive performances (the litter weight and the number of alive piglets at birth) some peculiar metabolites can contribute to piglets' survival and growth parameters (Table 4). In particular, the litter weight at birth and the concentration of acetate significantly entered in the model for BWG ($P < 0.0001$ and $P = 0.003$, respectively); the higher number of alive piglets at birth and the increased concentration of colostrum cis-aconitate resulted to be important variables affecting positively the number of weaned piglets ($P < 0.0001$ and $P =$

0.019, respectively), while dimethylamine (P = 0.0002) and taurine (P = 0.013) entered as variables in the model for the number of dead piglets per litter at day 3. There was no influence of farrowing season and parity order on BWG, the number of weaned pigs or the number of dead piglets at day 3.

The outcomes of the stepwise regression analysis were then tested with the GLM, and the results reported in Table 5. Both the higher average piglets' weight at birth (P < 0.0001) and the colostrum acetate concentration (P = 0.003) affected positively BWG (Table 5). The number of dead piglets at day 3 was mainly influenced by the concentration of dimethylamine (P = 0.001) and taurine (P = 0.027) in colostrum and partially by the litter size at birth (P < 0.1). In addition, the litter size at birth (P = 0.001) and a lower level of cis-aconitate in colostrum (P = 0.010) and the sow breed showed a trend (P = 0.021) were significantly associated with the number of weaned piglets.

Discussion

This is the first study based on ^{1}H-NMR metabolomics approach describing in three pig breeds the colostrum metabolome profile, the factors underlying its composition and the associations between colostrum metabolites and litter's fitness during suckling.

The three breeds showed different reproductive abilities in accordance with literature (Blasco et al., 1995; Sonderman and Luebbe, 2008), with L and LW sows exhibiting a higher average number of piglets alive at birth compared to D sows. These differences between breeds are also visible at the colostrum composition level (Simmen et al., 1990): considering the whole spectrum, the three breeds display clustering tendency, with the colostrum lactose amount explaining most of the colostrum composition differences between breeds. In particular, L and LW breed samples presented higher values of lactose. Lactose rate in cow milk is commonly associated with the health status of mammary gland, as higher lactose concentrations are positively correlated to healthier mammary glands and low amounts of this compound indicate the existence of intrammamary infections (Park et al., 2007). Considering the data available for the present work, it is not possible to support the same consideration in lactating sows, due to the absence of reference value for sow milk.

Furthermore, differences on UDP-n-acetylglucosamine. UDP-glucose and UDP-galactose were observed between breeds. UDP-sugars are intermediate products in cellular protein glycosylation and in the synthesis of lactose and other sugars, and are known to have also autocrine/paracrine signalling functions (Lazarowski and Harden, 2015). UDP-sugars are detected by the purinergic receptor P2Y14, a G Protein Coupled Receptor, found in many epithelia and in immune and inflammatory cells (Lazarowski and Harden, 2015). In weaned pigs, P2Y14 is expressed in the gastro-intestinal tract, particularly in the pyloric area (Colombo et al., 2014). We are not aware of the entity of the gastrointestinal expression of this receptor in piglets included in the present study; however it can be hypothesized that the presence of UDP-sugars may be involved in the activation/maturation of the neonate pig gastrointestinal immunity.

The breed effect on colostrum composition is not exclusively confined to sugars profile, which for sure are the most abundant metabolites between the identified, but also to other compounds, including alanine, citrate, succinate, creatine phosphate, creatinine, dimethylamine, cis-aconitate, myo-inositol and o-acetilcholine. Furthermore, the obtained colostrum spectra were affected also by the farrowing season: the samples gathered during winter and autumn exhibited differences in colostrum compositions respect to colostrum secreted during spring and summer. These differences could be ascribed to the environmental conditions affecting sows' performances: compounds such as acetate, which showed to be more abundant during cold seasons, may reflect the nutritional state of sows during cold months. Acetate in sows, typically fed high fibre diets in gestation, is the main product of hindgut fermentations; acetate is also a precursor for mammary synthesis of fat milk. Thus the season effect may reflect a reduced presence of fibre in diet during the warmer seasons or a different mammary usage of acetate in favour of the more efficient use of glucose (Linzel et al., 1969). Indeed, during cold seasons, the higher feed intake makes available larger amounts of energy and nourishing compounds respect to the lower daily feed intake characterizing sows living at higher environmental temperatures (Gourdine et al., 2006). In addition, farrowing season affected also the creatine pathway: in particular, creatine and creatine-phosphate amounts during the period ranging from September to February were significantly lower than in spring and summer; on the contrary creatinine was higher during the same period. Creatine is an important nutrient for the newborn (Brosnan

and Brosnan, 2007), thus variations in creatine content of colostrum may have nutritional relevance. In mice, it has been shown that milk creatine is extracted from the circulating plasma by the mammary gland, which conversely has little or no capacity to synthesize creatine (Lamarre et al., 2010). No research data is available for sow colostrum, but it can be assumed that also in this case variations in colostrum may reflect variations in blood creatine concentration. Here the variations in the ratio creatine and creatine-phosphate to creatinine may have resulted from a higher degradation of the first two compounds into creatinine, but there is no evidence that this reaction occurs in the mammary gland. The increasing amount of creatinine level is in general associated with a higher mobilization of stored proteins and indirectly with fat and lean levels in the body mass (Van Niekerk et al., 1963). A recent study of Decaluwé et al. (2014a) associated an increased amount of blood creatinine on 1st day of lactation with lower feeding levels in sows during late gestation period. However, we could not control feed intake in the days before farrowing and do not know if it changed with season. Thus further research is necessary to explain variations of creatine and related compounds in colostrum.

Some of the identified compounds were associated with litter weight gain during the first three days of life and to piglets' survival rates at day 3 and at weaning. In particular, we suppose that the positive effect of acetate on BWG was linked to the role of this compound in *de novo* synthesis of lipids and glucose (den Besten et al., 2013) and adipogenesis stimulation (Hong et al., 2005). Additionally, taurine colostrum concentration showed a positive correlation with piglets' survival rate at three days of life. Taurine was already proven to play a critical role in neonatal development, including the development of central nervous system and other tissues (Bryson et al., 2001; Aerts and Van Assche, 2002). Due to the essential role of this compound in neonatal period, it can be easily understood why a higher concentration of taurine in sows' colostrum exerted positive effects on piglets' survival during first days of life, independently from sows' breed. As regards the number of dead piglets at three days of life, this performance was associated with the concentration of dimethylamine secreted in colostrum. Dimethylamine is a biogenic amine, synthesised by bacterial action, known for its mutagenic, irritative and barrier-disrupting properties (Galli et al., 1993; Fluhr et al., 2005). The irritative effect of this compound can explain the significant negative association found in the present research between the colostrum dimethylamine

amount and the piglets' survival capacity. The observed increase in dimethylamine secretion could be a direct effect of the higher occurrence of bacteria fermentations in animal feeds (Juszkiewicz et al., 1980), then ingested by sows and secreted in colostrum and milk.

Similarly to dimethylamine, also cis-aconitate was negatively associated to piglets' survival capacity from birth to weaning. As cis-aconitate is an intermediate compound synthesized by several enteric bacteria, comprising *Salmonella* enterica (Lewis and Escalante-Semerena, 2006) and *Escherichia coli* (Shimizu, 2013), the increased amount of this tricarboxylic acid in colostrum may be a marker of the presence of pathogenic bacteria strains infecting maternal gut and mammary gland. Maternal microbiome is proved to affect newborn digestive tract, and pathogenic strains can easily pass from sow's gut to piglets' enteric tract. Thus, the increase of cis-aconitate in colostrum secretion might be the first sign of gut dysbiosis, which may then affect also piglets' microbiome and digestive tract homeostasis.

In conclusion, this study demonstrates that colostrum metabolome is greatly affected by breed and, in particular, Duroc sows showed colostrum compositions unlike any other. This result agrees with the generally accepted view that the differences among Duroc and white coated pig breeds may originate from distinct genetic origins, and consequently suggests that further genetic studies may help explaining the variations found among breeds in colostrum compositions. From the observation of the results obtained it can be suggested that the different temperatures occurring during the year affect sows' metabolism and, in turn, can also affect colostrum composition. Among the identified metabolites, acetate and taurine showed their positive effects on piglets' performances and survival rate, while dimethylamine and cis-aconitate exerted a negative influence on new-borns capacity to survive. This research represents a preliminary step towards the knowledge of pig colostrum composition and it is one of the first studies focusing on the associations between different swine colostrum compositions and litter performances using the ^{1}H-NMR technique. Further investigations are needed to extend the identification of the different compounds in swine colostrum and to elucidate their effects on new-borns. Furthermore, the possible interaction between sows' feeding and microbiota in the modulation of colostrum metabolome deserves further investigations.

References

Aerts L & Van Assche FA (2002) Taurine and taurine-deficiency in the perinatal period. *J Perinat Med* 30:281–286.

den Besten G, Lange K, Havinga R, van Dijk TH, Gerding A, van Eunen K, Müller M, Groen AK, Hooiveld GJ, Bakker BM & Reijngoud DJ (2013) Gut-derived short-chain fatty acids are vividly assimilated into host carbohydrates and lipids. Am. J. Physiol. *Gastrointest Liver Physiol* 305:G900–910.

Bailey NJ, Oven M, Holmes E, Nicholson JK & Zenk MH (2003) Metabolomic analysis of the consequences of cadmium exposure in Silene cucubalus cell cultures via 1 H NMR spectroscopy and chemometrics. *Phytochemistry* 62:851-858.

Blasco A, Bindanel JP & Haley CS (1995) Genetic and neonaltal survival. In *The neonatal pig: Development and survival*. [Varley MA editor], pp. 17-37, CAB International,Wallingford, UK.

Brosnan JT & Brosnan ME (2007) Creatine: endogenous metabolite, dietary, and therapeutic supplement. *Annu Rev Nutr* 27:241–61.

Bryson JM, Jackson SC, Wang H & Hurley WL (2001) Cellular uptake of taurine by lactating porcine mammary tissue. *Comp Biochem Physiol - B Biochem Mol Biol* 128:667–673.

Colombo M, Priori D, Trevisi P & Bosi P (2014) Differential Gene Expression in the Oxyntic and Pyloric Mucosa of the Young Pig. *PLOS ONE* 9:e111447.

Decaluwé R, Maes D, Cools A, Wuyts B, De Smet S, Marescau B, De Deyn PP & Janssens GPJ (2014a) Effect of peripartal feeding strategy on colostrum yield and composition in sows. *J Anim Sci* 92:3557–3567.

Decaluwé R, Maes D, Wuyts B, Cools A, Piepers S & Janssens GPJ (2014b) Piglets' colostrum intake associates with daily weight gain and survival until weaning. *Livest Sci* 162:185–192.

Devillers N, Le Dividich J & Prunier A (2011) Influence of colostrum intake on piglet survival and immunity. *Animal* 5:1605–12.

Dieterle F, Ross A, Schlotterbeck G & Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal Chem* 78:4281-4290.

Ferrari CV, Sbardella PE, Bernardi ML, Coutinho ML, Vaz IS, Wentz I & Bortolozzo FP (2014) Effect of birth weight and colostrum intake on mortality and performance of piglets after cross-fostering in sows of different parities. *Prev Vet Med* 114:259–266.

Fluhr JW, Kelterer D, Fuchs S, Kaatz M, Grieshaber R, Kleesz P & Elsner P (2005) Additive impairment of the barrier function and irritation by biogenic amines and sodium lauryl sulphate: A controlled in vivo tandem irritation study. *Skin Pharmacol Physiol* 18:88–97.

Foisnet A, Farmer C, David C & Quesnel H (2011) Farrowing induction induces transient alterations in prolactin concentrations and colostrum composition in primiparous sows. *J Anim Sci* 89:3048–3059.

Galli A, Paolini M, Lattanzi GG, Cantelli-Forti G & Bronzetti G (1993) Genotoxic and biochemical effects of dimethylamine. *Mutagenesis* 8:175-8.

Gopal PK & Gill HS (2000) Oligosaccharides and glycoconjugates in bovine milk and colostrum. *Br J Nutr* 84:S69–74

Gourdine JL, Bidanel JP, Noblet J & Renaudeau D (2006) Effects of season and breed on the feeding behavior of multiparous lactating sows in a tropical humid climate. *J Anim Sci* 84:469–480.

He Y, Liu S, Leone S & Newburg DS (2014) Human colostrum oligosaccharides modulate major immunologic pathways of immature human intestine. *Mucosal Immunol* 7:1–14.

Hong YH, Nishimura Y, Hishikawa D, Tsuzuki H, Miyahara H, Gotoh C, Choi KC, Feng DD, Chen C, Lee HG, Katoh K, Roh SG & Sasaki S (2005) Acetate and propionate short chain fatty acids stimulate adipogenesis via GPCR43. *Endocrinology* 146:5092–5099.

Juszkiewicz T, Kowalski B & Pastucha E (1980) Occurrence and formation of nitrosamines in animal feeds. *IARC Sci Publ* 31:589-94.

Korhonen HJ (2013) Production and properties of health-promoting proteins and peptides

from bovine colostrum and milk. *Cell Mol Biol* 59:12–24.

Lamarre SG, Edison EE, Wijekoon EP, Brosnan ME & Brosnan JT (2010) Suckling rat pups accumulate creatine primarily via de novo synthesis rather than from dam milk. *J Nutr* 140:1570-1573.

Lazarowski ER & Harden TK (2015) UDP-Sugars as Extracellular Signaling Molecules : Cellular and Physiologic Consequences of P2Y 14 Receptor Activation. *Mol Pharmacol* 88:151-60.

Lewis JA & Escalante-Semerena JC (2006) The FAD-dependent tricarballylate dehydrogenase (TcuA) enzyme of Salmonella enterica converts tricarballylate into cis-aconitate. *J Bacteriol* 188:5479–5486.

Van Niekerk BD, Reid JT, Bensadoun A & Paladines OL (1963) Urinary creatine as an index of body composition. *J Nutr* 79:463–473.

Noblet J, Dourmad JY, Etienne M & Le Dividich J (1997) Energy Metabolism in Pregnant Sows and Newborn Pigs. *J Anim Sci* 75:2708–2714.

Ogawa S, Tsukahara T, Imaoka T, Nakanishi N, Ushida K & Inoue R (2016) The effect of colostrum ingestion during the first 24 hours of life on early postnatal development of piglet immune systems. *Anim Sci J* 87:1511-1515.

Park YK, Koo HC, Kim SH, Hwang SY, Jung WK, Kim JM, Shin S, Kim RT & Park YH (2007) The analysis of milk components and pathogenic bacteria isolated from bovine raw milk in Korea. *J Dairy Sci* 90:5405–5414.

Rooke JA & Bland IM (2002) The acquisition of passive immunity in the new-born piglet. *Livest Prod Sci* 78:13–23.

Schlimme E, Martin D & Meisel H (2000) Nucleosides and nucleotides: natural bioactive substances in milk and colostrum. *Br J Nutr* 84:S59–68.

Shimizu K (2013) Metabolic Regulation of a Bacterial Cell System with Emphasis on Escherichia coli Metabolism. *ISRN Biochem* 2013:1–47.

Simmen FA, Whang KY, Simmen RCM, Peterson GA, Bishop MD & Irvin KM (1990) Lactational variation and relationship to postnatal growth of insulin-like growth factor-I in

mammary secretions from genetically diverse sows. *Domest Anim Endocrinol* 7:199–206.

Sonderman JP & Luebbe JJ (2008) Semen production and fertility issues related to differences in genetic lines of boars. *Theriogenology* 70:1380–1383.

Vallet JL, Calderon-Diaz JA, Stalder KJ, Phillips C, Cushman RA, Miles JR, Rempel LA, Rohrer GA, Lents CA, Freking BA & Nonneman DJ (2016) Litter of origin trait effects on gilt development. *J Anim Sci* 94:96–105.

Vallet JL, Miles JR, Rempel LA, Nonneman DJ & Lents CA (2015) Relationships between day one piglet serum immunoglobulin immunocrit and subsequent growth, puberty attainment, litter size, and lactation performance. *J Anim Sci* 93:2722–2729.

Vallet JL, Miles JR & Rempel LA (2013) A simple novel measure of passive transfer of maternal immunoglobulin is predictive of preweaning mortality in piglets. *Vet J* 195:91–97.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Wu J, Domellöf M, Zivkovic AM, Larsson G, Öhman A & Nording ML (2016) NMR-based metabolite profiling of human milk: A pilot study of methods for investigating compositional changes during lactation. *Biochem Biophys Res Commun* 469:626-632.

Table 1. Mean and standard deviation of the measured parameters for each breed and for the total population.

	[1] D	[2] L	[3] LW	Total
Number of sows	12	15	31	58
Order of parturition	2.750	4.067	4.161	3.845
Number of piglets born alive per litter				
Mean	8.917	12.600	11.903	11.466
SD	2.275	1.724	2.256	2.494
[4]Average LBW at birth, kg				
Mean	14.033	17.207	16.890	16.381
SD	3.679	2.081	2.828	3.060
Average piglet's weight at birth, kg				
Mean	1.588	1.376	1.432	1.450
SD	0.226	0.153	0.155	0.184
Number of alive piglets per litter at 3 days				
Mean	8.250	12.133	10.871	10.655
SD	1.865	1.767	2.291	2.453
Number of dead piglets per litter at 3 days				
Mean	1.600	1.750	1.583	1.619
SD	1.342	0.500	0.996	0.974
[4]Average LBW at 3 days, kg				
Mean	16.400	21.167	19.758	19.428
SD	3.993	3.059	4.264	4.211
Average piglet's weight at 3 days, kg				
Mean	2.001	1.761	1.826	1.846
SD	0.275	0.240	0.213	0.244
[5]BWG, kg				
Mean	3.085	4.453	4.103	3.983
SD	0.864	1.461	1.771	1.598
Number of weaned piglets				
Mean	7.333	11.133	10.032	9.759

SD	1.875	1.552	2.316	2.423
Incidence of diarrhoea				
Mean	3	1	4	8

[1] D stands for Duroc.

[2] L stands for Landrace.

[3] LW stands for Large White.

[4] LBW stands for litter Body Weight.

[5] BWG stands for the average litter Body Weight Gain from birth to day 3.

Table 2. Assignment table of the identified metabolites present in the ^{1}H-NMR spectra of colostrum at pH 7.420 and considered in the C-dataset. Chemical shift values are referenced to TMSP proton signals at 0.00 ppm. Chemical shift values are referenced to TMSP proton signals at 0.00 ppm. Glycerol (3.568, 3.661 and 3.793 ppm has not been listed as it has not been included in the PC analysis)

Assigned number	^{1}H chemical shift (ppm) [a]	Compound
1	1.332 (d)	Lactate
2	1.486 (d)	Alanine*
3	1.923 (s)	Acetate
4	2.028 (s)	N-Acetylglutamate*
5	2.063 (s)	N-Acetylglucosamine*
6	2.089 (s) - 5.552(dd) - 5.967 (d) - 7.944 (d) - 8.287 (d)	UDP-N-Acetylglucosamine
7	2.147 (s) -3.222 (s)	O-Acetylcholine
8	2.408 (s)	Succinate*
9	2.539 (d) - 2.667 (d)	Citrate
10	2.720 (s)	Dimethylamine
11	3.039 (s)	Creatine
12	3.046 (s)	Creatine phosphate
13	3.050 (s)	Creatinine
14	3.119 (d) - 5.712 (m)	cis-Aconitate
15	3.204 (s)	Choline
16	3.231(s) - 4.330 (m)	sn-Glycerophosphocholine
17	3.272(t) - 3.532 (dd) - 4.073 (t)	Myo-Inositol
18	3.259 (t) - 3.428 (t)	Taurine
19	3.302 (t) -3.684:3.906 (m), 3.980 (d) 4.461 (d) - 4.679 (d) - 5.243 (d)	Lactose
20	3.480 (s) - 4.142:4.278 (m) -5.607 (dd) -5.967 (m) -7.940 (d)	UDP-glucose
21	3.935 (s)	Glycolate
22	4.142:4.278 (m) - 4.379 (m) - 5.664 (dd)- 5.990 (m) - 7.942 - 7.995(d)	UDP-Galactose

23	5.917 (d) - 7.879 (d)	Uridine
24	8.406 (s)	Formate
25	4.423 (t) - 5.990 (m) - 8.102 (d)	UMP

[a] d, doublet; dd, doublet of doublets; m, multiplet; s, singlet; t, tripleTable 3. Effects of sow breed and season on identified colostrum metabolites.

* These compounds were included in the C- dataset for their specific biological role.

Table 3. Effects of sow breed and season on identified colostrum metabolites

Metabolite	Breed[1]			SEM	*P*-value	Season[2]				SEM	*P*-value
	D	L	LW			1	2	3	4		
Acetate	9.57	11.17	9.90	0.91	0.673	13.59	7.55	5.95	13.77	0.89	<0.0001
Lactate	5.38	6.70	8.88	1.85	0.456	4.93	10.18	8.71	4.13	1.81	0.137
Alanine	1.77	2.20	2.44	0.17	0.004	1.65	2.50	2.51	1.88	0.17	0.004
Citrate	209	301	257	13	<0.0001	246	286	265	228	12	0.006
Succinate[3]	2.21	3.26	3.50	0.19	<0.0001	2.59	3.26	3.55	2.55	0.21	0.024
Dimethylamine	2.86	4.44	4.51	0.49	0.029	2.30	5.19	5.39	2.85	0.48	0.0001
Creatine	39.8	59.9	58.7	3.4	<0.0005	40.9	59.0	64.7	46.6	3.3	<0.0005
Creatine Phosphate	3.40	7.99	8.01	0.88	0.003	1.90	8.50	10.78	4.69	0.86	<0.0001
Creatinine	13.6	16.7	16.5	1.1	0.400	19.9	11.8	11.6	19.2	1.1	<0.0001
Cis-Aconitate	1.41	1.93	1.74	0.14	0.030	1.10	2.14	2.28	1.26	0.14	<0.0001
O-Acetylcholine	77.1	196.9	156.6	13.9	<0.0001	101.7	148.4	171.7	152.3	13.6	0.093
sn-Glycerophosphocholine	446	543	414	38	0.018	430	457	507	477	37	0.816
Choline	7.82	10.62	9.91	1.19	0.455	8.86	11.64	9.43	7.87	1.16	0.340
N-Acetilglutamate	6.35	9.90	14.03	3.77	0.134	7.94	9.69	8.53	14.21	3.69	0.683
N-Acetilglucosamine	10.9	15.4	11.7	1.6	0.072	10.6	14.0	13.8	12.3	1.6	0.621
UDP-N-Acetilglucosamine	22.4	34.4	33.9	2.2	0.001	26.2	33.4	33.8	27.4	2.1	0.072
Lactose	458	579	811	30	<0.0001	535	644	667	619	30	0.168
Taurine	1.57	4.10	6.05	0.87	0.002	0.98	6.23	6.19	2.21	0.85	0.001

Myo-Inositol	63.53	76.95	82.86	3.76	0.001	62.58	73.17	81.13	80.91	3.68	0.070
UDP-Glucose	6.07	9.78	6.23	0.54	<0.0001	6.08	7.30	8.09	7.97	0.54	0.280
Glycolate	28.1	39.8	45.8	2.4	<0.0001	28.4	41.8	45.9	35.6	2.3	0.001
Uridine	3.14	3.72	3.23	0.38	0.551	3.05	3.55	3.49	3.37	0.38	0.936
UDP-Galactose	32.6	74.0	42.5	4.0	<0.0001	39.3	54.2	61.4	44.0	3.9	0.006
Formate	4.49	4.43	4.03	0.35	0.380	6.27	3.04	2.29	5.67	0.34	<0.0001
UMP	13.30	24.20	21.60	1.70	0.001	21.82	16.04	16.84	24.08	1.91	0.009

Mean of the identified metabolites are expressed as absolute area

[1] The Breed is assigned as D for Duroc, L for Landrace and LW for Large White.

[2] The seasons were assigned as follows: 1= if the parturition was included in the period between the 1st of December and the 28th of February; 2= between the 1st of March and the 31st of May; 3= between the 1st of June and the 31st of August; 4= between the 1st of September and the 30th of November.

[3]Succinate is the only metabolite showing an effect of the parity order (*P* value = 0.039), in particular with a parity order less than 4 succinate has a mean of 2.77, while the same metabolite with a parity order more than or equal to 4 has a mean of 3.20, a SEM of 0.15.

Table 4. Results of the stepwise regression analysis.

Variables	Coefficient	SE coefficient	F value	*P*-value
Model for BWG[1] (R^2 = 0.4286; C(p) = 0.8735)				
Intercept	0.715	0.533	1.79	0.186
Acetate	0.10433	0.03348	9.71	0.0029
Average piglet's weight at birth	0.00846	0.00164	26.63	<.0001
Model for the number of weaned piglets (R^2 = 0.4343; C(p) = 2.0849)				
Intercept	4.801	1.385	12.01	0.001
Cis- Aconitate	-0.90181	0.37157	5.89	0.0185
The number of alive piglets at birth	0.58395	0.09863	35.05	<.0001
Model for the number of piglets dead per litter at day 3 (R^2 = 0.2304; C(p) = 29.1881)				
Intercept	-0.333	0.27	1.44	0.2352
Dimethylamine	0.33082	0.08318	15.82	0.0002
Taurine	-0.11423	0.04432	6.64	0.0127

[1]BWG stands for the litter Body Weight Gain from birth to day 3.

Table 5. Results of the GLM analysis with the significant variables affecting litter body weight gain at day 3, the number of dead piglets at day 3 and the number of weaned pigs.

Variables	Coefficient	SE	*P*-value
GLM for BWG[1]			
Acetate[2]	0.104	0.033	0.003
Average piglet's weight at birth	0.008	0.002	<.0001
GLM for Number of dead piglets at day 3			
Intercept	-1.219	0.557	0.033
Dimethylamine[2]	0.296	0.084	0.001
Taurine[2]	-0.100	0.044	0.027
Number of alive piglets at birth	0.084	0.046	0.074
GLM for Number of weaned pigs			
Breed:		0.46	0.021
LW[3]	9.92		
L[4]	10.59		
D[5]	8.31		
Cis-Aconitate[2]	-0.952	0.354	0.010
Number of alive piglets at birth	0.413	0.111	0.001

[1]BWG stands for litter Body Weight Gain from birth to day 3 of life.

[2]Metabolites concentrations were considered in area arbitrary unit.

[3]LW stands for Large White.

[4]L stands for Landrace.

[5]D stands for Duroc.

Figures

Figure 1. Typical ^{1}H-NMR spectrum of aqueous extract of colostrum. ^{1}H-NMR spectrum registered on a colostrum sample. The spectrum has been split into three parts for the sake of clarity. Some resonances have been assigned by using Chenomx software and listed in table 2: A) Aliphatic or upfield region; B) Carbohydrate or midfield region, characterized by the presence of signals belonging to sugars and glycerol and C) Aromatic region or downfield region.

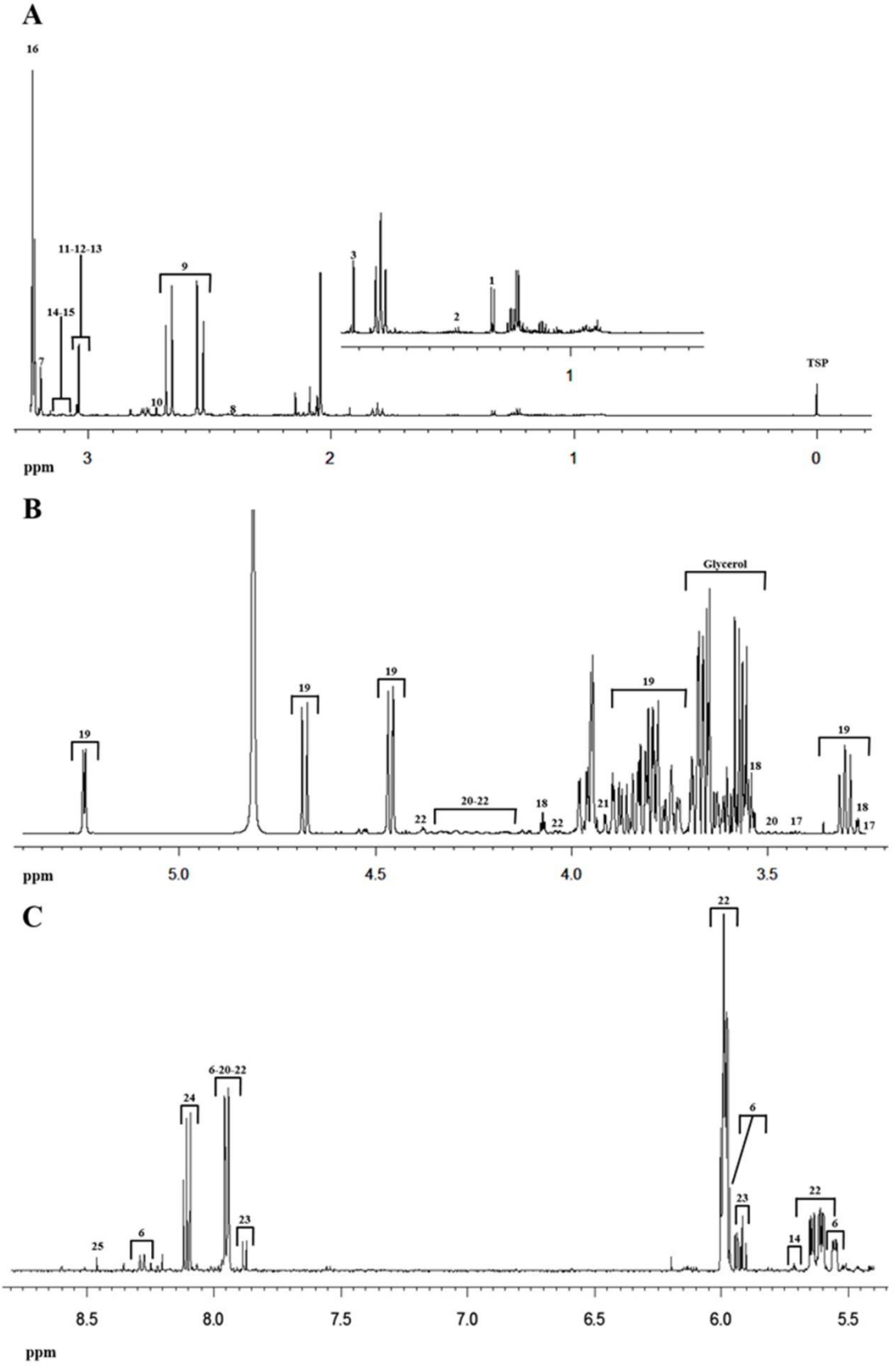
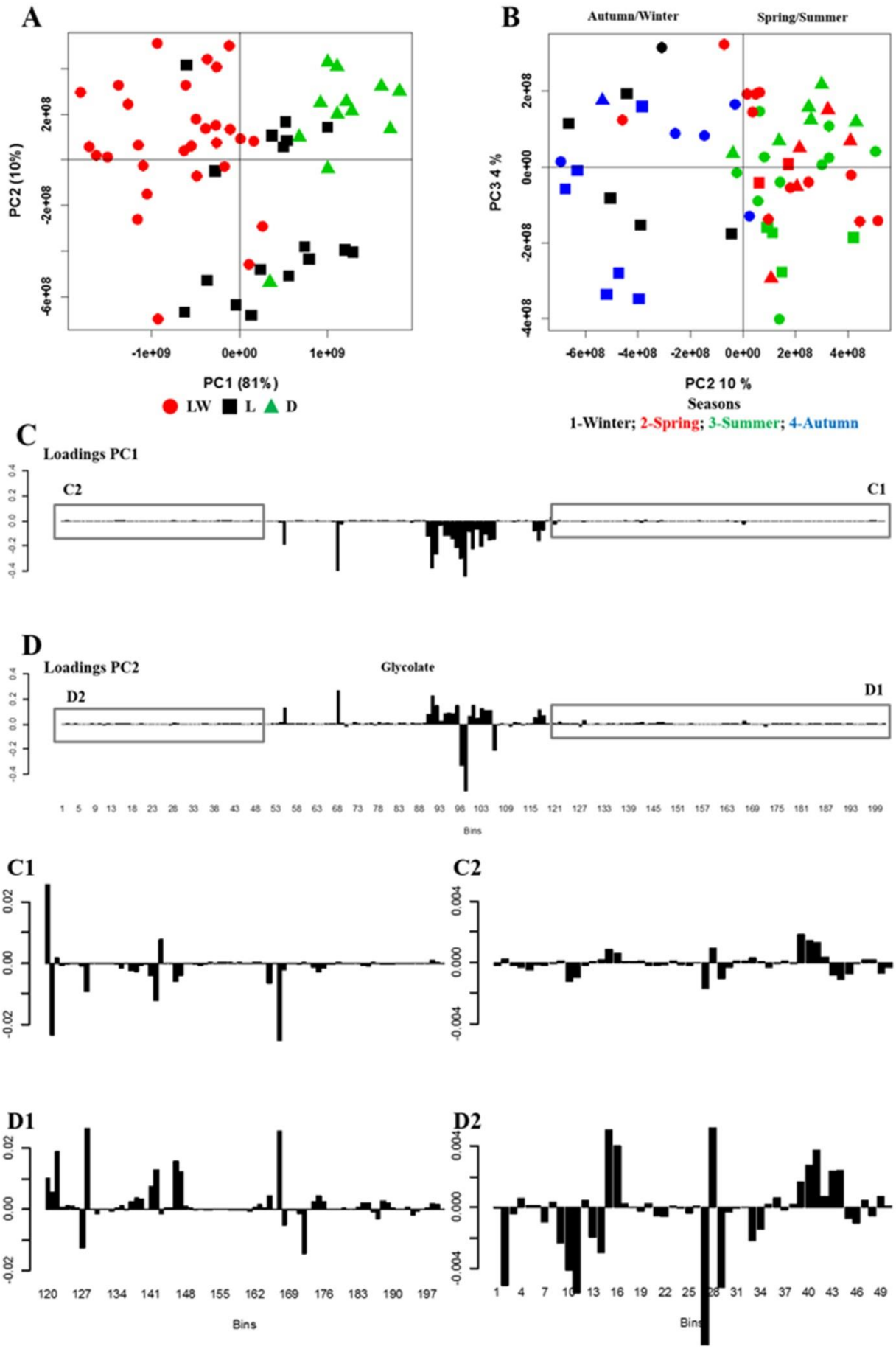

A
16
11-12-13
9
14-15
7
10
8
3
2
1
TSP
ppm
3
2
1
0
B
19
Glycerol
20-22
18
22
21
20
17
ppm
5.0
4.5
4.0
3.5
C
22
6-20-22
24
6
23
25
14
ppm
8.5
8.0
7.5
7.0
6.5
6.0
5.5

Figure 2. Score plots of PCA on ^{1}H-NMR binned spectra of colostrum obtained from different sow breed: A) PC1 vs PC2 and B) PC2 vs PC3. The first two PCs represent the 91% of the total variance. C-D) Loadings bar-plot for spectral bins along PC 1 and 2 respectively. Downfield (C1 and D1) and upfield (C2 and D2) regions of C and D loadings bar-plot were expanded on the vertical scale to appreciate the presence of small bar plot.


A
PC2 (10%)
2e+08
-2e+08
-6e+08
-1e+09
0e+00
1e+09
PC1 (81%)
LW
L
D
B
Autumn/Winter
Spring/Summer
PC3 4 %
2e+08
0e+00
-2e+08
-4e+08
-6e+08
-4e+08
-2e+08
0e+00
2e+08
4e+08
PC2 10 %
Seasons
1-Winter; 2-Spring; 3-Summer; 4-Autumn
C
Loadings PC1
C2
C1
D
Loadings PC2
Glycolate
D2
D1
Bins
C1
C2
D1
D2
Bins
Bins

Supplementary Table S1. The main bins accounting for the spectral differentiation and their relative chemical shift.

	Bin Number	ppm Interval	Metabolite[1]	Metabolite Name
PC1	10-11	8.134-8.071	25	UMP
	15	7.980-7.936	20 and 22	UDP- glucose, UDP-galactose.
	27-29	6.026-5.939	6, 22 and 23	UDP-N-Acetylglucosamine, UDP- galactose and Uridine
	39-41	5.675-5.588	22	UDP- galactose
	43-45	5.587-5.500	6	UDP-N-Acetylglucosamine
	49	5.412-5.368	NA	
	54-56	5.280-5.192	19	Lactose
	58-60	5.192-5.083	NA	
	67-70	4.507-4.397	19,20 and 22	Lactose, UDP- glucose and UDP- galactose
	72-75	4.396-4.288	22	UDP- galactose
	77-84	4.287-4.090	18 and 22	Myo-Inositol and UDP-galactose
	86-108	4.090-3.475	17,18,19 and 20	Taurine, Myo-Inositol, Lactose and UDP-glucose
	110-112	3.475-3.388	NA	
	114-123	3.387-3.146	7,16,17,18 and 19	O-Acetylcholine, sn-Glycerophosphocholine, Taurine, Myo-Inositol and Lactose
	127-128	3.080-3.0149	11,12 and 13	Creatine, Creatinine phosphate and Creatinine
	135	2.839-2.795	NA	
	137-139	2.795-2.770	10	Dimethylamine
	141-143	2.707-2.619	9	Citrate
	146-147	2.578-2.510	9	Citrate
	165	2.114-2.071	NA	
	167-168	2.070-2.005	NA	
	174-176	1.851-1.763	NA	
	184-185	1.367-1.3021	1	Lactate
	198	0.973-0.929	NA	
PC2	2	8.485-8.441	24	Formate
	7	8.221-8.178	NA	
	9-11	8.156-8.068	25	UMP

13-16	8.024-7.914	6,20 and 22	UDP-N-Acetylglucosamine, UDP- glucose and UDP-galactose
27-29	6.026-5.939	6,22 and 23	UDP-N-Acetylglucosamine, UDP- galactose and Uridine
33-34	5.895-5.807	NA	
39-46	5.674-5478	6 and 22	UDP-N-Acetylglucosamine and UDP- galactose
53-56	5.302-5.192	19	Lactose
59	5.170-5214	NA	
64	4.973-4.929	NA	
68-73	4.485-4.331	22	UDP- galactose
75-122	4.331-3.168	7,16,17,18,19,20,21 and 22	O-Acetylcholine, sn-Glycerophosphocholine, Taurine, Myo-Inositol, Lactose, UDP- glucose, Glycolate and UDP-galactose
124-125	3.168-3.102	14 and 15	cis-Aconitate and Choline
127-128	3.080-3.014	11,12 and 13	Creatine, Creatinine phosphate and Creatinine
130	3.014-2.971	NA	
135	2.839-2.795	NA	
137-139	2.795-2.770	10	Dimethylamine
141-143	2.707-2.619	9	Citrate
146-148	2.578-2.488	9	Citrate
163	2.1587-2.114	NA	
165	2.114-2.071	NA	
167-168	2.070-2.005	NA	
171-172	1.9612-1.895	3	Acetate
174-176	1.851-1.763	NA	
184-189	1.367-1.193	1	Lactate
194	1.127-1.083	NA	
198-199	0.973-0.907	NA	

[I] Metabolite assigned number referred to Table 2.

Chapter 8: Detection of differentially expressed genes in broiler pectoralis major muscle affected by White Striping – Wooden Breast myopathies.

P. Zambonelli, M. Zappaterra, F. Soglia, M. Petracci, F. Sirri, C. Cavani, R. Davoli

Department of Agricultural and Food Sciences (DISTAL), Bologna University.

Published in

Poultry Science

Zambonelli et al. *Poult Sci.* 2016 Sep;95(12):2771-2785.

https://doi.org/10.3382/ps/pew268

Summary

White Striping and Wooden Breast (WS/WB) are abnormalities increasingly occurring in the fillets of high breast yield and growth rate chicken hybrids. These defects lead to consistent economic losses for poultry meat industry, as affected broiler fillets present an impaired visual appearance that negatively affects consumers' acceptability. Previous studies have highlighted in affected fillets a severely damaged muscle, showing profound inflammation, fibrosis and lipidosis. The present study investigated the differentially expressed genes and pathways linked to the compositional changes observed in WS/WB breast muscles, in order to outline a more complete framework of the gene networks related to the occurrence of this complex pathological picture. The biochemical composition was performed on 20 Pectoralis major samples obtained from high breast yield and growth rate broilers (10 affected *vs*. 10 normal) and 12 out of the 20 samples were used for the microarray gene expression profiling (6 affected *vs.* 6 normal). The obtained results indicate strong changes in muscle mineral composition, coupled to an increased deposition of fat. In addition, 204 differentially expressed genes (DEG) were found: 102 up-regulated and 102 down-regulated in affected breasts. The gene expression pathways found more altered in WS/WB muscles are those related to muscle development, polysaccharide metabolic processes, proteoglycans synthesis, inflammation and calcium signaling pathway. On the whole, the findings suggest that a multifactorial and complex etiology is associated with the occurrence of WS/WB muscle abnormalities, contributing to further define the transcription patterns associated to these myopathies.

Introduction

In the past few decades, genetic selection for high breast yield and growth rate hybrids led to a dramatic increase in the incidence of several muscle myopathies and abnormalities (Dransfield and Sosnicki, 1999; Sandercock et al., 2009; Petracci and Cavani, 2012). Recently, it was found that up to 40% of broiler hybrids raised under commercial conditions were affected by white-striping (WS) with in some of those cases also the contextual occurrence of wooden breast (WB) (Lorenzi et al., 2014). As a consequence of the impaired visual appearance and consumers' acceptability severe WS and WB fillets are downgraded leading to considerable economic losses for poultry meat industry (Kuttappan et al., 2012; Petracci et al., 2015). Furthermore, these myopathies exert

a detrimental effect on quality traits and, as a result of the altered composition and reduced protein functionality (i.e. ability to hold/bind water, gel formation), may negatively affect both nutritional value and technological traits of meat (Mudalal et al., 2014; 2015). The emerging issue of WS/WB chicken fillets has to date no solution, as, despite the increasing number of studies aimed at investigating this defect, the knowledge concerning WS/WB occurrence is still incomplete and no certainty has been reached about the causes of this myopathy. Nevertheless, WS incidence is probably related to the animal genetics as its occurrence showed a heritability ranging from 0.338 (Bailey et al., 2015) to 0.65 (Alnahhas et al., 2016) depending on the studies.

The studies that have already been assessed on WS and WB affected breast muscles of high breast yield broilers identified through the biochemical and histological analysis an acutely inflamed and damaged muscle, with the presence of lipidosis and fibrosis (Mazzoni et al*.,* 2015; Soglia et al*.,* 2016).

Considering this scenario, the present study was aimed to compare the genomic transcription profile of WS/WB abnormal (ABN) *Pectoralis major* muscles respect to the normal ones (NORM) and obtain additional information about the gene networks possibly leading to the phenotypes typically related to WS/WB abnormalities.

Materials and Methods

Sample Selection and Preparation

Twenty boneless and skinless *Pectoralis major* muscles were selected from the same flock of Ross 708 broilers (males, weighting around 3.7 kg) in the deboning area of a commercial processing plant within 2 hours post-mortem. Birds belonging to this flock were farmed and slaughtered under commercial conditions according to Italian and European law for broiler chicken production. At slaughterhouse the birds were electrically stunned in agreement with the Council Regulation (EC) No 1099/2009 on the protection of animals at the time of killing. All slaughter procedures were monitored by the Veterinary team appointed by the Italian Ministry of Health. Fillets were selected evaluating the presence/absence of muscle abnormalities and classified as NORM and ABN according to the criteria proposed by Kuttappan et al. (2012) and Sihvo et al. (2014). In

particular, 10 NORM fillets without any hardened area and white striations and 10 ABN samples exhibiting diffused hardened areas and pale-bulging caudal end coupled with superficially white-striations in the cranial part were packaged and transported to the laboratory under refrigerated conditions (0-2°C). At the slaughterhouse, for the gene expression analysis, pieces of the 12 fillets showing the most extreme WS/WB phenotype among the 20 samples have been chosen (6 NORM and 6 ABN, the), immediately frozen in liquid nitrogen and then stored at -80°C until RNA extraction. RNA was extracted using TRIZOL reagent (Invitrogen Corporation, Carlsbad, California), as described in Davoli et al. (2011).

Quality and Technological Traits

At 24 hours *post-mortem*, fillets were individually weighted and colour was measured in triplicate on the ventral surface using a Chroma Meter CR-400 (Minolta Corp., Milan, Italy). Furthermore, the morphometric measurements (length, width and height, expressed in mm) were assessed with an electronic calliper as previously described by Mudalal et al. (2015) and ultimate pH evaluated according to the procedure described by Jeacocke (1977). Then, a parallelepiped sample (8 x 4 x 3 cm weighting approximately 80 g) was cut from the cranial part of each fillet and used to assess drip loss (percentage of weight lost as a consequence of refrigerated storage), cooking loss (after 45 min heating treatment at 80°C in a water bath) and Allo-Kramer shear force by using the procedures described in Petracci et al. (2013). Another sample (8 x 4 x 2 cm, weighting approximately 60 g) was excised from the middle section of the fillet, labelled and tumbled with 20% (wt/wt) sodium chloride (6%) and sodium tripolyphosphate (1.8%) marinade solution. Then, marinade uptake, cooking loss (after 25 min heating treatment at 80°C in a water bath), processing yield and Allo-Kramer shear force were assessed (Petracci et al., 2013).

Composition of Breast Fillets

Proximate (moisture, protein, fat, ash and collagen) and mineral composition of ABN and NORM fillets was determined on each *Pectoralis major* muscle applying standard methods. In particular, moisture and ash contents were calculated as the percentage of weight lost after drying 5 g of sample in oven (105°C for 16 hours) and incineration in muffle furnace (at 525°C), respectively

(AOAC, 1990). Crude proteins were determined according to Kjeldahl method by using copper sulphate as catalyst whereas total lipid amount was estimated by diethyl ether extraction performed with Soxhlet apparatus (AOAC, 1990). In addition, the colorimetric method proposed by Kolar (1990) was applied in order to determine hydroxyproline content and calculate the total amount of collagen (considering 7.5 as conversion factor). Inductively Coupled Plasma-Optic Emission Spectrometry technique (ICP-OES) was used in order to quantify minerals (Mg, K, P, Na and Ca) following the procedure suggested by the Environmental Protection Agency (EPA, 1996; 2007). Blanks were run to check for chemicals purity and reference materials (CRM GBW 09101, human hair control, Shanghai Institute of Nuclear Research Academia Sinica; CRM 201505 and 201605 Trace Element Whole Blood, Seronorm, Billingsad, Norway) were used to verify the accuracy. Finally, minerals content in tissues was calculated and expressed as mg 100 g^{-1} breast muscle.

SDS-PAGE Analysis of Muscle Proteins

One-dimensional SDS-PAGE analysis was carried out in order to evaluate the myofibrillar and sarcoplasmic proteins profile of ABN and NORM fillets following the procedure described by Liu et al. (2014) and removing the interfering substances such as salts, detergents, denaturants or organic solvents by READYPREP 2-D cleanup kit (Bio-Rad). Before loading, protein concentration was determined according to the Bradford assay (Bradford, 1976) and runs performed following the procedure described by Soglia et al. (2016). A molecular weight marker (Precision plus Standard protein, all blue pre-stained, Bio-Rad) was loaded on the first well of each gel and used to calculate the molecular weight of the separated bands. Each band was identified based on purified sarcoplasmic and myofibrillar proteins identified by mass spectrometry from literature (Zapata et al., 2012; Li et al., 2015) and the concentration expressed as relative abundance (%).

Statistical analysis of meat quality evaluation

Differences on meat quality and technological traits, composition as well as myofibrillar and sarcoplasmic protein profiles for ABN and NORM samples were tested by two-tailed Student's t

test. Statistical analyses were performed with SAS version 9.4 (SAS 9.4, Cary, NC. SAS Institute Inc.) and the nominal *P* value ≤0.05 was considered as significance threshold.

Microarray Expression Profiling

Each extracted RNA was checked for integrity and quality using an Agilent BioAnalyzer 2100, retro transcribed, amplified, labelled and applied to Affymetrix GeneChip Chicken Gene 1.1 ST v1 expression array by an outsource company (Cogentech Microarray Unit, Milan, Italy). All analytic procedures performed on microarray data were carried out using Partek Genomics Suite software, version 6.6 Copyright 2014 (Partek Inc., St. Louis, MO, USA). Gene expression profiles from the six ABN biological replicates were compared to the six NORM biological ones in order to identify differentially expressed genes (DEGs) indicated with the genes fold changes (FC) values between ABN an NORM broiler. FC filtering criteria combined with statistical T-test with FDR applied for multiple testing corrections were used to identify DEGs between the two conditions. The expression data obtained have been submitted to NCBI GEO database with the accession number GSE79276.

Validation by Quantitative Real Time-PCR

The results of the array expression analysis were validated by quantitative real-time PCR (qPCR). After DNAse treatment (TURBO DNA-freeTM, Ambion, Applied Biosystems), 1 µg of total RNA was reverse transcribed using the iScript cDNA Synthesis kit (Bio-Rad) according to the manufacturers' instructions. QPCR was performed on Rotor Gene™ 6000 (Corbett Life Science, Concorde, New South Wales) using 5 µl of SYBR® Premix Ex Taq™ (TAKARA Bio INC, Olsu, Shiga, Japan), 5 pmol of each primer, 2 µl of cDNA template diluted 1:10 and then was made up to the total volume of 10 µl with water. Rotor Gene™ 6000 protocol was optimised using specific annealing temperatures for each primer couple (Supplementary Table S1). The samples were first used to assess the expression level of three candidate-normalizing genes: glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), ribosomal protein L32 (*RPL32*), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*). Primers and PCR conditions used are reported in Supplementary Table S1. The expression levels of these three genes were evaluated

using NormFinder and *GAPDH* and *YWHAZ*, the two most stably expressed normalizing genes, used as reference genes. For each gene selected for validating the results of the expression array, an external primer pair to obtain the amplicon for the standard curve and an internal primer pair for the qPCR (Supplementary Table S1) were designed. For the validation of the microarray results, five genes have been chosen (crystallin alpha B (*CRYAB*), myoglobin (*MB*), glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase (*GNE*), utrophin (*UTRN*), prostaglandin F receptor (*PTGFR*)).

Threshold cycles obtained for the samples were converted by Rotor Gene 6000 to mRNA molecules/µl using for each gene the relative standard curve (Pfaffl, 2004; Zambonelli et al., 2016). Moreover, the average mRNA molecules/µl for each sample was normalized dividing the gene mRNA molecules/µl by the geometric average of *GAPDH* and *YWHAZ* mRNA molecules/µl in the given sample, as described in Zambonelli et al. (2016). Differences on the expression level calculated for ABN and NORM samples were tested by two-tailed Student's t test. Statistical analyses were performed with SAS version 9.4 (SAS 9.4, Cary, NC. SAS Institute Inc.) and the nominal *P* value ≤0.05 was considered as significance threshold.

Furthermore, in addition to the genes found differentially expressed, also ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 2 (*ATP2A2*) gene expression was tested through qPCR, and its absolute expression normalized using *GAPDH* and *YWHAZ* as normalizing genes.

Functional Characterization

Functional annotation, classification and annotation clustering of selected gene sets were carried out by DAVID Tools 6.7 (Huang et al., 2009a,b) using Biological Processes, Molecular Function gene ontology categories and KEGG pathways. A threshold for significance of $P < 0.05$ was considered to choose the significant functional categories.

The precise identification of the regulated snoRNAs was obtained by BLAST analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) using blastn algorithm and standard parameters.

In order to identify the codified name and the putative target genes of the differentially expressed micro-RNAs the miRBase website (http://www.mirbase.org/) was consulted in order to check the predicted chicken target genes resulting by the inspection of the correspondences obtained using miRDB (http://mirdb.org/miRDB/index.html) and TargetScan (http://www.targetscan.org/) databases available in miRBase.

Results

Quality and Technological Traits

The effects of muscle abnormalities on fillet weight and dimensions are displayed in Table 1. Overall, ABN samples exhibited higher ($P < 0.001$) weight coupled with increased length, width and middle height. As for raw meat quality (Table 2), ABN fillets were paler and revealed higher redness and ultimate pH values. Additionally, both raw and marinated ABN samples exhibited higher cooking losses ($P < 0.001$) and lower marinade uptake ($P < 0.001$). Besides, if compared with NORM group, a sharp increase in shear force value was measured in raw and marinated ABN fillets ($P < 0.001$).

Chemical and Mineral Composition

The occurrence of WS/WB abnormalities exerted a relevant impact on proximate and minerals compositions of the affected muscles (Table 3). In particular, ABN fillets exhibited higher ($P < 0.001$) moisture, fat and collagen contents coupled with reduced ash and proteins ($P < 0.001$) levels. As for mineral composition, no differences were found in potassium content whereas ABN samples had lower magnesium and phosphorus levels. Besides, an increased amount of sodium ($P < 0.001$) and calcium was observed in ABN fillets.

Identification of Differentially Expressed Proteins

With regard to myofibrillar proteins pattern (Table 4), nine bands having molecular weight from 16 to 220 kDa were identified. In detail, lower relative abundance of slow-twitch light chain myosin (LC1, 27.5 kDa) coupled with higher amount of 70 kDa MHC fragment were found in ABN samples.

As for sarcoplasmic protein pattern (Table 5), eleven bands, having molecular weight ranging from 25 to 114 kDa, were detected and almost all the enzymes involved in glucose metabolism differed between ABN and NORM groups. In particular, if compared with normal, ABN fillets exhibited lower amount of phosphoglycerate mutase (PGAM, 25 kDa), creatine kinase (KCRM, 43 kDa), phosphoglucose isomerase (GPI, 58 kDa) and pyruvate kinase (KPYM, 68 kDa) together with higher relative abundance of lactate dehydrogenase (LDH, 34 kDa), glyceraldehyde dehydrogenase (G3P, 36 kDa), aldolase (ALDO, 39 kDa) and glycogen phosphorylase (PYGL, 90 kDa). Besides, ABN samples exhibited higher Calcium-Transporting ATPase Sarcoplasmic Reticulum Type Slow Twitch Skeletal Muscle Isoform (AT2A2, 114 kDa, coded by *ATP2A2* gene).

Profiling of Differentially Expressed Genes

Comparing the gene expression profiles obtained for ABN *vs*. NORM samples, 207 differentially expressed genes (DEG) were found: 103 up-regulated and 104 down-regulated in ABN chickens (Supplementary Table S2). The overexpressed gene poly(A) binding protein, cytoplasmic 1 (*PABPC1*) (NM_001031597) was found twice on the list of DEG whereas the uncharacterized under expressed gene CD218879 was found three times among the probes included in the microarray then the number of unique DEG is 102 for the up-regulated and 102 also for the down-regulated. In both cases, the trend of expression was the same for the doublets or triplets of gene probe and the fold changes detected are identical or similar. Among the more expressed genes in ABN samples only three were currently not identified both in chickens and human genome. For the less expressed genes in ABN chickens, the situation found was more complex because we identified fifteen uncharacterized genes as described before but we found represented in this group also some short noncoding RNAs. In particular, we observed the presence of four small nucleolar RNAs (snoRNA, GGN5, GGN50, GGN73, GGN98) and of five micro-RNAs (miRNA; MIR196B, MIR205, MIR1600, MIR1716, MIR1805).

The differentially expressed genes mapped on all chicken chromosomes from 1 to 28 and one was mapped each on chromosome 30, mitochondrion genome (tRNA-Asn gene) and on the unassigned

linkage group LGE64. Furthermore, six genes are currently not located on the chicken genome sequence and four of them are not characterized yet. The intensity of the differential expression for the less expressed genes in ABN chickens varied between -1.07 and -2.71 FC with five genes showing a FC below -2. Among the overexpressed genes, the range of FCs was between 1.16 and 9.87 with 12 genes showing FC values between 2 and 3 and three genes showed higher values (nuclear factor, interleukin 3 regulated (*NFIL3*) 3.24, myoglobin (*MB*) 5.17, cysteine and glycine-rich protein 3 (cardiac LIM protein) (*CSRP3*) 9.87. In general, the moderate level of differences in expression can suggest the simultaneous involvement of several genes needed for the definition of the abnormal *Pectoralis major* muscle phenotypes.

Validation by qPCR

In order to confirm the results obtained with microarray, we selected five genes (*CRYAB*, *MB*, *GNE*, *UTRN*, *PTGFR*) that were analysed by qPCR and the data obtained were compared with the expression profiles obtained using microarrays (Figure 1A). The expression level of all validated genes was in agreement between the two analyses showing a very good correlation ($R^2 = 0.91$, Figure 1B). Moreover, we tested by qPCR also the expression level of *ATP2A2* mRNA, a gene not present among the probes of the utilized microarray as the protein coded by this gene was found differentially expressed between ABN and NORM samples using SDS-PAGE analysis (Table 2). The results of the qPCR experiment for the *ATP2A2* gene confirmed the difference of expression detected by SDS-PAGE analysis. According to this result, we also considered *ATP2A2* for the functional characterization of the differentially expressed genes.

Functional Characterization of the Differentially Expressed Coding Genes

A functional classification was carried out using DAVID tools (Table 6). The significant annotation clusters represented are related to several categories of the Biological Processes section according to the Gene Ontology classification such as developmental processes (in particular, muscle organ development), polysaccharide metabolic processes, response to reactive oxygen species and immune response, regulation of cell cycle, mRNA export from nucleus, blood vessels morphogenesis, intracellular transport, sensory perception of light stimulus. Interestingly, the

unique KEGG pathway represented was the calcium signalling pathway. By analysing this regulated pathway, we observed that five DEG, all overexpressed, were involved in the calcium homeostasis within the cells. Furthermore, two down-regulated genes, 5-hydroxytryptamine (serotonin) receptor 4 (*HTR4*) and tachykinin receptor 2 (*TACR2*), were also included in the calcium signalling pathway as reported by DAVID analysis.

Functional Characterization of the Differentially Expressed Noncoding Genes

Among the less expressed genes detected in ABN samples, the presence of nine noncoding RNAs is remarkable.

The first type of this category of genes is snoRNAs that are noncoding transcripts active within the nucleolus of the cells and that guide molecules for site specific modifications of other RNAs. By BLAST analysis (Table 7), we found that the four detected DE snoRNAs belong to the H/ACA box family and are involved in the pseudouridynilation of rRNAs and snRNAs changing an uridine to the pseudouridine isomer (Zhang et al., 2009; Holley and Topkara, 2011) which plays a key role in the activation of ribosome function. Furthermore, other functions are demonstrated for the snoRNAs such as the implication in post-transcriptional or in directing alternative splicing regarding specific mRNAs coding for proteins (Makarova et al., 2013). These short RNA molecules can be also the precursors of other small regulative RNAs like miRNAs and sdRNAs (sno-derived RNAs). All these findings are quite recent and validated only in few experiments. If these regulatory mechanisms will be confirmed, a specific role in regulating the muscle development could be found.

The second group of noncoding regulated transcripts includes five miRNAs that have been precisely identified by searching on miRBase. Furthermore, on the same database we searched for the putative target genes using as entry the DE chicken miRNAs and the overexpressed coding genes, according to miRNA repressive mechanism of regulation, as indication of a direct regulation (Table 8). Four out of five miRNAs have been indicated as putative regulators of eight genes (ankyrin repeat domain 31 (*ANKRD31*), family with sequence similarity 64, member A (*FAM64A*), PQ-loop repeat-containing protein 2-like (*LOC425021*), N-myristoyltransferase 2 (*NMT2*), phospholamban (*PLN*), plexin A1 (*PLXNA1*), paired related homeobox 2 (*PRRX2*), WD repeat and FYVE domain containing 3 (*WDFY3*)).

Discussion

WS and WB are abnormalities of chicken breast skeletal muscle observed quite frequently in most of modern broiler hybrids (Petracci et al., 2015). In the present study, the gene expression and the meat quality traits of affected chickens *Pectoralis major* muscles were investigated, in order to evaluate the effects exerted by these pathologic conditions on the gene transcription levels and the fillets biochemical composition. On the whole, the results outlined the presence of a severely damaged *Pectoralis major* muscle, with ABN samples showing consistent changes in the expression level of genes related to muscle development, reactive oxygen species metabolism, oxidative stress and signal transduction, blood vessel morphogenesis and polysaccharide metabolism. Considering the data reported in the present study, together with the results identified in literature (Kuttappan et al., 2013; Sihvo et al., 2014, Mazzoni et al., 2015; Mutryn et al., 2015; Soglia et al., 2016), a consistent impairment of normal muscle metabolism is evident for WS/WB defects and these abnormalities appear to be linked to a complex pathogenesis.
The complex pathological framework characterizing WS/WB defects has made extremely difficult to date the identification of the underlying causes at the basis of the alterations. Despite the literature about WS and WB defects agrees on the large number of histological, biochemical and metabolic alterations accompanying the occurrence of these breast muscle abnormalities, there is no consensus on the causes leading to the insurgence of this complex pathological framework. The results obtained from the present research have been evaluated and discussed in the light of the knowledge reported to date in literature on WS and WB topics, aiming to outline an overall view of the pathological changes affecting *Pectoralis major* muscle showing in addition the gene networks and the biological evidences related to the occurrence of the WS/WB tissue alterations. Furthermore, the overall analysis of the variations obtained to date at the genetic, biochemical biological and histological levels allows defining some possible hypotheses on the mechanisms determining the onset of these myopathies.

Oxidative Stress

In the present research, the microarray data showed an overall increase in the expression level of genes involved in the response against the accumulation of hydrogen peroxide and reactive

oxygen species in ABN fillets, in agreement with the presence of an oxidative stress possibly linked to a muscle hypoxic condition. In particular, ABN samples presented increased expression levels of crystallin alpha B (*CRYAB*), adenosine deaminase (*ADA*), *MB* genes and a decreased activity of reactive oxygen species modulator 1 (*ROMO1*) gene, that are all involved in the response to reactive oxygen species (Table 6 and Supplementary Table 2). These results agree with Mutryn et al. (2015) that identified a set of DE genes involved in myofibers reaction to oxidative stress in muscle samples of high breast yield chickens affected by WS/WB. The cause of this oxidative stress is not clear, although past studies can help defining some possible causes. One of the possible hypotheses reported in literature indicates an inadequate breast muscle vascularisation as a possible key factor in WS/WB occurrence (Mutryn et al., 2015). It is worth to note that in 1999, Dransfield and Sosnicki reported an increased proportion of glycolytic fibers with enlarged diameter in chicken lines selected for high growth rate and breast yield (Dransfield and Sosnicki, 1999). Moreover, Hoving-Bolink et al. (2000) showed an intense reduction in both vascularization and capillary to fiber ratio in chicken hybrids selected for high growth rate and breast yield compared to unselected chicken lines. Based on these results genetic selection in these chicken lines determined in muscles an inadequate blood vessel growth with the consequent impairment in oxygen supply and in the metabolic waste products displacement from breast myofibers. In agreement with Mutryn et al. (2015), also the present results indicated that an excessive accumulation of reactive oxygen species (ROS) within the muscle tissue of ABN samples might be involved in initiating the inflammatory mechanism typically associated with WS and/or WB muscle abnormalities.

Inflammation and Myofibers Degeneration

The profound alterations and the inflammatory status observed in previous researches were confirmed in the present study where an increased transcription of genes coding for proteins involved in biological processes related to tissue alteration was detected. More precisely, we have found that immunoglobulin superfamily, member 10 gene (*IGSF10*), heat shock 105kDa/110kDa protein 1 (*HSPH1*) and heat shock 60kDa protein 1 (Chaperonin) (*HSPD1*) were overexpressed in ABN samples (Supplementary Table S2), supporting the presence of the tissue inflammation. Moreover, the overexpression in ABNs of ADAM family metallopeptidase with thrombospondin type 1 motif 12 (*ADAMTS12*) and ADAM metallopeptidase with thrombospondin type 1 motif 19 (*ADAMTS19*; Supplementary Table S2) genes suggested the existence of a muscle tissue

inflammation, as *ADAMTS12* in particular is already known to be involved in the activation of inflammatory responses (Moncada-Pazos et al., 2012). Similarly, the higher transcription levels of nuclear factor, interleukin 3 regulated (*NFIL3*) and snail family zinc finger 2 (*SNAI2*; Supplementary Table S2), as genes encoding for hindering-cell death molecules (Keniry et al., 2014), could reveal the attempt to limit apoptotic processes and necrosis of the cells. Thus, the onset of a complex biological reaction aimed at contrasting the inflammatory process with the activation of anti-inflammatory responses was observed in ABN samples. These inflammatory and necrotic processes were previously found in WS and WB breast muscles (Kuttappan et al., 2013; Sihvo et al., 2014; Mutryn et al., 2015) and association to degenerative processes of muscular nerve growth was hypothesized. Indeed, the microarray analysis evidenced in WS/WB broilers reduced transcription levels of the genes deafness autosomal recessive 31 (*DFNB31*), syntaxin 3 (*STX3*), neurogenin 1 (*NEUROG1*), SLIT and NTRK-like family member 6 (*SLITRK6*), wingless-type MMTV integration site family member 7A (*WNT7A*), and EPH receptor A2 (*EPHA2*), involved in neuron genesis and differentiation as indicated by DAVID analysis.

Another differentially expressed gene found in the present study is interleukin 1 beta (*IL1B*), which encodes for a member of the interleukin 1 cytokine family. This protein exerts a central role as mediator of the inflammatory response and is involved in a variety of cellular activities (including cell proliferation, differentiation and apoptosis). Although the findings of the present study outlined an inflammatory process affecting ABN samples, a down-regulated transcription of *IL1B* gene was found (Supplementary Table S2). Despite the majority of the literature reports in muscles affected by inflammation an increased expression level of the *IL1B* gene (Dinarello, 1998; Li et al., 2008), in some cases this gene was found down-regulated during some chronic pathological situations (Karli et al., 2014). Moreover it is possible to hypothesize that the reduced transcription level of *IL1B* identified within the WS/WB samples could be linked to the low level of vascularization of the ABN samples as IL1B was also found to play a relevant role in the angiogenic processes promoting the emergence of new capillaries from pre-existing blood vessels (Dinarello, 1996; Voronov et al., 2003). Furthermore, as interleukin 1 has a pyrogenic role and its involvement in the pain sensation during inflammation has been evidenced, the down-regulation of *IL1B* transcription might be responsible for the lack of symptoms in chickens affected by WB/WS abnormalities during the broiler farming period.

Some genes identified as differentially expressed have been reported in literature associated to the development of myopathies. The over-expression of *PLN* observed in ABN samples

(Supplementary Table S2) was related in mice muscles to an altered phenotype similar to the centronuclear myopathy identified in human muscles (Fajardo et al., 2015). Centronuclear myopathy is a congenital myopathy characterized by centrally located nuclei, peculiarity that was already described in WS/WB affected muscles (Sihvo et al., 2014; Soglia et al., 2016). Moreover, *ATP2A2* gene and protein was found overexpressed in ABN samples. Soglia et al. (2016) reported that an impaired activity of AT2A2 protein might be involved in the WS/WB phenotype with a loss of adhesion among myocytes and the presence of abundant connective tissue replacing the muscle cells. This phenomenon is for some instances similar to a human pathology, Darier-White disease (Savignac et al., 2011) that is characterized by a loss of adhesion between epidermal cells and keratinization then leading to apoptosis of the same cells. The causative mutation of the human disease was found in one of the transcripts of *ATP2A2* gene.

Myofibers Regeneration

In the present study, several genes involved in muscle development and cell differentiation were found differentially expressed within the ABN cases (Table 7). In particular, the over-expression of *CSRP3* and *PTGFR* as well as the down-regulation of *P2RY1* gene identified in the present research, could be associated respectively to muscle fibers synthesis (Kong et al., 1997; Jansen and Pavlath, 2008) and myogenesis (Krasowska et al., 2014).

The cascade pathway of *PTGFR*, *PLN*, *GNAQ*, *PLCB2*, and *PLCD1* has been also related to mechanisms aimed at regenerating damaged muscle. The combined activity of these five DEG triggers several downstream metabolic pathways that increase muscle cells volume (Hindi et al., 2013; Horsley and Pavlath, 2003; 2004) and contribute to the regeneration of myofibers upon injury, suggesting a possible role of these genes in trying to repair the effect of severe *Pectoralis major* myopathies such as WB and WS.

Additionally, ABN fillets showed an increased mRNA level of the gene *FAM64A*. The expression of this gene was reported to be associated with rapidly proliferating tissues during mouse embryogenesis (Archangelo et al., 2008). Based on this previous finding, *FAM64A* over-expression may be associated to the regeneration processes taking place in the damaged breasts, and its up-regulation could be determined by gga-miR-196-5p that identifies FAM64 a as a specific target gene (Table 8). Similarly, also increased mRNA levels of *PLXNA1* and *PRRX2* genes, up-regulated in ABN samples, have been related to proliferating foetal fibroblasts and developing tissue (White et

al., 2003; Hota and Buck, 2012). As suggested for *FAM64A,* we hypothesize that *PLXNA1* and *PRRX2* expression changes may be respectively linked to the regulation exerted by gga-miR-205a and gga-miR-1600. Anyway, these hypotheses will need further studies to be proven.

On the whole, the results obtained from the microarray, with DE genes from pathways linked to muscle differentiation and development, can be interpreted as the evidences of an activation of muscle cells regenerative processes in response to the degenerative status. These tissue regenerative processes (nuclear rowing and cell multi-nucleation) have been evidenced within the WS/WB muscles in previous histological observations (Kuttappan et al., 2013; Sihvo et al., 2014, Soglia et al., 2016). Furthermore, a remarkable increase in the amount of AT2A2 (Table 5) coupled with an over-expression of *ATP2A2* gene and *MB* gene, was detected in ABN samples similar to what was observed by Mutryn et al. (2015) in muscles affected by WB abnormality. These Authors supposed that the higher expression of *ATP2A2* and *MB* might be the result of a shift from type IIb towards slow twitch type I fibers in abnormal breast muscles. This reorganization of the tissue might be explained considering that similar regenerative mechanisms, exhibiting an overall increase in slow-twitch fibers and the apoptosis of the fast-twitch ones, were previously observed in mice dystrophic muscles (Massa et al., 1997). In addition, the up-regulation in the transcription of acetyl-CoA acyltransferase 2 (*ACAA2*) gene, encoding for a protein exerting a relevant role in the metabolic pathway leading to mitochondrial beta oxidation of fatty acids, might support the hypothesized shift of ABN muscles towards oxidative metabolism.

Impaired Muscle Ion Homeostasis

ABN samples showed relevant changes in their chemical composition, with an overall modification in mineral content (Table 3). In particular, the increased sodium content may be related to the more elevated transcription of solute carrier family 9, subfamily A (NHE7, cation proton antiporter 7) member 7 (*SLC9A7*) gene, as it encodes a sodium and potassium/ proton antiporter (Kagami et al., 2008).

Additionally, also increased levels of calcium were identified in ABN samples (Table 3), supporting the existence of the intracellular calcium buildup already hypothesized by Mutryn et al. (2015) on the basis of the altered transcription levels identified for genes involved in calcium homeostasis. In agreement with this hypothesis, several genes linked to intracellular ion homeostasis were found differentially expressed in the present study (Figure 2), in particular the up-regulation of genes linked to the activation of G protein–coupled purinergic receptors was observed in cascade in ABN

samples. The over-expression of *PTGFR* coupled with the increase in guanine nucleotide binding protein (G protein) alpha 11 (*GNAQ*) and G protein-coupled receptor 1 (*GPR1*) might be related to the up-regulation of phospholipase C beta 2 (*PLCB2*) and phospholipase C delta 1 (*PLCD1*) genes (Figure 3), involved in the processes leading to the increase of Ca^{++} in the cells. According to Bucheimer and Linden (2004), the G protein–coupled receptors activate β phospholipase enzymes resulting in an increased intracellular Ca^{++} concentration and altered ions homeostasis. Within this framework, the reduced expression of calcium homeostasis modulator 3 gene (*CALHM3*; Supplementary Table S2), the higher synthesis of AT2A2 protein and the increased transcription of its relative gene *ATP2A2* and *PLN* gene can produce in ABN samples an overall alteration in calcium signalling pathway, contributing to the inflammatory processes. *ATP2A2* over-expression was noticed also by Mutryn et al. (2015) in breast muscles affected by WB abnormality. Between the differentially expressed miRNAs, gga-miR-1600 may be the putative regulator of *PLN* gene expression (Table 8), suggesting that for calcium signalling pathway a higher level of regulation might be responsible of relevant changes in the molecular mechanisms involved in the origin of the WB/WS myopathy.

Anyway, on the basis of the whole literature produced to date, it is not possible a certain assumption about the causes leading to this overall impairment in muscle ion homeostasis. The differential expression of genes related to the purinergic receptors pathways that we found in ABN samples may be one of the primary causes of the inflammation or, most likely, one of the effects of muscle tissue structural changes related to WS/WB abnormalities and the results of the activation of the purinergic receptors from the ATP released in the extracellular matrix spaces from damaged fibers (Bucheimer and Linden, 2004; Eltzschig et al., 2012).

Altered Glucose Metabolism, Lipidosis, Fibrosis and Proteoglycans Synthesis

The SDS-PAGE results revealed an intensified glycolytic activity in ABN samples, with higher amount of glycolytic enzymes such as lactate dehydrogenase (LDH), glyceraldehyde dehydrogenase (G3P), aldolase (ALDO) and glycogen phosphorylase (PYGL). Among the glycolytic enzymes, the magnesium-dependent enzymes phosphoglycerate mutase (PGAM), phosphoglucose isomerase (GPI) and pyruvate kinase (KPYM) were less expressed in ABN samples. The inadequate level of disposable magnesium in affected muscles (Table 3) could be linked to the lower translation of these magnesium-dependent enzymes in affected samples. A general

modification of the glycolytic enzymes expression was evident in ABN *Pectoralis major* muscles. Despite the increased synthesis of LDH enzyme in ABN samples, a higher ultimate pH in the affected breast muscles was observed (Table 2), suggesting that there was not an increase in the transformation of pyruvate into lactate, as normally expected in hypoxic conditions. In addition to the modifications in glycolytic enzymes synthesis, we found by microarray analysis that in ABN samples the genes *GNE*, glycogen branching enzyme (*GBE1*), UDP-glucose 6-dehydrogenase (*UGDH*), and protein phosphatase 1, catalytic subunit, beta isozyme (*PPP1CB*), involved in polysaccharide metabolic processes were overexpressed (Supplementary Table S2). GNE enzyme plays an essential role in hexosamine pathway, in particular for the biosynthesis of N-acetylneuraminic acid, a precursor of sialic acids. This evidence might suggest an alternative utilization of fructose 6-phosphate, produced by PGI enzyme during glycolysis. Indeed fructose 6-phosphate can undertake the glycolysis pathway or can be used as the initial substrate of the hexosamine and hexuronic acid pathways, resulting in collagen, proteoglycans and glycosaminoglycans synthesis. This shift towards hexosamine pathway was also described by Du et al. (2000) as a consequence of the accumulation of ROS species, which otherwise exerted an inhibitory effect on glycolysis. On the other hand, UGDH converts UDP-glucose to UDP-glucuronate and thereby participates in the biosynthesis of glycosaminoglycans such as hyaluronan (a common component of the extracellular matrix). The over-expressions of *UGDH, GNE*, *GBE1*, *PPP1CB* and interphotoreceptor matrix proteoglycan 2 (*IMPG2*) genes identified in the present research, together with the observed fibrosis described on the same samples by Soglia et al. (2016), could explain the increased presence of collagen and proteoglycans in the areas affected by WS/WB lesions. These findings agree with the evidences reported in literature for WS and/or WB abnormalities (Kuttappan et al., 2013; Sihvo et al., 2014; Mutryn et al., 2015; Velleman and Clark, 2015).

Additionally, the increased collagen and fat contents observed in ABN samples (Table 3) were in agreement with previous findings (Kuttappan et al., 2013; Mudalal et al., 2014; Soglia et al*.,* 2016). A similar situation was described by Lopes–Ferreira et al. (2001), who noticed in hypoxic conditions the development of both fibrosis and lipidosis within the skeletal muscle, with a replacement of the lost fibers with the collagen synthesis and lipid deposition.

The present results, combined with the existing knowledge, allowed to draw a scheme (Figure 4) describing a possible progression of the biological processes cascade hypothesized to be involved in the development of WS/WB myopathies.

On the whole, the findings of the present study show at the gene level that a complex etiology is associated with the occurrence of WS and WB muscle abnormalities. In WS/WB breast muscles, there is evidence of differentially expressed genes related to several functional categories: muscle development, polysaccharide metabolic processes, glucose metabolism, proteoglycans synthesis, inflammation, and calcium signaling pathway. By combining the functional roles for the differentially expressed genes, we hypothesized a network of biological changes that are acting simultaneously and are responsible of the phenotypic evidences of these myopathies.

Although the cause of these myopathies is still unclear, the majority of the results reported in literature suggests that selection criteria more and more addressed towards fast growing and high breast yield broilers could be involved in the occurrence of the breast oxidative stress that triggers the cascade of WS/WB related muscle alterations. The data obtained in the present research can be useful for the clarification of the WS/WB pathogenesis and further studies have to be planned to disentangle the complex etiology behind these myopathies.

References

Alnahhas N, Berri C, Chabault M, Chartrin P, Boulay M, Bourin MC & Le Bihan-Duval E (2016) Genetic parameters of white striping in relation to body weight, carcass composition, and meat quality traits in two broiler lines divergently selected for the ultimate pH of the pectoralis major muscle. *BMC Genet* 17:61.

AOAC (1990) Meat and meat products. In *Official Methods of Analysis*. pp. 931–948, 15th ed. Vol. 2. Assoc. Off. Anal. Chem., Washington, DC.

Archangelo LF, Greif PA, Hölzel M, Harasim T, Kremmer E, Przemeck GHK, Eick D, Deshpande AJ, Buske C, Hrabé de Angelis M, Olalla Saad ST & Bohlander SK (2008) The CALM and CALM/AF10 interactor CATS is a marker for proliferation. *Mol Oncol* 2:356–367.

Bagés S, Estany J, Tor M & Pena RN (2015) Investigating reference genes for quantitative real-time PCR analysis across four chicken tissues. *Gene* 561:82–87.

Bailey RA, Watson KA, Bilgili SF & Avendano S (2015) The genetic basis of pectoralis major myopathies in modern broiler chicken lines. *Poult Sci* 94:2870-2879.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254.

Bucheimer RE & Linden J (2004) Purinergic regulation of epithelial transport. *J Physiol* 555:311–321.

Dinarello CA (1996) Biologic basis for interleukin-1 in disease. *Blood* 87:2095–2147.

Dinarello CA (1998) Interleukin-1, interleukin-1 receptors and interleukin-1 receptor antagonist. *Int Rev Immunol* 16:457–499.

Dransfield E & Sosnicki AA 1999. Relationship between muscle growth and poultry meat quality. *Poult Sci* 78:743–746.

Du XL, Edelstein D, Rossetti L, Fantus IG, Goldberg H, Ziyadeh F, Wu J & Brownlee M (2000) Hyperglycemia-induced mitochondrial superoxide overproduction activates the hexosamine

pathway and induces plasminogen activator inhibitor-1 expression by increasing Sp1 glycosylation. *Proc Natl Acad Sci* 97:12222–12226.

Eltzschig HK, Sitkovsky MV & Robson SC (2012) Purinergic Signaling during Inflammation. *N Engl J Med* 367:2322–2333.

Environmental Protection Agency, Office of Solid Waste (1996) Method 3052. Microwave assisted Acid digestion of siliceous and organically based Matrices, Washington, DC: U.S. EPA, http://www.epa.gov/osw/hazard/testmethods/sw8467pdfs/3052.pdf

Environmental Protection Agency, Office of Solid Waste (2007) Method 6010C. Inductively Coupled Plasma-Atomic Emission Spectrometry. Washington, DC: U.S. EPA. http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/6010c.pdf

Fajardo VA, Bombardier E, McMillan E, Tran K, Wadsworth BJ, Gamu D, Hopf A, Vigna C, Smith IC, Bellissimo C, Michel RN, Tarnopolsky MA, Quadrilatero J & Tupling AR (2015) Phospholamban overexpression in mice causes a centronuclear myopathy-like phenotype. *Dis Model Mech* 8:999–1009.

Hindi SM, Tajrishi MM & Kumar A (2013) Signaling Mechanisms in Mammalian Myoblast Fusion. *Sci Signal* 6:re2–re2.

Holley CL & Topkara VK (2011) An Introduction to Small Non-coding RNAs: miRNA and snoRNA. *Cardiovasc Drugs Ther* 25:151–159.

Horsley V & Pavlath GK (2003) Prostaglandin F2 (alpha) stimulates growth of skeletal muscle cells via an NFATC2-dependent pathway. *J Cell Biol* 161:111–118.

Horsley V & Pavlath GK (2004) Forming a multinucleated cell: molecules that regulate myoblast fusion. *Cells Tissues Organs* 176:67–78.

Hota PK & Buck M (2012) Plexin structures are coming: opportunities for multilevel investigations of semaphorin guidance receptors, their cell signaling mechanisms, and functions. *Cell Mol Life Sci* 69:3765–3805.

Hoving-Bolink AH, Kranen RW, Klont RE, Gerritsen CLM & de Greef KH (2000) Fiber area and capillary supply in broiler breast muscle in relation to productivity and ascites. *Meat Sci* 56:397–402.

Huang DW, Sherman BT & Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.

Huang DW, Sherman BT & Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucl Acids Res* 37:1–13.

Jansen KM & Pavlath GK (2008) Prostaglandin F2α promotes muscle cell survival and growth through upregulation of the inhibitor of apoptosis protein BRUCE. *Cell Death Differ* 15:1619–1628.

Jeacocke RE (1977) The temperature dependence of anaerobic glycolysis in beef muscle held in a linear temperature gradient. *J Sci Food Agric* 28:551-556.

Kagami T, Chen S, Memar P, Choi M, Foster LJ & Numata M (2008) Identification and biochemical characterization of the SLC9A7 interactome. *Mol Membr Biol* 25:436–447.

Karli P, Martlé V, Bossens K, Summerfield A, Doherr MG, Turner P, Vandevelde M, Forterre F & Henke D (2014) Dominance of chemokine ligand 2 and matrix metalloproteinase-2 and -9 and suppression of pro-inflammatory cytokines in the epidural compartment after intervertebral disc extrusion in a canine model. *Spine J* 14:2976–2984.

Keniry M, Dearth RK, Persans M & Parsons R (2014) New Frontiers for the NFIL3 bZIP Transcription Factor in Cancer, Metabolism and Beyond. *Discoveries* 2:e15.

Kolar K (1990) Colorimetric determination of hydroxyproline as measure of collagen content in meat and meat products: NMKL collaborative study. *J Assoc Off Anal Chem* 73:54-57.

Kong, Y, Flick MJ, Kudla AJ & Konieczny SF (1997) Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD. *Cell Mol Biol* 17:4750–4760.

Krasowska E, Róg J, Sinadinos A, Young CNJ, Górecki DC & Zabłocki K (2014) Purinergic receptors in skeletal muscles in health and in muscular dystrophy. *Postepy Biochem* 60:483-489.

Kuttappan VA, Lee YS, Erf GF, Meullenet JF, McKee SR & Owens CM (2012) Consumer acceptance of visual appearance of broiler breast meat with varying degrees of white striping. *Poult Sci* 91:1240-1247.

Kuttappan VA, Shivaprasad HL, Shaw DP, Valentine BA, Hargis BM, Clark FD, McKee SR & Owens CM (2013) Pathological changes associated with white striping in broiler breast muscles. *Poult Sci* 92:331–338.

Li L, Fei Z, Ren J, Sun R, Liu Z, Sheng Z, Wang L, Sun X, Yu J, Wang Z & Fei J (2008) Functional imaging of interleukin 1 beta expression in inflammatory process using bioluminescence imaging in transgenic mice. *BMC Immunol* 9:49.

Li K, Zhao YY, Kang ZL, Wanh P, Han MY, Xu XL & Zhou GH (2015) Reduced functionality of PSE-like chicken breast meat batter resulting from alterations in protein conformation. *Poult Sci* 94:111-122.

Liu J, Ruusunen M, Puolanne E & Ertbjerg P (2014) Effect of pre-rigor temperature incubation on sarcoplasmic protein solubility, calpain activity and meat properties in porcine muscle. *LWT-Food Sci Technol* 55:483-489.

Lopes-Ferreira M, Núñez J, Rucavado A, Farsky SH, Lomonte B, Angulo Y, Moura Da Silva AM & Gutiérrez JM (2001) Skeletal muscle necrosis and regeneration after injection of Thalassophryne nattereri (niquim) fish venom in mice. *Int J Exp Pathol* 82:55–64.

Lorenzi M, Mudalal S, Cavani C & Petracci M (2014) Incidence of white striping under commercial conditions in medium and heavy broiler chickens in Italy. *J Appl Poult Res* 23:754-758.

Makarova, JA, Ivanova SM, Tonevitsky AG & Grigoriev AI (2013) New functions of small nucleolar RNAs. *Biochemistry (Mosc)* 78:638–650.

Massa R, Silvestri G, Zeng YC, Martorana A, Sancesario G & Bernardi G (1997) Muscle Regeneration in mdx Mice: Resistance to Repeated Necrosis is Compatible with Myofiber Maturity. *Basic Appl Myol* 7:387-394.

Mazzoni M, Petracci M, Meluzzi A, Cavani C, Clavenzani P & Sirri F (2015) Relationship between pectoralis major muscle histology and quality traits of chicken meat. *Poult Sci* 94:123–130.

Moncada-Pazos A, Obaya AJ, Llamazares M, Heljasvaara R, Suárez MF, Colado E, Noël A, Cal S & López-Otín C (2012) ADAMTS-12 metalloprotease is necessary for normal inflammatory response. *J Biol Chem* 287:39554–39563.

Mudalal S, Babini E, Cavani C & Petracci M (2014) Quantity and functionality of protein fractions in chicken breast fillets affected by white striping. *Poult Sci* 93:2108–2116.

Mudalal S, Lorenzi M, Soglia F, Cavani C & Petracci M (2015) Implications of white striping and wooden breast abnormalities on quality traits of raw and marinated chicken meat. *Animal* 9:728–734.

Mutryn MF, Brannick EM, Fu W, Lee WR & Abasht B (2015) Characterization of a novel chicken muscle disorder through differential gene expression and pathway analysis using RNA-sequencing. *BMC Genomics* 16:399.

Petracci M & Cavani C (2012) Muscle growth and poultry meat quality issues. *Nutrients* 4:1-12.

Petracci M, Mudalal S, Bonfiglio A & Cavani C (2013) Occurrence of white striping under commercial conditions and its impact on breast meat quality in broiler chickens. *Poult Sci* 92:1670-1675.

Petracci M, Mudalal S, Soglia F & Cavani C (2015) Meat quality in fast-growing broiler chickens. *Worlds Poult Sci J* 71:363-374.

Pfaffl MW (2004) Quantification strategies in real-time PCR. In A–Z of quantitative PCR [Bustin SA editor] International University Line (IUL), La Jolla, CA, USA.

Sandercock DA, Barker Z, Mitchell MA & Hocking PH (2009) Changes in muscle cell cation regulation and meat quality traits are associated with genetic selection for high body weight and meat yield in broiler chickens. *Genet Sel Evol* 41:8.

Savignac M, Edir A, Simon M & Hovnanian A (2011) Darier disease : a disease model of impaired calcium homeostasis in the skin. *Biochim Biophys Acta* 1813:1111–1117.

Sihvo HK, Immonen K & Puolanne E (2014) Myodegeneration with fibrosis and regeneration in the pectoralis major muscle of broilers. *Vet Pathol* 51:619-623.

Soglia F, Mudalal S, Babini E, Di Nunzio M, Mazzoni M, Sirri F, Cavani C & Petracci M (2016) Histology, composition, and quality traits of chicken Pectoralis major muscle affected by wooden breast abnormality. *Poult Sci* 95:651-659.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ & von Mering C (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucl Acids Res* 43:D447–D452.

Velleman SG & Clark DL (2015) Histopathologic and Myogenic Gene Expression Changes Associated with Wooden Breast in Broiler Breast Muscles. *Avian Dis* 59:410–418.

Voronov E, Shouval DS, Krelin Y, Cagnano E, Benharroch D, Iwakura Y, Dinarello CA & Apte RN (2003) IL-1 is required for tumor invasiveness and angiogenesis. *Proc Natl Acad Sci* 100:2645–2650.

White P, Thomas DW, Fong S, Stelnicki E, Meijlink F, Largman C, and Stephens P (2003) Deletion of the homeobox gene PRX-2 affects fetal but not adult fibroblast wound healing responses. *J Invest Dermatol* 120:135–144.

Zambonelli P, Gaffo E, Zappaterra M, Bortoluzzi S & Davoli R (2016) Transcriptional profiling of subcutaneous adipose tissue in Italian Large White pigs divergent for backfat thickness. *Anim Genet* 47:306–323.

Zapata I, Reddish JM, Miller MA, Lilburn MS & Wick M (2012) Comparative proteomic characterization of the sarcoplasmic proteins in the pectoralis major and supracoracoideus breast muscles in 2 chicken genotypes. *Poult Sci* 91:1654-1659.

Zhang Y, Wang J, Huang S, Zhu X, Liu J, Yang N, Song D, Wu R, Deng W, Skogerbø G, Wang XJ, Chen R & Zhu D (2009) Systematic identification and characterization of chicken (Gallus gallus) ncRNAs. *Nucl Acids Res* 37:6562–6574.

Table 1. Effect of breast abnormalities on weight, dimension and texture of raw chicken fillets

Trait	Breast meat category[1]		SEM	*P value*
	NORM	ABN		
Weight (g)	218.27	301.51	11.262	<0.0001
Top height (H1) [2] (mm)	37.82	44.31	0.959	0.0001
Middle height (H2) [3] (mm)	23.43	34.52	1.478	<0.0001
Bottom height (H3) [4] (mm)	7.60	12.70	0.808	0.0003
Length (mm)	185.03	194.10	2.815	0.1088
Length (mm/g)	85.42	64.76	2.738	<0.0001
Width (mm)	73.95	77.01	1.173	<0.0001
Width (mm/g)	34.37	25.64	1.279	0.0001

[1] NORM = without any abnormalities; ABN = with both white striping and wooden breast abnormalities. [2] H1 measured at the thickest point in the cranial part; [3] H2 measured at half distance of the breast length; [4] H3 measured as the vertical distance far from the end of the caudal part by

1 cm in a dorsal direction.

SEM = standard error of mean.

Table 2. Effect of breast abnormalities on quality traits of chicken meat

Trait	Breast meat category[1]		SEM	*P value*
	NORM	ABN		
Raw meat				
pHu	5.87	6.06	0.03	0.0039
CIE-L*	54.50	52.52	0.51	0.0486
CIE-a*	0.82	1.41	0.16	0.0568
CIE-b*	3.92	3.62	0.21	0.4800
Drip loss (%)	1.07	1.06	0.06	0.9166
Cooking loss (%)	21.45	34.04	1.73	<0.0001
Allo-Kramer shear force (kg/g)	4.26	7.54	0.68	0.0110
Marinated meat				
Uptake (%)	18.33	7.44	1.34	<0.0001
Cooking loss (%)	14.53	21.94	1.12	0.0001
Yield (%)	101.12	83.83	2.12	<0.0001
Allo-Kramer shear force (kg/g)	2.13	4.08	0.28	<0.0001

[1] NORM = without any abnormalities; ABN = with both white striping and wooden breast abnormalities.

*** = P < 0.001; * = P < 0.05; NS = not significant.

SEM = standard error of mean.

Table 3. Effect of breast abnormalities on chemical composition of chicken meat

Trait	Breast meat category[1]		SEM	*P value*
	NORM	ABN		
Moisture (%)	74.64	76.82	0.39	0.0014
Fat (%)	0.79	1.79	0.17	0.0007
Ash (%)	1.46	1.19	0.06	0.0138
Collagen (%)	1.16	1.35	0.04	0.0163
Protein (%)	23.37	18.45	0.69	<0.0001
Magnesium, Mg (mg/100g)	35.99	32.59	0.80	0.0272
Potassium, K (mg/100g)	359.31	362.96	6.59	0.7927
Phosphorus, P (mg/100g)	222.63	207.30	3.70	0.0326
Sodium, Na (mg/100g)	37.82	75.06	5.76	0.0001
Calcium, Ca (mg/100g)	7.81	11.32	0.69	0.0059

[1] NORM = without any abnormalities; ABN = with both white striping and wooden breast abnormalities.

SEM = standard error of mean.

Table 4. Effect of breast abnormalities on myofibrillar proteins composition of chicken meat

Protein (1)	Mol. Wt.	Breast meat category[1]		SEM	*P value*
		NORM	ABN		
LC3	16 kDa	12.41	15.68	1.07	0.1290
LC2	19 kDa	3.58	3.50	0.40	0.9208
LC1	27.5 kDa	13.71	8.01	1.09	0.0043
30 kDa troponin T fragment	29 kDa	4.00	4.96	0.29	0.1019
troponin T	34 kDa	4.81	4.67	0.30	0.8200
actin	42 kDa	33.81	36.88	1.77	0.4062
desmin	53 kDa	5.25	6.29	0.40	0.2022
70 kDa MHC fragment.	70 kDa	4.77	6.91	0.47	0.0161
MHC	220 kDa	16.19	13.12	1.77	0.4062

[1] NORM = without any abnormalities; ABN = with both white striping and wooden breast abnormalities.

SEM = standard error of mean.

(1) MHC = Myosin heavy chain; LC = Myosin light chain.

Table 5. Effect of breast abnormalities on sarcoplasmic proteins composition of chicken meat

Protein (1)	Mol. Wt.	Breast meat category[1]		SEM	*P value*
		NORM	ABN		
PGAM	25 kDa	7.44	5.54	0.33	0.0010
TPIS	26.4 kDa	7.16	6.50	0.24	0.1771
CA	31.8 kDa	9.42	8.27	0.58	0.3388
LDH	34 kDa	18.76	22.85	0.85	0.0104
G3P	36 kDa	10.33	13.26	0.62	0.0116
ALDO	39 kDa	7.06	10.42	0.62	0.0025
KCRM	43 kDa	11.23	9.89	0.32	0.0327
GPI	58 kDa	8.31	4.86	0.67	0.0053
KPYM	68 kDa	6.06	4.17	0.36	0.0037
PYGL	90 kDa	13.28	15.79	0.60	0.0301
AT2A2	114 kDa	0.00	1.86	0.35	0.0029

[1] NORM = without any abnormalities; ABN = with both white striping and wooden breast abnormalities.

SEM = standard error of mean.

(1) Abbreviations obtained from www.uniprot.org: PGAM = Phophoglycerate mutase; TPIS = Triosephosphate isomerase 1; CA = carbonic anhydrase; LDH = Lactate dehydrogenase; G3P = Glyceraldehyde dehydrogenase; ALDO = Aldolase; KCRM = Creatine kinase; GPI = Phosphoglucose isomerase; KPYM = Pyruvate kinase; PYGL = Glycogen phosphorylase; AT2A2 = Calcium-Transporting ATPase Sarcoplasmic Reticulum Type Slow Twitch Skeletal Muscle Isoform.

Table 6. DAVID functional clustering obtained considering all the differentially expressed genes

Category	Term	Count	P Value	Genes
Annotation Cluster 1	Enrichment Score: 1.63			
GOTERM_BP_ALL	GO:0032502~developmental process	43	2.39E-03	CHERP, PLXNA1, AGFG1, FHL1, UTRN, UGDH, ROMO1, PRRX2, MYT1, ADA, HEMGN, DFNB31, IGSF10, MUSK, BHLHA15, RASGRP1, PPL, IL1B, AP3D1, PLCD1, DOPEY2, ODF1, MB, APC, STX3, CRYAB, NEUROG1, SNAI2, DBH, CSRP3, UBP1, PPP1CB, EPHA2, SMTN, ATP2A2, GNAQ, PLN, USP22, HSPD1, GADD45B, SLITRK6, WNT7A, TMOD1
GOTERM_BP_ALL	GO:0007275~multicellular organismal development	38	8.33E-03	CHERP, PLXNA1, AGFG1, FHL1, UTRN, UGDH, PRRX2, MYT1, ADA, HEMGN, DFNB31, IGSF10, MUSK, PPL, IL1B, PLCD1, DOPEY2, ODF1, MB, APC, STX3, CRYAB, NEUROG1, SNAI2, DBH, CSRP3, UBP1, PPP1CB, EPHA2, SMTN, ATP2A2, GNAQ, PLN, HSPD1, USP22, GADD45B, SLITRK6, WNT7A
GOTERM_BP_ALL	GO:0030154~cell differentiation	24	1.65E-02	STX3, AGFG1, FHL1, UTRN, NEUROG1, MYT1, CSRP3, ADA, EPHA2, HEMGN, IGSF10, DFNB31, MUSK, GNAQ, BHLHA15, PPL, RASGRP1, ODF1, SLITRK6, GADD45B, WNT7A, TMOD1, APC, MB
GOTERM_BP_ALL	GO:0048869~cellular developmental process	24	2.55E-02	STX3, AGFG1, FHL1, UTRN, NEUROG1, MYT1, CSRP3, ADA, EPHA2, HEMGN, IGSF10, DFNB31, MUSK, GNAQ, BHLHA15, PPL, RASGRP1, ODF1, SLITRK6, GADD45B, WNT7A, TMOD1, APC, MB
GOTERM_BP_ALL	GO:0048856~anatomical structure development	32	3.22E-02	CHERP, AGFG1, FHL1, UTRN, UGDH, PRRX2, MYT1, ADA, DFNB31, IGSF10, MUSK, PPL, IL1B, PLCD1, MB, APC, STX3, CRYAB, NEUROG1, SNAI2, DBH, CSRP3, UBP1, EPHA2, SMTN, ATP2A2, GNAQ, PLN, HSPD1, SLITRK6, WNT7A, TMOD1
GOTERM_BP_ALL	GO:0032501~multicellular organismal process	48	4.83E-02	CHERP, PLXNA1, AGFG1, TACR2, FHL1, UTRN, UGDH, CNGB1, PRRX2, MYT1, ADA, KIFC3, HEMGN, DFNB31, IGSF10, MUSK, RAX2, PPL, P2RY1, IMPG2, IL1B, PLCD1, DOPEY2, ODF1, PPAP2A, PLCB2, MB, APC, STX3, CRYAB, KRT12, NEUROG1, SNAI2, DBH, CSRP3, UBP1, PPP1CB, EPHA2, SMTN, EPS8, ATP2A2, GNAQ, PLN, USP22,

				HSPD1, GADD45B, SLITRK6, WNT7A
Annotation Cluster 2	Enrichment Score: 1.18			
GOTERM_BP_ALL	GO:0007517~muscle organ development	7	1.12E-02	MUSK, SMTN, CRYAB, FHL1, PLN, UTRN, CSRP3
Annotation Cluster 3	Enrichment Score: 1.07			
GOTERM_BP_ALL	GO:0005976~polysaccharide metabolic process	5	1.70E-02	GBE1, GNE, UGDH, PPP1CB, CHST1
GOTERM_BP_ALL	GO:0006066~alcohol metabolic process	9	3.69E-02	GPD1L, ACAA2, GBE1, GNE, CRYAB, UGDH, DBH, PPP1CB, CHST1
GOTERM_BP_ALL	GO:0005996~monosaccharide metabolic process	6	4.74E-02	GBE1, GNE, CRYAB, UGDH, PPP1CB, CHST1
Annotation Cluster 4	Enrichment Score: 1.04			
GOTERM_BP_ALL	GO:0048514~blood vessel morphogenesis	6	3.96E-02	IL1B, PLCD1, PRRX2, DBH, CSRP3, UBP1
Annotation Cluster 5	Enrichment Score: 1.01			
GOTERM_BP_ALL	GO:0000302~response to reactive oxygen species	4	2.88E-02	CRYAB, ROMO1, ADA, MB
Annotation Cluster 6	Enrichment Score: 0.98			
GOTERM_BP_ALL	GO:0051726~regulation of cell cycle	8	2.78E-02	SKP2, IL1B, NEUROG1, USP22, OBFC2A, GADD45B, PPP1CB, APC

Annotation Cluster 7	Enrichment Score: 0.95			
KEGG_PATHWAY	hsa04020:Calcium signaling pathway	8	1.03E-03	GNAQ, ATP2A2, TACR2, PLN, HTR4, PLCD1, PTGFR, PLCB2
GOTERM_MF_ALL	GO:0004629~phospholipase C activity	3	2.89E-02	GNAQ, PLCD1, PLCB2
Annotation Cluster 8	Enrichment Score: 0.94			
GOTERM_BP_ALL	GO:0006406~mRNA export from nucleus	3	3.44E-02	ZC3H3, AGFG1, SMG7
GOTERM_BP_ALL	GO:0006913~nucleocytoplasmic transport	5	4.99E-02	ZC3H3, AGFG1, SMG7, SPTBN1, TOB1
Annotation Cluster 9	Enrichment Score: 0.93			
GOTERM_BP_ALL	GO:0046907~intracellular transport	14	5.44E-03	ZC3H3, SYNRG, STX3, AGFG1, AP1G1, SMG7, ATP2A2, BHLHA15, RANBP3, SPTBN1, AP3D1, GOSR2, DOPEY2, TOB1
GOTERM_BP_ALL	GO:0051649~establishment of localization in cell	15	1.85E-02	ZC3H3, SYNRG, STX3, AGFG1, AP1G1, SMG7, ATP2A2, BHLHA15, RANBP3, AP3D1, SPTBN1, GOSR2, DOPEY2, WNT7A, TOB1
GOTERM_BP_ALL	GO:0051641~cellular localization	15	3.51E-02	ZC3H3, SYNRG, STX3, AGFG1, AP1G1, SMG7, ATP2A2, BHLHA15, RANBP3, AP3D1, SPTBN1, GOSR2, DOPEY2, WNT7A, TOB1
Annotation Cluster 10	Enrichment Score: 0.74			
GOTERM_BP_ALL	GO:0002824~positive regulation of adaptive immune response based on somatic recombination of immune receptors built from	3	2.88E-02	IL1B, HSPD1, ADA

	immunoglobulin superfamily domains			
GOTERM_BP_ALL	GO:0050870~positive regulation of T cell activation	4	2.98E-02	AP3D1, IL1B, HSPD1, ADA
GOTERM_BP_ALL	GO:0002821~positive regulation of adaptive immune response	3	3.06E-02	IL1B, HSPD1, ADA
Annotation Cluster 11	Enrichment Score: 0.69			
GOTERM_BP_ALL	GO:0002026~regulation of the force of heart contraction	3	8.61E-03	ATP2A2, PLN, CSRP3
GOTERM_BP_ALL	GO:0008016~regulation of heart contraction	4	3.09E-02	ATP2A2, PLN, CSRP3, ADA
Annotation Cluster 12	Enrichment Score: 0.65			
GOTERM_BP_ALL	GO:0050953~sensory perception of light stimulus	6	4.30E-02	DFNB31, RAX2, IMPG2, KRT12, CNGB1, KIFC3
GOTERM_BP_ALL	GO:0007601~visual perception	6	4.30E-02	DFNB31, RAX2, IMPG2, KRT12, CNGB1, KIFC3
Annotation Cluster 13	Enrichment Score: 0.60			
GOTERM_BP_ALL	GO:0051726~regulation of cell cycle	8	2.78E-02	SKP2, IL1B, NEUROG1, USP22, OBFC2A, GADD45B, PPP1CB, APC
GOTERM_BP_ALL	GO:0043085~positive regulation of catalytic activity	10	4.14E-02	PSMB4, GNAQ, P2RY1, BIRC7, IL1B, HSPD1, PPAP2A, GADD45B, PPP1CB, PLCB2

Annotation Cluster 14	Enrichment Score: 0.55			
GOTERM_BP_ALL	GO:0003012~muscle system process	6	1.68E-02	SMTN, TACR2, CRYAB, UTRN, IL1B, MB

Table 7. Identification of the type of the regulated small nucleolar RNAs

Name GG	GB GG	nt GG	Name HS	GB HS	nt HS	Coverage (%)	Identity (%)	Full name
GGN5	EU240224	135	SNORA55	NR_002983	137	99	81	small nucleolar RNA, H/ACA box 55
GGN50	EU240266	132	SNORA22	NR_002961	134	99	82	small nucleolar RNA, H/ACA box 22
GGN73	EU240289	137	SNORA46	NR_002978	135	57	84	small nucleolar RNA, H/ACA box 46
GGN98	EU240313	129	SNORA28	NR_002964	126	38	84	small nucleolar RNA, H/ACA box 28

GG = *Gallus gallus*

HS = *Homo sapiens*

Table 8. Predicted overexpressed target genes found in miRBase for the differentially expressed miRNAs

gga-miR-196-5p	gga-miR-205a	gga-miR-1600	gga-miR-1716	gga-miR-1805-3p
FAM64A	*LOC425021*	*ANKRD31*	-	*WDFY3*
	NMT2	*PLN*		
	PLXNA1	*PRRX2*		

Figures

Figure 1. Validation by qPCR of five differentially expressed genes obtained by microarray analysis.

(A) Fold changes values obtained from microarrays, (black bars) and from qPCR data (white bars), for the five tested genes; (B) scatterplot showing the good correlation between the fold changes values calculated with the two experimental methods.

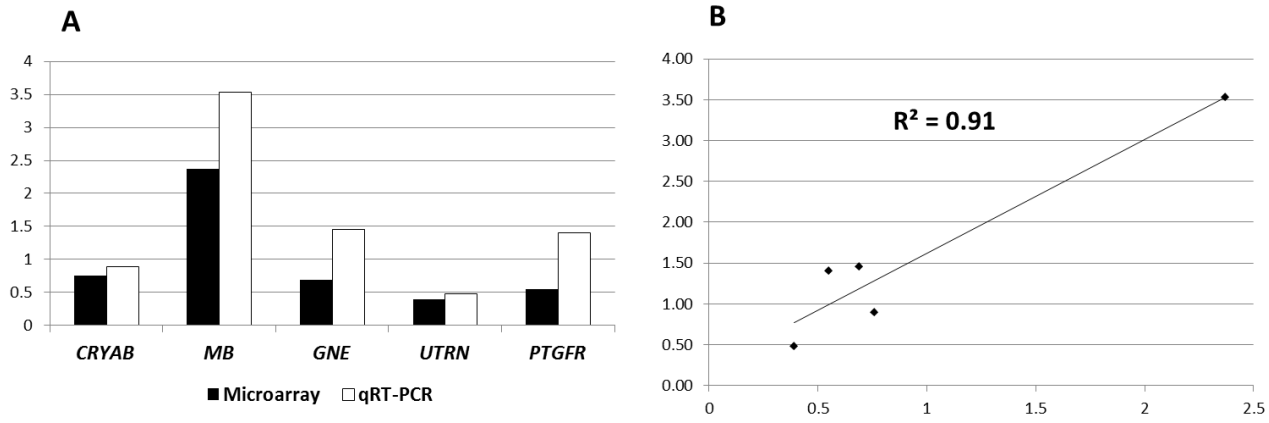
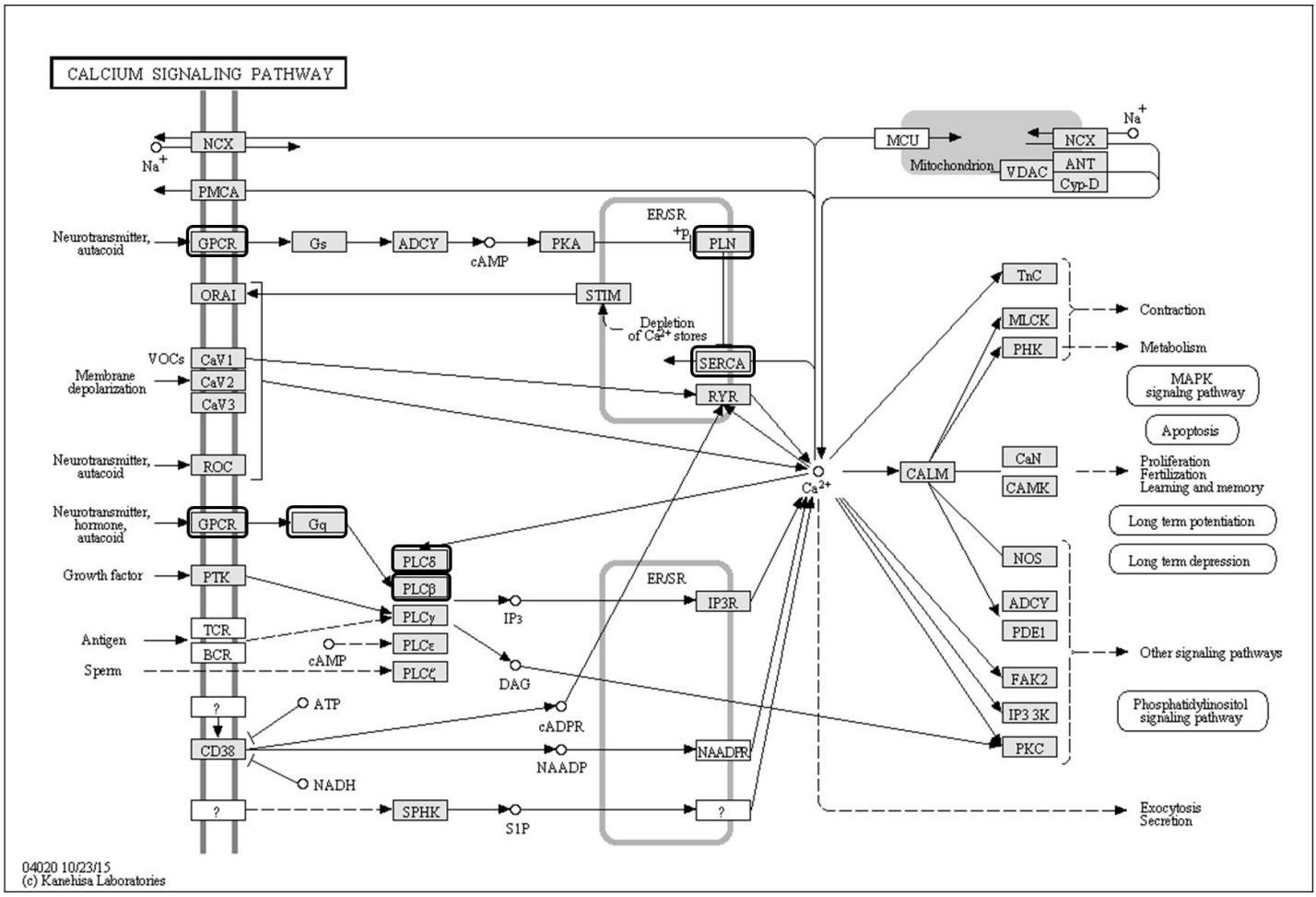


Figure 2. KEGG table representing the calcium signalling pathway. The differentially expressed genes found are indicated with black circles.



GPCR = *PTGFR*; Gq = *GNAQ*; PLCδ = *PLCD1*; PLCβ = *PLCB2*; PLN = *PLN*; SERCA = *ATP2A2*

Figure 3. A putative mechanism explaining some of the changes occurring in abnormal muscle samples deduced from the activities of the differentially expressed genes involved in calcium signalling pathway.

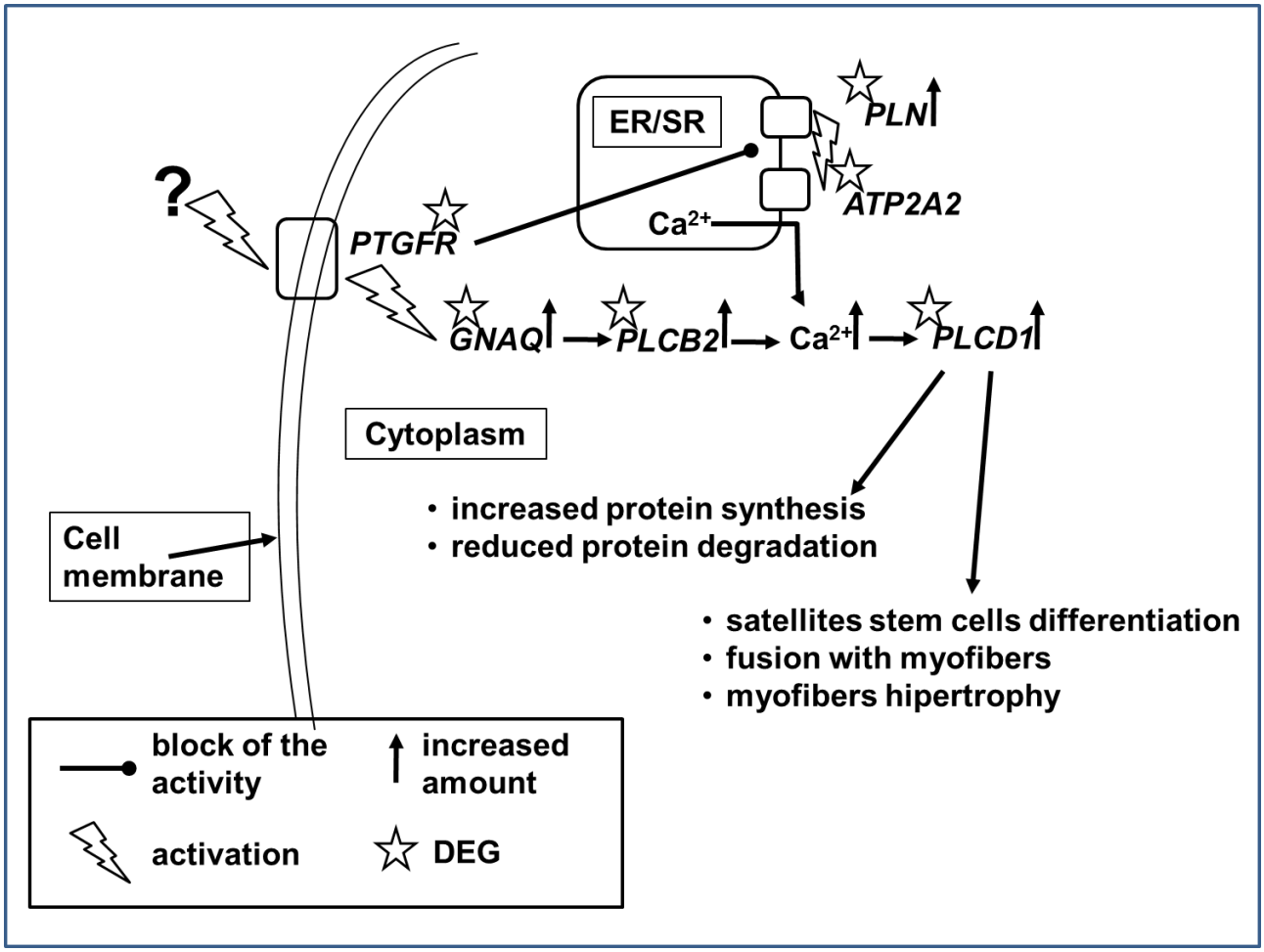
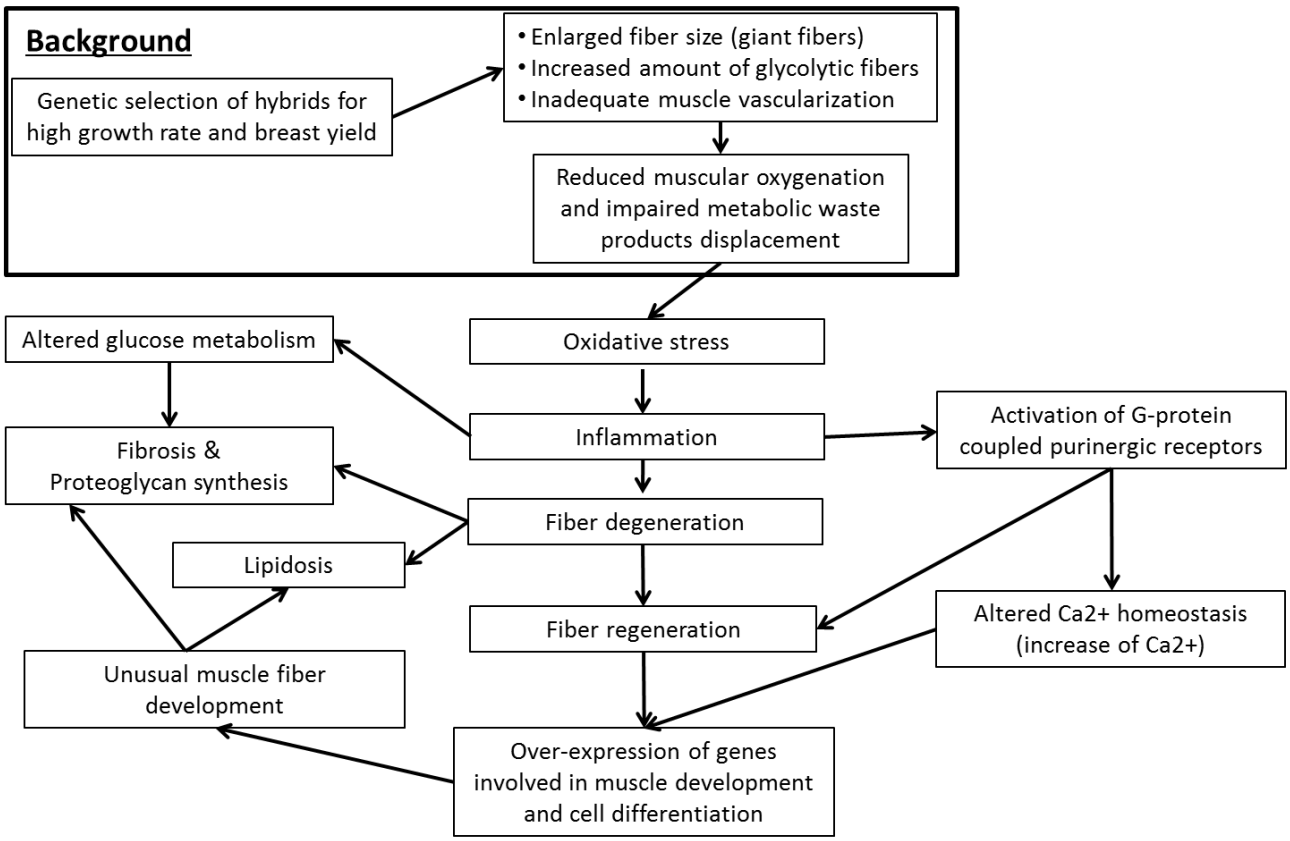


Figure 4. A schematic representation of one of the possible aetiologies at the basis of white striping and wooden breast abnormalities.



Supplementary Table S1. Primers and PCR conditions used for the validation performed using qPCR.

Gene (symbol)		Primer Forward	Primer reverse	Amplicon (nt)	Annealing temp. (°C)	Mg (mM)
Normalizing genes						
glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*)	EXT	GTCTGGAGAAACCAGCCAAG	AACTGAGCGGTGGTGAAGAG	344	67	3
	INT	TGACAGCCATTCCTCCAC	TGGACCATCAAGTCCACAAC	126	66	-
ribosomal protein L32 (*RPL32*)	EXT	AACAGAGTTCGCAGGAGGTT	TTGGTGATCTTGATGGCGAG	245	66	3
	INT(*)	ATGGGAGCAACAAGAAGACG	TTGGAAGACACGTTGTGAGC	139	66	-
tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*)	EXT	TTGATCCCCAATGCTTCGC	CCGATGTCCACAATGTCAAGT	379	66	2
	INT(*)	TTGCTGCTGGAGATGACAAG	CTTCTTGATACGCCTGTTG	60	66	-
Differentially expressed genes						
ATPase, Ca++ transporting, cardiac muscle, slow twitch 2 (*ATP2A2*)	EXT	GTGCCCCTGAAGGTGTAATC	CTTTGGTAGAAACGTCCTCATC	408	65	1.5
	INT	GTTGGAAACGCCAAAATACC	CGCAGTGTGTCTCTACCAGTTC	92	66	-
crystallin, alpha B (*CRYAB*)	EXT	CGCCTGAGGAGCTAAAAGTG	CTGGCATCCAATGAGAAGTG	331	65	1.5
	INT	GGAGTTCAGCAGGAAATACAGG	CAGGACACCATCCAGAGAGAG	82	66	-
glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase (*GNE*)	EXT	TTCGGGTGATGAGGAAGAAG	GTCCCCAGGTTGATGACAGG	175	66	2
	INT	TCCCATTTGACCAGTTCATTC	ATGCTCCCACTTCTCTGACC	87	66	-
myoglobin (*MB*)	EXT	GGCAACAAGTCCTCACCATC	TTGACTGGGATTTTGTGCTTC	276	66	3

prostaglandin F receptor (FP) (*PTGFR*)	INT	CCATCTGGGGAAAAGTGGAG	TCCAAAGTCTCAGGGTGGTC	80	66	-
	EXT	TGAGTCACCCGTAAGTTCCA	AGTGCATCTTCTGGAGTAGC	222	63	3
	INT	AAATCTACAGCCGCTCAGC	GCCCGTCTATGAGCATTGC	85	66	-
utrophin (*UTRN*)	EXT	GTAGGGATGCCGCAGATTT	GCAGGAAGTCCCTCTTTTCC	374	66	2
	INT	CCAGCGGATAGTGAAGCAA	CCTGAATGGCAGCTTTTGA	88	66	-

(*) These primer pairs are from Bagés et al., 2015.

EXT: primer pairs used for the amplification of a larger PCR product

INT: primer pairs used for the creation of the standard curve and for the qPCR analysis

Supplementary Table S2. List of the differentially expressed genes obtained by microarray analysis.

Supplementary data is available at PSA Journal online.

General conclusions

The studies reported in the present thesis were addressed towards the application of OMICs tehnologies (mainly genomics, functional genomics and transcriptomics) to deepen the knowledge of the chromosomal regions, transcripts and candidate genes affecting complex traits in livestock species. The main topic of the research was to investigate the genetic causes determining variations in porcine fat traits. To achieve this objective, different strategies have been used: OMICs technologies showed to be a useful tool for the detection of markers and transcripts related to phenotypic variations in porcine backfat. The results of Genome-Wide Association study (GWAS) identified more than thirty different genomic regions associated with backfat composition and thickness. About half of these regions were consistent with (or directly related to) QTLs found in previous studies, and six regions were new and not yet described in QTLdb. Among the identified genomic *loci*, one region located on porcine chromosome 8 showed the strongest associations with backfat fatty acid composition. This region harbours the gene *ELOVL elongase 6*. Through this approach, it was possible to strengthen the hypothesis that *ELOVL6* plays an important role in determining pig backfat fatty acid composition, affecting in particular the content of medium and long chain fatty acids. Furthermore, a transcriptome analysis on backfat tissue was performed, and revealed 86 differentially expressed transcripts, 72 more highly expressed in animals with increased fat deposition and 14 up-regulated in lean individuals. Almost 80% of the detected transcripts corresponded to non-annotated isoforms. The results of the transcriptome analysis also indicated *Perilipin 2* (*PLIN2*) as a candidate gene for backfat deposition: three new transcripts of *PLIN2* gene have been detected in our study, and one of them showed to be expressed two times more in fat than in lean pigs. This outcome suggests that *PLIN2* is strongly involved in the metabolic cascade related to subcutaneous fat tissue deposition. For the study of some candidate genes obtained from OMICs approaches, the GWAS and transcriptome information was integrated with focused association and expression studies. In particular, *Perilipins* were further investigated in order to deepen the information about these genes, poorly studied in pigs, but known in other species to be very important for the modulation of intracellular fat storage. The results obtained on this gene family can contribute to the knowledge of perilipin proteins role in the control of pig fat traits and to better comprehend their connections with other genes involved in lipid metabolism.

Mutations on *PLIN1* and *PLIN2* gene sequences have been studied in Spanish Duroc pigs, and the obtained associations highlighted an additive effect on backfat thickness at different ages for the SNP *GU461317:g.98G>A* located in *PLIN2* 3' untranslated region. On the whole, these evidences confirmed the important role that *PLIN2* plays in adipocytes metabolism, suggesting that this gene may be one of the main actors controlling adipose tissue energy storage. On the other hand, the research on porcine *PLIN5* confirmed the strong relationship existing between *PLIN5* and lipases: a SNP located in *PLIN5* gene downstream region was studied and resulted to be associated with changes in *LIPE* gene expression. These results may be of particular interest considering that in another study we found LIPE protein related to changes in pig *Semimembranosus* intramuscular fat content. If confirmed, the results imply that mutations in *PLIN5* gene sequence may affect lipases gene activity and indirectly the intracellular content of stored lipids. The association found between the SNP lying in the downstream region of *PLIN5* gene and the transcription levels of *LIPE* gene was observed in both Italian Duroc and Italian Large White pigs, while *PLIN5* gene expression levels differed between the two breeds. On the whole, samples belonging to Italian Duroc and Italian Large White breeds presented differences in the expression levels of several genes involved in energy and fat metabolism: in particular, *PLIN5* and *ACLY* showed transcription levels differing between Italian Duroc and Italian Large White *Semimembranosus* muscle, and *FASN*, *ACLY* and *ACACA* genes had distinct transcription levels in backfat samples. Therefore, our results confirm the evidences obtained in other studies: adipogenesis and fat metabolism differ among distinct breeds, suggesting that there may be breed-specific genetic and metabolic backgrounds affecting fatness. The different metabolic aptitude between breeds to deposit lean and fat tissues can be used for the understanding of the molecular pathways leading to a switch in the use of energy from adipogenesis towards muscle deposition. On the whole, the application of genomics and transcriptomics to the study of pig fat traits has proved to be an important tool for increasing the knowledge of the fat-related genes and can help improving livestock performances and animal products quality.

OMICs technologies were also used to investigate colostrum composition in a preliminary study. This explorative research demonstrated that *colostrum* metabolome is affected by breed, with Duroc sows showing *colostrum* compositions unlike any other. This evidence represents a preliminary result and suggests that variations in *colostrum* composition could have a genetic basis. The comprehension of the genetic determinism affecting this secretion would provide genes and markers extremely important for the deciphering of the genetic background of pig maternal

traits. However, further studies to investigate whether breed could affect also *colostrum* fat composition are in progress.

Furthermore, in addition to transcriptomics, a gene expression profiling was also assessed using microarray technology to investigate the onset of white striping wooden breast myopathies in chickens. The results returned 204 differentially expressed genes, mainly belonging to pathways associated with muscle development, polysaccharide metabolic processes, proteoglycans synthesis, inflammation and calcium signaling pathway. These data suggest that the considered myopathies have a complex multifactorial etiology; anyway the microarray analysis highlighted that one possible explanation for the occurrence of this pathology may be an insufficient vascularization of breast muscle, likely consequent to abnormal activities of angiogenic processes and genes (such as *IL1B*).

On the whole, the application of genomics, transcriptomics and metabolomics showed to be an effective tool to identify genetic variations associated to complex traits in different livestock species, and to detect genes and markers involved in phenotypic variations. The combined approach of OMICs technologies and single gene studies can be considered a useful strategy to overcome the limits of traditional genetic selection in livestock, permitting a deeper deciphering of genetics of polygenic and low heritable traits. Nevertheless, further studies are needed to achieve a deeper comprehension of the molecular pathways leading to the different productive performances in livestock species, since new information could provide tools useful for a more effective genetic selection.