UAB
Universitat Autònoma de Barcelona

UNIVERSITAT AUTÒNOMA DE BARCELONA
Departament de Ciència Animal i dels Aliments, Facultat de Veterinària

CENTRE DE RECERCA EN AGRIGENÒMICA
Plant and Animals Genomics Program

Exploring the Genome and Links Between Transcriptome and Fatty Acid Profiling of Muscle in Pigs

Jesús Valdés Hernández

Doctoral thesis to obtain the PhD degree in Animal Production of the Universitat Autònoma de Barcelona, March 2023

Supervisors:

Dr. Josep Maria Folch Albareda   **Dr. Yuliaxis Ramayo Caldas**

El Dr. Josep M. Folch Albareda, professor del

Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de

Barcelona, i el Dr. Yuliaxis Ramayo Caldas, investigador del Departament de Genètica i

Millora Animal, Institut de Recerca i Tecnologia

Agraroalimentàries (IRTA),

fan constar

que el treball de recerca i la redacció de la memòria de la tesi doctoral titulada "Exploring the genome and links between

transcriptome and fatty acid profiling of

muscle in pigs" han estat

realitzats sota la seva direcció per

JESUS VALDES HERNANDEZ

i certifiquen

que aquest treball s'ha dut a terme al Departament de Ciència Animal i del Aliments de la

Facultat de Veterinària de la Universitat Autònoma de Barcelona i al Programa de Plant and

Animal Genomics del Centre de Recerca en Agrigenòmica,

i consideren

que la memòria resultant es apta per optar al grau de Doctor en Producció Animal per la

Universitat Autònoma de Barcelona.

I perquè en quedi constància, signen aquest document a Bellaterra, a 16 de març del

2023.

Dedicada a mis padres, pues aunque ellos ya no están en el planeta tierra, y aún desde el más

allá, siempre me alumbran y me gian por el buen camnino.

"La vida no es lo que uno se representa, sino lo que se presenta"

— Jesús Valdés Hernández

Content

Summary

The intramuscular fat content and fatty acid (FA) composition are important parameters that affect meat quality, influencing its organoleptic characteristics and nutritional value. FAs provided by the diet or derived from *de novo lipogenesis* are transformed by the action of several genes, like those encoding desaturases and elongases of FAs in lipogenic tissues. In the current thesis, several studies were performed from the pig muscle transcriptome data (by RNA-Seq) and its FA composition (by gas chromatography) in Iberian × Duroc backcross (BC1_DU) pigs. Hence, in the context of multifactorial design, both univariate (work I and II) and multivariate (work III) statistical approaches are applied. Complementary analyses of gene ontology (GO), concordance, correlation, networks and regulatory impact factor (RIF) were performed.

In the first study, we identified 81 differentially expressed (DE) genes between the muscle transcriptome of divergent pigs (16 high and 16 low, H vs. L) for intramuscular oleic-to-stearic (C18:1n-9/C18:0) ratio. These DE genes included 57 down-regulated and 24 up-regulated, and functional analyses revealed that the PPAR signaling pathway (*PPARG, SCD, PLIN1,* and *FABP3*) was over-represented. In addition, other DE genes related to energy, lipid and carbohydrate metabolism, and mitochondrial FA beta-oxidation (*ACAD10, ACADVL, ECHDC3, FADS3, ILVBL,* and *MMAA*) were identified. Both, *SCD* and *PPARG* genes showed a higher expression in H animals and are relevant candidate genes affecting the biosynthesis and Δ9-desaturation of FAs. SCD is a ratio-limiting enzyme that catalyzes the conversion of C18:0 to C18:1n-9, while the induction of *SCD* expression appears to be regulated by the nuclear receptor PPARG (i.e., master regulator of adipogenesis). Moreover, PUFA was suggested to inhibit the transcription of *PPARG* gene, and was associated with the C18:1n-9/C18:0 ratio. Also, two haplotypes of seven SNPs located in the *SCD* gene were associated with the C18:1n-9/C18:0 ratio variation and the *SCD* gene expression.

The second study was focussed to the global association between muscle transcriptome and intramuscular FA profiles. Here, we used 129 BC1_DU animals with the expression values of 14,870 transcripts and 36 FA-related traits, including individual FA composition, sum of FAs (SFA, MUFA, and PUFA), FA ratios, and FA metabolic indices. We detected a variable number of associated genes across FAs (1,022 genes in total with 21 of 36 FA traits). However, only

53.52% of gene expression values (547/1022 genes) significantly correlated (range [-0.19 to 0.51]) with FA phenotypic values (21 FA traits). Our analysis revealed several candidate genes linked to FA composition in muscle, including well-known (e.g., *ACSL1, ELOVL6, FBP1, GOT1, LEP, LGALS12, LPL, MDH1, PLIN1, SC5D, SLC27A1*, and *TFRC*) and novel candidate genes (e.g., *TRARG1,* and long non-coding RNAs as *TANK, ENSSSCG00000011196,* and *ENSSSCG00000038429*). Furthermore, we identified gene regulators including transcription factors (e.g., *ESRRA* and *LBX1*) and co-factors (*NCOA2* and *LPIN1*) involved in FA metabolism. Together, FA-associated genes were present in 75 over-represented GO terms, including, signal transduction, oxidative phosphorylation, citrate cycle (TCA cycle), long-chain fatty acid import, and PPAR signaling pathway.

Finally, in the third study, we performed a multivariate analysis for the integration of intramuscular FA and gene expression profiles. In this study, we used 129 BC1_DU pigs with RNA-Seq expression data, but only 15 individual FA phenotypes. Our results revealed a subset of 378 canonical variables (13 FAs and 365 genes) that maximize the correlation (range [-0.39 to 0.41]) between the two datasets. In particular, six FAs (C20:4n-6, C18:2n-6, C20:3n-6, C18:1n-9, C18:0, and C16:1n-7) were among the most interconnected variables in the network. Among FA-correlated genes identified, several were related to lipid and/or carbohydrate metabolism (e.g., *ADIPOQ, CYCS, CYP4B1, ELOVL6, FBP1, G0S2, HMGCR, LEP, LGALS12, LPIN1, PLIN1, PNPLA8, PPP1R1B, SDHD, SDR16C5, SFRP5, SOD3,* and *TFRC*), meat quality (*GALNT15, GOT1, MDH1, NEU3,* and *PDHA1*), and transport of molecules (e.g., *EXOC7* and *SLC44A2*). Functional analysis with FA-correlated genes indicated 55 GO terms over-represented, including several biological processes and pathways, including mitochondrial gene expression, TCA cycle, regulation of lipolysis in adipocytes and insulin signaling pathways. Lastly, RIF analysis suggested the relevance of six transcription factors (*CARHSP1, LBX1, MAFA, PAX7, SIX5,* and *TADA2A*) as putative regulators of gene expression and intramuscular FA composition.

Resumen

El contenido de la grasa intramuscular y la composición de ácidos grasos (AG) son parámetros importantes que afectan a la calidad de la carne, influyendo tanto en sus características organolépticas como sobre su valor nutricional. Los AG provenientes de la dieta o derivados de la lipogénesis *de novo* son transformados por la acción de varios genes, por ejemplo, los genes que codifican desaturasas y elongasas de AG en tejidos lipogénicos. En la presente tesis, se realizaron varios estudios a partir de los datos del transcriptoma del músculo porcino (mediante RNA-Seq) y su composición de AG (mediante cromatografía de gases) en cerdos retrocruzados (Ibérico $\times$ Duroc, BC1_DU). Se aplicaron enfoques estadísticos tanto univariantes (trabajos I y II) como multivariantes (trabajo III). Asimismo, se realizaron análisis complementarios como los de ontología génica (OG), concordancia, correlación, redes, e identificación de posibles reguladores (FRI).

En el primer estudio identificamos 81 genes diferencialmente expresados (DE) entre el transcriptoma de músculo de cerdos divergentes (16 altos y 16 bajos, A vs. B) para la ratio intramuscular entre los AG oleico y esteárico (C18:1n-9/C18:0). Los análisis funcionales revelaron que la vía de señalización PPAR (*PPARG, SCD, PLIN1* y *FABP3*) fue sobrerrepresentada. Además, se identificaron otros genes DE relacionados con la energía, el metabolismo de lípidos y carbohidratos, y la betaoxidación mitocondrial del AG (*ACAD10, ACADVL, ECHDC3, FADS3, ILVBL* y *MMAA*). Tanto los genes *SCD* como *PPARG* mostraron una mayor expresión en los animales del grupo A y ambos son genes candidatos relevantes que afectan a la biosíntesis y la Δ9-desaturación de los AG. SCD es una enzima ratio-limitante que cataliza la conversión de C18:0 en C18:1n-9, mientras que la inducción de la expresión de *SCD* puede estar regulada por el receptor nuclear PPARG (máster regulador de la adipogénesis). Asimismo, los PUFA pueden inhibir la transcripción de *PPARG* y también se han asociado con el ratio C18:1n-9/C18:0. Además, dos haplotipos formados por siete polimorfismos localizados en el gen *SCD* se asociaron con la variación de la ratio C18:1n-9/C18:0, así como con la expresión del *SCD*.

El segundo estudio se enfocó en la asociación global entre los perfiles del transcriptoma de músculo y de AG intramuscular. Para ello utilizamos 129 animales BC1_DU con los valores de expresión de 14.870 transcritos y 36 caracteres relacionados con la composición de AG,

incluyendo la composición individual de los AG, la suma de los AG (saturados, monoinsaturados y poliinsaturados), los ratios entre los AG y los índices metabólicos de AG. Se identificaron un número variable de genes asociados a los caracteres de AG (1.022 genes en total con 21 de 36 caracteres de AG). Sin embargo, sólo el 53,52% de los valores de expresión génica (547/1.022 genes) se correlacionaron significativamente (rango [-0,19 a 0,51]) con valores fenotípicos de AG (21 caracteres de AG). Nuestro análisis reveló varios genes candidatos relacionados con la composición del AG en el músculo, incluyendo genes conocidos (por ejemplo, *ACSL1, ELOVL6, FBP1, GOT1, LEP, LGALS12, LPL, MDH1, PLIN1, SC5D, SLC27A1* y *TFRC*) y nuevos genes candidatos (por ejemplo, *TRARG1,* y ARN largos no codificantes como *TANK, ENSSSCG00000011196* y *ENSSSCG00000038429*). Además, identificamos reguladores génicos que incluyeron factores de transcripción (por ejemplo, *ESRRA* y *LBX1*) y cofactores (*NCOA2* y *LPIN1*) implicados en el metabolismo de AG. En conjunto, los genes asociados a los AG forman parte de 75 términos de OG sobrerrepresentados, incluyendo la transducción de señales, fosforilación oxidativa, ciclo del ácido tricarboxílico (ciclo TCA), entrada de AG de cadena larga y la vía de señalización PPAR.

Finalmente, en el tercer estudio realizamos un análisis multivariado para integrar los perfiles de AG intramusculares y de expresión génica. En este estudio utilizamos los mismos 129 cerdos BC1_DU con datos de expresión de RNA-Seq pero centrado sólo en 15 AG. Nuestros resultados revelaron un subconjunto de 378 variables canónicas (13 AG y 365 genes) que maximizan la correlación (rango [-0.39 a 0.41]) entre ambos conjuntos de datos. En particular, seis AG (C20:4n-6, C18:2n-6, C20:3n-6, C18:1n-9, C18:0 y C16:1n-7) fueron las variables más interconectadas de la red. Entre los genes asociados con los AG identificados varios relacionados con el metabolismo de los lípidos y/o los carbohidratos (por ejemplo, *ADIPOQ, CYCS, CYP4B1, ELOVL6, FBP1, G0S2, HMGCR, LEP, LGALS12, LPIN1, PLIN1, PNPLA8, PPP1R1B, SDHD, SDR16C5, SFRP5, SOD3* y *TFRC*), la calidad de la carne (*GALNT15, GOT1, MDH1, NEU3* y *PDHA1*) y el transporte de moléculas (por ejemplo, *EXOC7* y *SLC44A2*). El análisis funcional basado en los genes relacionados a los AG indicó 55 términos de OG sobrerrepresentados, entre ellos varios procesos y rutas biológicas, como la expresión génica mitocondrial, ciclo TCA, y las vías de regulación de la lipólisis en adipocitos y de señalización de la insulina. Por último, el análisis RIF sugirió la relevancia de seis factores de transcripción (*CARHSP1, LBX1, MAFA, PAX7, SIX5* y *TADA2A*) como posibles reguladores de la expresión génica y de la composición intramuscular de AG.

List of Tables

List of Figures

List of Papers and Studies

The present thesis is based on the work contained in the list of articles below:

Paper I. **Valdés-Hernández J.**, Ramayo-Caldas Y., Passols M., Criado-Mesas L., Castelló A., Esteve-Codina A., Sánchez A., and Folch JM. (2023). "Identification of Differentially Expressed Genes and Polymorphisms Related to Intramuscular Oleic-to-Stearic Fatty Acid Ratio in Pigs". Manuscript in preparation.

Paper II. **Valdés-Hernández J.**, Ramayo-Caldas Y., Passols M., Sebastià C., Criado-Mesas L., Crespo-Piazuelo D., Esteve-Codina A., Castelló A., Sánchez A., and Folch JM. (2023). "Global analysis of the association between pig muscle fatty acid composition and gene expression using RNA-Seq". Scientific Reports, 13, 535. https://doi.org/10.1038/s41598-022-27016-x.

Paper III. **Valdés-Hernández J.** Folch JM., Crespo-Piazuelo D., Passols M., Sebastià C., Criado-Mesas L., Castelló A., Sánchez A., and Ramayo-Caldas Y. (2023). "Identification of candidate regulatory genes for intramuscular fatty acid composition by transcriptome analysis in pigs". Genetics Selection Evolution (No. GSEV-D-23-00032, Under Review).

Abbreviations

BC1_DU	25% Iberian × 75% Duroc backcross
BC1_LD	25% Iberian × 75% Landrace backcross
BC1_PI	25% Iberian × 75% Pietrain backcross
3BCs	The three backcrosses together
LD muscle	*Longissimus dorsi* muscle
cDNA	Complementary DNA
mRNA	Messenger RNA
mtRNA	Mitochondrial RNA
PCR	Polymerase-chain reaction
RT-qPCR	High-throughput real-time quantitative PCR
NQ	Normalized quantity values of mRNA expression applying the relative standard curve method
RNA-Seq	RNA-sequencing
NGS	Next generation sequencing
TGS	Third-generation sequencing
SSC	*Sus scrofa* chromosome
SNP	Single nucleotide polymorphism
GWAS	Genome-wide association study
WGS	Whole genome shotgun sequencing
QTL	Quantitative trait *loci*
eQTL	Expression quantitative trait *loci*
DESeq2	Differential gene expression analysis based on the negative binomial distribution
ELMSeq	Algorithm referred with the abbreviation ELMSeq, an extended linear model for RNA-Seq data analysis
rCCA	Regularized canonical correlation analysis implemented in mixOmics package
FC	Fold change

P-adj.	Multiple testing adjusted *P*-value from DESeq2
q-value	Multiple testing corrected *P*-value
r	Pearson's correlation coefficient
PCA	Principal components analysis
CPM	Counts per million normalization
PCIT	Partial correlation coefficient with information theory
RIF	Regulatory Impact Factor
DW	Differential wiring as the correlation of the TF with all possible target genes
FAs	Fatty acids
SFA	Saturated fatty acids
MUFA	Monounsaturated fatty acids
PUFA	Polyunsaturated fatty acids
DE genes	Differentially expressed genes
TFs	Transcription factors
lncRNA	Long non-coding RNAs

_____

Chapter 1. General Introduction

1.1. Background and importance of pork production

An important source of animal protein worldwide is the production of pork. Nowadays, world meat consumption is increasing, being poultry and pork meats the most commonly consumed worldwide (FAOSTAT, 2022). According to Mote & Rothschild (2020), pork accounts for nearly 43% of all red meat consumed globally. However, such consumption differs between countries, regions, and human societies. This is perhaps due, for example, to people's income and population growth, among other factors. For instance, in Europe and North America the increase in meat consumption has stabilized while in other regions there is an increase (e.g., in Central and South America and rest of Asia), especially in China that is the top of meat consumption (Godfray et al., 2018).

From the current health crisis caused by *coronavirus disease* (COVID-19) pandemic, i.e., recognized by the World Health Organization (WHO, 2019), the importance of the pork industry, together with poultry and beef in meeting the food needs of an expanding demography has been further reaffirmed. Both the general public and scientific communities have become even more aware, to address the situation of increasing demand for animal protein, to improve the quality of health in general, as well as to reduce the environmental impact of livestock production. Because of this, there is a growing interest and concern for animal products that must respond to these region-specific challenges.

In 2020, approximately 109,835,405 tonnes (tons) of pork meat were produced worldwide, where Europe production contributed an average of 30,315,325 tons of pork meat (**Figure 1**). Worldwide, the main producers of pork are Asia (48.50%), followed by Europe (27.60%) and America (22%) (**Figure 2**). In the European Union (EU-27 list), the main producers are Germany, Spain, and France with 5,118,000, 5,003,430, and 2,201,110 tons produced, respectively (FAOSTAT, 2022). Likewise, Spain is among the top 10 meat pig producers worldwide, being the fourth largest producer after China, USA, and Germany. The pork sector is of key importance to the Spanish economy, accounting for around 14% of final agricultural production. Within livestock, the pig sector ranks first according to economic importance, accounting around 39% of final agricultural production (MAPA, 2021).

_____
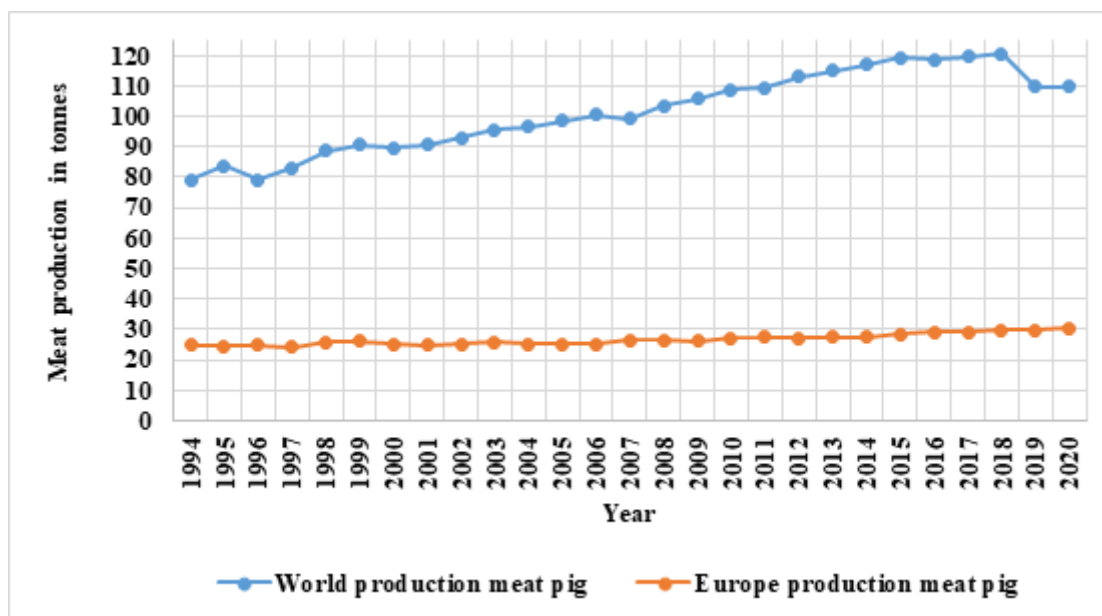
Figure 1. Evolution of pig-meat production both worldwide and Europe since 1994 until 2022. Source: FAOSTAT: <https://www.fao.org/faostat/en/#data/QCL/visualize> (Jul 13, 2022).
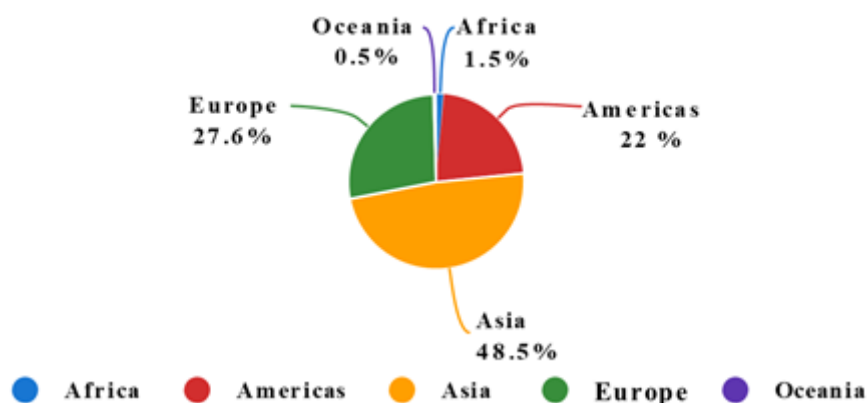
Figure 2. Production share of meat pig by region in 2020. Source: FAOSTAT: <https://www.fao.org/faostat/en/#data/QCL/visualize> (Jul 13, 2022).

According to these official data (MAPA, 2021), Spain is the first country in the EU in terms of census, with about 21% of the EU census. Furthermore, it is the second largest pork producer and exporter in the EU, after Germany, consolidating its position in the world market with a positive commercial balance. Likewise, Spain has spectacularly increased exports to third countries, especially to China and other Southeast Asian countries. At the national level, provisional data from the livestock slaughter survey for 2021, put pork production at record

levels, with more than 58.30 million animals slaughtered and around 5.20 million tons of meat produced (MAPA, 2021).

Within Spain, the higher pork production is concentrated in the autonomous communities of Catalunya (19.80%), Aragon (17.70%) and Castile-Leon (12.20%). Therefore, these three regions together represent 49.70% of the national pork production distribution (**Figure 3**). In addition, in Catalunya the latest data compiled by IDESCAT (*Instituto de Estadística de Catalunya*) in 2020 indicated a total of 8,072,041 head of porcine, which include the sum of the categories of piglets, pigs < 50 kg, fattening pigs and breeding pigs (IDESCAT, 2020).
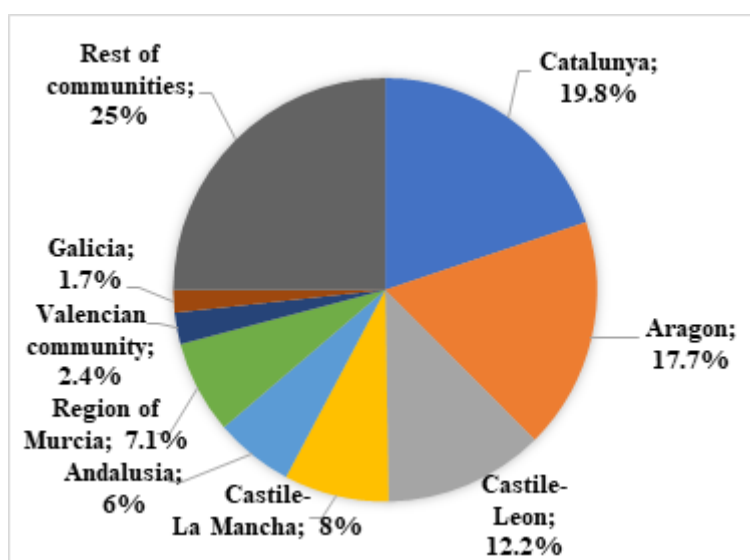
Figure 3. Distribution of total pork production by autonomous community in 2021. Source: MAPA:<https://www.mapa.gob.es/va/ganaderia/temas/produccion-y-mercados-ganaderos/indicadoressectorporcino2021_tcm39-564427.pdf> (Jul 16, 2022).

1.1.1. Progress in animal production

In general, animal production is supported by four important pillars: genetics, nutrition, management and health, which have significantly contributed to improve productivity and economic indicators. The optimization of feed production and factory farming methods have also contributed to the increased productivity. For instance, nutrition and genetics have been used mainly as independent disciplines. Utilizing animal breeding and feeding was one of the important advancements in animal production. Nutrition is a powerful tool to modify animal

_____

growth and the partition of energy between lean and fat tissues, thus affecting body and tissue composition (Amills et al., 2020), which can also be used to improve carcass leanness, meat's nutritional value, and overall production efficiency.

Concerning animal breeding, selection allowed a great increase in the productive efficiency of modern pig breeds, as well as crosses between breeds and synthetic lines. Among the modern pig breeds **(Figure 4)**, the Iberian breed is one of the scarce local swine breeds (Lopez-Bote, 1998), it is a fatty breed that has a higher feed intake, a higher genetic capacity to fat deposition (adipogenic potential), and backfat thickness; whereas Duroc breed have higher ham weights. For instance, Iberian pigs compared to Duroc pigs have been considered to have a more persistent ability of desaturation during periods of fasting. Inhibition of adipogenic and lipogenic genes after of fasting (24 h) resulted more intense in Duroc vs. Iberian pigs (Benítez et al., 2017). According to Liu et al. (2015), Duroc pigs are fattier and have lower muscle mass but a higher percentage of slow-twitch oxidative muscle fibers compared to Pietrain pigs, which are more muscular and have a higher lean meat percentage and more fast-twitch glycolytic fibers. Meanwhile, pig production centers on the utilization of crossbred animals taking advantage of the heterosis (Knol et al., 2016). All of which are part of the genetic background that is routinely used in current pig selection and breeding programs.

Although modern pig breeds have been genetically improved to achieve fast growth and a lean meat deposition, local pig breeds (like Basque, Alentejano, Iberian, and Meishan) also represent a valuable genetic resource as they are more highly adapted to their specific environmental conditions and feed resources, but also accumulate greater amounts of fat. Therefore, both (modern and local) pig breeds differ with respect to fat deposition and fat metabolic, among others traits (Poklukar et al., 2020).

_____
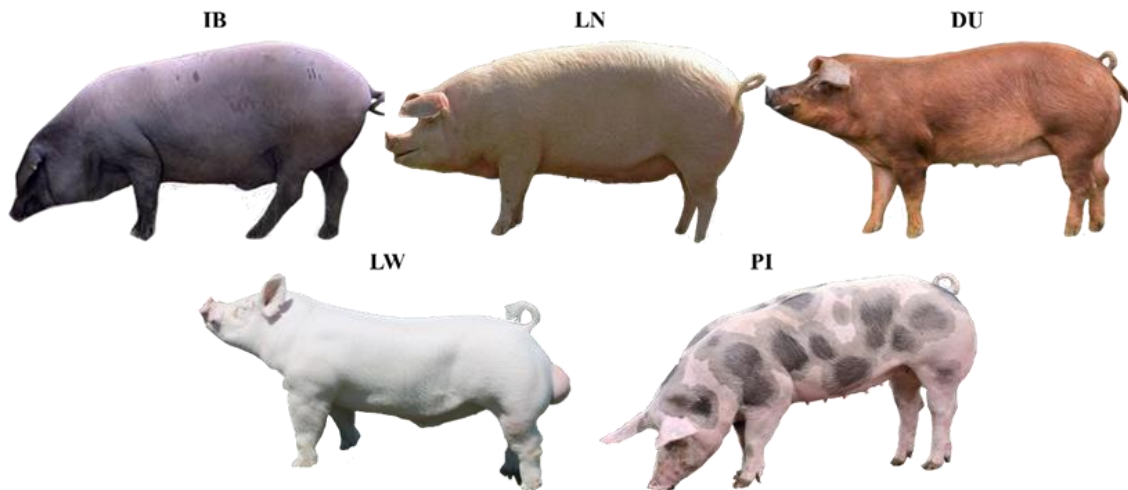
Figure 4. Illustration of six modern pig breeds used in pig production. IB = Iberian; DU = Duroc; LN = Landrace; LW = Large White (alias Yorkshire); and PI = Pietrain. Image source: <https://nobowa.com/pig-breeds/ and https://www.britishpigs.org.uk/pietrain> (Sep 22, 2022).

To accompany the aforementioned, pig breeding goals were initially focused on animal carcass yield, backfat thickness, lean carcass percentage and loin muscle area, thus also on indicators of fertility, fatness, feed efficiency, disease resistance and behaviour. However, genetic companies have increasingly focused on parameters associated with meat quality, like those summarized in the Pork Information Gateway "PIG" resource (Nold, 2006). In addition to the technological and sensory qualities of products that have undergone dry curing (Ruiz-Carrascal et al., 2000).

1.2. Pork meat quality

According to Cameron (1993), the term "meat quality" refers to a wide group of fresh meat processing and sensory qualities. The concept applies to both fresh meat and cured products.

1.2.1. Measurements of meat quality

A variety of criteria and perceptions are used to assess the meat quality, including sensory (eating) quality, nutritional quality, technological quality, hygienic quality, and nutritional quality, which constitute an important target for any farm animal production (Sellier, 1998; Gagaoua and Picard, 2020). In addition, each perception of meat quality implies different and

_____

related traits. Some of the most common characteristics related to the sensory, nutritional, and technological quality of pork meat are shown in **Figure 5**.
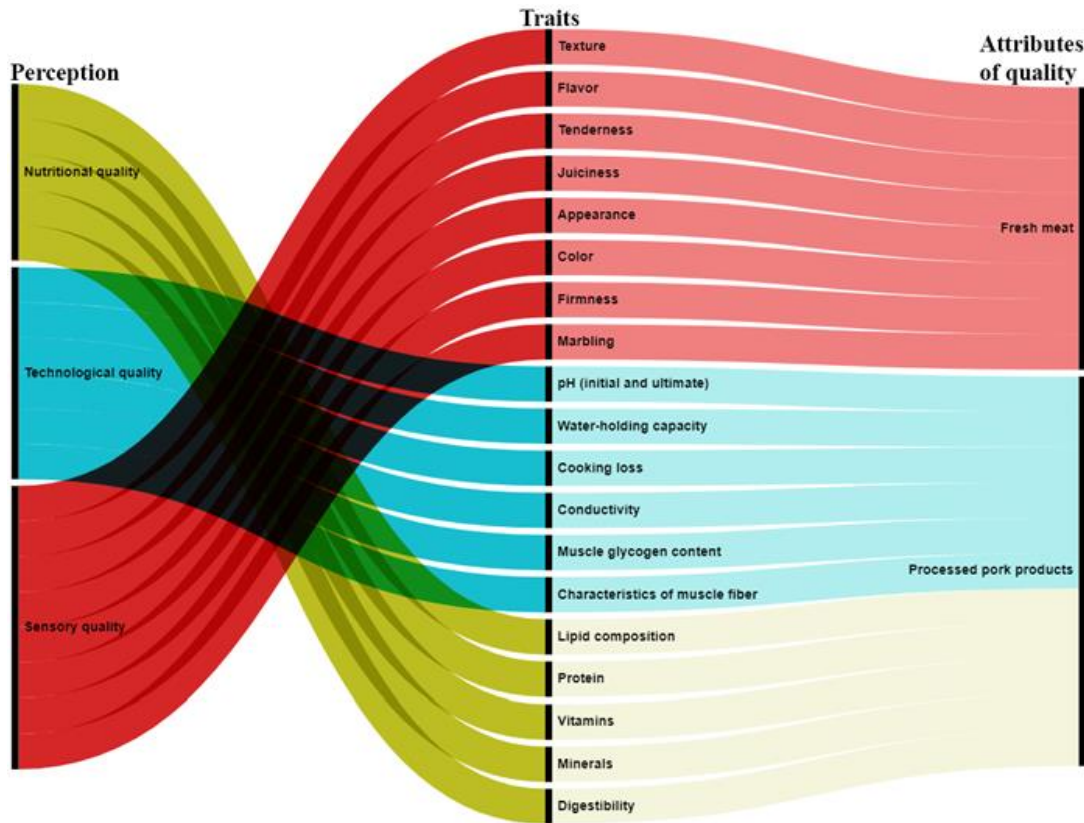
Figure 5. Summary of traits related with meat quality measurements of pork and delimitation according to quality attributes. Source: adapted from Sellier (1998).

Together with healthiness (nutritional quality), three other general terms of quality have been described: security (hygienic quality), satisfaction (organoleptic quality), and serviceability (ease of use, ability to be processed, and prices) (Listrat et al., 2016). Thus, meat quality can be influenced by key issues such as traceability, attributes as food safety (microbiological safety and absence of residues) and social value (animal welfare and environmentally friendly) (Kyriazakis & Whittemore, 2006). In addition, other factors that affect meat quality include carrying the stress gene, and extreme leanness and muscling (Nold, 2006).

However, customers require the production of high-quality pig meat produced under farming systems, which are safe and environmentally sound, and welfare-friendly (with good ethical practices) (Kyriazakis & Whittemore, 2006). In fact, quality characteristics of meat play an

_____

integral role in consumer acceptability. Consumers particularly enjoy traditional dry-cured ham, which, depending on quality, can have a significant added value. Therefore, producing high-quality meat is essential to maintaining the market and public support for pork production. In this section we will not describe in detail each of the categories for attributes of meat quality. Instead, we will illustrate some examples according to the interest of this work. In this way, some quality traits, such as intramuscular fat percentage (IMF%), water-holding capacity, and tenderness are important for customer acceptability of meat (Guo & Dalrymple, 2017). Fat has a characteristic flavor and is one of the major components of meat flavor (Miller, 2014). Also, in pork, beef and lamb, C18:0 concentration has been considered as closely related to the melting point of lipid and the firmness or hardness of carcass fat (Wood et al., 2003). Instead, C18:0 content has been significantly associated with sensory attributes of pork (Tikk et al., 2007). For instance, as breeders have intensively selected animals for increased leanness (Barton-Gade 1990; Cameron, 1990), the IMF (alias marbling) content has also decreased, producing a negative impact on pork flavor and meat tenderness, even though this relationship is not always strong. As the level of fat or marbling increases, consumers tend to like the flavor of beef and pork.

The mechanisms by which the IMF might improve the organoleptic quality are not entirely known. In this regard, it has been speculated by Wood et al. (2003), that the presence of fat cells between muscle fiber fascicles might separate them physically and promote tenderness. Lipids that trap moisture could improve juiciness (Wood et al., 2003) and promote salivation (Ruiz-Carrascal et al., 2000). FAs are converted into volatile compounds by lipid oxidation and Maillard reactions during cooking, which contribute to the characteristic aroma of pork (Gandemer, 2002; Mottram, 1998).

There are several types of fat deposits (e.g., internal, intermuscular, subcutaneous and IMF) that may be present and are of commercial interest. **Figure 6A-B** illustrates an example of fat deposition including that of IMF. On a macroscopic scale, skeletal muscle is made up of 90% muscle fibers and 10% fat and connective tissues (Listrat et al., 2016).
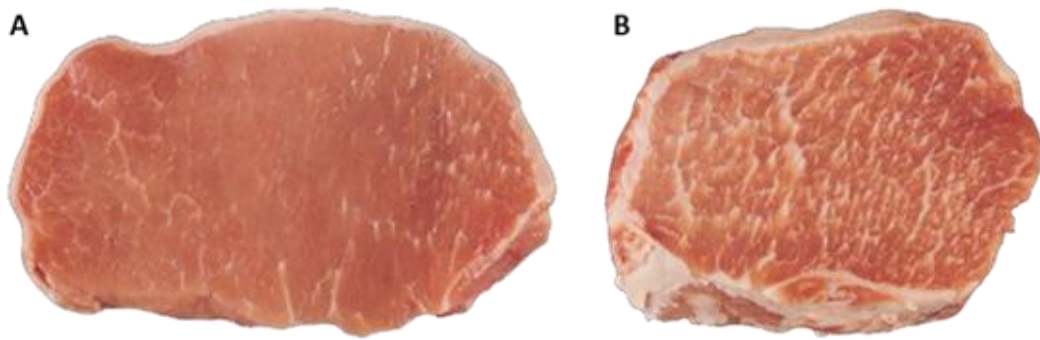
Figure 6. Illustration of type of fat deposition in *longissimus dorsi* muscle. Superindex A indicates a normal marbling, and B a high marbling.

The peak of IMF deposition (within the muscles) in the live animal occurs after maximization of other types of fat deposition. As animals grow and develop, fat is deposited sequentially into different fat depots (Miller, 2014), for example, internal fat in the abdominal or mesenteric cavity, followed by intermuscular (between muscles) and subcutaneous. However, depending on the species and age of the animal, as well as the energy intake, the proportion of each fat depot varies. According to Miller (2014), most of the pig fat is held in subcutaneous fat depots, ranging the IMF content from 0.5% in lean pigs to about 3% in fat pigs. Besides, subcutaneous fat was also the predominant fat depot over the growth range studied in Large White × Landrace castrated male pigs (Davies & Pryor, 1977). In addition, IMF is the only type of fat that cannot be removed from the meat at the time of consumption. However, the presence of IMF favors the main sensory characteristics perceived by the consumer at the time of purchase (color) and, above all, at the time of consumption (tenderness, juiciness and flavor (taste + aroma). In turn, if the meat is consumed hot, many of the FAs will change from solid to liquid (oil) state, thus facilitating chewing due to its lubricating effect and, therefore, the perception of greater tenderness.

However, together with IMF, muscle fibers and intramuscular connective tissue appear as the three main components of muscle, which are involved in the determination of various meat quality dimensions. It is interesting to note that the relative independence among the properties of these three major muscle constituents, suggests that it is possible to independently modulate these characteristics by genetic, nutritional, or environmental factors to achieve production efficiency and improve meat/flesh quality, as well as in order to control the quality of products and thus better fulfill the expectations of producers, meat processors, and consumers (Listrat et

_____

al., 2016). A compilation of the multifactorial background that affects meat quality is described below.

1.2.2. Factors that affect meat quality

The processes that contribute to meat quality are complex, and are in turn influenced by the interaction of multiple factors, as for example genetic background, environment, pre-slaughter handling, and slaughtering procedures (Sellier, 1994; Nürnberg et a., 1998; Eggen & Hocquette 2004; Lebret, 2008; Ngapo & Gariépy, 2008; Chriki et al., 2012; Rosenvold & Andersen 2003; Pena et al., 2013; Guo & Dalrymple, 2017). In addition, other factors affecting fatness or excess weight (like the age of animals), FA composition and gene expression play an important role in meat quality (**Figure 7**). Moreover, meat quality, as a product of many different physiological processes occurring in muscle, may involve a large number of genes associated with both structural and metabolic features of muscle (Eggen & Hocquette, 2004; Chriki et al., 2012; Pena et al., 2014).
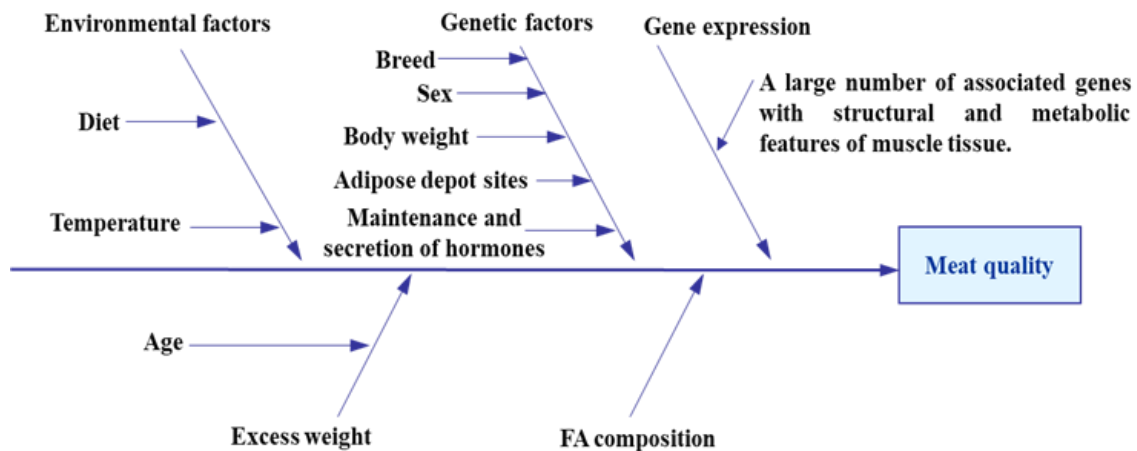
Figure 7. Compilation of multiple factors affecting meat quality in pigs.

Overall, it is widely assumed that the most significant effect modulating meat quality are the environmental factors. Genetic effects are expressed through several cascades of biological processes that influence and react to environmental exposures, which can give rise to gene–environment interactions (Abdellaoui et al., 2022). In the following sections, we will focus on two relevant factors for meat quality, such as FA composition and gene expression, which are also two important issues for both the pig industry and the consumer's health.

_____

1.3. Fatty acid composition studies in pig

1.3.1. Factors that affect fatty acid composition

Fatty acid (FA) composition varies across tissues (e.g., between adipose and muscle), and its relative proportion is influenced by numerous factors. Similar to the multifactorial scheme of meat quality (**Figure 7**), FA composition also includes the influence of diet, fatness, age/body weight, gender, genetics, environmental temperature, depot site, and maintenance (Nürnberg et al., 1998), among other factors. In addition, FA composition is closely related to the nutritive value and the taste of meat. In this way, the content and intramuscular composition of FAs are parameters that affect meat quality, influencing its organoleptic characteristics and nutritional value.

1.3.2. Definition, classification, origin, and nomenclature of fatty acids

FAs are a common component of the lipids, present in body fat and in the food we eat. While in biochemistry terminology, a FA is a carboxylic acid with an aliphatic chain, which could be either saturated or unsaturated. Commonly, natural FAs have a chain of 4 to 28 carbons (usually unbranched and even-numbered). Furthermore, FAs that comprise a carboxylic acid with hydrocarbon chains can vary in its length, ranging from 4 to 36 carbons long (Nelson et., 2008). In some FAs, this chain is unbranched and fully saturated (contains no double bonds); in others the chain contains one or more double bonds. Further, FAs are referred to as the main building blocks of lipids and contribute to the lipid diversity present in mammalian cells (Cockcroft, 2021). Interestingly, FAs may act as signaling molecules to regulate transcription of target genes encoding proteins involved in muscle lipid metabolism (Fritzen et al., 2020).

FAs are classified in many ways, for instance, FAs constitute one of the 4 main sub-groups of lipids (i.e., FAs: saturated & unsaturated; glycerids: glycerol containing lipids; nonglyceride: sphingolipids, steroids, waxes; and complex lipids: lipoproteins). For simplification, eukaryotic cells contain hundreds of many different lipid species that can be categorized into three main classes, such as the lipids present in the cell membrane (glycerophospholipids, sphingolipids, and sterols) (Cockcroft, 2021).

Most commonly, FAs can be classified as saturated (SFA, no double bond is present) or unsaturated (with one or more double bonds present). They can also be referred in two sub-groups, monounsaturated (MUFA) and polyunsaturated (PUFA). In turn, PUFA can be further catalogued into two series (i.e., omega-6 and omega-3) (Johnston & Sobhi 2018). In addition, FAs may be divided into two groups considering if they are derived directly from the diet or if they can be synthesized *de novo* through lipogenesis. FAs derived directly from the diet are known as essential (e.g., PUFA: C18:2*n*-6 and C18:3*n*-3), while those that can be synthesized through *de novo* lipogenesis are known as non-essential (e.g., SFA: C16:0 and C18:0, and MUFA: C16:1*n*-7 and C18:1*n*-9) (van Son et al., 2017).

Essential FAs like C18:2*n*-6 and C18:3*n*-3 can be used as precursors for the synthesis of other PUFA, including the C20:3*n*-6 and subsequently C20:4*n*-6 (Guillou et al., 2010; Bond et al., 2016). Likewise, *de novo* endogenously synthesized FAs represented ∼86% of the total non-essential FA deposition (including SFA as C14:0, C16:0, and C18:0, and MUFA such as C16:1*n*-7, and C18:1*n*-9 acids) (Kloareg et al., 2007). In FA composition of pork, the predominant SFAs are C14:0, C16:0 and C18:0, which together account for about 38% of the total FAs (Zhang et al., 2019). The most common MUFA in pig muscle is C18:1*n*-9, being C16:0 and C18:0 the predominant SFA, and C18:2*n*-6 is the most abundant PUFA (Wood et al., 2008). Meanwhile the most abundant PUFA repertoire includes C18:3*n*-3 and C20:4*n*-6, and the most common SFA found in many phospholipids are C16:0 and C18:0 acids (Cockcroft, 2021).

The nomenclature of FAs according to their common names, symbol abbreviation, and double bond positions is presented in **Table 1**. The FAs description also includes Δ nomenclature and ω (or n) nomenclature (Bond et al., 2016). As such, Δ and ω (or n) denote the position of the double bond counting from the carboxylic acid carbon and the terminal methyl carbon, respectively. The ω (or n) terminology, for instance, is useful for describing PUFA. Likewise, the nomenclature for any additional functions on the principal chain follows the rules of the International Union of Pure and Applied Chemistry (IUPAC) convention of 1957 (Brondz, 2005).

_____

Table 1. Summary of the common nomenclature and bond positions of major long-chain FAs (i.e., adapted from Bond et al., 2016). All bonds indicate a cis geometric configuration.

Common name	Symbol abbreviation	Double bond positions
Palmitic acid	C16:0	---
Palmitoleic acid	C16:1*n*-7	Δ9
Sapienic acid	C16:1*n*-10	Δ6
Stearic acid	C18:0	---
Oleic acid	C18:1*n*-9	Δ9
Vaccenic acid	C18:1*n*-7	Δ11
Linoleic acid	C18:2*n*-6	Δ9,12
α-Linolenic acid	C18:3*n*-3	Δ9,12,15
γ-Linolenic acid	C18:3*n*-6	Δ6,9,12
Stearidonic acid	C18:4*n*-3	Δ6,9,12,15
Arachidic acid	C20:0	---
Paullinic acid	C20:1*n*-7	Δ13
Gondoic acid	C20:1*n*-9	Δ11
Dihomo-γ-linolenic acid	C20:3*n*-6	Δ8,11,14
Meadic acid	C20:3*n*-9	Δ5,8,11
Arachidonic acid	C20:4*n*-6	Δ5,8,11,14
Eicosapentaenoic acid	C20:5*n*-3	Δ5,8,11,14,17
Behenic acid	C22:0	---
Erucic acid	C22:1*n*-9	Δ13
Adrenic acid	C22:4*n*-6	Δ7,10,13,16
n-6 Docosapentaenoic acid	C22:5*n*-6	Δ4,7,10,13,16
n-3 Docosapentaenoic acid	C22:5*n*-3	Δ7,10,13,16,19
Docosahexaenoic acid	C22:6*n*-3	Δ4,7,10,13,16,19
Lignoceric acid	C24:0	---
Nervonic acid	C24:1*n*-9	Δ15

_____

1.3.3. Physiological functions of fatty acids

Multiple lines of evidence suggest that FAs regulate many important physiological functions. For example, FAs are crucial for living organisms with a variety of roles, including cell signaling, signal transduction, cellular differentiation, metabolic homeostasis, body energy homeostasis, protection of digestive tract mucosa, as well as being essential cellular components and a major source of energy for animals. Among PUFAs, C20:4n-6 is utilized for eicosanoid synthesis and is a constituent of membrane phospholipids involved in signal transduction (Spector, 1999). C20:4n-6 plays an important role in the control and regulation of lipid metabolism, particularly in relation to the immune and inflammatory system (Santos et al., 2017). C18:3n-3 is the essential precursor of n-3 FAs series, while C18:2n-6 is the precursor in n-6 FAs series (Santos et al., 2017). C18:2n-6 has been reported to affect cell proliferation and lipid catabolic gene expression in mammals.

On the other hand, MUFA are mediators of signal transduction, cellular differentiation, and metabolic homeostasis. MUFA has the potential to affect a variety of key physiological variables, which include insulin sensitivity, metabolic rate and adiposity (Bond et al., 2016). C16:1n-7 and C18:1n-9 MUFA are the most predominant in triacylglycerols (TAGs), cholesteryl esters, wax esters and membrane phospholipids (Bond et al., 2016). Currently, there is also a greater interest in SFA and MUFA in the context of their effects on human health and disease prevention. According to Mensink (2013), C18:0 SFA does not seem to increase the low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol concentrations. As described for MUFA, increasing the content of C18:1n-9 could improve the organoleptic properties and overall acceptability parameters of meat, because they are positively correlated (Cameron et al., 2000). C18:1n-9 FA was linked to consumer acceptability of high-quality cured products, which represent a niche market of added value (Pena et al., 2016). Whilst emerging evidence in mice suggested C16:1n-7 as an adipose tissue-derived lipid hormone that strongly stimulates muscle insulin action, as well as has shown that adipose tissue uses C16:1n-7 lipokine to communicate with muscle and liver tissues, thereby regulating systemic metabolic homeostasis (Cao et al., 2008).

Regarding SFA, from a nutritional point of view it is recommended to limit its consumption due to its association with an increase in the cholesterol levels and with cardiovascular disease

_____

(CVD) risk or outcomes, and type 2 diabetes. Hence, dietary guidelines advice of limiting the intake of saturated fat in favour of monounsaturated and polyunsaturated fats (McLean et al., 2015). However, the benefits of replacing the consumption of SFA with PUFA remain controversial. Astrup et al. (2021), have highlighted a lack of rigorous evidence to support continued recommendations either to limit the consumption of SFA or to replace them with PUFA. Furthermore, in most comparative studies, no conclusive benefit on heart disease were found by replacing SFA by PUFA. On the other hand, SFA also play a role as energy stores, necessary for hormone production, cellular membranes, and for organ padding. In addition, certain SFA (e.g., C14:0 and C16:0) are also needed for important signaling and stabilization processes in the body.

1.3.4. Fatty acid metabolism

As described above, FAs can be provided by the diet or derived from *de novo* lipogenesis. The major dietary constituents (e.g., carbohydrates, lipids, amino acids, and proteins) undergo digestion by enzymes in the upper gastrointestinal tract. Among the dietary nutrients, FAs are primary sources for oxidation and mobilization from cellular stores. However, the utilization of modulated dietary FAs requires that they be absorbed through the intestine.

In all animals, most of the primary dietary lipids are in the form of several lipid classes, including phospholipids (PLs), sterols (e.g., cholesterol), triglycerides, and many other lipids (e.g., fat-soluble vitamins) (Iqbal & Hussain, 2009). The intestinal digestion of these dietary fat molecules undergoes with their emulsification (or solubilization process) mainly in the intestine, with peristalsis as a major contributing factor. The emulsions then mix with liver-derived bile salts and pancreatic juice to undergo marked changes in their chemical and physical form. However, for dietary FAs to be absorbed, transported, and taken up by body cells, they require the presence of enzymes, transporters and chaperone proteins, among other molecules. For instance, dietary FAs are metabolized by various hydrolytic enzymes including lingual, gastric, and pancreatic lipases. These enzymes produce free FAs and a mixture of monoglycerol and diglycerides from dietary triglycerides. Free FAs released from triacylglycerol (TAG) diffuse into intestinal epithelium cells, which are derived by cleavage of ester bonds due to the action of lipase, high temperature, and moisture (Chew & Nyam, 2020).

Following absorption of pancreatic lipase products by intestinal enterocytes, triglyceride re-synthesis occurs from monoglycerol and free FAs. The triglycerides are then solubilized into lipoprotein particles (complexes of lipid and protein) termed chylomicrons, which are formed from triglycerides together with cholesterol, phospholipids and proteins. Chylomicrons from the intestine are released into the lymphatic system, and then into the blood for delivery of dietary FAs to the various target tissues, such as liver to obtain energy, adipose tissue for energy storage or skeletal muscle for FA uptake and production of energy through oxidation. Therefore, FAs as important energy sources, they can yield adenosine triphosphate (ATP) in huge quantities. In accordance, adipose tissue is a highly active metabolic and essential endocrine organ for the circulation of free FA and regulation of lipid metabolism (Xing et al., 2016). In muscle, myocytes do not store TAGs and instead take up FAs that are released into the blood by adipocytes. While, in both adipose and muscle tissue, FAs can be supplied from lipolysis of stored TAG in these tissues (Nelson et al., 2008).

The dietary or the modified FAs can then enter a targeted pathway that promotes stepwise processing of lipids. Metabolic pathways are overlaid with complex regulatory controls that are exquisitely sensitive to changes in metabolic circumstances (Nelson et al., 2008). As an example, **Figure 8** illustrates an integration of metabolic regulation (either with FAs provided by diet or derived from lipogenesis), including the main intermediary metabolites, a set of genes encoding enzymes (e.g., elongases and desaturases), and anabolic and catabolic pathways. As complex as the metabolism of fats and FAs is closely linked to that of carbohydrates. Nonetheless, the regulation of carbohydrate metabolism in muscle reflects differences (e.g., in the enzymatic machinery) from liver and adipose tissue. Under certain nutritional conditions (e.g., the fed state), *de novo* lipogenesis converts excess carbohydrate into FAs (i.e., in lipogenic tissues) that are then esterified to storage triacylglycerols (TAGs), which could later provide energy via β-oxidation (Ameer et al., 2014). It is important to note, also, that FA biosynthesis and degradation reactions occur by different pathways, are catalyzed by different sets of enzymes, and take place in different parts of the cell (i.e., cytosol and mitochondrial, respectively).

_____
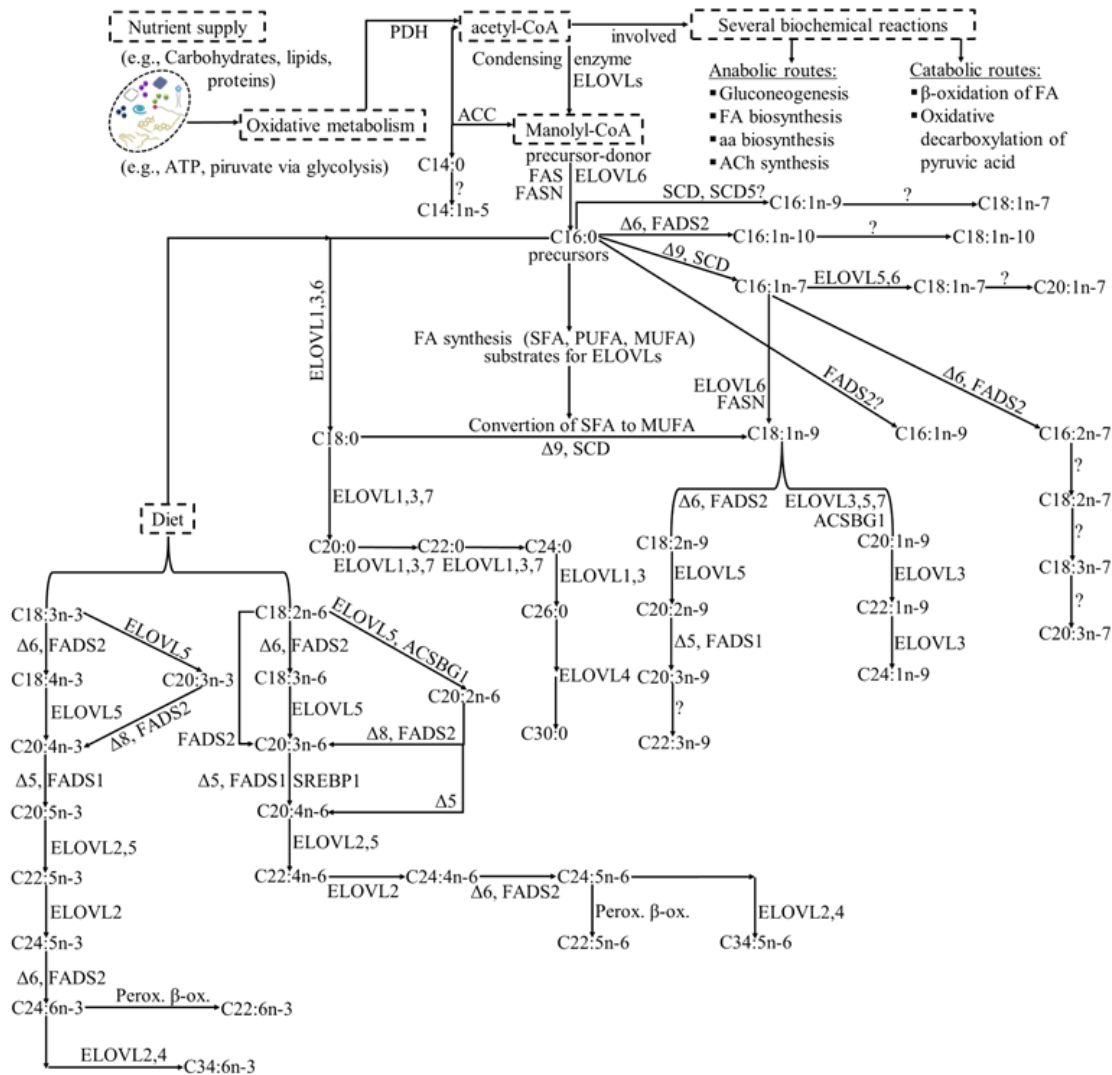
Figure 8. Compilation of FA metabolism pathway and related candidate genes in mammals (adapted from Cook & McMaster 2002; Guillou et al., 2010; Bond et al., 2016; Zhang et al., 2016). This scheme includes long chain and very long-chain FA synthesis, either of saturated and unsaturated FAs of the *n*-10, *n*-7 and *n*-9 series, which can be synthesized from palmitic acid (or C16:0) produced by the FA synthase (FAS). In particular, long chain FAs of the *n*-6 and *n*-3 series can only be synthesized from precursors obtained from the diet.

In mammals, a distinction must be made for omega-3 and omega-6 PUFA, i.e., they cannot be synthesized *de novo* (PUFA cannot be generated from MUFA due to lack of desaturases), thus must be provided by the diet (e.g., essential dietary FAs as linoleic (C18:2*n*-6) and α-linolenic (C18:3*n*-3) acids). Of particular note, they can be further elongated by addition of acetyl groups by elongase enzymes and double bonds added by desaturase enzymes (Cockcroft, 2021).

_____

Collectively, PUFAs are metabolized by a series of enzymes giving rise to physiologically active metabolites that can activate cells by binding to specific cell surface receptors.

1.3.4.1. Fatty acid synthesis

De novo lipogenesis is the synthesis of FAs. At least minimally, FA biosynthesis takes place in most tissues. Although the liver is the central site for *de novo* FA synthesis in rodents and humans, in other species adipose tissues are the primary sites of lipogenesis (Bergen & Mersmann 2005). For instance, in pig, adipose tissue plays a primary, if not a nearly exclusive, role in FA synthesis (O'Hea & Leveille 1969). FA biosynthesis requires several key steps, including the participation of a three-carbon intermediate, acetyl coenzyme A (acetyl-CoA), malonyl-CoA, and sequential reactions with an array of enzymes (**Figure 9**). Among them, acetyl-CoA is a vital molecule that conveys carbon atoms to the tricarboxylic acid (TCA) cycle (also called the Krebs cycle) to be oxidized for energy production. In addition, it turns out that the products from glycolysis (e.g., the pyruvate obtained from the conversion of glucose), and key cytosolic enzymes (acetyl-CoA carboxylase and FA synthase) are used to initialize *de novo* lipogenesis (Bond et al., 2016).
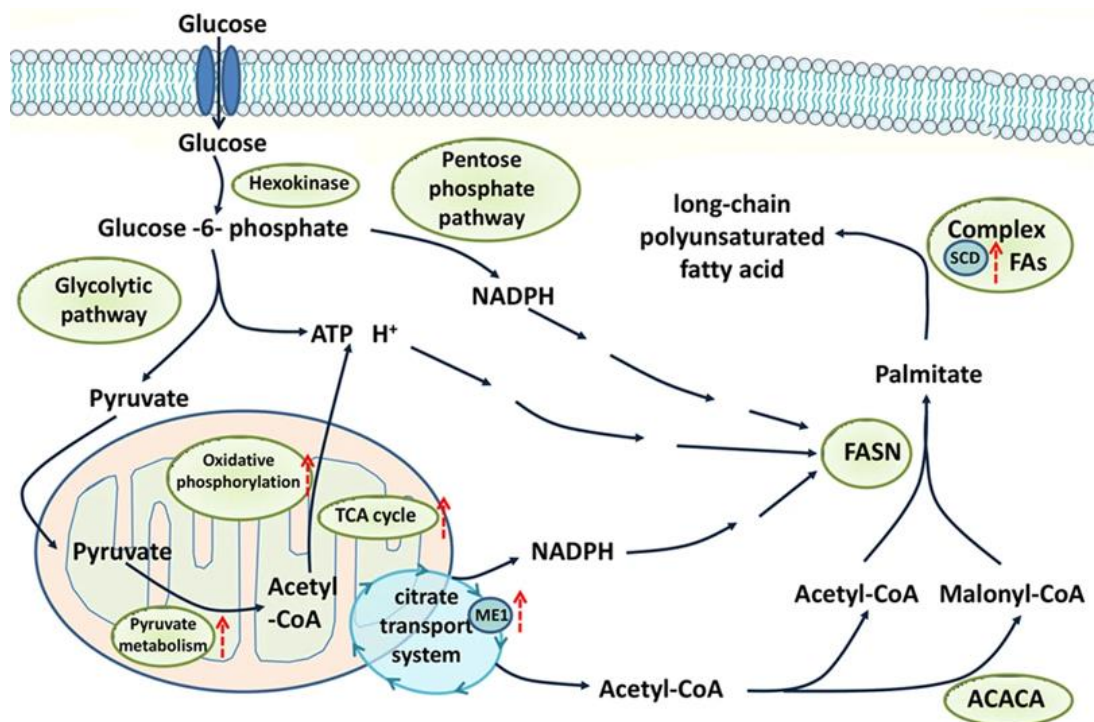
Figure 9. Schematic diagram with pathways in *de novo* FA synthesis (original illustration presented by Xing et al., 2016).

In the cytoplasm of cells, citrate derived from the TCA cycle is converted to acetyl-CoA by the action of ATP citrate lyase. Malonyl-CoA is then produced from the carboxylation of acetyl-CoA by acetyl-CoA carboxylase (*ACC*) (Bergen & Mersmann, 2005), which exists in two known isoforms (*ACC1* and *ACC2*) (Melmed, 2020). On the one hand, malonyl-CoA act as an inhibitor of ketogenesis and carnitine palmitoyltransferase 1 (*CPT1*) reaction, which transports FAs to the mitochondria for β-oxidation (McGarry et al., 1977) and, thereby malonyl-CoA prevents fat catabolism under physiological conditions, such as those in which energy is stored as fat through FA biosynthesis. On the other hand, malonyl-CoA molecules also serve as a substrate for FA synthase (FAS), which sequentially connects two carbon fragments to generate SFA such as palmitic (C16:0) acid (Melmed, 2020). As FAS catalyzes the conversion of acetyl-CoA and malonyl-CoA to C16:0 in the presence of NADPH (Wakil et al., 1983), it increases the number of carbons of the acetyl-CoA molecule through successive incorporations of two carbons of the malonyl-CoA group.

Palmitic (C16:0) synthesized *de novo* or derived from the diet is transported to the endoplasmic reticulum (ER) and on its membrane two major FA enzymatic modifications (i.e., elongation and desaturation) occur to yield longer chain SFA and unsaturated FA (Bond et al., 2016). The elongation of FA chain up to C_{26} is achieved through elongase enzymes (ELOVLs), which increase the length of FAs adding two carbons of the malonyl-CoA group. Also, the introduction of double bonds into FAs is mediated by desaturase enzymes. As an example, the enzyme stearoyl-CoA desaturase (*SCD*) (also known Δ9 desaturase) creates a double bond between the 9th and 10th carbon counting from the carboxyl end (Guillou et al., 2010; Cockcroft, 2021). *SCD* plays a function in the biosynthesis of unsaturated FA, resulting in the conversion of C16:0 acid to stearic (C18:0) acid through the action of a long-chain FA elongase. Subsequently, *SCD* can convert SFA as C16:0 and C18:0 into MUFA as palmitoleic (C16:1*n*-7) and oleic (C18:1*n*-9) acids, respectively (Guillou et al., 2010).

_____

1.3.4.2. Fatty acid beta oxidation

The β-oxidation cycle is the primary pathway used by the most organisms for FA degradation (also known as lipolysis). In mammals, β-oxidation occurs in both mitochondria and peroxisomes. Despite the similarity of several of the enzymes involved in this pathway in both organelles, some distinct physiological roles have been uncovered (Poirier et al., 2006). FAs as the main energy substrate providing the majority of cofactors necessary for mitochondrial oxidative phosphorylation, and contributing to ATP production (Fillmore et al., 2014). In addition to heart, liver and skeletal muscle are also very active tissues for mitochondrial FA β-oxidation (Wajner & Amaral 2016), as well as FAs are released from the adipose tissue for transport to the energy demanding tissues.

In relation to nutritional status, in fasting conditions or times of elevated energy demands the FA metabolism works in the direction of lipolysis (FA breakdown). Under these situations, in adipocytes the hydrolysis of TAG to produce FAs and glycerol is tightly regulated by neuroendocrine signals resulting in the activation of lipolytic enzymes (Frühbeck et al., 2014). However, such signals must be fine-tuned to precisely adapt the balance between TAG synthesis and breakdown to meet physiological needs. Concerning long-chain FAs, these are derived primarily from adipocyte lipolysis, where they enter the circulation bound to albumin, and from lipoprotein lipase-mediated breakdown of TAG contained in very low-density lipoproteins (LDL) and chylomicrons.

Selective FA transporters located on the cell membrane serve as an intermediates for the entry of FAs into the cell, including the tissue-specific FA transporter proteins (carriers) [FA translocase (FAT)/CD36], the plasma membrane isoform of FA-binding protein (FABPpm), and FA transport protein (FATP) 1/6 (Lopaschuk et al., 2010). In the cytosol, free FAs are activated by acyl-CoA synthetases from fatty acyl-CoA esters, which are subsequently transported into the mitochondria via the carnitine transport system (Sharpe & McKenzie, 2018). While long-chain FAs must first be shortened in peroxisomes before further oxidation in mitochondria, the breakdown of FAs up to 18 carbons in length occurs in the mitochondria. Subsequently, *CPT1* catalyzes the addition of carnitine to fatty acyl-CoA esters to form acylcarnitines, which are transported across the inner mitochondrial membrane (IMM) via carnitine acylcarnitine translocase (*CACT*). Inside the mitochondrial matrix, carnitine is

_____

removed by carnitine palmitoyltransferase 2 (*CPT2*) to regenerate fatty acyl-CoA esters and free carnitine, which is then recycled back across the IMM by *CACT* entering the β-oxidation cycle (**Figure 10**) (Sharpe & McKenzie, 2018).
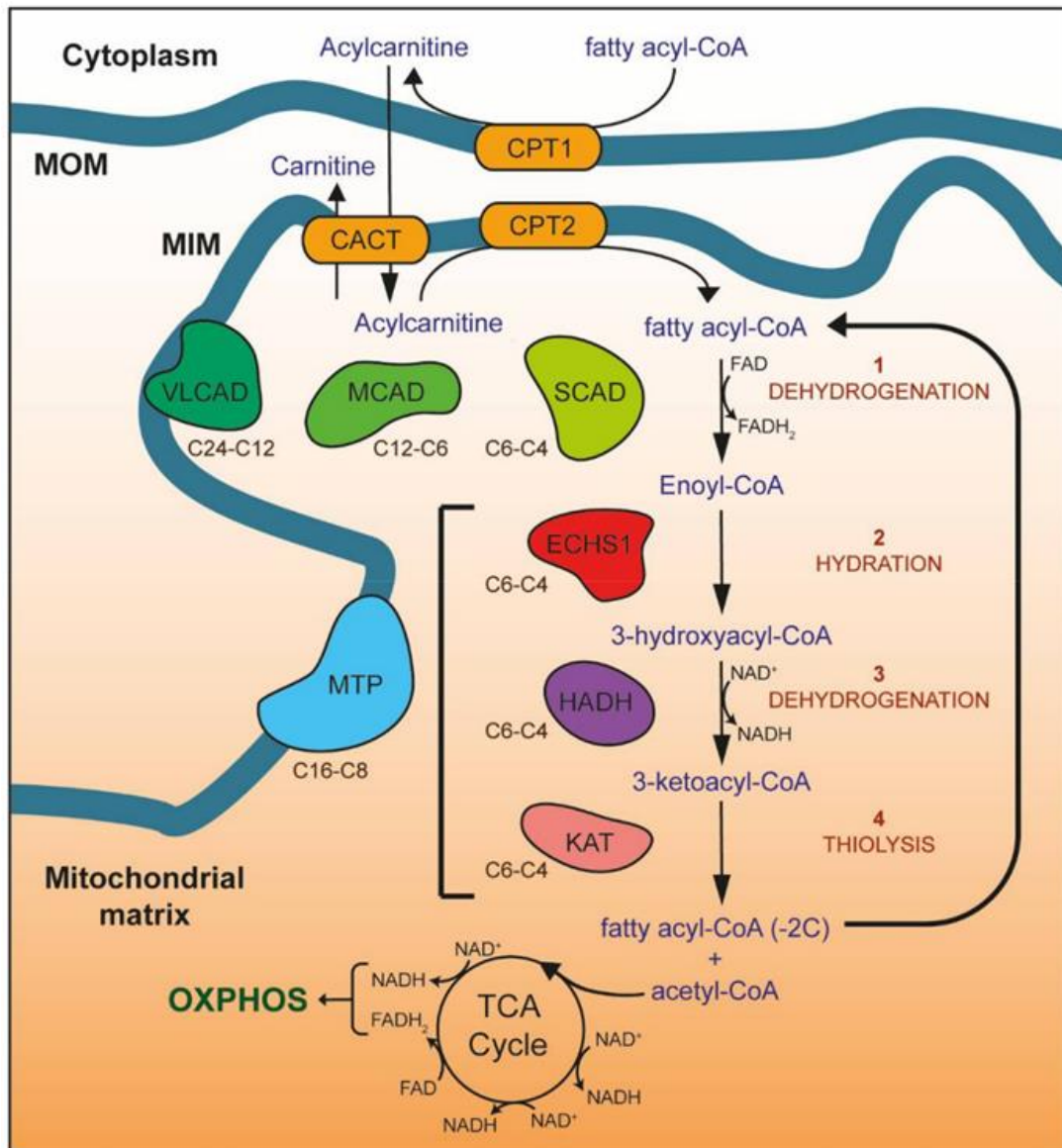
Figure 10. Mitochondrial FA β-oxidation (original illustration from Sharpe & McKenzie, 2018).

The fatty acyl-CoA chains within the mitochondrion are processed to yield one molecule of acetyl-CoA, two electrons and a two-carbon shortened fatty acyl-CoA, which are obtained through a series of four enzymatic reactions (dehydrogenation, hydration, a second dehydrogenation and thiolysis cleavage) (Bartlett & Eaton 2004; Sharpe & McKenzie, 2018).

_____

As shown in **Figure 10**, these reactions are then performed repeatedly until only two acetyl-CoA molecules remain. In addition to the production of acetyl-CoA, each β-oxidation cycle results in the production of nicotinamide adenine dinucleotide hydrogen (NADH), and flavin adenine dinucleotide H2 (FADH2) (Bacle & Ferreira, 2019), which are subsequently oxidized by the oxidative phosphorylation (OXPHOS) complexes I and II, respectively (Sharpe & McKenzie, 2018), as well as latter used by the electron transport chain for the production of ATP (Fillmore et al., 2014).

Furthermore, FA β-oxidation-related enzymes exhibit chain length specificity (**Figure 10**). The first dehydrogenation step of chain-length fatty acyl-CoAs (ranged at C24-C12, C12-C6, and C6-C4) is catalyzed by very long-, medium-, and short-chain acyl-CoA dehydrogenases (*VLCAD*, *MCAD*, *SCAD*), respectively. On the other three reactions, longer acyl-CoA chains (C16-C8) are catalyzed by the multi-domain mitochondrial trifunctional protein (MTP), which harbors long-chain enoyl-CoA hydratase, long-chain 3-hydroxyacyl-CoA dehydrogenase and 3-ketoacyl-CoA thiolase activities. Particularly, for medium- and short-chain FAs, the last three steps of FA β-oxidation are catalyzed by short-chain enoyl-CoA hydratase (*ECHS1*), hydroxyacyl-CoA dehydrogenase (*HADH*) and 3-ketoacyl-CoA thiolase (*KAT*) (Sharpe & McKenzie, 2018). Additionally, DeLany et al. (2000) sustained that diversity in FA structure (i.e., from differences in chain length, degree of unsaturation, and position and stereoisomeric configuration of the double bonds) can influence the rate of FA oxidation. For instance, research conducted on rat and human models, suggested that rats exhibited a greater ability to metabolize linoleic than palmitic (Cenedella & Allen, 1969). By contrast, in rats fed a fat-free meal or just the labeled FAs, the dietary rates of oxidation of linoleate and palmitate were similar (Toorop et al., 2016). In humans, SFA was highly oxidized, while the PUFA (as linolenic) and MUFA (as oleic) FAs were fairly oxidized. Meanwhile, linoleic FA appeared to be conserved.

1.3.5. Fatty acid determination by gas chromatography methodology

The gas chromatography (GC) analytical technique is a powerful tool to analyze FA determinations. In fact, a significant advance of FA analysis by gas GC has been the understanding of the importance of dietary FAs for human health (Seppänen-Laakso et al. 2002). In research with animals including livestock, GC methodology has also been commonly used to analyze lipids (or FA composition) derived from samples of animal tissues (e.g., adipose

_____

tissue or backfat, liver, and muscle) (Folch et al., 1957; Pérez-Enciso et al., 2000; García-Olmo et al., 2002; Clop et al., 2003; O'Fallon et al., 2007; Ntawubizi et al., 2010; Zhang et al., 2017). Depending on the study, a known amount of some FA (e.g., pentadecanoic acid, C15:0; margaric acid, C17:0; or heneicosanoic acid, C21:0) is added as an internal standard (it is mixed with the sample but does not interfere with any FA in the sample) for subsequent quantification of the peaks (and consequently lipids) could be performed after the transesterification procedure.

In pigs, the procedures of Folch et al. (1957) and O'Fallon et al. (2007) are commonly used as the standard methods for fat extraction and preparing FA methyl ester (FAME), either from adipose and muscle samples, respectively. In brief, such preparations can then be analyzed in a gas chromatograph (e.g., Thermo, Milan, Italy) equipment with a split injection method (or splitless mode) and a flame ionization detector (Torres-Pitarch et al., 2014). Subsequently, FAME peaks for individual FA are identified by comparing their retention times with those of authentic standards (Segura and Lopez-Bote, 2014), provided by a chromatography analytical products company (e.g., by Sigma–Aldrich, Alcobendas, Spain or Supelco–PA, USA).

Usually, in most of the FA studies the relative values (percentages) for major FAs are computed (Fiego et al., 2005; Ramayo-Caldas et al., 2012; Muñoz et al., 2013; Segura & Lopez-Bote, 2014; Benítez et al., 2015, 2017), where the fat content per gram (g) of fresh tissue, the FA content per gram of fat, and the proportion of FAs (in %) over total FAs are quantified. Thus, the results can be expressed as g/100 g of total lipids, total lipids are g/100 g of tissue, grams per 100 g of detected FAMEs, and g FA/100 g total FAs (in context of intramuscular fat), as well as by groups of FAs including the total sum of SFA, MUFA and PUFA FAs, respectively. However, only specialized works with detailed analytics allow the absolute quantification of lipids, in which to obtain a good concentration of minority FAs without major distortions, fat fractionation should be considered, but also an independent analysis of membrane lipids and triglycerides. For instance, Domaradzki et al. (2022) have reported the FAs profile of lipids with details on transesterification and calibration procedures and quality assurance from adipose and muscle tissues in pigs, including e.g., C20:1cis-11 or C20:1*n*-11, and nutritionally important long-chain *n*-3 PUFA (e.g., C20:5*n*-3, C22:5*n*-3, and C22:6*n*-3 acid).

_____

1.4. Pig Genomics

Advances in molecular genetics techniques provided an interesting array of integrated strategies, including traditional and modern approaches across several "omics" such as genomics, transcriptomics, and epigenomics. These technologies can be applied to livestock to dissect and decipher the molecular and gene regulatory networks underlying the complex quantitative traits (Mondal & Singh, 2020).

The beginning of the genomics era, allowed swine geneticists to implement genetic markers to eliminate defects, select for features that would improve feed efficiency, growth, meat quality, and increased litter size traits. In this regard, sequencing of the swine genome has made it possible to genotype animals for thousands of SNPs to identify markers linked to relevant traits, as well as the use of genomic selection to predict the genomic breeding value (Mote & Rothschild, 2020). Thus, genomic information can facilitate the selection of relevant traits, increasing the accuracy of breeding value prediction and obtaining earlier evaluations.

1.4.1. Pig genome and its sequencing

The pig (*Sus scrofa*) is a species of even-toed ungulate in the family Suidae (pigs). Since the early 1990s, through the European International Pig Gene Mapping Project (PiGMaP), the pig was one of the first farm animal species for which the scientific community decided to generate a genetic map. Initially, with microsatellites markers evenly spaced through the genome and other landmark *loci* physically assigned to chromosomes (Chr.) (Haley et al., 1990). The authors described the multiple advantages of the pig for genetic mapping, including a well-defined karyotype, large full-sib families, short generation interval, and the availability of diverse genetic stocks. The PIGMaP consortium allowed the implementation of multipoint linkage analysis, which was used for genetic mapping of quantitative trait *loci* (QTLs) for growth and fatness in pigs (Andersson et al., 1994). In addition, allowing the mapping of genes and microsatellite markers in porcine Chr. 14 (Kapke et al., 1996).

In September 2003, the Swine Genome Sequencing Consortium (SGSC), composed by academic, government and industry representatives, was created (Schook et al., 2005). The SGSC's mission was to obtain the pig genome reference sequence, developing a roadmap for

_____

creating the required scientific resources and to integrate existing genetic and physical maps in pigs (Groenen et al., 2011; Raudsepp & Chowdhary, 2011). The strategy of hierarchical shotgun Sanger sequencing of bacterial artificial chromosome (BAC) clones was used (Humphray et al., 2007). The map provided a template for the generation and assembly of high-quality anchored sequences across the genome, just as the physical map integrated previous landmark maps with restriction fingerprints and BAC end sequences from over 260,000 BACs derived from 4 BAC libraries, and takes advantage of alignments to the human genome to improve the continuity and local ordering of the clone contigs. Such efforts were further integrated by hybrid approach, combining hierarchical shotgun sequencing of BAC clones and whole genome shotgun sequencing (WGS) (Archibald et al., 2010). In addition, WGS data were generated using both Sanger capillary sequencing and Illumina/Solexa sequencing techniques. Thus, in September 2009, the annotated genome assembly (Sscrofa9) was released with Ensembl 56, as well as a revised assembly (Sscrofa10) was under construction and will incorporate WGS data providing $> 30 \times$ genome coverage. In 2012, Groenen et al. (2012) presented the assembly with analysis of the genome sequence of a female domestic Duroc pig, and a comparison with the genomes of wild and domestic pigs from Europe and Asia. The authors provided a high-quality draft of the pig genome, whose assembly (Sscrofa10.2, Ensembl release 80) comprises 2.60 gigabases (Gb) assigned to chromosomes with a further 212 megabases (Mb) in unplaced scaffolds. This upgraded version of the pig genome sequence (called Sscrofa10.2) was available in the Ensembl database (http://www.ensembl.org). Likewise, the completion of the pig draft genome sequence marks a milestone in 20 years of pig genome studies. In this way, the Pig Genome Database (PGD) was created to facilitate information mining and integration within the pig and across species (https://www.animalgenome.org /pig/genome/db/).

However, annotation and assembly of pig genome was constantly improving. In 2017, a new version was compiled (i.e., Sscrofa 11.1 by SGSC), which was also made available in the Ensembl database. This new assembly was created using data generated by PacBio RSII long reads, a third generation sequencing (TGS) technology that generated a 65 x genome coverage over a total sequence length of 2.5 Gb. Nowadays, the genomes of hundreds of pigs of several breeds have been re-sequenced, and are available in open access databases like (https://www.ncbi.nlm.nih.gov/) or FAANG (https://www.faang.org/). The last reference genome Sscrofa11.1 available on the Ensembl database is the release 109 [Feb 2023 © EMBL-EBI]. A summary of the whole pig genome including an illustration of the karyotype, assembly

_____

(SScrofa11.1, database version 109.111), gene counts, and single nucleotide polymorphism (SNP) counts is available at http://www.ensembl.org/Sus_scrofa/Location/Genome.

Massive sequencing or next-generation sequencing (NGS) techniques were revolutionary in genomic studies. NGS have allowed the massive detection of SNPs and the development of high-throughput genotyping arrays (described below) in livestock. Novel sequencing platforms make it possible to read millions of molecules simultaneously. The four most widely used platforms nowadays are *Illumina* and *Ion Torrent* (NGS), and *PacBio* and *Nanopore* (TGS). For instance, there is an interactive learning resource for NGS techniques named SequencEnG (Sequencing Techniques Engine for Genomics), which is supported by the institutions of University of Illinois Urbana-Champaign and NIH big data to knowledge (BD2K). SequencEnG is part of the project KnowEnG (Knowledge Engine for Genomics) (Zhang et al., 2019). In addition, a compilation of the main NGS techniques that can be applied to pigs is presented in **Figure 11**. However, broadly outlined, the most common applications are whole genome sequencing (WGS), whole exome sequencing (WES), messenger RNA sequencing (RNA-Seq), ChIP-sequencing (ChIP-Seq), methylation-sensitive restriction enzyme sequencing (Methyl-Seq), and metagenomics. According to the topic of this thesis, our focus is to outline some of the current applications of RNA-Seq as summarized below.
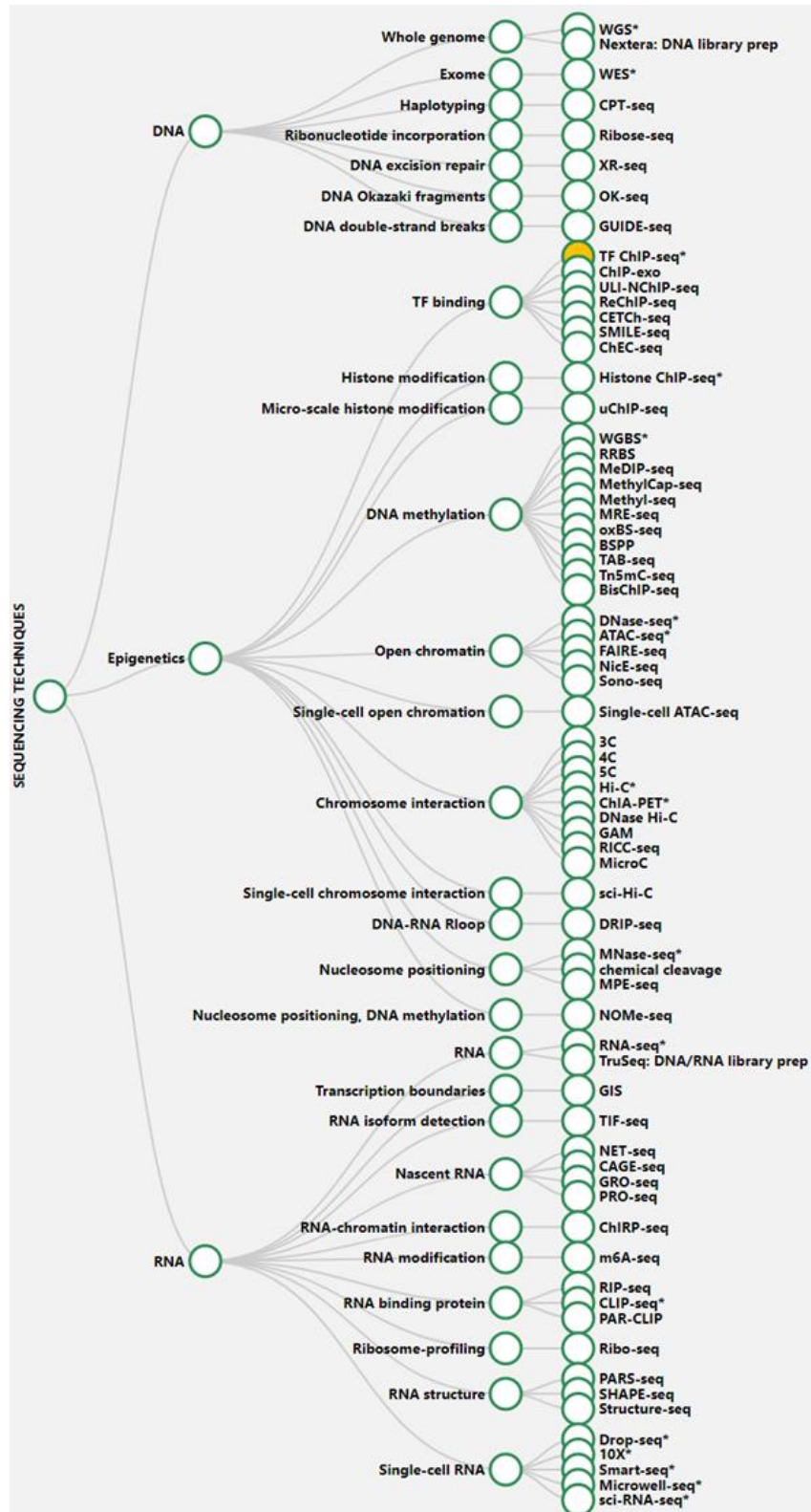
Figure 11. Sequencing tree of analysis pipelines for DNA, RNA and epigenetics techniques. Interactive analysis pipelines are available for sequencing techniques with "*". Source: <https://github.com/KnowEnG/SequencEnG>.

_____

1.4.2. Pig genome sequence variation

Genome sequencing provides important information for the study of genetic variability. Among the genetic markers, microsatellites (tandem repeats of 2 to 5 bases) and SNPs (*single nucleotide polymorphisms*, a variation that affects a nucleotide in the DNA sequence) are well-known in annotated genomes. Particularly, SNP information has been used to build for each species dense panels of markers evenly distributed throughout the entire genome. Several biotechnological tools have been developed to genotype these panels, with two main companies (*Illumina* and *Affymetrix*) leading the market for genotyping platforms for livestock animals (Blasco & Pena, 2018). A range of SNP-based arrays (also known as SNP chips) are commercially available for genotyping at varying densities of SNP panels (**Table 2**). In general, the use of chips with 50,000-80,000 SNPs in a genome is common, and there are high-density chips with 850,000 SNPs or more (Blasco, 2017). Nowadays, sequencing costs are substantially decreasing, and may eventually be comparable to genotyping costs. Ramos et al. (2009) developed the first commercial SNP panel for high-throughput genotyping in pig, which was commercially available from *Illumina* (*PorcineSNP60 BeadChip*, San Diego, CA). Groenen (2015) by *Affymetrix* technology developed a chip containing ~650,000 SNPs and including a large percentage of the SNPs present in the *PorcineSNP60 v2 BeadChip*.

Table 2. Summary of the commercial genotyping chips and its densities (low-density, LD; and high-density, HD), which are currently available for pigs (adapted from Blasco & Pena, 2018), and SNPchimp v.3 database (<https://webserver.ibba.cnr.it/SNPchimp/index.php/data-source>).

Chip name	No. of SNPs	Supplier		Technology
PorcineSNP60 v1 BeadChip (HD)	62,163	*Illumina*		Infinium chemistry
PorcineSNP60 v2 BeadChip (HD)	64,232	*Illumina*		Infinium chemistry
GGP-Porcine (LD)	10,241	*GeneSeek-Neogen*		Infinium chemistry

_____

Table 2. Summary of the commercial genotyping chips and its densities (high-density, HD) *(Continued)*.

Chip name	No. of SNPs	Supplier		Technology
GGP Porcine (HD)	>65,000 (e.g., both of 68,528 and 70,231)	*GeneSeek-Neogen*		Infinium chemistry
Axiom Porcine 650K (HD)	658,692	*Affymetrix*		Axiom assay

1.4.3. Gene expression studies

The expression of a eukaryotic gene is a complex process involving a series of steps prior to the actual synthesis of a protein. These steps include the transcription of the gene into the primary RNA product, processing of this initial gene transcript to remove intron sequences and create the mature 3' terminus, transport of the processed mRNA transcript to the cytoplasm, and then, the translation of the messenger RNA into protein. With very few exceptions, all of the genes that encode proteins follow this pathway. In a typical mammalian cell, messenger RNA (mRNA) which is translated into proteins makes up around 4% of the total RNA mass and, apart from 80% ribosomal RNA (rRNA), other operational RNAs make up the rest (Wu et al., 2014). Nevertheless, the speeds and efficiencies of transcriptional step, RNA processing, nuclear export, and protein translation may differ considerably from gene to gene. In addition, expression of a gene can be controlled at many levels, including transcription, mRNA splicing, mRNA stability, translation and post-translational events such as protein stability and modification (Day & Tuite, 1998). In fact, an important aspect of gene expression is that it can be regulated by other elements such as some types of proteins (transcription factors) or short-chain RNAs.

On the other hand, gene expression has been studied for specific genes or proteins, for multiple genes simultaneously, and at the general level in a tissue or cell type using different technologies. For example, classical single gene measurement using reverse transcription quantitative real-time PCR (RT-qPCR) for the evaluation of candidate genes like *ACSL4* (liver and backfat) for meat quality in pigs (Corominas et al., 2012). In addition to single gene expression, several papers have been reported studying candidate genes for specific traits

_____

related to lipid metabolism, for example, *DGAT1* and *DGAT2* (Cui et al., 2011), *APOA2* (Ballester et al., 2016), *ELOVL6* (Corominas et al., 2015), *FABP4* and *FABP5* (Ballester et al., 2017), *FADS2* (Gol et al., 2018), and *DGAT2* (Solé et al., 2021). By real-time quantitative PCR (qPCR) in a dynamic array chip (Fluidigm) using the BioMark system, the simultaneous expression analysis of 48 genes was performed for lipid metabolism in pigs (Puig-Oliveras et al., 2016, Ballester et al., 2017a; Revilla et al., 2018; Criado-Mesas et al., 2020).

Regarding other techniques for serial analysis of gene expression, microarrays allow the massive measurement and comparison of gene expression covering up to 23,000 probes/features simultaneously. Microarrays in pigs have been used to analyze LD muscle transcriptome using the *Porcine GeneChip array* (*Affymetrix, Inc.*), and underlined as one tool widely used in published experiments interrogating diverse pig muscle phenotypes (Pena et al., 2013, 2014). In pig muscle tissue, microarrays were also used to identify differentially expressed genes (DEGs) among animals with divergent phenotypes for fatness traits (Cánovas et al., 2010), with contrasting levels of fat (Liu et al., 2009), with tenderness and IMF content in pork (Hamill et al., 2012), between genetic types based in Iberian pig production and its crossbreeding with the Duroc during the early stages of development (Óvilo et al., 2014), as well as with different backfat thicknesses in energy metabolism-related tissues (subcutaneous adipose (fat), liver, and LD muscle) (Kojima et al., 2018). Other studies based on microarrays have been conducted to investigate the existence of common determinants of gene expression in the porcine liver and skeletal muscle (González-Prendes et al., 2019), as well as to detect predictors of feed efficiency in growing pigs with extremely low or high residual feed intake (RFI) (Messad et al., 2019).

The microarray technology is limited by their hybridization-based nature (e.g., low sensibility and the high background) and has been progressively replaced by other sequencing methods. Hence, advances in NGS technologies have enabled the whole transcriptome sequencing (RNA-Seq). Unlike microarrays, RNA-Seq technology has better resolution and higher reproducibility. It also allows to determine the abundance of transcripts with a greater dynamic range of expression levels, with higher precision estimates, and the detection of novel transcripts (Nookaew et al., 2012; Black et al., 2014), as well as to study RNA in individual cells including its composition, abundance and variability (Conesa et al., 2016). As a revolutionary tool, RNA-Seq is massively used to explore the transcriptome with an

_____

unprecedented detail, but it also helps to understand molecular mechanisms and then it tells about signaling pathways. In fact, RNA-Seq can be used in combination with other functional genomics methods to enhance gene expression analysis (**Figure 12**) (Conesa et al., 2016).
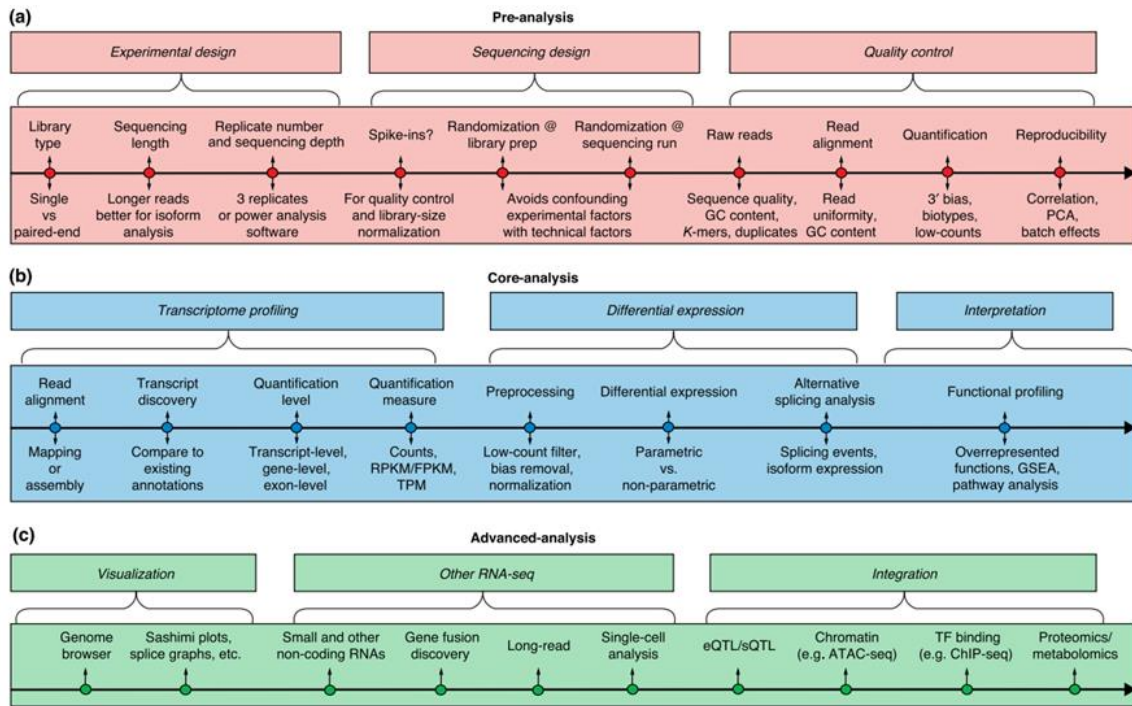
Figure 12. Illustration depicting a generic roadmap for the design and analysis of RNA-Seq experiments using Illumina standard sequencing (i.e., original illustration presented by Conesa et al. 2016). Legend: ChIP-seq: Chromatin immunoprecipitation sequencing; eQTL: Expression quantitative loci; FPKM: Fragments per kilobase of exon model per million mapped reads; GSEA: Gene set enrichment analysis; PCA: Principal component analysis; RPKM: Reads per kilobase of exon model per million reads; sQTL: Splicing quantitative trait loci; TF: Transcription factor; and TPM: Transcripts per million. The key analysis for each step are: (a) Preprocessing includes experimental design, sequencing design, and quality control steps. (b) Core analyses include transcriptome profiling, differential gene expression, and functional profiling. (c) Advanced analysis includes visualization, other RNA-seq technologies, and data integration.

RNA-Seq experiments yield very large, complex data sets that demand fast, accurate and flexible software to reduce the raw read data to comprehensible results. Consistently, Pertea et al., (2016) have described a standard protocol that comprises all steps necessary to process a

_____

large set of raw sequencing reads and create lists of gene transcripts, expression levels, and DEGs and transcripts. Focusing on open-source software and alternative tools available for bioinformatics analysis of RNA-Seq, among the most popular are the STAR, HISAT, StringTie, and Ballgown or Rsubread package (Dobin et al., 2013; Frazee et al., 2015; Kim et al., 2015; Pertea et al., 2015; Liao et al., 2019). Together, they allow scientists to align reads to a genome, assemble transcripts including novel splice variants, compute the abundance of these transcripts in each sample and compare experiments to identify differentially expressed genes (DEGs) and transcripts.

Until March 2023, 32,933 publications containing RNA-Seq data appeared in PubMed. For instance, one of the first applications of RNA-Seq in pigs was the study of the transcriptome profile in liver, adipose and muscle tissues from animals with extreme phenotypes for the intramuscular FA composition (Ramayo-Caldas et al., 2012a; Corominas et al., 2013; Puig-Oliveras et al., 2014; Lim et al., 2017). Besides, several RNA-Seq studies in pigs have reported DEGs related to sex, breed, growth rate, fatness, back fat thickness, lipid profile, meat quality, immunity, and thermogenic role of muscle tissue (Zhao et al., 2011; Esteve-Codina et al., 2011; Pérez-Montarelo et al., 2012; Jiang et al., 2013; Sodhi et al., 2014; Ghosh et al., 2015; Cardoso et al., 2017, 2018; Gao et al., 2019; Albuquerque et al., 2021; Zhang et al., 2022; Zhang et al., 2023). In addition, by machine learning applications, the identification of genes associated with RFI (including the prediction of lipid-related genes) has been reported using RNA-Seq data from pigs classified as phenotypically extreme (i.e., high or low RFI groups) (Piles et al., 2019).

Other examples include several comparative studies in muscle tissue which compared the transcriptome profile between lean (Landrace, Large-White) and fatty breeds (mainly Chinese as Lantang native, and other local obese breeds, e.g., Korean native, Italian Casertana, and Basque) (Wimmers et al., 2007; Kim et al., 2010; Zhao et al., 2011; Damon et al., 2012; Timperio et al., 2013; Yu et al., 2013).

1.5. Genomic studies of meat quality-related traits in pigs

In this session, we emphasize the main approximations that have allowed the identification of genomic regions associated with meat quality and FA composition, including quantitative trait *loci* (QTL) and candidate genes responsible for QTLs, or even regulating the expression (eQTL)

_____

of lipid-target genes, as well as genome wide association studies (GWAS) in pigs. Identification of genes and variants linked to production traits is one of the main goals of QTL and GWAS in livestock animals. Quantitative phenotypes and molecular markers (like microsatellites or SNPs) distributed along the genome are utilized to search QTLs. In light of this, QTL mapping is a powerful approach to detect genomic regions co-segregating with a specific trait by linkage analysis. In addition, GWAS test the association of thousands of genetic variants across the genome with a specific trait or disease. Likewise, this methodology has generated a myriad of robust associations for a range of traits and diseases (Uffelmann et al., 2021).

An extensive summary of QTLs can be found in the database called PigQTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/SS/index) (Hu et al., 2019), which is a valuable tool containing pig QTLs and association data curated from published data. The current release of the PigQTLdb (no. 49, Dec 28, 2022) contains 36,725 QTLs/associations from 788 publications. Those QTLs/associations represent 698 different traits. Among them, there are numerous QTL associated with meat quality, for instance, 1092 QTLs are associated with drip loss, 898 QTLs are associated with intramuscular fat content, and 892 QTLs are associated with meat color, as well as 6,490 QTLs are associated with fatty acid content traits. Our research group based on the projects called IBMAP and many others have reported pioneering research with a relevant contribution in the context of genomic studies. A summary of main candidate genes of QTLs identified for the traits of interest is presented in **Table 3**.

Table 3. Summary of functional candidate genes identified in QTLs or GWAS analysis for pig production traits.

Traits	Chr.	Functional candidate genes	References
Growth, carcass and fatness	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, X	*ADIPOQ, FASN, FTO, IGF2, LEP, LEPR, FABP4, FAT1, MC4R, MRF, MSTN, MYPN, POU1F1, PLAG1, TAS2R39, TAS2R4, PCK1, RYR1, ACACA, SREBF1, PGM2L1, PLBD2, HMGA1, DHCR7, FGF23, MEDAG, DGKI, PTN, APOA2, RETSAT, COPA, RNMT, PALMD, NR3C1, ACSM2B,*	(Óvilo et al., 2000); (Pérez-Enciso et al., 2000); (Kim et al., 2000); (Roberts et al., 2001); (Kennes et al., 2001); (Varona et al., 2002); (Van Laere et al., 2003); (Mercadé et al., 2005, 2006); (Óvilo et al., 2005); (Pérez-Enciso et al., 2005);

Table 3. Summary of functional candidate genes identified in QTLs or GWAS analysis for pig production traits *(Continued)*.

Traits	Chr.	Functional candidate genes	References
Growth, carcass and fatness	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, X	*COG3, PDE10A, DHCR7, MFN2, CCNY, PANK3, TAS2R38, GABRB3, XKR4, MGAM*	(Stachowiak et al., 2013); (Latorre et al., 2016); (Fontanesi et al., 2009, 2010); (Óvilo et al., 2010); (Ballester et al., 2016); (Martínez-Montes et al., 2017, 2018); (Xu et al., 2020); (Gozalo-Marcilla et al., 2021); (Wang et al., 2022)
Meat quality	2, 3, 4, 6, 8, 12, 14, 15, 17, X	*ACACA, FASN, ACSL4, CAPNS1, CAST, CA3, CYBSA, CYP2E1, ELOVL6, FABP4, FABP5, PCK1, PHKG1, PPARGC1A, PRKAG3, RYR1, SCD, TTN, AGPAT5*	(Yu et al., 1999); (Óvilo et al., 2000); (Ovilo et al., 2002); (Milan et al., 2000); (Ciobanu et al., 2001); (Pérez-Enciso et al., 2002); (Mercadé et al., 2005); (Gandolfi et al., 2011); (López et al., 2015); (Zhang et al., 2019); (Piórkowska et al., 2020); (Molinero et al., 2022)
Fatty acid content and composition	2, 5, 6, 7, 8, 9, 12, 14, 16, X	*SCD, ELOVL3, DGAT2, LEPR, ELOVL5, ELOVL6, ELOVL7, FADS2, ABCD2, FASN, IGF2,*	(Pérez-Enciso et al., 2000); (Clop et al., 2003); (Mercade et al., 2006); (Muñoz et al., 2007); (Uemoto et al., 2012); (Renaville et al., 2013); (Corominas et al., 2013, 2015); (Estany et al., 2014);

_____

Table 3. Summary of functional candidate genes identified in QTLs or GWAS analysis for pig production traits *(Continued)*.

Traits	Chr.	Functional candidate genes	References
Fatty acid content and composition	2, 5, 6, 7, 8, 9, 12, 14, 16, X	*ACSL4, CPN1, PAX2, PKD2L1, ENTPD7, SEMA4G, SREBF1, PLA2G7, THRSP*	(Ros-Freixedes et al., 2016); (Zhang et al., 2016); (van Son et al., 2017); (Viterbo et al., 2018); (Criado-Mesas et al., 2019); (Zhang et al., 2019); (Zappaterra et al., 2019); (Crespo-Piazuelo et al., 2020); (Solé et al., 2021)

Among several interesting examples for pigs, *ELOVL6* gene was associated with a major QTL effect on fatty acid composition in an Iberian × Landrace backcross (Corominas et al., 2013, 2015) and expression differences were associated with the abundance of C16:0 and C16:1n-7 FAs in backfat and muscle tissues. Another example is an intronic variant of the *IGF2* gene, which was described as the causal mutation of a QTL for muscle growth and back-fat depth in the pigs (Van Laere et al., 2003). Another example is the *SCD* gene, which has been pinpointed in several GWAS studies as a functional candidate gene for meat quality and fatty acid composition traits. Interestingly, in a purebred Duroc line, a polymorphism in the *SCD* promoter (AY487830:g.2228T>C SNP, rs80912566) was proposed as a causal mutation partially explaining the genetic variance of fat desaturation ratios (e.g., the 18:1/18:0 ratio) in muscle and subcutaneous fat (Estany et al., 2014). In addition, *UNC93A* gene has also been reported as a candidate gene in a QTL study for meat quality and disease resistance in the Chinese Jiangquhai pig breed (Oyelami et al., 2020).

On the other hand, several eQTL studies have been conducted to reveal genome locations that regulate the expression of a target gene. For instance, with expression data obtained by real-time qPCR in a Fluidigm chip in the muscle of Iberian × Landrace backcrossed animals, eQTLs regulating the expression of several lipid-related genes were identified. However, no eQTLs for *SCD* gene expression were found (Puig-Oliveras et al., 2016). Moreover, in Duroc pigs (Lipgen population), a panel of SNPs were associated with microarray data of genes mapping to QTL-SSC14 and involved in IMF traits of the porcine muscle (González-Prendes, et al., 2019). The

_____

authors detected the absence of any eQTL regulating *SCD* mRNA expression. They also found no significant differences in *SCD* expression between pigs with different genotypes of the SNP rs80912566 (Estany et al., 2014). Finally, in a large eQTL study of genes associated with lipid metabolism in muscle (Criado-Mesas et al., 2020) did not detect any eQTL for both the *SCD* and the *PPARG* genes.

_____

Chapter 2. Objectives

This PhD thesis was done under the framework of the research project "Functional genomics, systems biology, and microbiomics applied to the identification of genetic regulators of growth, fatness, and meat quality traits in pigs", funded by the Spanish Ministery of Science and Innovation (AGL2017-82641-R). In addition, the animal material used in this thesis was generated in several projects of the IBMAP consortium, involving members of the INIA, IRTA and UAB institutions.

2.1. Global aim

The main aim of this thesis was to identify candidate genes, biological processes and pathways linked with the intramuscular fatty acid profile in pigs.

2.2. Specific aims

1. To identify genes involved in $\Delta 9$ fatty acid desaturation process by the analysis of the muscle transcriptome in pigs with divergent oleic-to-stearic (C18:1n-9/C18:0) fatty acid ratio in muscle.

2. To explore the association between the porcine intramuscular fatty acid profile and the muscle transcriptome, in order to identify genes and gene networks involved in meat fatty acid composition.

3. To perform a multivariate integrative analysis between the intramuscular fatty acid and gene expression profiles of pig muscle to pint-point key regulators of lipid metabolism.

Papers and Studies

_____

Chapter 3

Identification of Differentially Expressed Genes and Polymorphisms Related to Intramuscular Oleic-to-Stearic Fatty Acid Ratio in Pigs

Jesús Valdés-Hernández [1,2]*, Yuliaxis Ramayo-Caldas [3], Magí Passols [1], Lourdes Criado-Mesas [1], Anna Castelló [1,2], Anna Esteve-Codina [4],Armand Sánchez [1,2], and Josep M. Folch [1,2]*

[1]Plant and Animal Genomics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra, Spain

[2]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain

[3]Departament de Genètica i Millora Animal, Institut de Recerca y Tecnologia Agraroalimentàries (IRTA), Caldes de Montbui, Spain

[4]CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain

*   Correspondence: jesus.valdes@cragenomica.es (J.V.-H.); JosepMaria.Folch@uab.cat (J.M.-F.); Tel.: +34 935636600 ext. 3345 (J.M.-F)

Manuscript in preparation

_____

Simple Summary: Muscle tissue is important in many physiological and metabolic processes, but also for its relevance in the pork meat industry. The transcriptome analysis, by RNA-Seq, of muscle samples from pigs with divergent intramuscular oleic-to-stearic fatty acid ratio (C18:1n-9/C18:0) allowed the identification of 81 differentially expressed genes (DEGs). Among the DEGs, *SCD* is a promising candidate gene to explain the intramuscular C18:1n-9/C18:0 acid ratio differences.

Abstract: The C18:1n-9/C18:0 ratio is an important indicator of the biosynthesis and desaturation of fatty acids in muscle. By using an RNA-Seq approach in muscle samples from 32 BC1_DU (25% Iberian and 75% Duroc) pigs with high (H) and low (L) values of C18:1n-9/C18:0 fatty acids ratio, a total of 81 DEGs were identified. Functional analyses of DEGs indicate that mainly PPAR signaling pathway (*PPARG, SCD, PLIN1,* and *FABP3*) was enriched. *SCD* is directly involved in the conversion of C18:0 to C18:1n-9, and PPARG is a transcription factor regulating lipid metabolism genes, including *SCD*. Seven genetic variants within the *SCD* are associated in two haplotypes. Although the haplotypes are segregating at different frequencies in the H and L groups, they not fully explain the desaturation ratios or the *SCD* expression levels. A more complex model including PUFA is suggested to explain the regulation of the C18:1n-9/C18:0 desaturation ratio in porcine muscle.

Keywords: desaturation index; swine; muscle transcriptome; lipid metabolism

1. Introduction

Fatty acid (FA) content and composition play an important role in meat quality determination affecting flavor, oxidative stability, color, firmness [1,2] and its nutritional value. In pork, the FA composition varies across fat tissues and muscles and it is greatly influenced by several factors including diet, fatness, body weight, gender, age, breed, environment, and secretion of hormones [3,4].

FAs can be classified into three groups (saturated FAs: SFA, monounsaturated FAs: MUFA, and polyunsaturated FAs: PUFA) that are provided by the diet, or derived from *de novo* lipogenesis [5]. In pigs, the essential PUFA (linoleic or C18:2*n-6* and α-linolenic or C18:3*n-3*) are only delivered by the diet and can be used for the synthesis of other PUFA (eicosadienoic or C20:2*n-6*, eicosatrienoic or C20:3*n-3*, and arachidonic or C20:4*n-6* acids). Conversely, *de novo* endogenously synthesized FAs represented ∼86% of the total non-essential FA deposition (including SFA as myristic or C14:0, palmitic or C16:0, and stearic or C18:0, and MUFA such as palmitoleic or C16:1*n-7*, and oleic or C18:1*n-9* acids) [6]. The most important MUFA in pig muscle is C18:1*n-9*, being C16:0 and C18:0 the major SFA, and C18:2*n-6* the most abundant PUFA [2].

The intramuscular fat (IMF) oleic acid content has been positively correlated with pork meat flavour [7], while the percentage of stearic acid has been associated with fat firmness [1]. A high IMF and oleic acid content has been reported both in raw meat and dry-cured products of Iberian pigs feeded by grass and acorns during the fattening period, being a criterion of its quality [8].

The biosynthesis of MUFA from SFA is mainly mediated by stearoyl-CoA desaturase (*SCD*), a regulatory enzyme catalyzing the rate-limiting step in the overall *de novo synthesis* of C18:1*n-9* and C16:1*n-7* acids [9]. These two FAs are produced by the Δ9-desaturation of their respective precursors, the C18:0 and C16:0 acids, respectively [10]. Hence, the desaturation ratio C18:1*n-9*/C18:0 can be used to infer the *SCD* activity [11]. Interestingly, in a purebred Duroc line, a polymorphism in the *SCD* promoter (AY487830:g.2228T>C SNP, rs80912566) was proposed as a causal mutation partially explaining the genetic variance of fat desaturation ratios (e.g., the 18:1/18:0 ratio) in muscle and subcutaneous fat, and the T allele was also associated with a higher S*CD* expression in muscle [11]. In addition, the dry-cured hams from Duroc pigs carrying the T allele showed an increased C16:1, C18:1*n-9*, C18:1*n-7*, and MUFA

content and decreased C18:0 and SFA content [12]. However, a Genome-Wide Association Study (GWAS), performed for the IMF C18:1n-9/C18:0 ratio in the cross BC1_DU (25% Iberian and 75% Duroc) [13], suggested a different causal variant for this trait. Further, a recent association analysis between the muscle transcriptome and its FA profile was performed in 129 Iberian × Duroc backcrossed pigs [14], but the *SCD* expression was not associated with the C18:1n-9/C18:0 ratio.

The present work is focused in the analysis of the muscle transcriptome of pigs with divergent values for the C18:1n-9/C18:0 ratio, with the aim of identifying genes and enriched biological pathways associated with the Δ9-desaturation process. In addition, we simultaneously provided evidence for *SCD* gene polymorphisms that affect its mRNA expression and may emerge as a source of variation in muscle fatty acid desaturation.

2. Materials and Methods

2.1. Ethics statement

All procedures involving animals were conducted in accordance with relevant guidelines and regulations of the animal care and use committee of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). This study followed ethical principles in animal research, according to the Ethical Committee of the IRTA. Our research is also reported in full compliance with ARRIVE guidelines (https://arriveguidelines.org/).

2.2. Animal material and sample collection

A total of 129 animals from an experimental backcross, named BC1_DU (25% Iberian and 75% Duroc), were analised. All animals, 59 females and 70 males (non-castrated), were maintained under the same intensive conditions and fed *ad libitum* with a cereal-based commercial diet and free access to water as described in Martínez-Montes et al. [15]. Animals were slaughtered in five batches, in a commercial abattoir with an average age of 190 days, and 73.70 kg of carcass weight. *Longissimus dorsi* (LD) skeletal muscle samples were collected immediately after slaughter and snap frozen in liquid nitrogen and stored at -80 °C until analysis.

FAs composition in LD samples (n = 129) was determined using a gas chromatography of methyl esters protocol [16]. In brief, a sample of ~200 g of LD muscle was used to measure the FA profile. Then, the quantity of each individual FA methyl ester was calculated and expressed

as relative percentage out of the total amount of FAs. Total percentages of SFA, MUFA, and PUFA were obtained through the sum of the individual FAs. In the present work we focused our attention on the product/substrate ratio of oleic (C18:1n-9) and stearic (C18:0) acids, frequently called as desaturation index. Notwithstanding, we also calculated the desaturation ratio of C16:1n-7 to C16:0, MUFA to SFA and the respective global desaturation index for these FAs (C18:1n-9+C16:1n-7/C18:0+C16:0 ratio). From the 129 BC1_DU animals, 32 samples with extreme values for C18:1n-9/C18:0 ratio were selected for the RNA-Seq analysis, 16 with high values (H) and 16 with low values (L), using balanced sex and slaughterhouse batch factors in both groups. In order to determine significant differences in the FA profile for H and L groups (n = 32 extreme pigs), the non-parametric Mann-Whitney U test via R-base stats v3.6.2 package was used. In addition, the effect size (r) with its magnitude (small: (r) $\geq$ 0.10 & < 0.30, moderate: (r) > 0.30 & < 0.50, and large: (r) $\geq$ 0.50) was calculated using the "wilcox_effsize" function of the rstatix v0.7.0 package [17].

2.3. Total RNA isolation

Total RNA was isolated from the LD muscle of 32 animals using the RiboPure™ Isolation kit for High Quality Total RNA (Ambion®; Austin, TX) following the manufacturer's recommendations. RNA quantification and purity was determined with a NanoDrop ND-1000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). RNA integrity was checked by Agilent Bioanalyzer-2100 equipment (Agilent Technologies, Inc., Santa Clara, CA, USA), and samples with an RNA integrity number (RIN) greater than 7 were used for the RNA-Seq experiment.

2.4. RNA library preparation and sequencing

Library preparation and sequencing was performed at CNAG institute (Centro Nacional de Análisis Genómico, Barcelona, Spain). For each sample, one paired-end library was prepared using TruSeq Stranded mRNA kit (Illumina, Inc.; San Diego CA, USA). To discriminate among samples, libraries were labeled by barcoding and pooled to be run in an Illumina HiSeq 3000/4000 instrument (Illumina, Inc.; San Diego CA, USA). In this study, 2×75 bp reads and a mean of 42.14 million of paired-reads per sample were generated.

2. 5. Single-nucleotide polymorphism (SNP) genotyping

Nine single nucleotide polymorphisms (SNPs) and one Indel located in the *SCD* gene were genotyped in the 32 pigs using Taqman OpenArray™ genotyping plates custom-designed in a QuantStudio™ 12K flex Real-Time PCR System (ThermoFisher Scientific). Of these polymorphism, four were located in the promoter region, one in the 5'UTR region, and four in the 3'UTR region (Table S1).

2.6. Bioinformatic analyses

Quality control and reads statistics were determined by using MultiQC v0.7 program [18]. Sequencing reads were mapped employing the STAR alignment software (v2.7.9a) with default parameters [19], and using the Sscrofa11.1 pig genome assembly as reference (annotation database Ensemble Genes 97). Then, gene expression was quantified by RSEM v1.2.28 program [20].

All statistical analyses were implemented in R v4.1.0 program [21]. Filtering was performed to keep only genes with a minimum of 32 read counts and 14,921 genes were retained. Subsequently, analysis of differentially expressed genes (DEGs) between H and L groups (with L as reference level) was performed using the DESeq2 v1.32.0 package [22]. The statistical negative binomial model to analyze gene expression (variable response) included three predictors or fixed factors, the sex (2 levels), slaughterhouse batch (4 levels) and extreme phenotypes (2 levels: H and L). All genes with a fold change (FC) value between H and L groups of at least 1.2 ($\log_2 FC \geq |0.26|$) and adjusted *P*-value (*P*-adj) < 0.05 for multiple testing correction [23] were selected as DEGs. A principal component analysis (PCA) was performed using normalized count matrix by DESeq2 regularized logarithm (*rlog*) transformation in the factoextra v1.0.7 package [24]. With this step we explored the gene expression variation between the H and L groups.

Finally, the list of DEGs were submitted to the ClueGO v2.5.8 plug-in [25] in Cytoscape v3.8.2 software [26] for GO term and KEGG pathways enrichment analysis. Gene Enrichment significance was assessed with a hypergeometric test using the method of Benjamini-Hochberg [23] for multiple testing correction (corrected *P*-value < 0.05).

_____

2.7. Validating RNA-Seq results with RT-qPCR

Three of the DEGs, fatty acid binding protein 3 (*FABP3*), peroxisome proliferator activated receptor gamma (*PPARG*) and stearoyl-CoA desaturase (*SCD*), were selected due to its biological relevance for gene expression profiling by reverse transcription quantitative real-time PCR (RT-qPCR) in the LD muscle of the 32 extreme pigs. All the steps for Multiplex RT-qPCR were described in Puig-Oliveras et al. [27], including the housekeeping genes (i.e., Actin beta: *ACTB* and TATA-box binding protein: *TBP*). According to Criado-Mesas et al. [28], the gene expression levels were expressed as normalized quantity values (NQ) applying the relative standard curve method. Besides, a $\log_2$ transformation for data normalization was applied. Significant differences (*P*-value < 0.05) in RT-qPCR gene expression levels between H and L groups were determined using a t-test approach. Pearson correlation between RT-qPCR ($\log_2$NQ) and RNA-Seq ($\log_2$CPM via edgeR v3.40.2 package) [29] expression values was calculated using the cor.test function.

3. Results and Discussion

3.1. Description of the intramuscular fatty acid profile

The intramuscular FA profile of the 32 animals selected for RNA-Seq (Table 1) are in agreement with the range of values previously reported in pig muscle [2]. As expected, higher desaturation indexes were observed in H group in comparison to L animals, including the C18:1*n*-9/C18:0 ratio (trait of interest, $P = 1.54E\text{-}06$ with CV = 22.33%), the C16:1*n*-7/C16:0 ratio ($P = 2.64E\text{-}05$), the C18:1*n*-9+C16:1*n*-7/C18:0+C16:0 ratio ($P = 3.33E\text{-}09$), and the MUFA/SFA ratio ($P = 3.33E\text{-}09$). In addition, the proportion of the products of the desaturation process (C18:1*n*-9, $P = 2.23E\text{-}07$; C16:1*n*-7, $P = 1.47E\text{-}05$; and MUFA, $P = 1.50E\text{-}07$) was higher in H animals. Conversely, no significant differences were detected for the substrates of *SCD* gene (C16:0, $P = 0.10$; C18:0, $P = 0.24$; and SFA, $P = 0.81$). *SCD* is the rate-limiting enzyme catalyzing the synthesis of MUFA, mainly C18:1*n*-9 and C16:1*n*-7 [9]. The two major products of *SCD* reaction are C18:1*n*-9 and C16:1*n*-7 formed by Δ9-desaturation [10] of C18:0 and C16:0, respectively. Hence, the higher FAs desaturation indices (C18:1*n*-9/C18:0 and C16:1*n*-7/C16:0) of H animals can be explained by a higher expression of the *SCD* gene.

In addition, a high variability (CV) in all percentages of PUFA traits was observed (Table 1), being lower its content in H animals. The comparison between both groups indicates a lower

proportion of the main PUFA (C18:2n-6, P = 7.07E-05; C18:3n-3, P = 2.98E-03; C20:4n-6, P = 5.73E-05; and PUFA, P = 7.07E-05) in H animals. Interestingly, both n-3 and n-6 PUFA have been described as powerful inhibitors of *SCD* gene expression [30]. However, in the 32 animals analyzed here, no significant correlation was observed between PUFA and *SCD* expression, suggesting a different molecular mechanism connecting the lower PUFA with the higher MUFA contents observed in H-animals muscle. In addition, the IMF content is also higher in H animals (*P = 2.59E-03*), while no differences in carcass weight (CW) were observed.

Table 1. Relative percentage of intramuscular FA composition and differences between H and L groups of BC1_DU pigs.

Trait	Parameter		Mean by group ± SEM		Significance	
	Mean global ± SEM	CV (%)	H	L	P-value[t]	Effsize
IMF (%)	3.608±0.229	35.847	4.187±0.385	3.028±0.151	2.59E-03	0.520[L]
CW (kg)	76.759±1.242	9.155	76.006±2.012	77.512±1.502	6.65E-01	0.080[S]
Fatty acid composition (%)						
C14:0	1.235±0.050	22.900	1.354±0.048	1.115±0.078	1.70E-02	0.420[M]
C16:0	23.287±0.345	8.369	23.814±0.436	22.761±0.513	1.02E-01	0.293[S]
C18:0	13.691±0.382	15.788	13.220±0.350	14.162±0.672	2.39E-01	0.213[S]
C16:1n-7	2.746±0.118	24.227	3.216±0.102	2.275±0.131	1.47E-05	0.706[L]
C18:1n-9	34.314±1.316	21.699	40.302±0.847	28.327±1.290	1.82E-06	0.800[L]
C18:2n-6	14.120±1.301	52.112	9.485±1.177	18.755±1.657	7.07E-05	0.660[L]
C18:3n-3	0.444±0.025	31.952	0.379±0.027	0.509±0.037	2.98E-03	0.513[L]
C20:4n-6	3.288±0.455	78.240	1.622±0.310	4.955±0.623	5.73E-05	0.666[L]
SFA	38.213±0.679	10.044	38.389±0.730	38.038±1.169	8.38E-01	0.040[S]
MUFA	37.060±1.399	21.349	43.518±0.866	30.602±1.332	1.50E-07	0.806[L]
PUFA	17.853±1.752	55.515	11.486±1.497	24.219±2.246	7.07E-05	0.660[L]
C16:0+ C16:1n-7	26.033±0.425	9.231	27.030±0.507	25.036±0.597	2.34E-02	0.400[M]
C18:0+ C18:1n-9	48.006±1.471	17.329	53.522±1.160	42.489±1.88	8.69E-05	0.653[L]
C16:1n-7/C16:0	0.117±0.004	20.478	0.135±0.003	0.100±0.005	2.64E-05	0.746[L]
C18:1n-9/C18:0	2.538±0.100	22.336	3.060±0.042	2.016±0.058	1.54E-06	0.853[L]
MUFA/SFA	0.969±0.032	18.643	1.134±0.012	0.803±0.020	3.33E-09	0.853[L]
MUFA/PUFA	3.042±0.350	65.047	4.448±0.395	1.635±0.293	1.87E-05	0.700[L]
PUFA/SFA	0.493±0.057	65.837	0.315±0.056	0.672±0.079	1.06E-03	0.560[L]
C16:1n-7+ C18:1n-9/ C16:0+ C18:0	1.001±0.034	18.931	1.176±0.012	0.827±0.021	3.33E-09	0.853[L]

[t]Value calculated with the non-parametric Mann-Whitney U test using "wilcox.test()", SEM = standard error of the mean, CV = coefficient of variation, Effsize = effect size (r) calculated via "wilcox_effsize()". The superscripts [S, M, L] for (r) are the magnitudes commonly interpreted in the published literature as: small effect (0.10 to < 0.3), moderate effect (0.30 to < 0.5) and large

_____

effect (>= 0.5). SFA (C14:0 + C16:0 + C18:0), MUFA (C16:1n-7 + C18:1n-9), and PUFA (C18:2n-6 + C18:3n-3 + C20:4n-6) = sum of main saturated, monounsaturated and polyunsaturated FAs, respectively.

3.2. Transcriptome analysis of the LD muscle tissue in pigs

RNA-Seq data of LD muscle from 32 pigs with extreme C18:1n-9/C18:0 ratio (H and L) was generated in the sequencing experiment. After quality control, a mean of 42.14 million of 75 bp-long paired-reads per sample were obtained. In total, 137.8 gigabases (Gb) of sequence was generated, and with a range of 2.50 to 8.90 Gb per sample. Approximately 88.40% (ranging from 74.4 to 95.56%) of the reads were uniquely mapped to the Sscrofa11.1 pig genome assembly. On the other hand, a total of 25,880 genes corresponding to 22,342 annotated protein coding genes were identified. Of these, 14,958 were known genes with symbol ID.

3.3. Differential gene expression analysis

After filtering the matrix of counts, a total of 14,921 genes were subjected to DEGs analysis between the two groups of extreme pigs (H and L). According to the employed cutoff (absolute FC ≥ 1.2 and P-adj ≤ 0.05), 79 protein-coding genes and two pseudogenes were identified as DEGs (Table S2). The volcano plot (**Figure 1**) shows twenty-four genes that had a higher expression (29.63%) in the H group in comparison with L group (i.e., the reference level) and fifty-seven with lower expression (70.37%).
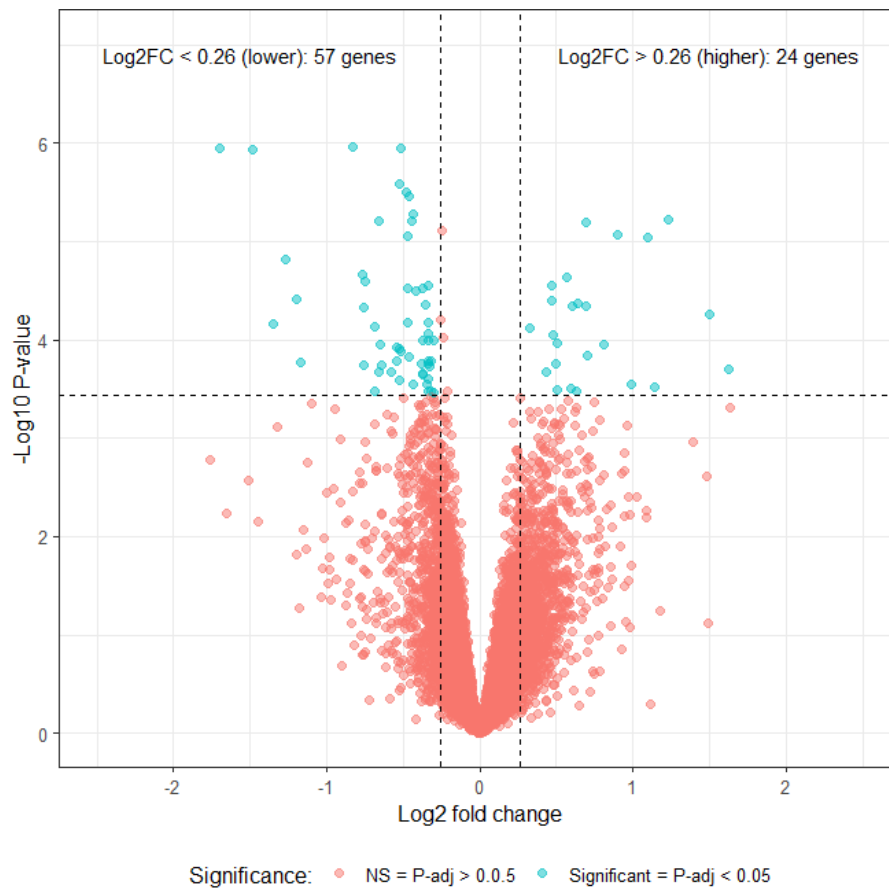
Figure 1. Muscle RNA-Seq differentially expressed genes between H and L groups for C18:1*n*-9/C18:0 ratio in the LD muscle of BC1_DU pigs. Each dot represents a gene (14,921 genes). On the x-axis the log$_2$FC values with thresholds of −0.26 and 0.26 and limits of −2.5 and 2.5 were represented. On the y-axis, the significance in gene expression is represented as -log10 of the *P*-value with limits of 0 to 7. Cyan dots indicate differentially-expressed protein-coding genes by FC and *P*-adj (81 DEGs), while red dots correspond to the rest of the genes non DEGs. Black dashed lines indicate a log fold-change threshold of -1.2 and 1.2 and an adjusted *P*-value (*P*-adj) threshold of 0.05 (the –Log$_{10}$*P*-value of 3.43 corresponds to the *P*-adj threshold).

A principal component analysis (PCA) for the expression values of the 81 DEGs showed a clear separation between samples of the H and L groups. But also, PCA results indicate similarities among samples subjected to the same group (i.e., samples that cluster together were reasonably well-separated along the x-axis) (**Figure 2**).
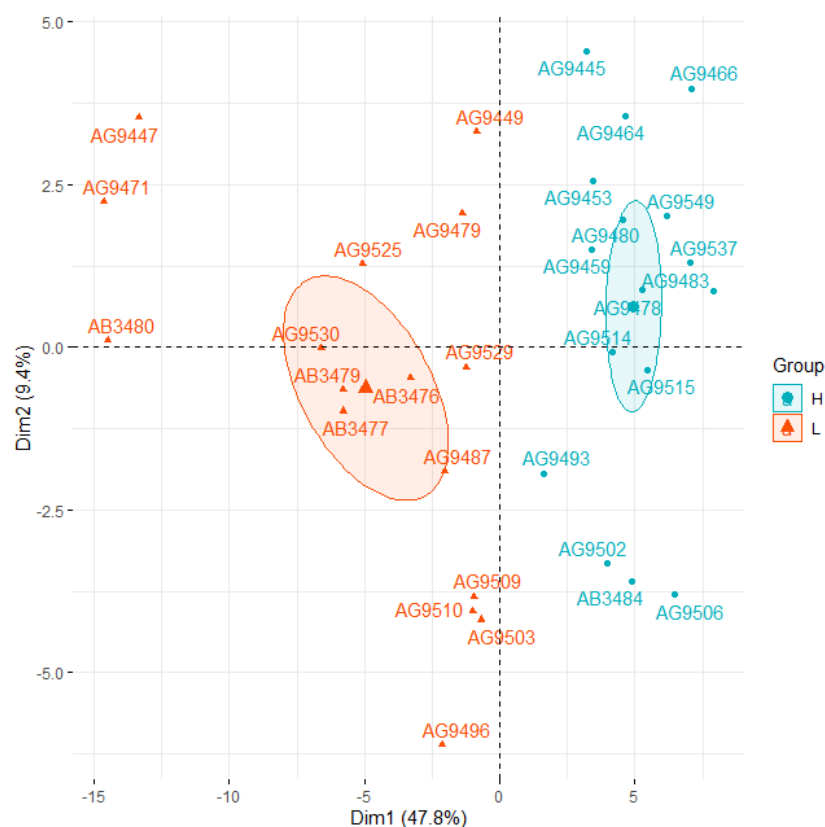
Figure 2. Principal component analysis summarizing the separation and similarities among groups (H and L) using *rlog* values of 81 DEGs in factoextra R package. This PCA plot illustrates the separation between H and L animals, differing in the C18:1*n*-9/C18:0 ratio. The two first principal components (PC1: x-axis and PC2: y-axis, respectively) are shown.

In general, the scientific literature search of the 81 DEGs (Table S2) revealed that ten genes (*ACAD10*, *ACADVL*, *ECHDC3*, *FABP3*, *FADS3*, *ILVBL*, *MMAA*, *PLIN1*, *PPARG*, and *SCD*) are well-known for their role in energy, lipid and carbohydrate metabolism, and mitochondrial fatty acid beta-oxidation; which are important pathways to modulate the FA composition. Three of these genes (*ACADVL*, *PLIN1* and *SCD*) and four other DEGs (*DZANK1*, *GOT1*, *PANK1*, and *PPP1R1B*) were previously detected as DEGs in the LD muscle transcriptome of phenotypically extreme animals (H vs. L groups, only females) for intramuscular FA composition from an Iberian × Landrace backcross population (BC1_LD) [31]. In addition, the hepatocyte growth factor activator (*HGFAC*, *P*-adj = 4.34E-07), which is the most significant DE gene, has been reported as a ChREBP-regulated hepatokine in mice and humans. *HGFAC* codes for a protease that activates the pleiotropic hormone hepatocyte growth factor [32]. This gene enhances lipid and glucose homeostasis, mediated in part by the activation of hepatic peroxisome proliferator-activated receptor gamma (*PPARG*).

3.3.1. Genes with increased expression in H compared with L pigs

Among the up-regulated genes in H-group animals, the stearoyl-CoA desaturase (*SCD*) was one of the most significant DEGs (log$_2$FC = 1.10, *P* = 9.1298E-06, *P*-adj = 6.24E-03, Table S2). This is the expected result according to the *SCD* gene function, which converts SFA (16-carbon 16:0 acid and 18-carbon 18:0 acid) to MUFA (16:1*n*-7 and 18:1*n*-9 acids, respectively) [33]. Hence, a higher *SCD* gene expression is likely determining the increased desaturation ratio values (C18:1*n*-9/C18:0, C16:1*n*-7/C16:0, C18:1*n*-9+C16:1*n*7/C18:0+C16:0, and MUFA/SFA) in muscle in H animals.

Complementarily, we also investigated the correlation between the expression of *SCD* and phenotypic value of the desaturation ratios in muscle. Here we found that the correlation was moderate, in detail r = 0.65 for C18:1*n*-9/C18:0, r = 0.68 for C16:1*n*-7/C16:0, and r = 0.54 for C18:1*n*-9+C16:1*n*-7/C18:0+C16:0; in all three cases with *P* ≤ 0.001.

A polymorphism in the promoter region of the *SCD* gene (AY487830:g.2228T>C SNP, rs80912566) was described as the causal mutation for muscle and fat desaturation ratios (18:1/18:0 ratio) in purebred Duroc pigs, being the T allele associated with higher desaturation ration and increased *SCD* expression in muscle [11]. However, in a GWAS performed for the IMF C18:1*n*-9/C18:0 ratio in animals of the BC1_DU backcross (the 32 RNA-Seq animals belong to this backcross) [13], the rs80912566 SNP was not the most significantly associated to IMF C18:1*n*-9/C18:0 ratio. Here, we genotyped ten additional genetic variants, including this putative causal SNP, located in the promoter, 5'UTR and 3'UTR regions of the *SCD* gene (Table S1) in the 32 RNA-Seq animals. Animals of the H group had a higher allelic frequency of the T allele of the rs80912566 SNP (0.91) in comparison to L group animals (0.59). Seven SNPs, from the promoter to the 3'UTR region, are associated in two haplotypes (from rs80813866 at SSC14:111460852 to rs334462984 at 14:111475110), being three SNPs of the 3'UTR excluded (rs338494000, rs710198292, and rs331969256). Haplotype H1 (C-C-C-T-A-C-G) contains the T allele of the rs80912566 SNP, while haplotype H2 (A-G-T-C-G-T-C) has the alternative C allele. Among the H animals 13 where homozygous for haplotype H1 (H1H1) and three heterozygous (H1H2). Conversely, among L animals four where homozygous for haplotype (H1H1), 11 heterozygous (H1H2) and one homozygous for haplotype H2 (H2H2). Hence, H animals showed a higher frequency of the H1H1 genotype than L animals, but the haplotype is not explaining completely the classification in H and L animals, nor the *SCD*

_____

expression or the C18:1n-9/C18:0 ratio (**Figure 3**). These results suggest that other causal variant in linkage disequilibrium with rs80912566 may be explaining the IMF C18:1n-9/C18:0 ratio and *SCD* expression differences. Hence, further studies will be required to find and validate the causal mutation in the BC1_DU population, as well as in other pig populations.
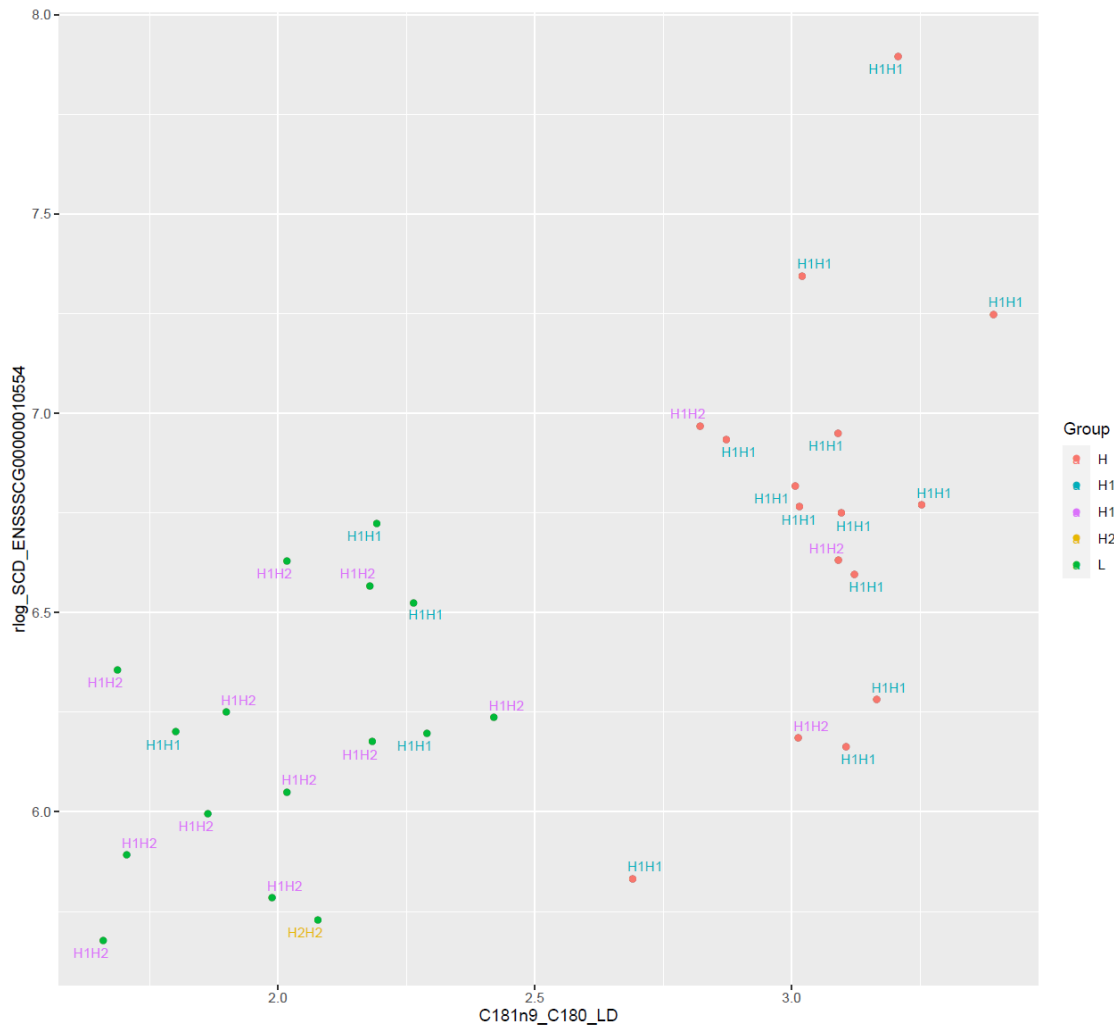
Figure 3. Distribution of the haplotypes of single nucleotide polymorphisms (SNPs, rs80813866, rs80956173, rs323081995, rs80912566, rs342182479, rs331901016, and rs334462984) located in the *SCD* gene. Such haplotypes are represented according to the classification in H and L animals, SCD expression, and C18:1n-9/C18:0 ratio.

Several transcription factors (TFs) binding sites were found in the pig *SCD* gene promoter [11], including the PUFA response element (PUFA-RE), the sterol regulatory element (SRE) where the SREBP-1C transcription factor binds, the peroxisome proliferator-activated receptor gamma (*PPARG*) motif, and the retinoic receptor and the retinoic acid receptor α (RXR:RARα) binding elements, suggesting a fine regulation of its expression. In our RNA-Seq study, the

PPARG gene expression was increased in the H group (log$_2$FC = 0.63, *P* = 3.28E-04, *P*-adj = 4.94E-02, Table S2). *PPARG* is a member of peroxisome proliferator-activated receptors (PPARs), which are a family of transcription factors involved in the regulation of lipid metabolism genes [34], including the induction of the *SCD* gene expression [35]. Hence, a higher expression of *PPARG* in H animals is associated with higher *SCD* levels in pig muscle, but we cannot conclude that there is a direct effect. Moreover, PUFA (including long chain FAs, LCFA) has been reported to repress the induction of lipogenic genes (e.g., *SCD*) through different mechanisms [36–38], including binding to the PUFA-RE element and suppression of *SREBP-1c* gene transcription and proteolytic processing, and enhancing mRNA *SREBP-1c* degradation. For instance, *SREBP-1c* has been reported to activate *SCD* gene expression by binding to the SRE element of the promoter. The lower PUFA content in the muscle of H animals (*P* = 7.07E-05) might repress the *PPARG* gene expression (r $_{PUFA-PPARG}$ = -0.52, *P* = 2.01E-03), and consequently modulate the *SCD* gene expression in LD muscle.

As shown in Table S2, another attractive DEG between the two groups of animals was the perilipin 1 gene (*PLIN1*) which showed higher expression (log$_2$FC = 1.63, *P* = 2.01E-04, *P*-adj = 0.03) in the H group. *PLIN1* is a member of the PAT family of proteins that are able to bind lipid droplets [39] and plays a role in regulating intracellular lipid storage and mobilization, controlling lipid homeostasis. *PLIN1* has been identified among the LD muscle overexpressed genes in pigs with a high IMF content, both in Iberian purebred pigs [40] and in commercial hybrid pigs [41]. Similarly, H animals, with higher *PLIN1* expression, showed a higher IMF content in LD muscle (*P* = 2.59E-03, Table 1).

The fatty acid desaturase 3 gene (*FADS3*) (log$_2$FC = 0.51, *P* = 1.07E-04, *P*-adj = 0.02) was overexpressed in the H-group animals (Table S2). *FADS3* is a member of the FADS family of fatty acid desaturases, including *FADS1* and *FADS2* genes, clustered together on porcine chromosome 2 [42]. *FADS3* function is unknown, but its high sequence homology to functional *FADS1* and *FADS2* suggest that it is involved in long chain PUFA biosynthesis [43]. In addition, the expression of *FADS3* has been shown to be up-regulated markedly in mice lacking *FADS2* [44], which suggested that *FADS3* is a homologous gene to *FADS2*.

Our results also showed an overexpression of nerve growth factor receptor (*NGFR*) (log2FC = 1.23, *P* = 6.09E-06, *P*-adj = 5.65E-03) in H group animals (Table S2). *NGFR* gene is regulating the translocation of *GLUT4* to the cell surface in adipocytes and muscle cells in response to

insulin [36]. In addition, it has been involved in the regulation of insulin-dependent glucose uptake and the regulation of glucose homeostasis [45].

3.3.2. Genes with decreased expression in H compared with L pigs

The fatty acid binding protein 3 (*FABP3*) ($\log_2$FC = -0.66, *P* = 2.11E-04, *P*-adj = 0.03) had a lower expression in the H vs. L groups (Table S2). The *FABP3*, also known as *H-FABP* (heart-type cytoplasmic fatty acid-binding protein), belongs to a family of FA binding proteins (FABPs, with nine members), which are involved in intracellular transport of long-chain FAs in muscle cells and the regulation of FA uptake [46,47]. FABPs are intracellular lipid chaperones which dictate the destiny of lipids and coordinate lipid trafficking and signaling in cells. As a member of this gene family, *FABP3* is expressed in a large number of tissues, including heart, skeletal muscle and adipose tissue among others, but also it has been reported to be involved in FAs uptake and transport to mitochondria for β-oxidation in muscle [48]. Furthermore, FABPs binding FAs (overall to long-chain FAs) can potentially enter the nucleus to target members of the PPARs family (including *PPARG*) of transcription factors [48].

3.4. Gene ontology and pathways analysis

To better understanding the underlying biological processes, molecular functions and metabolic pathways affected by DEGs (81 genes), functional gene ontology (GO) and KEGG enrichment analyses were performed. The 81 DEGs were significantly enriched (Kappa score = 0.4 and *P*-adj < 0.05) in 2 GO and 1 KEGG pathway (Table 3). The main GO term detected was the *PPAR signaling pathway* (*FABP3, PLIN1, PPARG,* and *SCD*). In concordance, the peroxisome proliferator-activated receptor (*PPAR*) *signaling pathway* was significantly overrepresented in the muscle transcriptome analysis of pigs (Iberian × Landrace backcross) with divergent muscle FA composition [31]. PPARs are important upstream transcription factors regulating fatty acid metabolism and modulating gene expression of target genes, depending on the presence of co-repressors or co-activators [49]. *PPARG* is one of the three known isoforms of PPARs, which are ligand-activated transcription factors. Likewise, *PPARG* is involved in the regulation of adipogenesis, lipogenesis, lipid storage, and insulin sensitivity. In addition, a CRISPR/Cas9 transgenic pig with increased expression of *PPARG* in skeletal muscle showed a higher expression of genes involved in adipocyte differentiation and fat deposition as *FABP4*, *LPL*, and *PLIN1* genes [50].

_____

Table 2. Functional analysis for the 81 differentially expressed genes between H and L groups in BC1_DU pigs.

GO term	*P*-value	*P*-value by B-H	Associated genes
Activation of cysteine-type endopeptidase activity involved in apoptotic process	1.52E-03	1.52E-03	*BOK, NGFR, PPARG*
Negative regulation of ERK1 and ERK2 cascade	1.09E-03	1.64E-03	*CAMK2N1, EPHB2, VRK3*
PPAR signaling pathway	1.96E-04	5.87E-04	*FABP3, PLIN1, PPARG, SCD*

Footnotes: These results of gene ontology analysis were obtained via ClueGO plug-in in Cytoscape software, B-H is Benjamini-Hochberg method indicating the type of correction for *P*-value.

3.5. Validating RNA-Seq results with RT-qPCR

We selected three protein coding genes (*FABP3*, *PPARG* and *SCD*) with an important role in lipid metabolism to validate the differential expression RNA-Seq results via RT-qPCR. There was a good concordance between the expression values measured by RNA-Seq ($\log_2$CPM) and the RT-qPCR ($\log_2$NQ) methods, showing significant correlations ($P \leq 0.005$) of 0.776 for *FABP3*, 0.693 for *PPARG*, and 0.834 for *SCD* genes. Since PPARG is a transcription factor which binds to the *SCD* gene promoter [11], the correlation between gene expression of both genes was screened. A strong correlation between the expression of *SCD* and *PPARG* genes (r = 0.75, *P* = 9.295E-07) measured by RT-qPCR was observed. A similar correlation value between the expression of these genes (r = 0.78, $P < 1.00 \times 10^{-16}$) has also been reported in Iberian × Landrace backcrossed pigs [27].

Comparison of the H and L groups RT-qPCR expression levels indicated significant differences ($P < 0.05$) for *SCD* and *PPARG* genes and suggestive differences ($P < 0.1$) for *FABP3* gene (Table 3). Hence, confirming the expression findings of the RNA-Seq analysis.

_____

Table 3. Validations of *SCD*, *PPARG*, and *FABP3* gene expression in H and L animals obtained by Real-Time quantitative PCR in the LD muscle of BC1_DU pigs.

Gene	n	Mean ± SD	Mean by group in $\log_2$NQ scale ± SEM		Significance	
			H	L	CI95%	*P*-value
SCD	31[¥]	0.27±0.53	-1.98±0.32	-3.40±0.20	0.63, 2.20	0.00101
PPARG	31[¥]	0.27±0.22	-1.89±0.21	-2.48±0.15	0.032, 1.14	0.03891
FABP3	31[¥]	0.96±0.62	-0.56±0.19	-0.02±0.21	-1.14, 0.05	0.07193

[¥]Of the total of 32 samples, only one NA value was observed in the three genes, n = number of data, Mean ± SD = raw means and SD = standard deviation, CI95% = confidence interval at 95 percent, SEM = standard error of the mean, and NQ = normalized quantity values. In the comparison of means between H and L groups, the function "t.test()" was used.

4. Conclusions

In summary, we identified 81 DEGs and the *PPAR signaling pathway* overrepresented from the LD muscle transcriptome analysis in pigs with extreme values of the intramuscular C18:1*n*-9/C18:0 ratio. Among the DEGs, *SCD* showed a higher expression in H animals in agreement with their increased desaturation indices (C18:1*n*-9/C18:0 and C16:1*n*-7/C16:0). Seven SNPs located in the *SCD* gene are segregating in two haplotypes with different frequencies in the H and L animals. However, other additional factors are required to explain the differences in desaturation ratios. Other DEGs may be associated with the differences between H and L animals, not only in C18:1*n*-9/C18:0 ratio, but also in IMF and PUFA like C18:2*n*-6, C18:3*n*-3, and C20:4*n*-6. A model including *SCD* gene expression and PUFA is suggested to explain the C18:1*n*-9/C18:0 ratio differences.

Supplementary Materials: Table S1: Genotyping of SNPs located in the *SCD* gene by Taqman OpenArray™ plates in LD muscle from BC1_DU pigs. Description: Information of primers used for the genotyped of ten additional genetic variants, including the putative causal SNP (rs80912566), located in the promoter, 5'UTR and 3'UTR regions of the *SCD* gene; Table S2: Differentially expressed genes between H and L groups for C18:1n-9/C18:0 ratio in LD muscle

_____

from BC1_DU pigs. Description: FC: Fold change. *P*-value calculated with Wald test. *P*-adj: adjusted *P*-values by Benjamini-Hochberg method.

References

1. Wood, J.D.; Richardson, R.I.; Nute, G.R.; Fisher, A. V; Campo, M.M.; Kasapidou, E.; Sheard, P.R.; Enser, M. Effects of Fatty Acids on Meat Quality: A Review. *Meat Sci.* **2003**, *66*, 21–32.

2.   Wood, J.D.; Enser, M.; Fisher, A. V.; Nute, G.R.; Sheard, P.R.; Richardson, R.I.; Hughes, S.I.; Whittington, F.M. Fat Deposition, Fatty Acid Composition and Meat Quality: A Review. *Meat Sci.* **2008**, *78*, 343–358.

3.   Nürnberg, K.; Wegner, J.; Ender, K. Factors Influencing Fat Composition in Muscle and Adipose Tissue of Farm Animals. *Livest. Prod. Sci.* **1998**, *56*, 145–156.

4.   Lim, B.; Lee, K.T.; Park, J.E.; Chung, W.H.; Jang, G.W.; Choi, B.H.; Hong, K.C.; Kim, T.H. Age-Related Changes in Intramuscular and Subcutaneous Fat Content and Fatty Acid Composition in Growing Pigs Using Longitudinal Data. *Meat Sci.* **2012**, *91*, 358–363.

5.   Ameer, F.; Scandiuzzi, L.; Hasnain, S.; Kalbacher, H.; Zaidi, N. De Novo Lipogenesis in Health and Disease. *Metabolism* **2014**, *63*, 895–902.

6.   Kloareg, M.; Noblet, J.; Milgen, J. van. Deposition of Dietary Fatty Acids, de Novo Synthesis and Anatomical Partitioning of Fatty Acids in Finishing Pigs. *Br. J. Nutr.* **2007**, *97*, 35–44.

7.   Cameron, N.D.; Enser, M.; Nute, G.R.; Whittington, F.M.; Penman, J.C.; Fisken, A.C.; Perry, A.M.; Wood, J.D. Genotype with Nutrition Interaction on Fatty Acid Composition of Intramuscular Fat and the Relationship with Flavour of Pig Meat. *Meat Sci.* **2000**, *55*, 187–195.

8.   Ruiz-Carrascal, J.; Ventanas, J.; Cava, R.; Andrés, A.I.; García, C. Texture and Appearance of Dry Cured Ham as Affected by Fat Content and Fatty Acid Composition. *Food Res. Int.* **2000**, *33*, 91–95.

9.   Ntambi, J.M.; Miyazaki, M. Regulation of Stearoyl-CoA Desaturases and Role in Metabolism. *Prog. Lipid Res.* **2004**, *43*, 91–104.

10. Guillou, H.; Zadravec, D.; Martin, P.G.P.; Jacobsson, A. The Key Roles of Elongases and Desaturases in Mammalian Fatty Acid Metabolism: Insights from Transgenic Mice. *Prog. Lipid Res.* **2010**, *49*, 186–199.

11. Estany, J.; Ros-Freixedes, R.; Tor, M.; Pena, R.N. A Functional Variant in the Stearoyl-CoA Desaturase Gene Promoter Enhances Fatty Acid Desaturation in Pork. *PLoS One* **2014**, *9*, e86177.

12. Henriquez-Rodriguez, E.; Tor, M.; Pena, R.N.; Estany, J. A Polymorphism in the Stearoyl-CoA Desaturase Gene Promoter Increases Monounsaturated Fatty Acid Content in Dry-Cured Ham. *Meat Sci.* **2015**, *106*, 38–43.

13. Crespo-Piazuelo, D.; Criado-Mesas, L.; Revilla, M.; Castelló, A.; Noguera, J.L.; Fernández, A.I.; Ballester, M.; Folch, J.M. Identification of Strong Candidate Genes for Backfat and Intramuscular Fatty Acid Composition in Three Crosses Based on the Iberian Pig. *Sci. Reports 2020 101* **2020**, *10*, 1–17.

14. Valdés-Hernández, J.; Ramayo-Caldas, Y.; Passols, M.; Sebastià, C.; Criado-Mesas, L.; Crespo-Piazuelo, D.; Esteve-Codina, A.; Castelló, A.; Sánchez, A.; Folch, J.M. Global Analysis of the Association between Pig Muscle Fatty Acid Composition and Gene Expression Using RNA-Seq. *Sci. Reports 2023 131* **2023**, *13*, 1–12.

15. Martínez-Montes, Á.M.; Fernández, A.; Muñoz, M.; Noguera, J.L.; Folch, J.M.; Fernández, A.I. Using Genome Wide Association Studies to Identify Common QTL Regions in Three Different Genetic Backgrounds Based on Iberian Pig Breed. *PLoS One* **2018**, *13*, e0190184.

16. Mach, N.; Devant, M.; Díaz, I.; Font-Furnols, M.; Oliver, M.A.; García, J.A.; Bach, A. Increasing the Amount of N-3 Fatty Acid in Meat from Young Holstein Bulls through Nutrition. *J. Anim. Sci.* **2006**, *84*, 3039–3048.

17. Kassambara, A. Rstatix: Pipe-Friendly Framework for Basic Statistical Tests [R Package Rstatix Version 0.7.2] **2023**.

18. Ewels, P.; Magnusson, M.; Lundin, S.; Käller, M. MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report. *Bioinformatics* **2016**, *32*, 3047–3048.

19. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21.

20. Li, B.; Dewey, C.N. RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinformatics* **2011**, *12*, 1–16.

21. R Core Team R: A Language and Environment for Statistical Computing. R foundation for Statistical Computing. Vienna, Austria **2021**.

22. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21.

23. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300.

24. Kassambara, A.; Mundt, F. Extract and Visualize the Results of Multivariate Data Analyses [R Package Factoextra Version 1.0.7] **2020**.

25. Bindea, G.; Mlecnik, B.; Hackl, H.; Charoentong, P.; Tosolini, M.; Kirilovsky, A.; Fridman, W.-H.; Pagès, F.; Trajanoski, Z.; Galon, J. ClueGO: A Cytoscape Plug-in to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks. *Bioinformatics* **2009**, *25*, 1091–1093.

26. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504.

27. Puig-Oliveras, A.; Revilla, M.; Castelló, A.; Fernández, A.I.; Folch, J.M.; Ballester, M. Expression-Based GWAS Identifies Variants, Gene Interactions and Key Regulators Affecting Intramuscular Fatty Acid Content and Composition in Porcine Meat. *Sci. Reports 2016 61* **2016**, *6*, 1–12.

28. Criado-Mesas, L.; Ballester, M.; Crespo-Piazuelo, D.; Castelló, A.; Fernández, A.I.; Folch, J.M. Identification of eQTLs Associated with Lipid Metabolism in Longissimus Dorsi Muscle of Pigs with Different Genetic Backgrounds. *Sci. Rep.* **2020**, 1, *10*.

29. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* **2010**, *26*, 139–140.

30. Ntambi, J.M. Regulation of Stearoyl-CoA Desaturase by Polyunsaturated Fatty Acids and Cholesterol. *J. Lipid Res.* **1999**, *40*, 1549–1558.

31. Puig-Oliveras, A.; Ramayo-Caldas, Y.; Corominas, J.; Estellé, J.; Pérez-Montarelo, D.; Hudson, N.J.; Casellas, J.; Ballester, J.M.F.M. Differences in Muscle Transcriptome among Pigs Phenotypically Extreme for Fatty Acid Composition. *PLoS One* **2014**, *9*, e99720.

32. Sargsyan, A.; Doridot, L.; Hannou, S.A.; Tong, W.; Srinivasan, H.; Ivison, R.; Monn, R.; Kou, H.H.; Haldeman, J.M.; Arlotto, M.; et al. HGFAC Is a ChREBP-Regulated Hepatokine That Enhances Glucose and Lipid Homeostasis. *JCI Insight* **2023**, *8*, e153740.

33. Yee, J.K.; Mao, C.S.; Ross, M.G.; Lee, W.N.P.; Desai, M.; Toda, A.; Kjos, S.L.; Hicks, R.A.; Patterson, M.E. High Oleic/Stearic Fatty-Acid Desaturation Index in

Cord Plasma from Infants of Mothers with Gestational Diabetes. *J. Perinatol. 2014 345* **2014**, *34*, 357–363.

34. Paton, C.M.; Ntambi, J.M. Biochemical and Physiological Function of Stearoyl-CoA Desaturase. *Am. J. Physiol. - Endocrinol. Metab.* **2009**, *297*, 28–37.

35. Miller, C.W.; Ntambi, J.M. Peroxisome Proliferators Induce Mouse Liver Stearoyl-CoA Desaturase 1 Gene Expression. *Proc. Natl. Acad. Sci.* **1996**, *93*, 9443–9448.

36. Liang, R.; Han, B.; Li, Q.; Yuan, Y.; Li, J.; Sun, D. Using RNA Sequencing to Identify Putative Competing Endogenous RNAs (CeRNAs) Potentially Regulating Fat Metabolism in Bovine Liver. *Sci. Rep.* **2017**, *7,* 1–12.

37. Nakamura, M.T.; Cheon, Y.; Li, Y.; Nara, T.Y. Mechanisms of Regulation of Gene Expression by Fatty Acids. *Lipids* **2004**, *39*, 1077–1083.

38. Mater, M.K.; Thelen, A.P.; Pan, D.A.; Jump, D.B. Sterol Response Element-Binding Protein 1c (SREBP1c) Is Involved in the Polyunsaturated Fatty Acid Suppression of Hepatic S14 Gene Transcription. *J. Biol. Chem.* **1999**, *274*, 32725–32732.

39. Bickel, P.E.; Tansey, J.T.; Welte, M.A. PAT Proteins, an Ancient Family of Lipid Droplet Proteins That Regulate Cellular Lipid Stores. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2009**, *1791*, 419–440.

40. Muñoz, M.; García-Casco, J.M.; Caraballo, C.; Fernández-Barroso, M.Á.; Sánchez-Esquiliche, F.; Gómez, F.; Rodríguez, M. del C.; Silió, L. Identification of Candidate Genes and Regulatory Factors Underlying Intramuscular Fat Content Through Longissimus Dorsi Transcriptome Analyses in Heavy Iberian Pigs. *Front. Genet.* **2018**, *9,* 608.

41. Li, B.; Weng, Q.; Dong, C.; Zhang, Z.; Li, R.; Liu, J.; Jiang, A.; Li, Q.; Jia, C.; Wu, W.; et al. A Key Gene, PLIN1, Can Affect Porcine Intramuscular Fat Content Based on Transcriptome Analysis. *Genes (Basel).* **2018**, *9*, 194.

42. Taniguchi, M.; Arakawa, A.; Motoyama, M.; Nakajima, I.; Nii, M.; Mikawa, S. Genomic Structural Analysis of Porcine Fatty Acid Desaturase Cluster on Chromosome 2. *Anim. Sci. J.* **2015**, *86*, 369–377.

43. Park, W.; Kothapalli, K.; Reardon, H.; Kim, L.; Brenna, J. Novel Fatty Acid Desaturase 3 (FADS3) Transcripts Generated by Alternative Splicing. *Gene* **2009**, *446*, 28–34.

44. Stroud, C.K.; Nara, T.Y.; Roqueta-Rivera, M.; Radlowski, E.C.; Lawrence, P.; Zhang, Y.; Cho, B.H.; Segre, M.; Hess, R.A.; Brenna, J.T.; et al. Disruption of FADS2 Gene in Mice Impairs Male Reproduction and Causes Dermal and Intestinal Ulceration. *J. Lipid Res.* **2009**, *50*, 1870–1880.

45. Baeza-Raja, B.; Akassoglou, K. Glucose Homeostasis and P75NTR: The Sweet Side of Neurotrophin Receptor Signaling. Cell Cycle **2012**, *11*, 3151–3152.

46. Chmurzyńska, A. The Multigene Family of Fatty Acid-Binding Proteins (FABPs): Function, Structure and Polymorphism. *J. Appl. Genet. 2006 471* **2006**, *47*, 39–48.

47. Glatz, J.F.C.; Schaap, F.G.; Binas, B.; Bonen, A.; Van Der Vusse, G.J.; Luiken, J.J.F.P. Cytoplasmic Fatty Acid-Binding Protein Facilitates Fatty Acid Utilization by Skeletal Muscle. *Acta Physiol. Scand.* **2003**, *178*, 367–371.

48. Furuhashi, M.; Hotamisligil, G.S. Fatty Acid-Binding Proteins: Role in Metabolic Diseases and Potential as Drug Targets. *Nat. Rev. Drug Discov. 2008 76* **2008**, *7*, 489–503.

49. Wagner, N.; Wagner, K.D. The Role of PPARs in Disease. *Cells 2020, 9*, 2367.

50. Gu, H.; Zhou, Y.; Yang, J.; Li, J.; Peng, Y.; Zhang, X.; Miao, Y.; Jiang, W.; Bu, G.; Hou, L.; et al. Targeted Overexpression of PPARγ in Skeletal Muscle by Random Insertion and CRISPR/Cas9 Transgenic Pig Cloning Enhances Oxidative Fiber Formation and Intramuscular Fat Deposition. *FASEB J.* **2021**, *35*.

Supplementary Materials

Table S1. Genotyping of SNPs located in the *SCD* gene by Taqman OpenArray™ plates in LD muscle from BC1_DU pigs.

Haplotype	SNPs						
	rs80813866	rs80956173	rs323081995	rs80912566	rs342182479	rs331901016	rs334462984
H1	C	C	C	T	A	C	G
H2	A	G	T	C	G	T	C

_____

Table S2. Differentially expressed genes between H and L groups for C18:1n-9/C18:0 ratio in LD muscle from BC1_DU pigs.

Ensembl ID	Symbol	log$_2$FC	FC	*P*-value	*P*-adj
ENSSSCG00000008700	*HGFAC*	-8.33	0.003	3.53E-11	4.34E-07
ENSSSCG00000009039	*MMAA*	-0.46	0.73	6.12E-08	3.77E-04
ENSSSCG00000005669	*MIGA2*	-0.83	0.56	1.10E-06	2.36E-03
ENSSSCG00000001550		-1.70	0.31	1.11E-06	2.36E-03
ENSSSCG00000010251	*SUPV3L1*	-0.51	0.70	1.12E-06	2.36E-03
ENSSSCG00000011195	*GALNT15*	-1.48	0.36	1.15E-06	2.36E-03
ENSSSCG00000025751	*MDH1*	-0.52	0.69	2.57E-06	4.53E-03
ENSSSCG00000011119	*ECHDC3*	-0.47	0.72	3.16E-06	4.75E-03
ENSSSCG00000011314	*LZTFL1*	-0.46	0.72	3.48E-06	4.75E-03
ENSSSCG00000009836	*ACAD10*	-0.43	0.74	5.36E-06	5.65E-03
ENSSSCG00000017548	*NGFR*	1.23	2.35	6.09E-06	5.65E-03
ENSSSCG00000017947	*ACADVL*	-0.44	0.74	6.25E-06	5.65E-03
ENSSSCG00000011197	*OXNAD1*	-0.65	0.63	6.26E-06	5.65E-03
ENSSSCG00000022989	*ZNF704*	0.69	1.62	6.42E-06	5.65E-03
ENSSSCG00000015281	*PLEKHA6*	0.89	1.86	8.41E-06	6.24E-03
ENSSSCG00000039715	*FAHD1*	-0.46	0.72	8.84E-06	6.24E-03
ENSSSCG00000010554	*SCD*	1.10	2.15	9.13E-06	6.24E-03
ENSSSCG00000015290	*CDK18*	-1.26	0.42	1.52E-05	9.84E-03
ENSSSCG00000010537	*GOT1*	-0.76	0.59	2.20E-05	1.34E-02
ENSSSCG00000023045		0.57	1.49	2.28E-05	1.34E-02
ENSSSCG00000007920		-0.74	0.60	2.52E-05	1.41E-02
ENSSSCG00000028856	*ABCF3*	-0.34	0.79	2.78E-05	1.44E-02
ENSSSCG00000038805		0.47	1.39	2.83E-05	1.44E-02
ENSSSCG00000006354	*TOMM40L*	-0.47	0.72	2.96E-05	1.44E-02
ENSSSCG00000005645	*TRUB2*	-0.36	0.77	3.03E-05	1.44E-02
ENSSSCG00000010860		-0.41	0.75	3.22E-05	1.47E-02
ENSSSCG00000027922	*UNC5CL*	-1.19	0.44	3.87E-05	1.67E-02
ENSSSCG00000022446	*SEL1L3*	0.47	1.39	3.94E-05	1.67E-02
ENSSSCG00000032684	*BOK*	0.64	1.56	4.30E-05	1.68E-02
ENSSSCG00000021746	*MRPL35*	-0.35	0.78	4.46E-05	1.68E-02
ENSSSCG00000021917	*HIPK4*	0.69	1.61	4.52E-05	1.68E-02
ENSSSCG00000029838	*FZD2*	0.60	1.52	4.58E-05	1.68E-02
ENSSSCG00000007094	*DZANK1*	-0.76	0.59	4.64E-05	1.68E-02
ENSSSCG00000017801		1.49	2.83	5.49E-05	1.93E-02
ENSSSCG00000000264	*MFSD5*	-0.47	0.72	6.57E-05	2.14E-02
ENSSSCG00000005007		-0.33	0.79	6.66E-05	2.14E-02

Table S2. Differentially expressed genes between H and L groups for C18:1*n*-9/C18:0 ratio in LD muscle from BC1_DU pigs *(Continued).*

Ensembl ID	Symbol	log$_2$FC	FC	*P*-value	*P*-adj
ENSSSCG00000003565	*NR0B2*	-1.35	0.39	6.78E-05	2.14E-02
ENSSSCG00000014148	*TMEM161B*	-0.68	0.62	7.25E-05	2.23E-02
ENSSSCG00000005041	*FERMT2*	0.32	1.26	7.61E-05	2.29E-02
ENSSSCG00000038615	*TAF9*	-0.33	0.79	8.63E-05	2.53E-02
ENSSSCG00000006345	*OLFML2B*	0.48	1.40	8.82E-05	2.53E-02
ENSSSCG00000016577	*ATP6V1F*	-0.37	0.77	9.99E-05	2.68E-02
ENSSSCG00000030174	*AP3S2*	-0.30	0.81	1.00E-04	2.68E-02
ENSSSCG00000003384	*DNAJC11*	-0.34	0.79	1.03E-04	2.69E-02
ENSSSCG00000013073	*FADS3*	0.50	1.42	1.07E-04	2.75E-02
ENSSSCG00000010529	*SFRP5*	0.81	1.75	1.11E-04	2.75E-02
ENSSSCG00000002778	*ZDHHC1*	-0.65	0.64	1.12E-04	2.75E-02
ENSSSCG00000013830	*ILVBL*	-0.53	0.69	1.21E-04	2.91E-02
ENSSSCG00000010068	*SLC2A11*	-0.52	0.69	1.23E-04	2.92E-02
ENSSSCG00000032605		-0.51	0.70	1.33E-04	3.08E-02
ENSSSCG00000037970	*SOX8*	0.70	1.63	1.44E-04	3.29E-02
ENSSSCG00000000083		-0.46	0.73	1.49E-04	3.34E-02
ENSSSCG00000037710	*ZNF473*	-0.54	0.69	1.61E-04	3.46E-02
ENSSSCG00000003204	*VRK3*	-0.34	0.79	1.62E-04	3.46E-02
ENSSSCG00000007780		-0.31	0.80	1.63E-04	3.46E-02
ENSSSCG00000014900	*RAB30*	-1.16	0.45	1.70E-04	3.46E-02
ENSSSCG00000008002	*CIAO3*	-0.34	0.79	1.72E-04	3.46E-02
ENSSSCG00000010576	*NOLC1*	-0.38	0.77	1.74E-04	3.46E-02
ENSSSCG00000024009		0.49	1.41	1.74E-04	3.46E-02
ENSSSCG00000010507	*TLL2*	-0.76	0.59	1.79E-04	3.48E-02
ENSSSCG00000015450	*ZNF786*	-0.63	0.64	1.81E-04	3.48E-02
ENSSSCG00000001881	*MAN2C1*	-0.33	0.79	1.84E-04	3.48E-02
ENSSSCG00000001844	*PLIN1*	1.62	3.09	2.01E-04	3.75E-02
ENSSSCG00000036883	*FABP3*	-0.66	0.63	2.11E-04	3.78E-02
ENSSSCG00000015742	*TMEM177*	-0.58	0.67	2.11E-04	3.78E-02
ENSSSCG00000007864	*GPRC5B*	0.43	1.35	2.12E-04	3.78E-02
ENSSSCG00000034124	*CIAO2A*	-0.37	0.77	2.17E-04	3.82E-02
ENSSSCG00000001659	*KLC4*	-0.36	0.77	2.28E-04	3.95E-02
ENSSSCG00000022915	*DDX27*	-0.33	0.79	2.48E-04	4.25E-02
ENSSSCG00000003621		-0.52	0.69	2.55E-04	4.30E-02

_____

Table S2. Differentially expressed genes between H and L groups for C18:1*n*-9/C18:0 ratio in LD muscle from BC1_DU pigs *(Continued)*.

Ensembl ID	Symbol	log$_2$FC	FC	*P*-value	*P*-adj
ENSSSCG00000001431	*FKBPL*	-0.43	0.74	2.81E-04	4.63E-02
ENSSSCG00000003527	*EPHB2*	0.99	1.99	2.84E-04	4.63E-02
ENSSSCG00000017890	*PITPNM3*	-0.34	0.79	2.86E-04	4.63E-02
ENSSSCG00000017498	*PPP1R1B*	1.13	2.20	3.05E-04	4.87E-02
ENSSSCG00000007572	*LFNG*	0.59	1.51	3.14E-04	4.94E-02
ENSSSCG00000007586	*FSCN1*	0.50	1.42	3.23E-04	4.94E-02
ENSSSCG00000011579	*PPARG*	0.63	1.55	3.28E-04	4.94E-02
ENSSSCG00000031847	*RBM28*	-0.31	0.80	3.32E-04	4.94E-02
ENSSSCG00000013391		-0.33	0.79	3.34E-04	4.94E-02
ENSSSCG00000010456	*PANK1*	-0.68	0.62	3.37E-04	4.94E-02
ENSSSCG00000036944	*TACO1*	-0.30	0.81	3.44E-04	4.98E-02

Footnotes: FC: Fold change, *P*-value calculated with Wald test, *P*-adj: adjusted *P*-values by Benjamini-Hochberg method.

Paper II

scientific reports

Jesús Valdés-Hernández[1,2*,] Yuliaxis Ramayo-Caldas[3], Magí Passols[1], Cristina Sebastià[1,2], Lourdes Criado-Mesas[1], Daniel Crespo-Piazuelo[3], Anna Esteve-Codina[4], Anna Castelló[1,2], Armand Sánchez[1,2] & Josep M. Folch[1,2]

[1]Plant and Animal Genomics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra, Spain.

[2]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra, Spain.

[3]Departament de Genètica i Millora Animal, Institut de Recerca y Tecnologia Agraroalimentàries (IRTA), Caldes de Montbui, Spain.

[4]CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain.

*email: jesus.valdes@cragenomica.es

_____

Abstract

Fatty acids (FAs) play an essential role as mediators of cell signaling and signal transduction, affecting metabolic homeostasis and determining meat quality in pigs. However, FAs are transformed by the action of several genes, such as those encoding desaturases and elongases of FAs in lipogenic tissues. The aim of the current work was to identify candidate genes, biological processes, and pathways involved in the modulation of intramuscular FA profile from longissimus dorsi muscle. FA profile by gas chromatography of methyl esters and gene expression by RNA-Seq were determined in 129 Iberian × Duroc backcrossed pigs. An association analysis between the muscle transcriptome and its FA profile was performed, followed by a concordance and functional analysis. Overall, a list of well-known (e.g., *PLIN1, LEP, ELOVL6, SC5D, NCOA2, ACSL1, MDH1, LPL, LGALS12, TFRC, GOT1,* and *FBP1*) and novel (e.g., *TRARG1, TANK, ENSSSCG00000011196,* and *ENSSSCG00000038429*) candidate genes was identified, either in association with specific or several FA traits. Likewise, several of these genes belong to biological processes and pathways linked to energy, lipid, and carbohydrate metabolism, which seem determinants in the modulation of FA compositions. This study can contribute to elucidate the complex relationship between gene expression and FA profile in pig muscle.

_____

Introduction

Fatty acids (FAs) are crucial for living organisms, playing an essential role as mediators of signal transduction, cellular differentiation, and metabolic homeostasis. FAs can be classified into three groups (saturated: SFA, monounsaturated: MUFA, and polyunsaturated: PUFA) that are either provided by the diet or derived from de novo fatty acid (FA) synthesis[1,2]. FAs derived only from the diet are known as essential (e.g., C18:2*n*-6 and C18:3*n*-3) while those that can be synthesized through de novo lipogenesis are known as non-essential (e.g., C16:0, C16:1*n*-7, C18:0 and C18:1*n*-9)[3].

Studies on pork meat quality and nutritive values have received special attention over the last decade. In fact, intramuscular FA content and its composition are important indicators of meat quality in pigs. In pork, meat quality parameters are mainly evaluated by changes in the flavour, oxidative stability, firmness, color[4,5], and its nutritional value which is mainly determined by the FA profile. However, FA composition varies across tissues, and it is greatly affected by several factors such as environmental and host-factors including diet, fatness, body weight, gender, breed, and age, among others[6–8]. In addition, genetic background and gene expression are also able to influence the FA composition traits. In fact, the combination of methodologies such as gas chromatography and RNA sequencing (RNA-Seq), provides a powerful tool to analyze FA determinations[9], as well as global changes in muscle transcriptome[10], but also to discover genes contributing to intramuscular FA variation in lipogenic tissues from pig populations[11–13].

The relationship between FA composition and gene expression is bidirectional. For example, dietary FAs like PUFA can affect gene expression by regulating the activity of several families of transcription factors (TFs, including PPARs, LXRs and SREBPs)[14]. By switching the direction of the effect, FA traits can also be modulated by genes encoding enzymes like desaturases and elongases in lipogenic tissues[2,15]. Notwithstanding, the regulatory mechanisms beyond these reactions are complex and involve the combination of TFs[2], as well as several biological processes and pathways. In the context of animal and plant breeding, there is an increasing interest in identifying genes controlling the phenotypic variation of complex traits. That is the case of several studies focused on the genetic basis of intramuscular FA composition in pig muscle across several breeds[8,16–21].

The aim of this work was to study the association between the porcine *longissimus dorsi* (LD) muscle FA profile and its transcriptome, focusing on the identification of the most relevant candidate genes, biological processes and pathways related to intramuscular FA composition.

Methods

Animals, sample collection and phenotypic data. A total of 129 animals generated by an experimental backcross named as BC1_DU (25% Iberian and 75% Duroc) were employed. All pigs were maintained under the same intensive conditions and fed *ad libitum* with a commercial cereal-based diet and free access to water. A more detailed description of the backcross BC1_LD generation, experimental design, animal raising, and feeding is provided in Martínez-Montes et al.[22] Animal procedures were carried out according to the Spanish Policy for Animal Protection RD1201/05, which meets the European Union Directive 86/609 about the protection of animals used in experimentation. This study was conducted in accordance with relevant guidelines and regulations of the animal care and use committee of the *Institut de Recerca i Tecnologia Agroalimentàries* (IRTA), which adopts "The European Code of Conduct for Research Integrity" and "Good Experimental Practices". Likewise, the experimental protocol was approved by the Ethical Committee of the IRTA. Our study is also reported in full compliance with ARRIVE guidelines (https://arriveguidelines.org/). Animals were slaughtered in five batches in a commercial abattoir of Mollerussa (Spain). Samples of *longissimus dorsi* (LD) muscle were collected, immediately snap frozen in liquid nitrogen and stored at −80 °C until analysis. In addition, the distribution by sex was 59 females and 70 males, males were not castrated. At slaughter, pigs had an average age of 190 days (range 174–205 days), and 73.70 kg of carcass weight (range 46.10–109.20 kg).

FA composition in the C14-C22 range was determined using a gas chromatography of methyl esters protocol as described by Mach et al.[23] in intramuscular LD muscle (n = 129). In brief, 200 g of LD muscle samples of 129 BC1_DU pigs were homogenized and used to measure the FA profile in duplicate. Additional information on the LD muscle FA composition in BC1_DU population is indicated in Crespo-Piazuelo et al.[20].The FA composition (n = 17 FAs) was expressed as percentage of total identified FAs. Total percentages of SFA, MUFA, and PUFA were obtained through the sum of the individual FAs (Table 1). FA and metabolic ratios were

calculated from the ratio between individual FA percentages as it is shown in Table 1. In addition, we further calculated the following FA metabolic indices:

Average Chain Length (ACL) = $\sum$(percentage of FAs $\times$ carbon length),

Double bond index (DBI) = $\sum$(percentage of FAs $\times$ number of double bond),

Unsaturated Index (UI) = [(DBI $\times$ 100)/SFA], and

Peroxidability Index (PI) = $\sum$[(percentage of monoenoic acid $\times$ 0.025) + (percentage of dienoic acid $\times$ 1) + (percentage of trienoic acid $\times$ 2) + (percentage of tetraenoic acid $\times$ 4)].

Table 1. Summary of descriptive statistics on the FA composition traits and FA metabolic indices in the LD muscle from BC1_DU pigs. Fatty acid composition is expressed as percentage of total fatty acids.

Trait	Name	Mean	SD	Min	Max	SEM	CV
C14:0	Myristic acid	1.27	0.23	0.73	1.78	0.02	17.95
C16:0	Palmitic acid	23.91	1.65	18.69	27.59	0.15	6.90
C17:0	Margaric acid	0.25	0.11	0.12	0.90	0.01	42.27
C18:0	Stearic acid	14.38	1.69	8.81	19.90	0.15	11.78
C20:0	Arachidic acid	0.23	0.08	0.08	0.71	0.01	33.94
C16:1n-7	Palmitoleic acid	2.79	0.53	1.24	4.08	0.05	18.83
C16:1n-9	9-Hexadecenoic acid	0.30	0.12	0.16	0.92	0.01	38.62
C17:1	Heptadecenoic acid	0.19	0.09	0.10	0.71	0.01	46.53
C18:1n-9	Oleic acid	35.93	5.71	19.99	44.15	0.50	15.88
C18:1n-7	Vaccenic acid	3.82	0.30	3.02	4.83	0.03	7.97

Table 1. Summary of descriptive statistics on the FA composition traits and FA metabolic indices in the LD muscle from BC1_DU pigs. Fatty acid composition is expressed as percentage of total fatty acids *(Continued)*.

Trait	Name	Mean	SD	Min	Max	SEM	CV
C20:1n-9	Gondoic acid	0.73	0.16	0.35	1.48	0.01	22.38
C18:2n-6	Linoleic acid	12.13	5.82	4.81	29.34	0.51	48.01
C18:3n-3	α-Linolenic acid	0.40	0.13	0.15	0.89	0.01	32.77
C20:2n-6	Eicosadienoic acid	0.43	0.12	0.14	0.91	0.01	28.61
C20:3n-3	Eicosatrienoic acid	0.18	0.10	0.02	0.65	0.01	54.98
C20:3n-6	Dihomo-gamma-linolenic acid	0.45	0.29	0.09	1.49	0.03	63.29
C20:4n-6	Arachidonic acid	2.58	1.94	0.47	10.51	0.17	75.01
SFA	Saturated FAs	40.04	3.05	29.11	46.11	0.27	7.61
MUFA	Monounsaturated FAs	43.47	6.13	26.17	52.30	0.54	14.11
PUFA	Polyunsaturated FAs	16.00	8.11	6.00	38.62	0.71	50.71
ACL	Average chain length	17.50	0.09	17.33	17.79	0.01	0.49
MUFA/SFA	Ratio of MUFA to SFA	1.09	0.13	0.62	1.36	0.01	11.84
MUFA/PUFA	Ratio of MUFA to PUFA	3.54	1.85	0.71	8.54	0.16	52.32
PUFA/SFA	Ratio of PUFA to SFA	0.42	0.25	0.14	1.24	0.02	61.27
C16:1n-7/C16:0	Ratio of palmitoleic to palmitic	0.12	0.02	0.05	0.16	0.001	16.11

Table 1. Summary of descriptive statistics on the FA composition traits and FA metabolic indices in the LD muscle from BC1_DU pigs. Fatty acid composition is expressed as percentage of total fatty acids *(Continued)*.

Trait	Name	Mean	SD	Min	Max	SEM	CV
C18:1n-9/C18:0	Ratio of oleic to stearic	2.51	0.37	1.22	3.39	0.03	14.87
C20:1n-9/C20:0	Ratio of gondoic to arachidic	3.49	1.30	1.07	10.84	0.11	37.30
C20:4n-6/C20:3n-6	Ratio of arachidonic to dihomo-gamma-linolenic	5.37	1.11	2.62	8.66	0.10	20.65
C18:1n-7/C16:1n-7	Ratio of vaccenic to palmitoleic	1.41	0.29	1.02	3.37	0.03	20.46
C20:3n-6/C18:2n-6	Ratio of dihomo-gamma-linolenic to linoleic	0.04	0.01	0.02	0.08	0.0007	25.52
C20:4n-6/C18:2n6	Ratio of arachidonic to linoleic	0.19	0.06	0.06	0.41	0.01	33.07
C18:2n-6/C18:3n-3	Ratio of linoleic to α-linolenic	30.23	10.25	12.15	73.65	0.90	33.92
ω6/ω3	Ratio of omega-6 to omega-3	26.46	8.07	14.33	49.78	0.71	30.48
PI	Peroxidability index	26.06	14.19	9.48	73.21	1.25	54.44
DBI	Double-bond index	0.82	0.15	0.63	1.31	0.01	18.14
UI	Unsaturated index	2.10	0.58	1.39	4.18	0.05	27.72

SEM—standard error of the mean. CV—coefficient of variation (in percentage).

The following sums of fatty acids, ratios and indexes were calculated:

SFA = C14:0 + C16:0 + C17:0 + C18:0 + C20:0; MUFA = C16:1n-7 + C17:1 + C18:1n-7 + C18:1n-9 + C20:1n-9; PUFA = C18:2n-6 + C18:3n-3 + C20:2n-6 + C20:3n-6 + C20:4n-6; ω6/ω3 = (C18:2n-6 + C20:2n-6 + C20:3n-6 + C20:4n-6)/(C18:3n-3 + C20:3n-3); Average Chain Length (ACL) = (C14:0) × 14 + (C16:0 + C16:1n-9 + C16:1n-7) × 16 + (C17:0 + C17:1) × 17 + (C18:0 + C18:1n-9 + C18:1n-7 + C18:2n-6 + C18:3n-3) × 18 + (C20:0 + C20:1n-9 + C20:2n-6 + C20:3n-6 + C20:4n-6 + C20:3n-3) × 20)/100; Double Bond Index (DBI) = (C16:1n-9 + C16:1n-7 + C17:1 + C18:1n-9 + C18:1n-7 + C20:1n-9) × 1 + (C18:2n-6 + C20:2n-6) × 2 + (C18:3n-3 + C20:3n-6 + C20:3n-3) × 3 + (C20:4n-6) × 4)/100; and Peroxidability Index (PI) = (C16:1n-9 + C16:1n-7 + C17:1 + C18:1n-9 + C18:1n-7 + C20:1n-9) × 0.025 + (C18:2n-6 + C202n-6) × 1 + (C18:3n-3 + C20:3n-6 + C20:3n-3) × 2 + (C20:4n-6) × 4.

Total RNA isolation and sequencing. Total RNA was isolated from the LD muscle (100 mg) of 129 animals using the the RiboPure™ Isolation kit for High Quality Total RNA (Ambion®; Austin, TX) following the manufacturer's recommendations. RNA quantification and purity was done with a NanoDrop ND-1000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). RNA integrity was checked by Agilent Bioanalyzer-2100 equipment (Agilent Technologies, Inc., Santa Clara, CA, USA), using only those samples with an RNA integrity number (RIN) greater than 7 for the RNA-Seq experiment.

Library preparation and sequencing was performed at CNAG institute (*Centro Nacional de Análisis Genómico*, Barcelona, Spain). For each sample, one paired-end library was prepared using TruSeq Stranded mRNA kit (Illumina, Inc.; San Diego CA, USA). To discriminate among samples, libraries were labeled by barcoding and pooled to be run in Illumina HiSeq 3000/4000 instruments (Illumina, Inc.; San Diego CA, USA). In brief, in this study 2 × 75 bp reads, a mean of 45.09 million of paired-reads per sample, and an average of 90.06% (ranging from 80.51 to 96.09%) of uniquely mapped reads were generated.

Bioinformatic analyses. Quality control and basic statistics of reads were performed using FastQC (v0.11.9)[24] and MultiQC (v0.7)[25] programs, respectively. Sequencing reads were mapped employing the STAR software (v2.7.9a) with default parameters[26], and using the Sscrofa11.1 pig genome assembly as reference. Then, gene expression was quantified using RSEM (v1.2.28)[27] software with default parameters and annotation from pig Ensembl Genes

_____

97. Filtering was performed to keep rows of those genes with at least 129 reads in total, thus, 14,870 genes were retained. Normalization of the read counts was done as indicated below.

Association between whole-transcriptome and FA profile in muscle. The association study was conducted using the ELMSeq approach[28] on R v4.1.0[29]. We used the type I penalty function to test the association of gene expression with FA profile in LD muscle, while adjusting for all other covariates. The final model is described as follows:

$$\mathbf{Trait}_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\mathbf{Expression}_{1i} + \boldsymbol{\beta}_2\mathbf{Sex}_{2i} + \boldsymbol{\beta}_3\mathbf{Batch}_{3i} + \boldsymbol{\varepsilon}_i \ (1)$$

where $Trait_i$ represents each FA trait or FA index (n = 36) transformed into log2 base, i represents the individuals, β0 is the intercept, $Expression_i$ indicates the normalized gene expression value [log(counts + log.add)], here "counts" is a matrix of $m \times n$ sizes, where m is the number of genes and n is the number of samples (14,870 × 129). β_1, β_2 and β_3 are the beta coefficients, respectively. Sex_i and slaughterhouse $Batch_i$ represents fixed effects with 2 (female and male) and 5 levels (B1, B2, B3, B4 and B5), respectively. ε_i is the vector of residual effects. For ELMSeq, we used the sig_leval = 0.01, percentile = sig_leval, rho = 1, and log.add = 1e-3*min(counts[counts > 0]) parameters. Benjamini and Hochberg procedure[30] was used to correct the raw P-value (BH adjusted P-value < 0.05) in the post-processing stage.

Among the total associated genes with each trait, we counted the number of lipid related genes using a gene list manually curated and compiled by our research group based on gene functional annotation accordingly to GO and pathway terms available in gene ontology databases (e.g., KEGG, Reactome, WikiPathways, STRINGdb, AmiGO 2 and BioSystems of NCBI). The BioMart web tool (https://www.ensembl.org/) was used to extract attributes from Ensembl genome browser: (i) gene name, and (ii) human orthologues genes for those pig gene stable IDs without gene name. Afterwards, gene names were submitted (gene set augmentation) to the Geneshot tool using the ARCHS4 resource and the AutoRIF parameter[31]. Likewise, we compared the list of genes compiled in pig against the list of transcription factors (TFs) and TF co-factors from the AnimalTFDB v3.0 database http://bioinfo.life.hust.edu.cn/Animal TFDB/#!/).

Concordance analysis. In order to have a preliminary quality control based on the biological role of the results of gene expression, a study of overlap between the number of associated

genes with each trait was carried out. This analysis was organized by grouping the FA into five categories: (i) saturated FAs, (ii) monounsaturated FAs, (iii) polyunsaturated FAs, (iv) FA ratios, and (v) metabolic ratios and indices. Within each group, we explored the number of genes detected by each trait as well as the intersection size using the ComplexUpset package v1.3.3[32].

Gene ontology enrichment analysis. Gene Ontology (GO) enrichment analysis with the list of associated genes (by category among the trait class) were implemented using the ClueGO v2.5.8 plugin[33] in Cytoscape v3.9.0 software[34]. All genes expressed in LD muscle were employed as background in the overrepresentation analysis across biological processes, molecular functions and pathways (KEGG). The statistical significance was assessed with a hypergeometric test using Benjamini and Hochberg method[30] for multiple testing correction (BH corrected P-value < 0.05). In addition, a minimum k-score of 0.44 was used, GO tree interval levels set from three to eight, and a minimum of three genes per cluster with at least 4% in selected genes. Results with and without the fusion feature "GO Term Fusion" were generated to evaluate the redundant parent–child terms. To this end, ClueGo output was visualized using an R script that allowed to facet by categories via the ggplot2 package v3.3.5[35].

Correlation analysis. Pearson correlation analysis was carried out between the gene expression values and the FA phenotypic values using the R cor.test function. Here, the same normalization scale for the variables to be correlated was applied, as it has been described in the previous step of the ELMSeq algorithm. The Benjamini–Hochberg procedure[30] was used to correct the raw P-value corresponding to the correlation coefficients. The goal of this strategy was to provide information about the direction of the association reported by ELMSeq, and therefore to obtain interpretable results regarding positive or negative relationships between the FA composition and the gene expression.

Selection of candidate genes associated with the FA profile in muscle. The list of detected candidate genes were prioritized based on the following criteria: (i) candidate genes delimited into functional categories across biological processes, molecular functions and pathways (i.e., lipid metabolism, carbohydrate metabolism, amino acid metabolism, and nucleic acid metabolism terms); (ii) genes that correlate both specifically and simultaneously with FA

phenotypes; and (iii) literature mining plus novelty assessment via Geneshot (including: rare, uncommon, common and very common levels).

Results

Identification of candidate associated genes with FA profile in muscle. In the first step we performed an association analysis using the ELMSeq algorithm to identify genes (n = 14,870 genes) associated with the profile of 36 FA in 129 BC1_DU pigs. We observed a variable number of associated genes across FA (1022 genes in total, Supplementary Table S1). Among the 36 traits, 22 FA were associated with a range of 1–553 genes. For each trait, we estimated the distribution of the number of associated genes across metabolic processes, lipid metabolism, TFs or co-factors, and novel genes (including long non-coding RNAs) (Supplementary Table S2).

As shown in **Fig. 1**, it summarizes a description of the results corresponding to each of the FA categories: (1) saturated FAs, represented by three traits (C16:0, C18:0, and total SFA) with 2, 29 and 33 associated genes, respectively; (2) monounsaturated FAs, six traits (C16:1*n*-7, C18:1*n*-9, C18:1*n*-7, and C20:1*n*-9, C16:1*n*-9 and total MUFA) were linked to 8, 33, 15, 123, 1 and 34 associated genes, respectively; (3) polyunsaturated FAs, which includes two essential FAs (C18:3*n*-3 and C20:2*n*-6) and the total PUFA being associated to 4, 147 and 3 genes; (4) FA ratios (C16:1*n*-7/C16:0, C18:1*n*-7/C16:1*n*-7, C18:1*n*-9/C18:0, C18:2*n*-6/C18:3*n*-3, and ω6/ω3) linked to 4, 4, 3, 483 and 553 genes. Finally, on metabolic ratios and indices, five traits (MUFA/PUFA, MUFA/SFA, ACL, DBI, and UI) were linked with 30, 7, 22, 13 and 26 genes, respectively.
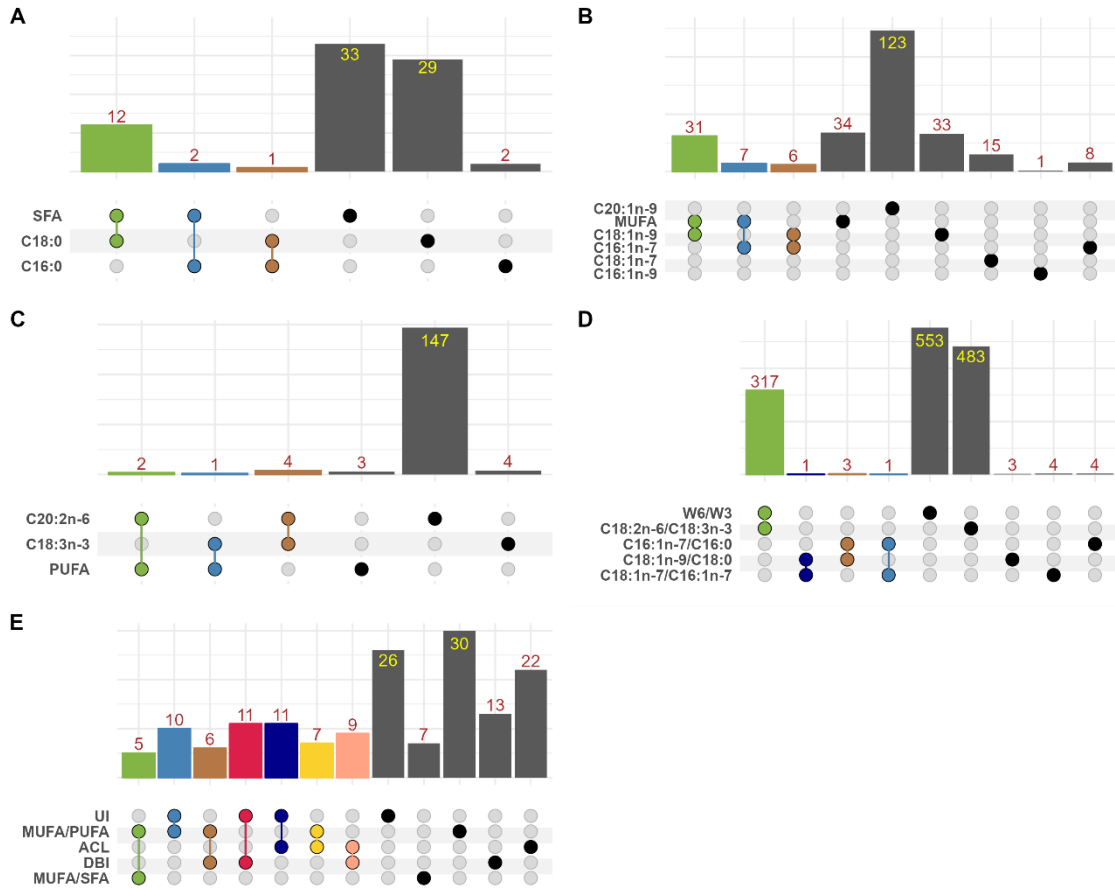
Figure 1. Analysis of overlapping and total number of associated genes by category of traits. The upset diagram (from left to right) shows the shared fraction (intersection size) of associated genes per pair of traits (colored bars), as well as the number of associated genes in each trait (gray bars). (A), (B), (C), (D) and (E) represents the group of saturated FAs, monounsaturated FAs, polyunsaturated FAs, FA ratios, and metabolic ratios and indices, respectively.

Concordance analysis. The concordance analysis across the 22 FAs revealed that among the saturated FAs, 36.36% and 100% of the genes associated to C18:0 and C16:0 were commonly linked to the total SFA (**Fig. 1A**). On the other hand, monounsaturated FAs showed an overlap of genes associated to C16:1n-7 and C18:1n-9 compared with total MUFA (20.58% and 91.17%, respectively), and 18.18% of shared genes was found between C16:1n-7 and C18:1n-9 FAs (**Fig. 1B**). Regarding the polyunsaturated FAs, there were 33.33% and 66.66% overlapping genes between C18:3n-3, C20:2n-6 and total PUFA (**Fig. 1C**). Finally, the ω6/ω3 and C18:2n-6/C18:3n-3 ratios showed a 57.32% of common genes because the ω6/ω3 ratio

contains C18:2n-6 and C18:3n-3 as major FAs, whereas the contrast between the rest of ratios showed a lower number of common genes (**Fig. 1D**).

Gene ontology enrichment analysis. In the functional analysis, no significant GO terms/pathways were found for the list of associated genes (i.e., 50 in total) with the SFA category. Instead, for the other categories a total of 75 terms (biological process, molecular functions and pathways) were significantly enriched (**Fig. 2**). The full results (including genes by each GO terms) with and without "GO Term Fusion" can be consulted in Supplementary Table S3. As shown in **Fig. 2**, terms related to lipid metabolism were found on some categories, for example, "KEGG:03320 PPAR signaling pathway", "GO:0008203 cholesterol metabolic process", "KEGG:00564 Glycerophospholipid metabolism", and "GO:0044539 long-chain fatty acid import". Moreover, GO terms involved in carbohydrate metabolism were also enriched such as "GO:0006086 acetyl-CoA biosynthetic process from pyruvate", "GO:0004738 pyruvate dehydrogenase activity, "KEGG:00020 Citrate cycle (TCA cycle)", and "KEGG:00,620 Pyruvate metabolism". Other GO terms were found related to carbohydrate and lipid homeostasis "KEGG:04310 Wnt signaling pathway" and control metabolic processes "KEGG:04150 mTOR signaling pathway", respectively. In addition, GO terms related to some biological process, such as "GO:0015980 energy derivation by oxidation of organic compounds", "GO:0006091 generation of precursor metabolites and energy", "KEGG:00190 Oxidative phosphorylation", "GO:2001256 regulation of store-operated calcium entry", and "GO:0032543 mitochondrial translation", were also enriched (**Fig. 2**). Finally, due to the low number of GO terms (n = 2, GO:0050880 and GO:0035150) found for the different metabolic ratios and indices, they were grouped into the single category term of "FA ratios and metabolic indices" (**Fig. 2**).
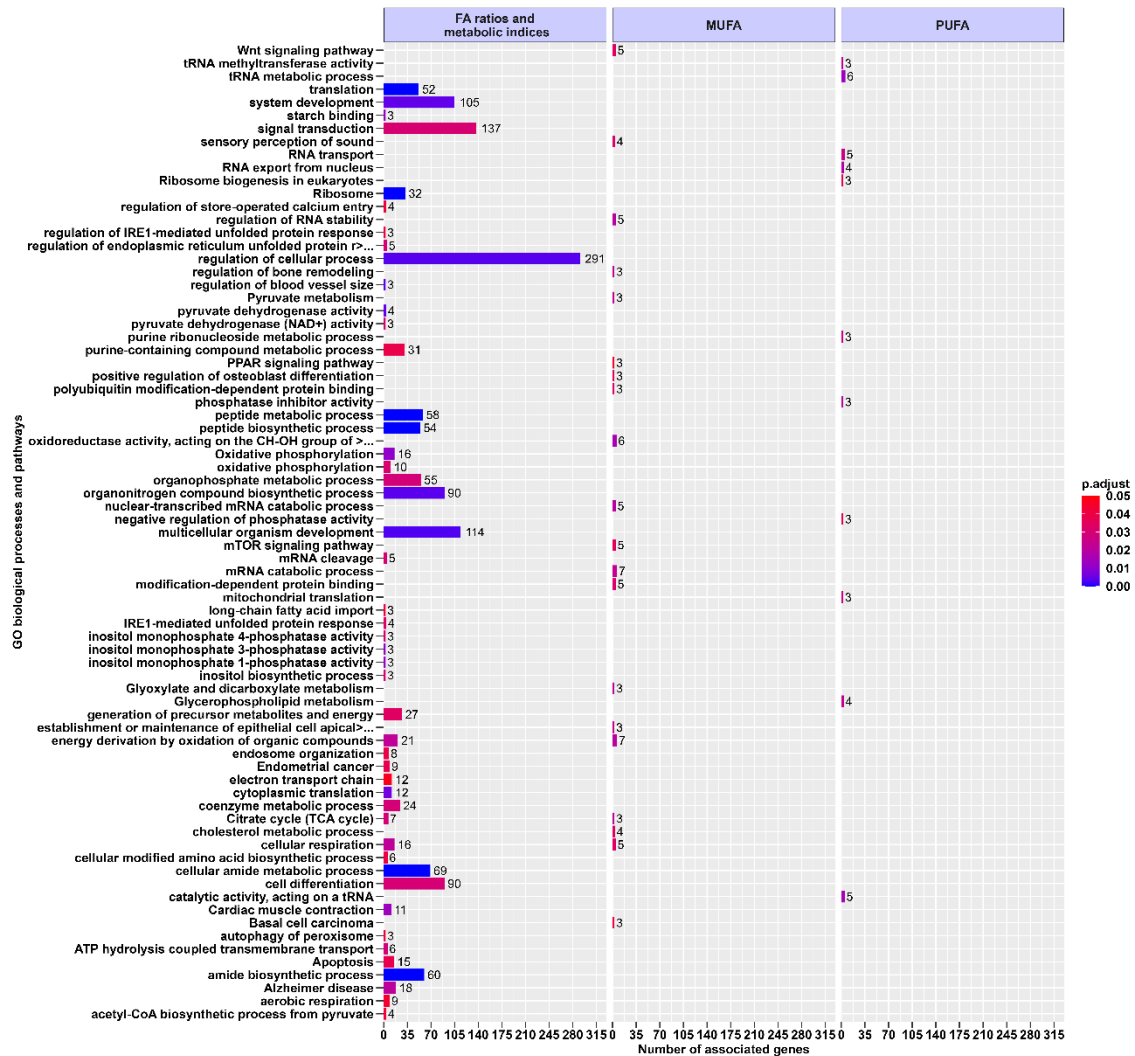
Figure 2. Functional enrichment analysis of the associated genes with FA profile and grouping by traits in BC1_DU pigs. The barplot with facets (from left to right) shows the enriched GO terms (BH adjusted P-value < 0.05) when grouping by traits. The traits grouped in this analysis are displayed in Fig. 1A-E. Legend: The full names to symbols ">..." are regulation of endoplasmic reticulum unfolded protein response; oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor and establishment or maintenance of epithelial cell apical/basal polarity. GO terms of each individual category are presented in **Table S3**.

Correlation analysis with candidate genes and representative traits. Pearson correlations between the 22 FAs and their associated genes were calculated. Among these traits, 21 (excluding the C16:1n-9 FA) showed significant correlations (BH adjusted P-value < 0.05) with 547 genes (Supplementary **Table S4**). Out of these, 57 genes were prioritized by our analytical approach (**Fig. 3**). In general, the absolute correlation between gene expression and FA profiles

_____

ranged from low to moderate (from - 0.19 to 0.51). Details about particular correlations can be found in **Fig. 3**. We noticed several genes simultaneously correlated with more than one trait (*FBP1, PLIN1, MDH1, LEP, ACSL1, CYCS, SFRP5, PPP1R1B, IDH3A, TFRC, GOT1, SC5D, G0S2, PDHA1, LBX1, LGALS12, UQCRC2, PNPLA8, ABHD5, ESRRA, GYG2, ZDHHC1, PLCD3, UNC93A*, *ENSSSCG00000017801* alias *TRARG1, ENSSSCG00000015889* alias *TANK,* and *ENSSSCG00000038429*). We also observed genes specifically correlated with a single trait, such as: *LPL, ELOVL6, LPIN1, NCOA2, SDHD, SLC27A4, SLC16A6, SLC27A1, THOC1, TFAM, CYP2B22, ZDHHC20, NUP35, HOXB6, CBR2* and *IGFBP5*). Remarkably, when we explored the existence of genes inversely correlated between traits, we observed: (1) genes positively correlated with MUFA but negatively correlated with PUFA (*PLIN1, SFRP5, PPP1R1B,* and *TRARG1*), and vice versa (*CYCS* and *TFRC*); (2) genes positively correlated with MUFA and metabolic indices but negatively correlated with FA ratios (*LEP*); (3) genes positively correlated with MUFA and FA ratios but negatively correlated with PUFA (*LGALS12*); (4) genes negatively correlated with SFA but positively correlated with FA ratios and/or metabolic indices (*NMNAT2*); and (5) genes positively correlated with SFA traits but negatively correlated with metabolic indices (*FBP1*).
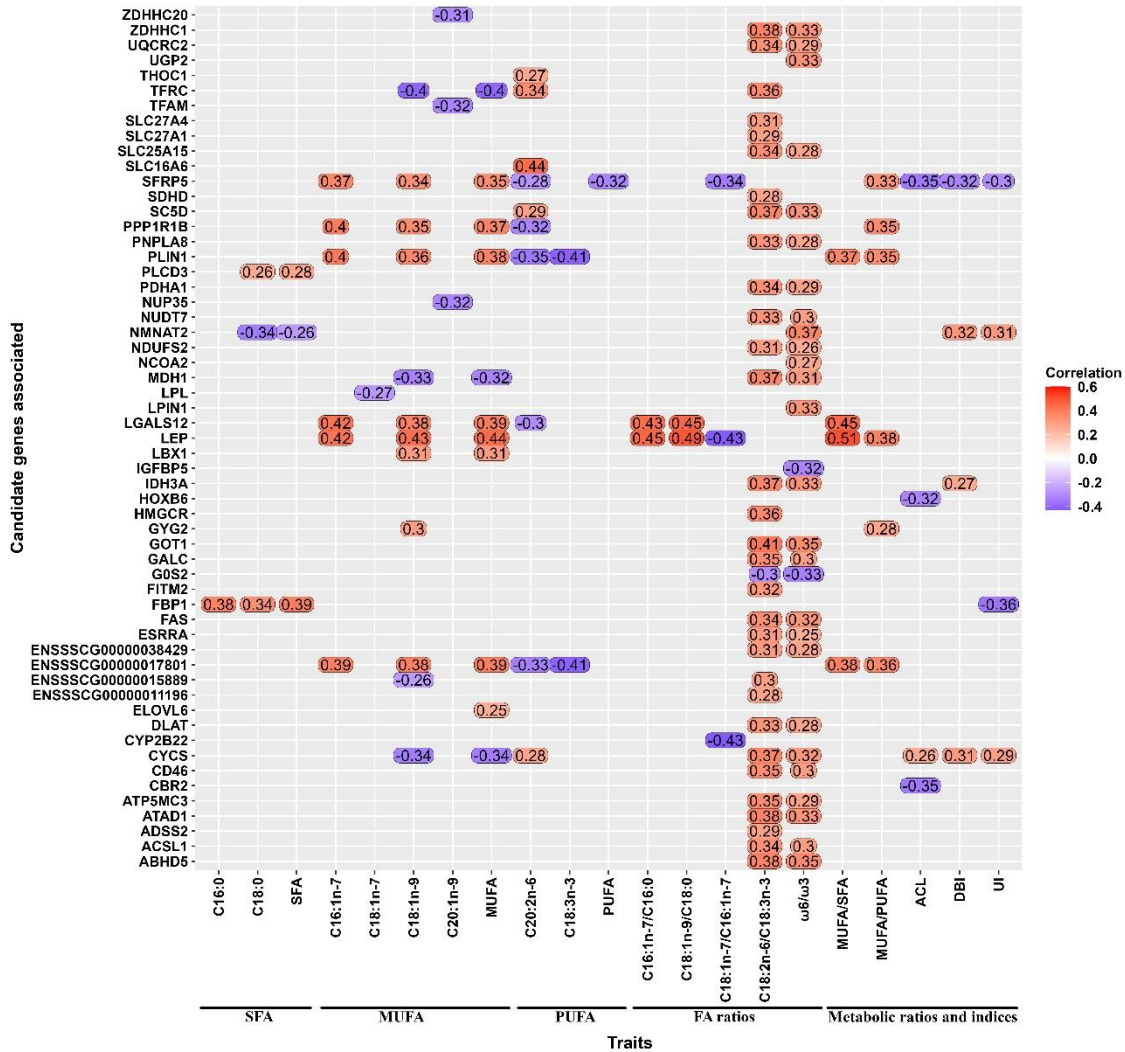
Figure 3. Heatmap of Pearson correlations of 21 representative traits with gene expression of 57 candidate genes related to FA profile traits in BC1_DU pigs. All these genes passed the correlation criteria according to a BH adjusted *P*-value < 0.05. The grouping of the FA profile traits is the same as that of the ELMSeq step.

Discussion

Studies on pork meat nutritive values and quality have received special attention over the last decade. Among other factors, meat nutritive value is determined by FA composition, playing also an important role in meat quality traits. Likewise, gene expression is assumed to be able to control the variation of such traits. In the present work, we aimed to study the association of muscle transcriptome with intramuscular FA profile from LD muscle to identify candidate genes, biological processes and pathways related with FA composition in pigs. Here, we used

_____

the ELMSeq approach[28], which has been extensively employed in the study of human diseases like cancer or metabolic disorders. However, to the best of our knowledge, this is the first study that reports the use of ELMSeq approach to test the association between muscle transcriptome and FA composition in pigs or other livestock species.

Here, we reported well-known candidate genes, but also promising novel genes associated to the intramuscular FA profile (Supplementary **Tables S1** and **S2**). In order to prioritize genes, the list of candidates was processed through an analytical pipeline including: (i) concordance analysis between the number of associated genes across categories of FAs (**Fig. 1**), (ii) ontology-based functional classification (**Fig. 2** and Supplementary **Table S3**), and (iii) correlation analysis of FA phenotypes and respective gene expression of the associated genes (**Fig. 3**). As results, we highlighted 57 candidate genes including protein-coding, transcriptional regulators, and long-non coding RNAs (**Fig. 3**). In the following sections, we discuss some of these genes accordingly to their biological function, as well as the novelty of our findings (Supplementary **Table S2**).

Genes related to biosynthesis and degradation of FAs. FAs are transformed by the actions of a vast array of enzymes and, we found a subset of genes that encode enzymes involved in the degradation and synthesis of FAs. This list of candidate genes included enzymes with lipolytic or lipase activity (*FBP1, LPL, MDH1*, *ACSL1,* and *GOT1*), and lipogenic activity (*ELOVL6*, *HMGCR* and *SC5D*). A detailed examination of the association study revealed that fructose-bisphosphatase 1 (*FBP1*) showed a synergistic effect (positively associated) with C16:0, C18:0 and SFA traits, but antagonistic (negatively associated) effect with UI trait (**Fig. 3**). *FBP1* catalyzes the hydrolysis of fructose 1,6-bisphosphate to fructose 6-phosphate and inorganic phosphate, acting as a gluconeogenesis regulatory enzyme (GeneCards ID: GC09M094603)[36]. The expression of ELOVL fatty acid elongase 6 (*ELOVL6*) was highly associated with MUFA (**Fig. 3**). *ELOVL6* encodes an enzyme responsible for condensation reaction, one of the four-step process (condensation, reduction, dehydration, and one further reduction) for elongation of very-long-chain (*ELOVL*) FAs. Further, *ELOVL6* preferentially utilizes SFA and MUFA as a substrate (e.g., C16:0 and C16:1*n*-7, respectively)[37]. To be noted, previous studies in Iberian × Landrace pigs (BC1_LD) provided evidence of the link between a polymorphism in the promoter region of *ELOVL6* (*ELOVL6*:c.-533C > T) and the percentage of C16:0 and C16:1*n*-7 FAs in LD muscle and backfat[38]. In addition, the *ELOVL6*:c.-394G > A

polymorphism was suggested as the causal mutation for the QTL on pig chromosome 8 (SSC8) that affects FA composition in BC1_LD pigs[39]. Despite this, our findings did not detect association between *ELOVL6* and SFA, and just a suggestive association with C18:1*n*-9 and the MUFA/PUFA ratio (Supplementary Table S1).

Among the genes with lipolytic effect, we detected malate dehydrogenase 1 (*MDH1*) associated to C18:1*n*-9, MUFA, and C18:2*n*-6/C18:3*n*-3 and ω6/ω3 ratios (**Fig. 3**). *MDH1* encodes an enzyme which catalyzes reversible oxidation of malate to oxaloacetate in many metabolic pathways (including the citric acid cycle), utilizing the NAD/NADH-dependent system (GeneCards ID: GC02P063557)[36]. The function of *MDH1* is primarily related to the production of aerobic energy for muscle contraction[40], and it was reported as candidate gene of meat quality traits in Chinese pig breeds[41]. The expression of another gene, lipoprotein lipase (*LPL*), was negatively associated with C18:1*n*-7 (**Fig. 3**). *LPL* belongs to the PPAR signaling pathway (**Fig. 2**), and has the dual functions of triglyceride hydrolase and ligand/bridging factor for receptor-mediated lipoprotein uptake (GeneCards ID: GC08P019901)[36]. Therefore, *LPL* hydrolyzes circulating triglyceride containing chylomicrons and very lowdensity lipoproteins to produce free FA. These free FAs can be assimilated by different tissues, such as muscle and adipose tissue[42].

Regarding the genes involved in lipogenesis, we detected sterol-C5-desaturase (*SC5D*) as positively associated with C20:2*n*-6, and with C18:2*n*-6/C18:3*n*-3 and ω6/ω3 ratios (**Fig. 3**). *SC5D* (also known as *SC5DL*) encodes an enzyme of cholesterol biosynthesis (GeneCards ID: GC11P121292)[36]. *SC5DL* has been found to be one of the downregulated genes related to cholesterol metabolism in various tissues (including muscle) of lambs[43]. In a similar way, 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR*) was positively associated with C18:2*n*-6/C18:3*n*-3 ratio (**Fig. 3**). *HMGCR* is a rate-limiting enzyme in cholesterol synthesis (GeneCards ID: GC05P075336)[36]. *HMGCR* is also related to several biological processes including pyruvate dehydrogenase activity, coenzyme metabolic process and signal transduction. A previous report in a Duroc population provided evidence of the links between the expression of *HMGCR* gene in muscle with several traits such as carcass lean percentage, C18:0 and C18:2*n*-6 contents, but also showed a positive correlation with cholesterol-related traits, intramuscular fat (IMF), and C18:1*n*-9 and C16:0 FA content[44].

Our results also reported the association of acyl-CoA synthetase long chain family member 1 (*ACSL1)* and glutamate oxaloacetate transaminase 1 (*GOT1*) genes with C18:2*n*-6/C18:3*n*-3 and ω6/ω3 ratios (**Fig. 3**). *ACSL1* participate in the long-chain fatty acid import and signal transduction biological process (**Fig. 2**). The protein encoded by *ACSL1* is an isozyme of the long-chain fatty-acid-coenzyme A ligase family (GeneCards ID: GC04M184755)[36]. *ACSL1* is involved in the synthesis of long-chain acyl-CoA esters, FA degradation and phospholipid remodeling[45]. Likewise, this gene was found to be associated with lipid metabolism and mitochondrial oxidation of FAs in pigs[46]. In relation to FA degradation, Zhou et al.[47], suggested that *GOT1* was crucial for providing oxaloacetate at low glucose levels, likely to maintain the redox homeostasis. In our results, *GOT1* was also related to organonitrogen compound biosynthetic process and signal transduction GO terms (**Fig. 2**).

Genes related to carbohydrate and lipid metabolism. We also identified several genes belonging to carbohydrate and lipid metabolism that were associated with various of the analyzed traits such as *PLIN1, CYCS, SFRP5, LEP* and *PPP1R1B* (**Fig. 3**). Among the well-known genes, we observed perilipin 1 (*PLIN1*) positively correlated to MUFA traits and metabolic ratios (MUFA/SFA and MUFA/PUFA), and negatively correlated with C20:2*n*-6 and C18:3*n*-3 traits (**Fig. 3**). *PLIN1* was also enriched in the PPAR signaling pathway (**Fig. 2**). *PLIN1* codifies for a protein belonging to the family of perilipins, which play a role in regulating intracellular lipid stor age and mobilization[48]. Association between the expression of *PLIN1* and porcine IMF deposition and adipocyte differentiation has already been eported[49]. In addition, a lipolytic function of *PLIN1* and increased expression in subcutaneous adipose biopsies from Iberian pigs fed a carbohydrate-enriched diet have been reported[50]. In our results, the expression of cytochrome c somatic (*CYCS)* was negatively correlated with MUFA and positively with correlated PUFA. *CYCS* encodes a small *heme* protein that functions as a central component of the electron transport chain in mitochondria (GeneCards ID: GC07M025118)[36].

The expression of leptin (*LEP*) was significantly correlated with more than one trait (**Fig. 3**) (i.e., positively correlated with MUFA traits, C16:1*n*-7/C16:0 and C18:1*n*-9/C18:0 ratios, and metabolic ratios as MUFA/SFA and MUFA/PUFA), but it was negatively with C18:1*n*-7/C16:1*n*-7 ratio. *LEP* encodes a protein that plays a major role in the regulation of energy homeostasis (GeneCards ID: GC07P128241)[36]. In addition, as a pleiotropic adipocytokine it can regulate several physiologic functions. *LEP* constitutes a circulating hormone that orchestrates behavioral and metabolic responses to nutrient intake[51]. Furthermore, *LEP*

_____

interacts with other hormonal mediators, regulators of energy status and metabolism (e.g., insulin or glucagon) to regulate growth and reproduction processes[52]. Indeed, *LEP* can regulates complex biological effects through its receptors. By using the Iberian pig as translational model for studies on metabolic syndrome, type 2 diabetes and nutrition-associated diseases, a polymorphism of the leptin receptor (LEPR c.1987C > T) has been informed to increase insatiability and obesity (i.e., state that in human medicine is called as leptin resistance)[53], and as such, the Iberian pigs would be resistant to leptin-induced lipolysis[50]. However, in our study the *LEPR* gene was not associated with any FA traits.

In the present work, we observed the transferrin receptor (*TFRC*) as negatively correlated with C18:1*n*-9 and MUFA traits, but positively correlated with C20:2*n*-6 and C18:2*n*-6/C18:3*n*-3 ratio (**Fig. 3**). *TFRC* (also known as *TFR1)* encodes a cell surface receptor necessary for cellular iron uptake by the process of receptormediated endocytosis (GeneCards ID: GC03M196027)[36]. In mammals, *TFR1* imports the transferrin-bound iron from the extracellular environment into cells. Interestingly, the intracellular labile free iron is indispensable for lipid peroxidation and ferroptosis execution[54]. On the other hand, protein phosphatase 1 regulatory inhibitor subunit 1b (*PPP1R1B*) was positively associated with MUFA traits but negatively correlated with PUFA traits (**Fig. 3**). *PPP1R1B* (also known as *DARPP-32)* encode a bifunctional signal transduction molecule, while stimulation of dopaminergic and glutamatergic receptors regulates its phosphorylation and function (GeneCards ID: GC17P039626)[36]. This gene is a potent inhibitor of protein phosphatase 1 (PPP1, previously known as PP1) when phosphorylated at Thr34 by cAMP-dependent protein kinase (PKA)[55], and this protein could be important in the control of glycogen metabolism and muscle contraction, among other activities. Thus, it suggests that phosphorylation sites are implicated in the fine regulation of *DARPP-32* function and a modulation of the regulation of *PPP1* via *DARPP-32*. In addition, *PPP1R1B* has been reported as differentially abundant in the LD muscle of phenotypically extreme pigs[13], and positively correlated with the IMF content of LD muscle in Duroc × Luchuan pigs[56].

FA and glucose transport genes. Five out of the 57 candidate genes were functional classified as FA and glucose transporters. Regarding FA transporters, our results suggested a significant association of four members of the solute carrier (SLC) gene superfamily with FA ratios and C20:2*n*-6 (**Fig. 3**). Members of SLC superfamily encode membrane-bound transporters, which play essential roles in transporting a variety of substrates across cellular membranes. These

include amino acids, glucose, inorganic cations and anions, FAs and lipids, acetyl coenzyme A, and vitamins, among others[57]. Out of all the SLC members detected, the solute carrier family 27 member 1 (*SLC27A1)* has also been reported as a transporter of the predominant substrates of long-chain FAs[58]. Regarding glucose transporter, *ENSSSCG00000017801* was identified among the novel genes reported in our study, which was significantly associated with MUFA and PUFA traits, and metabolic ratios (**Fig. 3**). However, *ENSSSCG00000017801* is a novel gene recently annotated in pigs and thus, there is limited information in the literature about its functions. Nevertheless, the human orthologous of this novel gene is the trafficking regulator of GLUT4 1 (*TRARG1*), which is predicted to be involved in endosome to plasma membrane protein transport and glucose import in response to insulin stimulus (NCBI Gene ID: 286,753)[59]. Likewise, *TRARG1* is of particular interest as it was previously recognized to be located in the glucose transporter type 4 (*GLUT4*) storage vesicles, and to positively regulate *GLUT4* trafficking or translocation[60].

Regulators including transcription factor and co-factors. In order to reinforce the biological significance of our study, we also focused the discussion on regulators including those TFs and co-factors found among the list of candidate genes. Ladybird homeobox 1 (*LBX1*) was positively associated with C18:1*n*-9 and MUFA (**Fig. 3**). *LBX1* is a member of the ladybird-like gene family which encodes a homeodomain transcription factor[61]. *LBX1* plays a putative regulatory role during the postnatal development of the porcine skeletal muscle in Meishan pigs[62]. In the present study, we also observed association between well-documented co-factors of FA metabolism. Two of such genes were the nuclear receptor coactivator 2 (*NCOA2*) and lipin 1 (*LPIN1*), which were positively associated with ω6/ω3 ratio (**Fig. 3**). *NCOA2* encodes a protein that acts as a transcriptional coactivator for nuclear hormone receptors, including steroid, thyroid, retinoid, and vitamin D. In fact, a key role of *NCOA2* as modulator of the intramuscular FA composition in pigs has been reported[63]. On the other hand, *LPIN1* encodes a regulatory enzyme that catalyzes the penultimate step in triglyceride synthesis, including the dephosphorylation of phosphatidic acid to yield diacylglycerol (GeneCards ID: GC02P011677)[36]. He et al.[64] reported a link between a polymorphism located in *LPIN1* gene with the percentage of leaf fat and IMF in pigs.

Another interesting co-factor was the galectin-related inhibitor of proliferation (*LGALS12*). We observed pleiotropic associations of *LGALS12*, which was positively associated with C16:1*n*-7, C18:1*n*-9, MUFA, MUFA/SFA, C16:1*n*-7/C:16:0 and C18:1*n*-9/C18:0, but negatively

correlated with C20:2*n*-6 (**Fig. 3**). This gene participates in signal transduction and regulation of cellular process (**Fig. 2**). *LGALS12* was found to perform a critical role in lipid metabolism in mice, functioning as an intrinsic negative regulator of lipolysis, regulating lipolytic protein kinase A signaling by acting upstream of phosphodiesterase activity to control cAMP levels[65]. Lastly, Wu et al.[66] reported that *LGALS12* knockdown could inhibit adipogenesis in porcine adipocytes by downregulating lipogenic genes and activating the PKA–Erk1/2 signaling pathway.

Conclusions

Taken together, our results identify candidate genes linked to intramuscular FA composition in muscle, including well-known and novel candidate genes involved in biological processes and pathways mainly related to energy, lipid, and carbohydrate metabolism that appear to be determinant in the modulation of intramuscular FA profile in pigs.

Data availability

All relevant data produced or evaluated in this research are disclosed in the paper as well as its supplementary information files. The RNA sequencing data used and analyzed in the current study are available from sequence read archive (SRA), NCBI BioProject under the accession number PRJNA882638 (https://www.ncbi.nlm.nih.gov/sra).

References

1. Ameer, F., Scandiuzzi, L., Hasnain, S., Kalbacher, H. & Zaidi, N. De novo lipogenesis in health and disease. *Metabolism* **63**, 895–902 (2014).

2. O'Neill, L. M. *et al.* Fatty acid desaturation and elongation in mammals. In *Biochemistry of Lipids, Lipoproteins and Membranes* 7th edn, 201–226 (Elsevier, 2021).

3. van Son, M. *et al.* Genome-wide association study confirm major QTL for backfat fatty acid composition on SSC14 in Duroc pigs. *BMC Genom.* **18**, 1–13 (2017).

_____

4. Wood, J. D. *et al.* Effects of fatty acids on meat quality: A review. *Meat Sci.* **66**, 21–32 (2003).

5. Wood, J. D. *et al.* Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci.* **78**, 343–358 (2008).

6. Bosch, L., Tor, M., Reixach, J. & Estany, J. Age-related changes in intramuscular and subcutaneous fat content and fatty acid composition in growing pigs using longitudinal data. *Meat Sci.* **91**, 358–363 (2012).

7. Nürnberg, K., Wegner, J. & Ender, K. Factors influencing fat composition in muscle and adipose tissue of farm animals. *Livest. Prod. Sci.* **56**, 145–156 (1998).

8. Zhang, W. *et al.* Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. *Sci. Rep.* **6**, 1–12 (2016).

9. García-Olmo, J. *et al.* Determination of the precision of the fatty acid analysis of Iberian pig fat by gas chromatography. Results of a mini collaborative study. *Meat Sci.* **60**, 103–109 (2002).

10. Ropka-Molik, K. *et al.* Comprehensive analysis of the whole transcriptomes from two different pig breeds using RNA-Seq method. *Anim. Genet.* **45**, 674–684 (2014).

11. Ramayo-Caldas, Y. *et al.* Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genom.* **13**, 1–18 (2012).

12. Corominas, J. *et al.* Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genom.* **14**, 1–14 (2013).

13. Puig-Oliveras, A. *et al.* Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. *PLoS ONE* **9**, e99720 (2014).

14. Jump, D. B. Dietary polyunsaturated fatty acids and regulation of gene transcription. *Curr. Opin. Lipidol.* **13**, 155–164 (2002).

15. Guillou, H., Zadravec, D., Martin, P. G. P. & Jacobsson, A. The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. *Prog. Lipid Res.* **49**, 186–199 (2010).

16. Óvilo, C. *et al.* Longissimus dorsi transcriptome analysis of purebred and crossbred Iberian pigs differing in muscle characteristics. *BMC Genom.* **15**, 1–24 (2014).

17. Zhang, Y. *et al.* Genetic correlation of fatty acid composition with growth, carcass, fat deposition and meat quality traits based on GWAS data in six pig populations. *Meat Sci.* **150**, 47–55 (2019).

_____

18. Puig-Oliveras, A. *et al.* Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. *Sci. Rep.* **6**, 1–12 (2016).

19. Criado-Mesas, L. *et al.* Identification of eQTLs associated with lipid metabolism in Longissimus dorsi muscle of pigs with different genetic backgrounds. *Sci. Rep.* **10**, 1–13 (2020).

20. Crespo-Piazuelo, D. *et al.* Identification of strong candidate genes for backfat and intramuscular fatty acid composition in three crosses based on the Iberian pig. *Sci. Rep.* **10**, 1–17 (2020).

21. Yang, B. *et al.* Genome-wide association analyses for fatty acid composition in porcine muscle and abdominal fat tissues. *PLoS ONE* **8**, e65554 (2013).

22. Martínez-Montes, Á. M. *et al.* Using genome wide association studies to identify common QTL regions in three different genetic backgrounds based on Iberian pig breed. *PLoS ONE* **13**, e0190184 (2018).

23. Mach, N. *et al.* Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. *J. Anim. Sci.* **84**, 3039–3048 (2006).

24. Andrews, S. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

25. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

26. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

27. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 1–16 (2011).

28. Liu, K., Ye, J., Yang, Y., Shen, L. & Jiang, H. A unified model for joint normalization and differential gene expression detection in RNA-seq data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **16**, 442–454 (2019).

29. R Core Team. A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria. https://www.R-project.org/ (2022).

30. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

31. Lachmann, A. *et al.* Geneshot: search engine for ranking genes from arbitrary text queries. *Nucleic Acids Res.* **47**, W571–W577 (2019).

_____

32. Krassowski, M. ComplexUpset: Create complex upset plots using ggplot2 components. https://cran.r-project.org/web/packages/ComplexUpset/index.html (2021).

33. Bindea, G. *et al.* ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).

34. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

35. Wickham, H. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org (2016).

36 Stelzer, G. *et al.* The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* https://doi.org/10.1002/cpbi.5 (2016).

37. Matsuzaka, T. *et al.* Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. *Nat. Med.* **13**, 1193–1202 (2007).

38. Corominas, J. *et al.* Polymorphism in the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. *PLoS ONE* **8**, e53687 (2013).

39. Corominas, J. *et al.* Epigenetic regulation of the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. *Genet. Sel. Evol.* **47**, 1–11 (2015).

40. Lo, A. S. Y. *et al.* Developmental regulation and cellular distribution of human cytosolic malate dehydrogenase (MDH1). *J. Cell. Biochem.* **94**, 763–773 (2005).

41. Gao, P. *et al.* Selection of candidate genes affecting meat quality and preliminary exploration of related molecular mechanisms in the Mashen pig. *Asian-Australas. J. Anim. Sci.* **32**, 1084–1094 (2019).

42. Goldberg, I. J. Lipoprotein lipase and lipolysis: Central roles in lipoprotein metabolism and atherogenesis. *J. Lipid Res.* **37**, 693–707 (1996).

43. Dervishi, E. *et al.* Gene expression and fatty acid profiling in longissimus thoracis muscle, subcutaneous fat, and liver of light lambs in response to concentrate or alfalfa grazing. *Front. Genet.* **10**, 1070 (2019).

44. Cánovas, A. *et al.* Functional and association studies on the pig HMGCR gene, a cholesterol-synthesis limiting enzyme. *Animal* **4**, 224–233 (2010).

45. Widmann, P., Nuernberg, K., Kuehn, C. & Weikard, R. Association of an ACSL1 gene variant with polyunsaturated fatty acids in bovine skeletal muscle. *BMC Genet.* **12**, 1–13 (2011).

46. Liu, X. *et al.* Muscle transcriptional profile based on muscle Fiber, mitochondrial respiratory activity, and metabolic enzymes. *Int. J. Biol. Sci.* **11**, 1348–1362 (2015).

47. Zhou, X., Curbo, S., Li, F., Krishnan, S. & Karlsson, A. Inhibition of glutamate oxaloacetate transaminase 1 in cancer cell lines results in altered metabolism with increased dependency of glucose. *BMC Cancer* **18**, 1–14 (2018).

48. Gol, S. *et al.* Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs. *J. Anim. Breed. Genet.* **133**, 24–30 (2016).

49. Li, B. *et al.* A key gene, PLIN1, can affect porcine intramuscular fat content based on transcriptome analysis. *Genes (Basel)* **9**, 194 (2018).

50. Benítez, R. *et al.* Modulatory effects of breed, feeding status, and diet on adipogenic, lipogenic, and lipolytic gene expression in growing iberian and duroc pigs. *Int. J. Mol. Sci.* **19**, 22 (2018).

51. Wang, J. *et al.* The effect of leptin on Lep expression is tissue-specific and nutritionally regulated. *Nat. Med.* **5**, 895–899 (1999).

52. Margetic, S., Gazzola, C., Pegg, G. G. & Hill, R. A. Leptin: A review of its peripheral actions and interactions. *Int. J. Obes.* **26**, 1407–1433 (2002).

53. Torres-Rovira, L. *et al.* Diet-induced swine model with obesity/leptin resistance for the study of metabolic syndrome and type 2 diabetes. *Sci. World J.* **2012**, 1–8 (2012).

54. Wang, L., Ouyang, S., Li, B., Wu, H. & Wang, F. GSK-3β manipulates ferroptosis sensitivity by dominating iron homeostasis. *Cell Death Discov.* **7**, 1–9 (2021).

55. Girault, J. A. & Nairn, A. C. Chapter four-DARPP-32 40 years later. *Adv. Pharmacol.* **90**, 67–87 (2021).

56. Liu, Y. *et al.* Trait correlated expression combined with eQTL and ASE analyses identified novel candidate genes affecting intramuscular fat. *BMC Genom.* **22**, 1–14 (2021).

57. He, L., Vasiliou, K. & Nebert, D. W. Analysis and update of the human solute carrier (SLC) gene superfamily. *Hum. Genom.* **3**, 195–206 (2009).

58. Stahl, A. A current review of fatty acid transport proteins (SLC27). *Pflugers Arch. Eur. J. Physiol.* **447**, 722–727 (2004).

59. TRARG1 trafficking regulator of GLUT4 (SLC2A4) 1 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/286753.

60. Duan, X. *et al.* Membrane topology of trafficking regulator of GLUT4 1 (TRARG1). *Eur. PMC Biochem.* **57**(26), 3606–3615 (2018).

61. Jagla, K. *et al.* Mouse Lbx1 and human LBX1 define a novel mammalian homeobox gene family related to the *Drosophila lady bird* genes. *Mech. Dev.* **53**, 345–356 (1995).

_____

62. Chao, Z. *et al.* Molecular characterization and expression patterns of Lbx1 in porcine skeletal muscle. *Mol. Biol. Rep.* **38**, 3983–3991 (2011).

63. Ramayo-Caldas, Y. *et al.* From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine. *BMC Genom.* **15**, 232 (2014).

64. He, X. P. *et al.* Investigation of Lpin1 as a candidate gene for fat deposition in pigs. *Mol. Biol. Rep.* **36**, 1175–1180 (2009).

65. Yang, R. Y. *et al.* Ablation of a galectin preferentially expressed in adipocytes increases lipolysis, reduces adiposity, and improves insulin sensitivity in mice. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18696–18701 (2011).

66. Wu, W., Yin, Y., Xu, K., Peng, Y. & Zhang, J. Knockdown of LGALS12 inhibits porcine adipocyte adipogenesis via PKA–Erk1/2 signaling pathway. *Acta Biochim. Biophys. Sin. (Shanghai)* **50**, 960–967 (2018).

Acknowledgements

_____

Author contributions

J.M.F. conceived the study and was the principal investigator of the project. J.M.F. collected the animal samples. L.C.M., D.C.P. and A.C. performed the total RNA isolation. J.M.F. and Y.R.C. designed the RNA-Seq experiment. A.E.C. performed the bioinformatic analysis to generate the gene expression data. D.C.P, M.P. and C.S. participated in data processing and bioinformatic methods. J.V.H. proposed the statistical methodology and performed the global association study and gene functional analysis. Y.R.C. reviewed the statistical methodology and conceived the structure of the paper. J.V.H., Y.R.C. and J.M.F. wrote the paper. All authors critically revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-27016-x.

Correspondence and requests for materials should be addressed to J.V.-H.

Reprints and permissions information is available at www.nature.com/reprints.

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

_____

_____

Paper III

Identification of candidate regulatory genes for intramuscular fatty acid composition by transcriptome analysis in pigs

Jesús Valdés-Hernández[1,2*], Josep M. Folch[1,2], Daniel Crespo-Piazuelo[3], Magí Passols[1], Cristina Sebastià[1,2], Lourdes Criado-Mesas[1], Anna Castelló[1,2], Armand Sánchez[1,2], & Yuliaxis Ramayo-Caldas[3*]

[1]Plant and Animal Genomics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra, Spain

[2]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra, Spain

[3]Departament de Genètica i Millora Animal, Institut de Recerca y Tecnologia Agraroalimentàries (IRTA), Caldes de Montbui, Spain

*Corresponding author: jesus.valdes@cragenomica.es; yuliaxis.ramayo@irta.cat

E-mail addresses:

JVH: jesus.valdes@cragenomica.es
JMF: josepmaria.folch@uab.cat
DCP: daniel.crespo@irta.cat
MP: magi.passols@cragenomica.es
CS: cristina.sebastia@cragenomica.es
LCM: lourdes.criado@cragenomica.es
AC: anna.castello@uab.cat
AS: armand.sanchez@uab.cat
YRC: yuliaxis.ramayo@irta.cat

_____

Abstract

Background

The intramuscular fat content and its fatty acid (FA) composition are typically controlled by the action of several genes, each with small effects. In the current study, we aimed to perform a multivariate integrative analysis between intramuscular FA and transcriptome profiles of porcine *longissimus dorsi* (LD) muscle to pinpoint candidate genes and putative regulators related to FA composition. We also employed a combination of network, regulatory impact factor, and functional analysis to reinforce the biological relevance.

Results

For this propose, we used LD RNA-Seq and intramuscular FA composition of 129 Iberian × Duroc backcrossed pigs. We identified 378 correlated variables (13 FAs and 365 genes), including six FAs (C20:4n-6, C18:2n-6, C20:3n-6, C18:1n-9, C18:0, and C16:1n-7) that were among the most interconnected variables in the network. FA-correlated genes were related to lipid and/or carbohydrate metabolism (e.g., *ADIPOQ*, *CYCS*, *CYP4B1*, *ELOVL6*, *FBP1*, *G0S2*, *HMGCR*, *LEP*, *LGALS12*, *LPIN1*, *PLIN1*, *PNPLA8*, *PPP1R1B*, *SDHD*, *SDR16C5*, *SFRP5*, *SOD3*, and *TFRC*), meat quality (*GALNT15*, *GOT1*, *MDH1*, *NEU3*, and *PDHA1*), and transport (e.g., *EXOC7* and *SLC44A2*). Functional analysis highlighted 55 gene ontology terms over-represented, including well-known biological processes and pathways regulating lipid and carbohydrate metabolism. Regulatory impact factor (RIF) analysis indicated the pivotal role of six transcription factors (*CARHSP1*, *LBX1*, *MAFA*, *PAX7*, *SIX5*, and *TADA2A*) as putative regulators of gene expression and intramuscular FA composition. Overall, *TADA2A* and *CARHSP1* are novel regulators in pigs, which presented extreme RIF scores, but also *TADA2A* expression correlated (either positive or negative direction) with C20:4n-6, C18:2n-6, C20:3n-6, C18:1n-9, and *CARHSP1* expression (positive direction) with the C16:1n-7 lipokine. The *in silico* prediction suggested the change in co-expression correlation for the six regulators and their target genes according to differences in the FA profile. For *TADA2A* and *CARHSP1*, shared target genes potentially involved in lipid metabolism (e.g., *GOT1*, *PLIN1*, and *TFRC*) were found.

_____

Conclusions

The current results contribute to elucidating the complex and bipartite relationship between intramuscular FA and gene expression profiles in pigs. Together, these findings highlight promising candidate genes and putative regulators linked to meat quality, lipid and carbohydrate metabolism, which could potentially control gene expression and modulate FA composition on pig LD muscle.

Background

Fatty acids (FAs) are crucial for living organisms, serving as important energy sources, but also playing an important role in human health. Furthermore, FA composition plays a significant role in determining meat quality parameters in pigs [1], including technological and sensorial quality of meat products [2]. FAs can be classified as saturated (SFA, no double bond is present) or unsaturated (with one or more double bonds present), they can also be referred as monounsaturated (MUFA) and polyunsaturated (PUFA).

Fat, liver, and muscle are important tissues for FA metabolism and show different FA compositions, which are affected by several factors including genetic, management, environmental, and gene expression. Notwithstanding, the relationship between FA composition and gene expression is complex and still not fully elucidated. For complex phenotypes such as intramuscular fat (IMF, also referred as marbling) content and its FA composition in porcine muscle [3], we have previously identified by hierarchical clustering analysis that lipogenic-related genes showed positive correlations with MUFA in general, while the lipolytic-related genes showed a positive correlation specifically with PUFA. Likewise, by conducting RNA sequencing (RNA-Seq) experiments, we also have reported numerous candidate genes with differential effect or global changes on intramuscular FA composition across several tissues like backfat, liver and muscle [4–7].

In the context of systems biology, there are several tools that allow the exploration and integration of biological data sets with a focus on variable selection and graphical representation. Among them, mixOmics includes a plethora of multivariate methodologies with extensive statistical approximations [8]. On the other hand, a better understanding of biological mechanisms involved in the determination of complex traits requires integrative approaches via

gene networks [9], as well as the identification of the main regulators driving gene-by-gene interaction [10].

The combination of the above-mentioned approaches provides a representation of the connection between data in a natural way, identification of patterns and hidden knowledge, and interpretation of the results to formulate biological hypotheses. In the current study, we performed a multivariate integrative analysis between intramuscular FA and gene expression profiles in pigs. Further, we aimed to identify representative FA phenotypes, regulators, candidate genes, and biological processes and metabolic pathways related to the FA composition in muscle.

Methods

Animals and phenotypic data

A total of 129 animals from an experimental backcross (25% Iberian and 75% Duroc, BC1_DU) were employed. All animals were maintained under the same intensive conditions and fed *ad libitum* with a cereal-based commercial diet and free access to water. The experimental BC1_DU generation, animal raising, and feeding was described in a previous paper [11]. Animal procedures were carried out according to the Spanish Policy for Animal Protection RD1201/05, which meets the European Union Directive 86/609 about the protection of animals used in experimentation. This study was conducted in accordance with relevant guidelines and regulations of the animal care and use committee of the *Institut de Recerca i Tecnologia Agroalimentàries* (IRTA), which adopts "The European Code of Conduct for Research Integrity". The experimental protocol was approved by the Ethical Committee of the IRTA. Our study is also reported in full compliance with ARRIVE guidelines (https://arriveguidelines.org/).

Animals were slaughtered in the same commercial abattoir of Mollerussa (Spain). Samples of the *longissimus dorsi* (LD) skeletal muscle (59 females and 70 non-castrated males) distributed in five slaughterhouse batches were collected, immediately snap frozen in liquid nitrogen and stored at -80 °C until analysis. At slaughter, the average age of pigs was 190 days (range 174 to 205 days), with an average carcass weight (CW) of 73.70 kg (range 46.10 to 109.20 kg).

FA composition with C14-C20 range in LD muscle (n = 129) was determined using a gas chromatography of methyl esters protocol [12]. Briefly, 200 g of muscle samples from each BC1_DU pig were homogenized and used to measure the FA profile per duplicate. Crespo-Piazuelo et al. [13] paper provides additional information on the muscle FA composition in the BC1_DU population. Then, the quantification of each individual FA methyl ester (n = 15 phenotypes of FAs) was calculated and expressed as a relative percentage of the total amount of FAs (Table 1).

Table 1 Summary of descriptive statistics on the FA composition phenotypes measured in relative values in the *longissimus dorsi* muscle from BC1_DU pigs.

Trait	Name	Mean	SD	Min	Max	SE	CV*
C14:0	Myristic acid	1.27	0.23	0.73	1.78	0.02	17.95
C16:0	Palmitic acid	23.91	1.65	18.69	27.59	0.15	6.90
C18:0	Stearic acid	14.38	1.69	8.81	19.90	0.15	11.78
C20:0	Arachidic acid	0.23	0.08	0.08	0.71	0.01	33.94
C16:1*n*-7	Palmitoleic acid	2.79	0.53	1.24	4.08	0.05	18.83
C16:1*n*-10	Sapienic acid	0.30	0.12	0.16	0.92	0.01	38.62
C18:1*n*-9	Oleic acid	35.93	5.71	19.99	44.15	0.50	15.88
C18:1*n*-7	Vaccenic acid	3.82	0.30	3.02	4.83	0.03	7.97
C20:1*n*-9	Gondoic acid	0.73	0.16	0.35	1.48	0.01	22.38
C18:2*n*-6	Linoleic acid	12.13	5.82	4.81	29.34	0.51	48.01
C18:3*n*-3	α-Linolenic acid	0.40	0.13	0.15	0.89	0.01	32.77
C20:2*n*-6	Eicosadienoic acid	0.43	0.12	0.14	0.91	0.01	28.61
C20:3*n*-3	Eicosatrienoic acid	0.18	0.10	0.02	0.65	0.01	54.98
C20:3*n*-6	Dihomo-gamma-linolenic acid	0.45	0.29	0.09	1.49	0.03	63.29
C20:4*n*-6	Arachidonic acid	2.58	1.94	0.47	10.51	0.17	75.01

*Coefficient of variation (CV) expressed in percentage.

Gene expression data

Total RNA was isolated from the muscle (100 mg) of 129 animals using the RiboPure™ Isolation kit for High-Quality Total RNA (Ambion®; Austin, TX) following the manufacturer's recommendations. RNA quantification and purity was done with a NanoDrop ND-1000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). While RNA integrity was

checked by Agilent Bioanalyzer-2100 equipment (Agilent Technologies, Inc., Santa Clara, CA, USA), and samples with an RNA integrity number (RIN) greater than 7 were used for the RNA-Seq experiment.

Library preparation and sequencing was performed at CNAG institute (*Centro Nacional de Análisis Genómico*, Barcelona, Spain). For each sample, one paired-end library with an approximately 300-bp insert size was prepared using TruSeq Stranded mRNA kit (Illumina, Inc.; San Diego CA, USA). To discriminate among samples, libraries were labeled by barcoding and pooled to be run in Illumina HiSeq 3000/4000 instruments (Illumina, Inc.; San Diego CA, USA). In brief, this study yielded a mean of 45.09 million of 2×75 bp paired-end reads per sample, resulting on an average of 90.06% (ranging from 80.51% to 96.09%) of uniquely mapped reads.

Bioinformatic and statistical analyses

Quality control and basic statistics of sequence data were performed using FastQC (v0.11.9) [14] and MultiQC v0.7 [15] programs. Sequencing reads were mapped employing the STAR software (v2.7.9a) with default parameters [16], and using the *Sscrofa11.1* pig genome assembly as reference. Then, gene expression was quantified by RSEM (v1.2.28) software [17] with default parameters and annotation from Ensembl Pig Genes 97.

Pre-processing of the raw count matrix was performed by filtering based on a minimum of 129 reads in total per row and 15,091 genes were retained for further analyses. We transformed this matrix data into counts per million (CPM) to normalize the values (i.e., with log = TRUE and prior.count = 1 arguments) using edgeR v3.38.1 package [18]. Later, the 15,091 genes retained were matched against the newer annotation from Ensembl Pig Genes 104 (*Sscrofa11.1*) using the biomaRt package [19] v2.52.0, remaining a total of 12,381 genes with a gene name or symbol.

Afterwards, a regularized canonical correlation analysis (rCCA) was performed using the expression dataset of the 12,381 genes and the 15 FA traits measured in the 129 individuals. The rCCA is a multivariate approach, implemented in mixOmics v6.14.1 package [8], that allows the identification of subsets of canonical variables that maximize the correlation between two datasets "X and Y" (respectively of sizes n × p and n × q) [20]. In our case, X was the matrix of 15 FA traits and Y was a matrix with the 12,381 genes expression values. The shrinkage

method was used to tune out the regularization parameters (λ1 and λ2) with values of λ1 = 0.05 and λ2 = 0.15, and ncomp = 3. Rather than considering all genes that were included in the first canonical component (CC1) and according to the previous estimate by Ramayo-Caldas et al. [21], we applied a conservative approach and only kept the genes for which the correlation between gene expression and FA traits was equal or higher than 0.29.

To display the rCCA results and improve its interpretation, three graphical outputs (circle plot, network, and Clustered Image Maps) were applied. All of them implemented in mixOmics [8]. First, a circle plot with subsequent dimensions via *plotVar* function was used in order to represent the nature or structure of the correlation between variables of interest. Secondly, a network plot was conducted as a robust way to evaluate the correlation structure by using the *network* function. Here, the generated output was exported to Cytoscape file format using igraph v1.3.2 package [22]. In the network [23] FAs were filled with colors to facilitate the identification of the group (i.e., SFA, MUFA and PUFA) and the connected genes. Likewise, the genes were filled with different colors according to the functional group or gene family using a pre-built list accordingly to their functions (see gene functional classification section). In addition, a complementary heatmap via ComplexHeatmap v2.14.0 package [24] to illustrate the different clusters of variables and the degree of correlation between them.

Regarding the functional analysis, genes included in the CC1 (cutoff of r ± |0.29|) were submitted to the ClueGo v2.5.4 plugin [25] in Cytoscape v3.7.1 software [23] with default parameters. Gene ontology (GO) significance was assessed with a hypergeometric test keeping only those GO terms (biological processes, molecular function, and pathways) with a Benjamini-Hochberg (BH) corrected *P*-value < 0.05 [26]. All genes expressed in muscle (12,381) were employed as background in this step. Furthermore, GO tree interval levels were set from three to eight and a minimum k-score of 0.44 and a minimum of three genes per cluster with at least 4% in selected genes were used. Results with and without the fusion feature "GO Term Fusion" were generated to evaluate the redundant parent-child terms. Additionally, we visualized the ClueGO output using an R script via ggplot2 v3.3.5 package [27], which allowed us to identify the intersection of significantly associated genes according to over-represented GO terms.

_____

Regulatory impact factors (RIF) analysis

The RIF analysis was conducted using the *runAnalysis* function of the CeTF v1.8.0 package [28]. The RIF algorithm is well described in the original paper by Reverter et al. [10]. Briefly, the RIF metric (RIF1 and RIF2) aims to identify relevant regulators (i.e., transcription factors, TFs) from gene expression data. This step calculates, for each condition, the co-expression correlation between the TFs and the differentially expressed (DE) genes. For the differential analysis, we created two conditions by classifying samples according to their FA profile through a principal component analysis (PCA). In the PCA, the *prcomp* function with scale = TRUE was used considering as input the composition data of the pre-selected FAs from the rCCA. In fact, we chose the extreme values from PC1 (condition 1 and condition 2 with n = 60, i.e., 30 samples per condition). Here, animals in condition 1 consisted of a profile of less SFA and MUFA, and more PUFA, while those in condition 2 were vice versa (more SFA and MUFA, and less PUFA). In addition, the *fviz_pca* function of the factoextra v1.0.7 package was used to extract and visualize the PCA results [29]. Significant differences (corrected *P*-value < 0.05) between the mean of FA conditions and by phenotypes selected in the rCCA were determined using a t-test approach, and the SEM (standard error of the mean) was calculated. In addition, transcription factors (TFs) present in the expression data were identified using the list of pig TFs available in the AnimalTFDB v3.0 database (http://bioinfo.life.hust.edu.cn/Animal TFDB/#!/).

Transcription factors target gene analysis

The list of TFs used in the RIF step was also examined to identify *in silico* putative target DE genes between the two conditions (n = 60 samples). This complementary analysis was conducted using the *SmearPlot* function [28]. Predicted target genes were extracted as identified by this approach. In addition, the generation of these calculations for DE genes and TFs included default parameters (lfc = 1.5 and padj. = 0.05).

Functional classification of FA-interconnected genes in the network

To facilitate the functional gene discovery, we classified the global list of rCCA-derived genes with a view to enhance biological interpretation. First, we used the information from ClueGO analysis that divided the genes into different functional groups, which contained the biological processes and pathways clustered together according to GO term similarities. Then, a manual

list (list 1, see Additional file 1: **Table S1**) of genes with functional annotation and plausible function in different aspects of FA, lipid and carbohydrate metabolism was compiled. Second, we created a trained list (list 2, see Additional file 2: **Table S2**) using GUILDify v2.0 tool [30]. This list included genes associated with predefined keywords such as: "adipokine", "amino acid metabolism", "electron transport chain", "enzyme", "fatty acid beta-oxidation", "fatty acid metabolism", "fatty acid synthesis", "gluconeogenesis", "glucose metabolism", "glycolysis", "tricarboxylic acid cycle", "lipid metabolism", "carbohydrate metabolism", "lipid degradation", "lipid synthesis", "nucleic acid", "nucleotide metabolism", "nutrient", "receptor", and "transporter"; as well as *homo sapiens* species and lipogenic tissues (adipose, liver and muscle-skeletal) options. Briefly, GUILDify uses BIANA knowledge base to create a species-specific interaction network for each gene detected. In the current study, the *netcombo* prioritization algorithm based on network topology, and the highest guild score for the top 100 gene products (with only seed) were considered to compose the trained list. Another list of TFs and cofactors in pigs (list 3, see Additional file 3: **Table S3**) was created according to the annotation of the aforementioned AnimalTFDB v3.0 database. Altogether, a match among gene lists (list 1 vs. list 2 and list 1 vs. list 3) was tested. Finally, the gene functional classification was based on the potential biological function compiled from the overlap between gene lists.

Results

Regularized canonical correlation analysis (rCCA)

Here, we explored the relationship between the muscle transcriptome (n = 12,381 genes) and the intramuscular FA composition (n = 15 FAs) of 129 BC1_DU pigs. The correlation structure between data sets of interest is shown in **Fig. 1**. In the correlation circle plot, the two circles (of radii 0.29 and 1) delimit the subsets of variables that define each component in the plot. After gene selection, the rCCA yielded 365 genes and 15 FAs (see Additional file 4: **Table S4**) that were included in the first canonical component (CC1). As expected, our results (**Fig. 1**) revealed that CC1 separated both SFA (C16:0) and MUFA (C18:1n-9) from PUFA (C18:3n-3, C18:2n-6, C20:2n-6, C20:3n-3, C20:3n-6 and C20:4n-6), while the second canonical component (CC2) differentiated C16:1n-7 MUFA from C18:0 SFA. Additional correlation circle plot from the PCA with contribution variables according to CC1 versus third canonical component (CC3) is presented in Additional file 5: **Figure S1**.

The most relevant variables participating in the definition of each component were C16:1*n*-7, C18:1*n*-9 and C18:0, as well as C18:2*n*-6, C20:3*n*-6, C20:4*n*-6, C20:3*n*-3 and C20:2*n*-6 (**Fig. 1**). Notably, the 365 genes were clustered in 4 groups at radii ~0.29 (with prioritized variables of gene expression) (**Fig. 1**). However, visual inspection of **Fig. 1** is not intuitive for the identification of candidate genes present in each of the 4 clusters. Instead, details on the genes correlated with each FA (with and without cutoff) can be found in Additional file 4: **Table S4**.
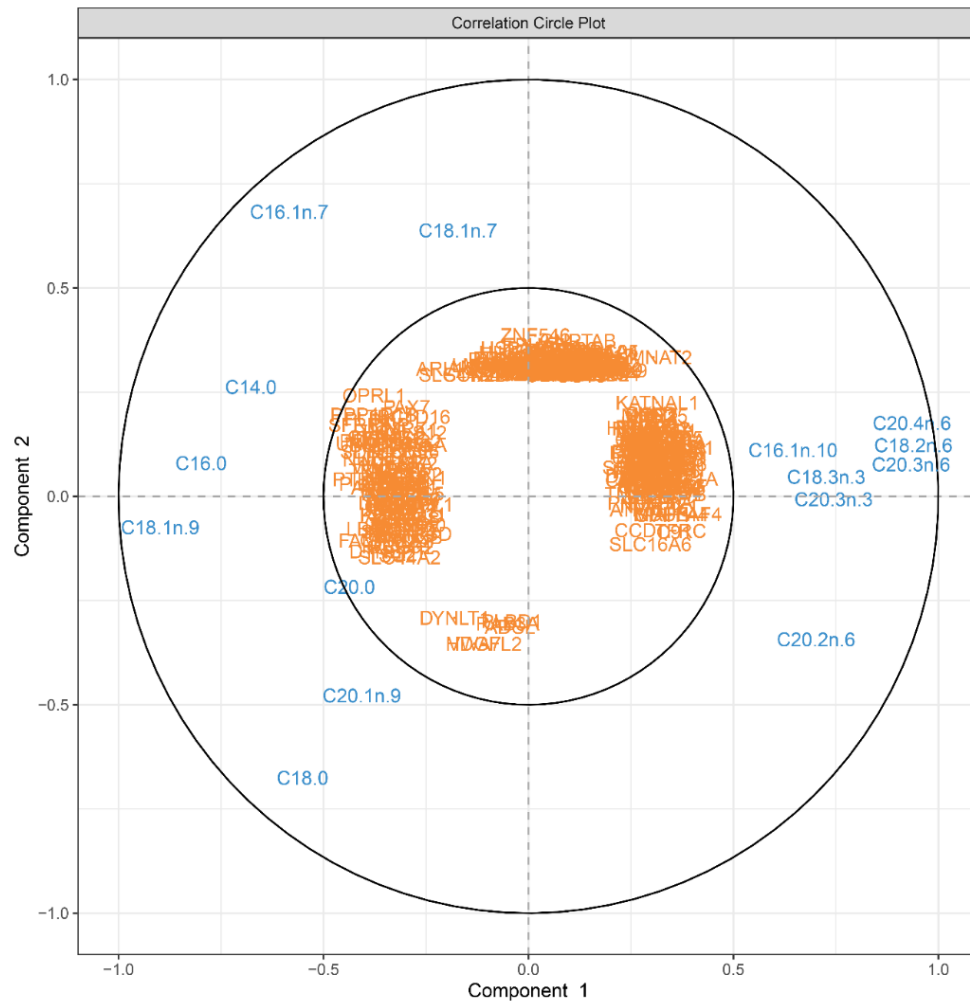
Figure 1 Correlation circle plot from the PCA applied to of the FA phenotypes and gene expression in muscle of BC1_DU pigs displayed from the first two rCCA dimensions (15 FAs and 365 genes selected in total).

A network approach was used to evaluate the structure of associations between variables (**Fig. 2**). As shown in **Fig. 2**, variables (FAs and genes) were drawn as the nodes (rectangles and circles, respectively), and the bipartite relationship between them were represented by the colored edges (i.e., lines in green and in red indicated positive and negative correlations,

respectively). Overall, we detected a complex correlation structure between the FA profile and gene expression data. Simultaneously, the representation of correlations suggested that positive relationships were more common than negative ones (**Fig. 2**). On the other hand, the network adds another layer of information allowing the visualization of groups of variables (**Fig. 2**). Our results showed that 13 FAs and 365 genes (378 variables in total) were selected on the first three rCCA dimensions. Among the 13 FAs, five (C16:1n-7, C20:1n-9, C18:0, C18:2n-6, and C20:4n-6) were correlated with a list of specific genes for each of them (see Additional file 4: **Table S4**). The fourth most interconnected FAs with multiple correlated genes (see Additional file 4: **Table S4**) were C20:4n-6, C18:2n-6, C20:3n-6 and C18.1n-9 (**Fig. 2**). Remarkably, both C20:4n-6 and C18:2n-6 showed the highest number of correlated genes, having 87% of shared genes between them. Likewise, these two FAs contain information of several connected genes with FAs from the periphery and the center of the network (**Fig. 2**).

At the bottom right, in particular, C18:1n-9 was grouped very close to seven other FAs (C14:0, C20:3n-3, C20:2n-6, C20:3n-6, C16:1n-10, C18:3n-3 and C16:0) (**Fig. 2**). In fact, C20:3n-6 and C18:1n-9 were the second pair of FAs that exhibited the highest number of correlated genes, having 80.54% of genes in common. To be noted, the subsets of connected genes suggested more complex relationships with the presence of shared or specific genes for FAs (**Fig. 2**).

It is worth mentioning that we observed a set of 10 rCCA-genes (*PLIN1*, *UNC93A*, *SFRP5*, *PPP1R1B*, *LEP*, *OPRL1*, *PTPRU*, *NUDT14*, *TFRC* and *CYP4B1*) that showed the highest number of connections with FAs (i.e., 7 to 10 connections out of 13 FAs). However, we also detected 107 genes uniquely correlated with a specific FA (**Fig. 2**, and see Additional file 4: **Table S4**). For instance, *CEBPG* and *GCLC* were negatively correlated with C18.0 while *EXOC7* was positively correlated. Genes such as *ADIPOQ*, *LGALS12*, *PAX7*, *CARHSP1* and *YBX2* were positively correlated with C16:1n-7. As other examples, *CEBPA*, *ELOVL6*, and *MCTP2* correlated negatively with C18:2n-6 and C18:3n-3, respectively; while *SDHD* and *LPIN1* correlated positively with C20:4n-6, and also *LEPROTL1* was positively correlated with C18:2n-6 and C20:4n-6. Meanwhile, the *PNPLA8* gene was negatively correlated with C18:1n-9, but it also correlated positively with C18:2n-6 and C20:4n-6 (**Fig. 2**).

Figure 2 Network plot for the longissimus dorsi muscle study in BC1_DU pigs. Green and red edges indicate positive and negative correlation. Legend: Output obtained for the first three rCCA dimensions (13 FAs and 365 genes were selected), showing the correlation structure for all bipartite relationships with a correlation cutoff of 0.29. Color legend: FAs: magenta = SFA members; royal blue = MUFA members; orange red = PUFA members; and genes: darkorange = enzyme; aquamarine = adipokine; chartreuse = TF; turquoise = TF cofactors; yellow = lipid metabolism-related genes; navajowhite = carbohydrate metabolism; crimson = glycolysis; gold = transporter; lightpink = fatty acid beta-oxidation; coral = amino acid metabolism; cornsilk = receptor family; deepskyblue = nucleic acid metabolism. Out of a total of 365 genes, the 176

colored genes were referring to a functional or gene family classification (while 189 genes were unclassified).

A complementary heatmap illustrated the different clusters of variables, and the degree of correlation between them (see Additional file 6: **Figure S2**). Meanwhile, details on correlation distribution of FAs with gene expression via density heatmap (with quantiles and mean values) are presented in Additional file 7: **Figure S3**. The heatmap and hierarchical clustering displayed the aforementioned 13 FAs and 365 genes, in which FA traits that are mainly correlated with the gene expression value, cluster together. In general, as shown in Additional file 6: Figure S2, two large clusters of FAs and four clusters of genes were observed. Cluster 1 was composed of six (C16:1n-7, C18:0, C20:1n-9, C18:1n-9, C16:0 and C14:0) FAs, while cluster 2 grouped the other seven (C20:2n-6, C16:1n-10, C18:3n-3, C20:3n-3, C20:3n-6, C20:4n-6, and C18:2n-6) FAs. Details on the grouping of genes contained in each cluster and its respective correlation with FAs can be accessed on Additional file 6: Figure S2 and Additional file 4: **Table S4**, respectively.

Functional analysis of genes correlated with FAs

After rCCA, the gene list (n = 365) was submitted to a GO analysis. A total of 55 GO terms (8 molecular functions, 9 pathways and 38 biological processes) (**Fig. 3**, and see Additional file 8: **Figure S4** with "GO Term Fusion", respectively) were significantly over-represented (BH corrected P-value < 0.05). In total, 125 genes were annotated into different functional groups, including an enrichment in GO terms of lipid and carbohydrate metabolism. Notably, some of the closely associated KEGG pathways were "regulation of lipolysis in adipocytes", "citrate cycle (TCA cycle)", "non-alcoholic fatty liver disease (NAFLD)", "oxidative phosphorylation" and "Insulin signaling pathway". Likewise, GO analysis suggested that these genes were significantly enriched in well-known biological processes, for example, "mitochondrial gene expression", "tricarboxylic acid cycle", "electron transport chain", "ATP hydrolysis coupled cation transmembrane transport", "regulation of response to nutrient levels", "magnesium ion transmembrane transport", and "respiratory electron transport chain" (see Additional file 8: **Figure S4**). The complete results of GO terms with and without "GO Term Fusion" are listed in Additional file 9: **Table S5**.
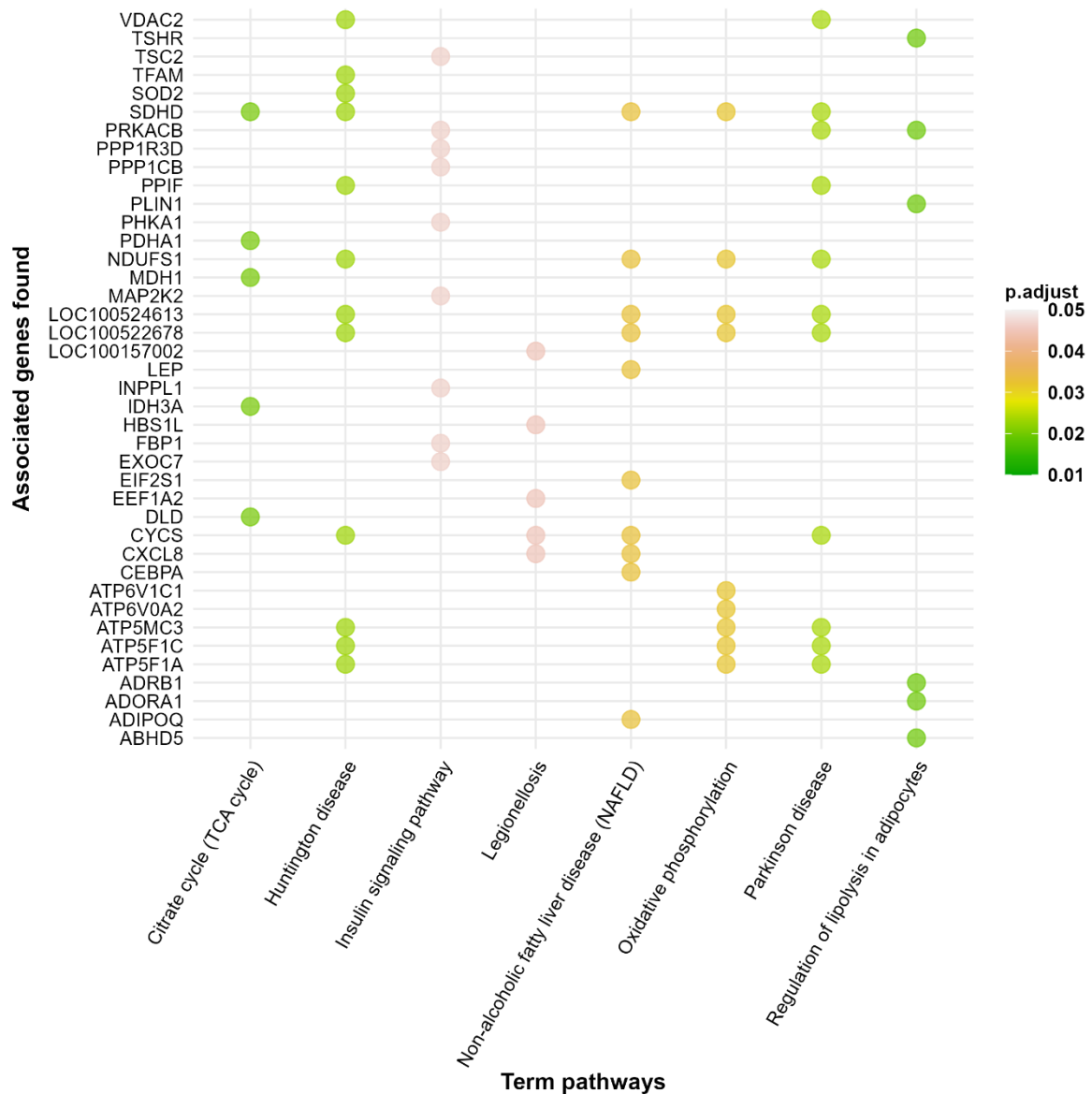
Figure 3 Functional analysis with the correlated genes from the rCCA approach that were significantly enriched in GO terms according to metabolic pathway delimitation. Legend: This output is a representation of the original results table generated with the ClueGO plugin in Cytoscape software.

Regulatory impact factors (RIF) analysis

Based on the output of the rCCA, relevant putative regulators (i.e., TFs) were identified from gene expression data (n = 365 genes including 22 TFs). For this purpose, the classification of animals into two conditions (30 individuals each) was determined based on a PCA considering the FA composition data (n = 13 FA, **Fig. 4**). Condition 1 consisted of a FA profile with less SFA and MUFA and more PUFA, while condition 2 considered the opposite (more SFA and

MUFA and less PUFA), with the exception of C16:1*n*-10, having a similar pattern to that of PUFA (see Additional file 10: Table S6). In addition, the PCA extracted two factors (with variables projected onto PC1 and PC2), that explained 77.40% of the total data variation (**Fig. 4**).
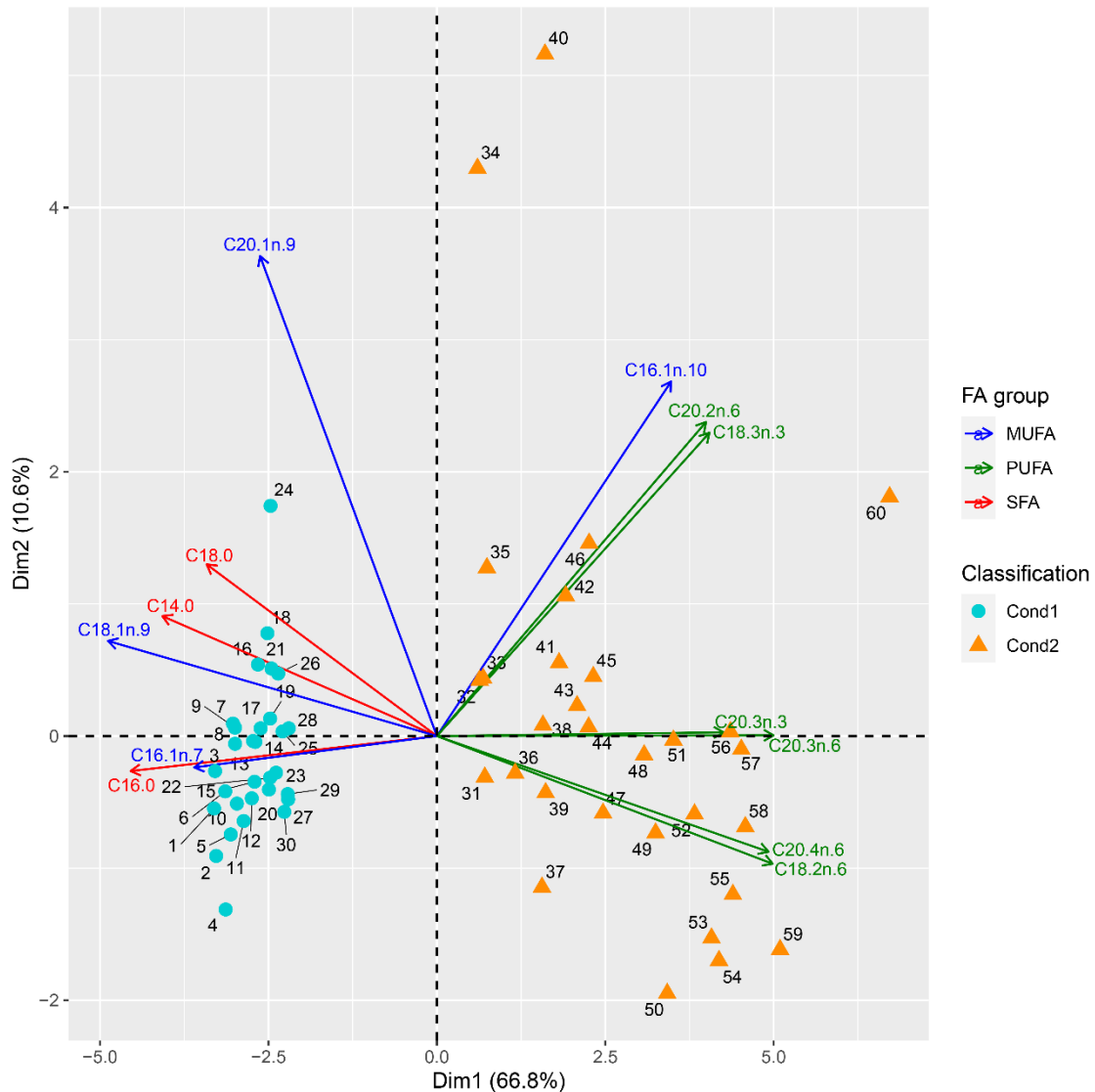
Figure 4 Principal component analysis (PCA) summarizing the separation and similarities among FA conditions with the extreme values. Legend: PC1 and PC2 explained 77.40% of the total variance. The visualization of the eigenvectors was implemented to delimit the FAs belonging to their respective group (SFA, MUFA, PUFA).

Through RIF, we identified 19 significant (lfc = 1.5 and padj = 0.05) and 3 non-significant (*CEBPA*, *DBX1*, and *HOXB6*) regulators out of the 22 TFs identified on the rCCA analysis,

whose RIF parameters are summarized in **Table 2**. For interpretability, it is important to note that information from RIF1 and RIF2 is complementary. RIF1 score classified the TFs most differentially co-expressed with the highly abundant and highly DE genes, whereas RIF2 score classified the TF with the most altered ability to act as predictors of the abundance of DE genes. Among them, here we described the top five TF with extreme values for RIF metrics (**Table 2**). RIF1 (positive score: *CARHSP1, TERF1*, *CEBPG, TFAM, MAF,* and negative score: *TADA2A, MBD2*, *HBP1, SIX5, FOXJ3*); and RIF2 (positive score: *KLF10, TADA2A, HBP1, FOXJ3, ZNF407,* and negative score: *MBD2, SIX5, TERF1, CEBPG, MAF*). On absolute values of RIF, we found the first and second more relevant regulators according to RIF1 (*TADA2A* and *CARHSP1*) and RIF2 (*KLF10* and *TADA2A*), respectively (Table 2). Likewise, **Table 2** illustrates that TFs were classified into ten different families (i.e., based on AnimalTFDB3.0 database). In addition, the results of the co-expression analysis between the TF and rCCA-derived genes (n = 343) in each experimental condition were presented in Additional file 11: **Table S7**. While we also computed the correlation differential (alias DW) written as the correlation of the TF with all possible target genes in conditions 1 and 2, respectively (see Additional file 12: **Table S8**).

Table 2 RIF analysis with 19 significant TFs identified by either RIF1 or RIF2 metrics according to FA conditions and gene expression under study in BC1_DU pigs.

TF	Gene description	Family	avgexpr	RIF1	RIF2
TADA2A	Transcriptional adaptor 2A	MYB	4.11	-2.09	1.44
MBD2	Methyl-CpG binding domain protein 2	MBD	4.90	-1.43	-1.44
SIX5	SIX homeobox 5	Homeobox	2.38	-0.83	-1.30
TERF1	Telomeric repeat binding factor 1	MYB	3.21	1.63	-1.28
CEBPG	CCAAT enhancer binding protein gamma	TF_bZIP	4.96	1.49	-1.15
HBP1	HMG-box transcription factor 1	HMG	6.12	-0.89	1.12
FOXJ3	Forkhead box J3	Fork_head	6.36	-0.54	1.09
CARHSP1	Calcium regulated heat stable protein 1	CSD	3.37	1.86	0.82
MAF	MAF bZIP transcription factor	TF_bZIP	7.19	0.39	-0.68
TFAM	Transcription factor A, mitochondrial	HMG	4.79	0.82	-0.52
KLF10	Kruppel like factor 10	zf-C2H2	6.24	-0.32	1.53
YBX2	Y-box binding protein 2	CSD	4.22	-0.38	-1.06

_____

Table 2 RIF analysis with 19 significant TFs identified by either RIF1 or RIF2 metrics according to FA conditions and gene expression under study in BC1_DU pigs *(Continued)***.**

TF	Gene description	Family	avgexpr	RIF1	RIF2
ZNF407	Zinc finger protein 407	zf-C2H2	3.74	0.10	0.95
ZNF524	Zinc finger protein 524	zf-C2H2	4.84	0.13	0.73
PAX7	Paired box 7	PAX	2.72	-0.28	0.46
ZNF326	Zinc finger protein 326	Others	5.56	0.35	-0.45
LBX1	Ladybird homeobox 1	Homeobox	3.53	0.38	-0.21
MAFA	MAF bZIP transcription factor A	TF_bZIP	3.23	-0.19	-0.09
ZNF146	Zinc finger protein 146	zf-C2H2	3.46	-0.20	0.03

Footnotes: The 19 TFs belong to 10 different families. The expression average (in log2 CPM) for each TF is indicated by avgexpr variable.

In silico **prediction of transcription factor target genes in the post-RIF stage**

After the RIF analysis, an *in silico* prediction was carried out based on the change in co-expression between each differentially expressed (DE) TF and its target genes for both FA conditions. Consistently, target genes whose expression was DE between conditions were identified. The *in silico* identification of target genes for each TF was observed in only 6 (lfc = 1.5 and padj = 0.05) out of the 19 significant TFs (*CARHSP1*, *LBX1*, *MAFA*, *PAX7*, *SIX5*, and *TADA2A*), which belonged to five (MYB, CSD, Homeobox, PAX and TF_bZIP) of the ten families identified (**Table 3**). Additional information on the distribution of DE genes and specific TF in both conditions, as well as the relationship between log(baseMean) and expression difference, is presented in see Additional file 13: Table S9.

Overall, based on functional information, several of the candidate targets (listed in **Table 3**) were involved in lipid metabolism (*ADIPOQ*, *DHRS3*, *FGFR4*, *NEU3*, *PLIN1*, *TFRC*, and *WNT5B*); carbohydrate metabolism (*ADIPOQ*, *FBXO6*, *GALNT15*, *GOT1*, *NEU3*, *NUDT14,* and *PPP1R3D*); glucose metabolism (*MAFA*); and ion binding (*ADIPOQ*, *FGFR4*, *GOT1*, *NUDT14*, *SOD3,* and *TADA2A*); among other categories.

Altogether, this novel approach has allowed us to discover promising targets of newly identified TF genes in pig muscle associated with FA metabolism, such as *TADA2A* and *CARHSP1*. As shown in **Table 3**, both genes were targets for each other, but also they shared targets FA-related genes (e.g., *GOT1*, *PLIN1,* and *TFRC*).

_____

Table 3 Distribution of differentially expressed genes and transcription factors in both conditions, including potential target genes by FA condition in BC1_DU pigs.

Regulatory genes	No. of potential targets	Differentially co-expressed target genes	
		Condition 1 (less SFA and MUFA, and more PUFA)	Condition 2 (more SFA and MUFA, and less PUFA)
TADA2A	15	*TBC1D16, TBKBP1, WNT5B, DHRS3, GADD45G, KCNT1, MAMSTR, PLPP7, SIX5*	*TFRC, APCDD1, CARHSP1, GOT1, PLIN1, SOD3*
CARHSP1	8	*DNAJB9, GOT1, MAFA, PLIN1, TFRC, ADIPOQ*	*FBXO6, TADA2A*
SIX5	5	*TADA2A, FBXO6, GADD45G*	*FGFR4, GALNT15*
PAX7	9	*PLPP7, TENT5C, GALNT15, GOT1, MAFA*	*WNT5B, ADIPOQ, FBXO6, NEU3*
MAFA	8	*PAX7, TFRC, CARHSP1, DHRS3, FBXO6, FGFR4, GADD45G*	*PLIN1*
LBX1	3	*FBXO6, GALNT15*	*TBKBP1*

Discussion

Identification of subsets of canonical variables that maximize the correlation between gene expression and FA profiles

Our study conducted with the rCCA multivariate approach [20] revealed a subset of 378 correlated variables, 13 FAs and 365 genes. By using correlation circle plot, the first component (CC1) clearly separated FAs with relevant contribution, C18:2n-6, C20:3n-6, C20:4n-6 (positive) and C18:1n-9, C16:0 (negative), while C16:1n-7 (positive) and C18:0 (negative) were the main contributors to the second component (CC2). Hence, CC1 differentiated the essential FAs taken from the diet of other FAs that can be *de novo* synthesized in the pig. Conversely, CC2 separated the MUFA derived from FA biosynthesis like C16:1n-7 or C18:1n-7 from the C18:0 SFA.

The exploration of the interaction network [23] revealed the five most interconnected FAs (C20:4n-6, C18:2n-6, C20:3n-6, C18.1n-9, and C18.0), which showed the highest number of associated genes. Among them, we found that n-6 FAs exhibited more than 51% of shared genes between them. Remarkably, we also found 100% and 80.54% overlap when comparing the C18.1n-9 versus C20:4n-6 and C18:2n-6, and against C20:3n-6 lists, respectively. It turns out that some of the genes correlated with C18.1n-9 also exhibited other associations including those with minority FAs. This means that C18.1n-9 is a key FA in muscle, capturing complex associations with the presence of shared or specific genes linked to FA metabolim. Consequently, five FAs (C14:0, C16:0, C18:3n-3, C20:2n-6, and C20:3n-3) were connected with 28 genes in total, indicating their lesser participation in the differences in FA composition observed in our population. Of particular note, for example, C20:2n-6 and C20:3n-3 only exhibited negative correlations with a list of specific genes. Furthermore, five out (C16:1n-7, C20:1n-9, C18:0, C18:2n-6, and C20:4n-6) of the 13 analyzed FAs were correlated with a list of specific genes for each of them.

Interestingly, the FAs were grouped into two large groups via the heatmap [24]. Essential FAs and their derivatives molecules (e.g., C18:2n-6, C18:3n-3, C20:2n-6, C20:3n-3, C20:3n-6, and C20:4n-6) were clustered apart from those that can be derived from biosynthesis (e.g., C14:0, C16:0, C18:0, C18:1n-9, and C20:1n-9). Multiple lines of evidence suggest that essential FAs of the omega-6 series, like arachidonic (C20:4n-6) and its dietary precursor linoleic (C18:2n-6), and the omega-3 series, like the α-linolenic (C18:3n-3) PUFA have important physiological functions [31]. Arachidonic FA is an important constituent of membrane phospholipids and it is involved in signal transduction [32]. In addition, essential PUFA are precursors for the synthesis of collectively known as eicosanoids (prostaglandins, thromboxanes, and leukotrienes) [33], and lipoxins. Meanwhile, omega-6 and omega-3 PUFA have antagonistic inflammatory functions, being arachidonic a pro-inflammatory and immunoactive FA [34]. Of the MUFA, palmitoleic (C16:1n-7) and oleic (C18:1n-9) are the most predominant in triacylglycerols, cholesteryl esters, wax esters and membrane phospholipids [35]. It has been informed that C18:1n-9 may improve meat organoleptic properties and overall acceptability parameters of meat [36], as well as linked to consumer acceptability of high-quality cured products [37]. Furthermore, emerging evidence in mice suggest that C16:1n-7 is an adipose tissue-derived lipid hormone that strongly stimulates muscle insulin action [38].

_____

Candidate genes that correlate with FA composition traits are mainly involved in lipid and/or carbohydrate metabolism and meat quality

Within the 365 FA-correlated genes we observed well-documented candidate genes for lipid and FA metabolism. The gene with the highest number of connections was *PLIN1* (perilipin 1), which was associated with 10 out of the 13 FAs. In addition, *PLIN1* was positively correlated with SFA and MUFA, but negatively correlated with PUFA. As a perilipin member, *PLIN1* plays a role in the regulation of intracellular lipid storage and mobilization [39]. Besides, *PLIN1* was already proposed as a candidate gene in another study carried out by our group, which analyzed the muscle transcriptome of Landrace backcrossed (25% Iberian and 75% Landrace, BC1_LD) pigs with divergent FA composition [6].

With the exception of C18:3n-3, other four genes (*LEP*, *PPP1R1B*, *SFRP5*, and *UNC93A*) were connected with the same 9 FAs as *PLIN1*, presenting a similar pattern of positive correlations with SFA and MUFA, and negative correlations with PUFA. The *LEP* gene encodes the protein leptin, an adipokine that acts as a major regulator for food intake and energy homeostasis [40], being also involved in the regulation of the neuroendocrine axis [41]. In muscle, *LEP* activates AMP-activated kinase (AMPK), thereby increasing the oxidation of FAs and the uptake of glucose [42]. In fact, significant associations of the *LEP* and leptin receptor (*LEPR*) polymorphisms with FA composition have been reported in subcutaneous fat in Iberian × Landrace pigs [43]. Furthermore, *LEP*, together with *SFRP5* (secreted frizzled related protein 5), were identified as IMF-correlated genes in the muscle of Duroc × Luchuan pigs [44]. In addition, *SFRP5* was also identified as a DE gene in the skeletal muscle of Duroc pigs with extreme lipid profiles [45]. In mice, a long non-coding RNA of the protein phosphatase 1 regulatory inhibitor subunit 1B (*PPP1R1B*) gene is involved in skeletal muscle development [46]; such data argue for an important role of *PPP1R1B*-lncRNA in promoting myogenic differentiation through competing for polycomb repressive complex 2 (*PRC2*) binding with chromatin of myogenic master regulators. While the relationship between *UNC93A* (unc-93 homolog A) and the lipid metabolism in muscle has yet to be explored, a study in mice determined that the expression of this putative solute carrier responded to nutrient availability [47]. In addition, *UNC93A* was also mentioned as a candidate gene in a QTL study for meat quality and disease resistance in the Chinese Jiangquhai pig breed [48].

Apart from the five genes mentioned earlier, there are other relevant genes involved in lipid metabolism that are worth mentioning, such as *ADIPOQ, ELOVL6, FBP1*, *LPIN1, G0S2, PNPLA8*, and *SOD3*. As another adipokine and candidate for meat quality, adiponectin (*ADIPOQ*) was positively correlated with C16:1*n*-7. The protein encoded by *ADIPOQ* enhances FA oxidation both in skeletal muscle and liver. In addition, *ADIPOQ* stimulates muscle glucose uptake and inhibits glucose production by the liver, thus, decreasing blood glucose levels [49].

On the regulation of lipogenesis, the ELOVL fatty acid elongase 6 (*ELOVL6*) gene is of particular interest, as it is directly involved in the elongation of FAs in mammals [50]. In our group, Corominas et al. [51] already reported a plausible effect of the expression levels of *ELOVL6* on of the abundance of C16:0 and C16:1*n*-7 FAs in backfat and muscle of Landrace backcrossed pigs. However, in our study, the BC1_DU pigs showed only a negative correlation with the essential FA C18:2*n*-6. As *ELOVL6* does not elongate C18:2*n*-6, this negative correlation may be due to the increase on the relative abundance of other SFA and MUFA caused by *ELOVL6*.

Our data indicated that lipin 1 (*LPIN1*) gene was positively correlated with the abundance of C20:4*n*-6. This gene participates in the metabolism of the arachidonic acid in *Caenorhabditis elegans* organism [52]. Likewise, this gene was also explored as a potential candidate gene on a previous study conducted by our group on the intramuscular FA composition of the Landrace backcrossed population [3]. Apart from *PLIN1*, other genes involved in lipolysis and adipogenesis were *FBP1*, *G0S2* and *SOD3*, which were positively correlated with C18:1*n*-9 and negatively correlated with C18:2*n*-6, C20:3*n*-6, and C20:4*n*-6. Fructose-bisphosphatase 1 (*FBP1*) is a gluconeogenesis regulatory enzyme whose loss induces glycolysis and results in increased glucose uptake, among other cytosolic catalytic activities [53]. The porcine G0/G1 switch 2 (*G0S2*) expression level increased significantly during adipogenesis (both *in vitro* and *in vivo* studies) [54]. In the same study, *G0S2* was suggested to be a negative regulator of adipose triglyceride lipase (ATGL)-mediated lipolysis and cell proliferation in adipose tissue, as well as being closely related to lipid accumulation [54]. As an antioxidant enzyme, *SOD3* (superoxide dismutase 3) is secreted by the adipocytes and has the potential to prevent oxidative stress. In mice, Cui et al. [55] suggested that the specific function of *SOD3* in blocking adipogenesis and that the overexpression of *SOD3* suppressed the expression of pro-inflammatory genes in adipose tissue, and increased the expression of anti-inflammatory genes.

Altogether, the positive correlation of *FBP1*, *G0S2* and *SOD3* with the oleic acid, and their negative correlation with the three other PUFA may be due to their involvement in the lipid metabolism of the host, rather than on the metabolism of the essential FAs and their derivatives.

On the other hand, we also found that the patatin-like phospholipase domain containing 8 (*PNPLA8*) and glutamate oxaloacetate transaminase 1 (*GOT1*) genes were negatively correlated with C18:1*n*-9, while they were positively correlated with C18:2*n*-6 and C20:4*n*-6. *PNPLA8* belongs to a family of phospholipases that catalyze the cleavage of fatty acids from membrane phospholipids [56]. Furthermore, *PNPLA8*, also participates in mitochondrial lipid remodeling, and its absence causes mitochondrial dysfunctions and increased oxidative stress [57]. As another candidate for meat quality and carbohydrate metabolism, *GOT1* controls cellular metabolism through coordinating the utilization of carbohydrates and amino acids to meet nutrient requirements [58], but also this gene is crucial to provide oxaloacetate at low glucose levels, likely to maintain the redox homeostasis.

On the other hand, we also found that the patatin-like phospholipase domain containing 8 (*PNPLA8*) gene was the strongest negatively correlated gene with C18:1*n*-9, while it was positively correlated with C18:2*n*-6 and C20:4*n*-6. *PNPLA8* belongs to a family of phospholipases that catalyze the cleavage of FAs from membrane phospholipids [56]. Furthermore, *PNPLA8* is involved in mitochondrial lipid remodeling, and its absence causes mitochondrial dysfunctions and increased oxidative stress [57].

It is also worth mentioning the potential role in lipid metabolism of two transporter genes, *EXOC7* and *SLC44A2*. The exocyst complex component 7 (*EXOC7)* gene was positively correlated with C18:0. As *EXOC7* is a component of the exocyst complex, which regulates free FA uptake by adipocytes [59], it is involved in diverse biological functions, including promoting the translocation of glucose transporter *GLUT4* in the cell. Although there were other members of the solute carrier family, the solute carrier family 44 member 2 (*SLC44A2)* presented the same correlation pattern (positive with MUFA but negative with PUFA) as *FBP1*, *G0S2* and *SOD3*. In mice, *SLC44A2* was discovered to mediate choline transport into mitochondria, where it regulates synthesis of ATP, platelet activation and thrombosis [60]. In addition, supporting information for *SLC44A2* has suggested its important role for normal homeostasis [61].

Identification of putative regulators and their potential gene targets

The implementation of RIF based on the rCCA results is a useful and complementary approach to better understand the complex regulation mediated by TFs and their potential target genes. Comparing the expression profiles between FA conditions, we identified six interesting TF-regulators with less-known (*TADA2A* and *CARHSP1*) but also with well-known (*MAFA*, *SIX5*, *LBX1,* and *PAX7*) functions in livestock species. In each of them, potential target genes were also identified. Remarkably, we found TFs that accordingly to RIF1 (*TADA2A* and *CARSHP1*) and RIF2 (*KLF10* and *TADA2A*) were scored as the first and second most relevant TF. To be noted, several target genes of *TADA2A* and *CARSHP1* were detected as functional related to lipid metabolism (e.g., *PLIN1* and *TFRC*) and/or meat quality (e.g., *GOT1*). However, no target genes were *in silico* detected for *KLF10*.

Considering the correlation analysis, our findings in muscle suggests that transcriptional adaptor 2A *(TADA2A) was* linked to the 4 most interconnected FAs (positively associated with C20:4*n*-6, C18:2*n*-6, and C20:3*n*-6 respectively, and negatively associated with C18:1*n*-9). The protein encoded by *TADA2A* is part of the ATAC (Ada-Two-A-Containing) complex that interacts with the TATA-binding protein for transcriptional activation [62]. In addition, *TADA2A* was suggested to be involved in *de novo* hepatic lipogenesis in chicken fed with different diets [63]. On the other hand, calcium regulated heat stable protein 1 (*CARHSP1*) was positively linked to the sixth most interconnected FA (C16:1*n*-7). In mice, *CARHSP1* is regulated by the nutrient status in the liver and was suggested to inhibit hepatic gluconeogenic gene expression via repression of transcriptional activity of the *PPARα* transcription factor [64]. Consequently, both regulators were targets for each other in animals of the condition 2 (i.e., more SFA and MUFA, and less PUFA contents), as well as having shared target genes, such as *GOT1*, *PLIN1,* and *TFRC*. Taken together, these findings suggest some of the transcriptional circuits through which key regulatory genes exert their impact on their targets and FAs. Conversely, FAs may act as well as signaling candidates to regulate the transcription of target genes encoding proteins involved in muscle lipid metabolism [65].

Functional analysis of FA-correlated genes

Results from the functional analyses showed significantly over-represented GO terms related to carbohydrate and lipid metabolism. Among them, one of the most important was the citrate cycle (TCA cycle or Krebs cycle), which is an important aerobic pathway for the final steps of

the oxidation of carbohydrates, FAs, and amino acids [66], as well as this cycle provides precursors for many biosynthetic pathways. It is worth noting that the *LEP*, *MDH1* and *CYCS* genes were also enriched on the insulin signaling pathway, which can affect lipid metabolism [67].

As indicated by another metabolic pathway, the regulation of lipolysis in adipocytes highlights the potential role of certain candidate genes in skeletal muscle lipolysis (e.g., *PLIN1*), as well as FAs derived from intramuscular lipolysis (e.g., C16:1n-7 and C18:1n-9). In skeletal muscle, PLIN proteins are likely involved in the hydrolysis of triglycerides stored in lipid droplets and the passage of FAs to the mitochondria for their oxidation in the skeletal muscle [68]. Furthermore, a previous study in porcine adipocytes has shown that *PLIN1* is a novel candidate gene for IMF deposition and adipocyte differentiation [69]. Also taking into account the exemplified FAs, activation of adipocyte lipolysis by C16:1n-7 acid treatment has been demonstrated while C18:1n-9 acid was chosen as a control FA in investigations in mice [70]. The results of the functional and RIF analyses, highlighted the importance of certain target genes (e.g., *PLIN1*, *ADIPOQ*, *GOT1*, *TFRC*, or *CARHSP1*) which were found to be enriched in various biological processes related to lipid metabolism and RNA binding, respectively. For instance, *PLIN1* and *GOT1* belonged to pathways related with fat deposition (e.g., *PLIN1* in oxidation, accumulation, synthesis, concentration and homeostasis of lipids, and *GOT1* in homeostasis of lipids).

Conclusions

This study suggests the relevance of key FA phenotypes, but also promising FA-correlated genes, putative regulators, biological processes, and pathways related to intramuscular FA composition in pigs. Several of the candidate genes identified in this work were consistent with functions related to lipid and/or carbohydrate metabolism and meat quality. Likewise, the current findings indicate the potential impact of two particularly interesting regulators (*TADA2A* and *CARHSP1*) of the gene expression, as well as their modulatory effect on the FA composition.

Declarations

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All relevant data produced or evaluated in this research are disclosed in the paper as well as its supplementary information files. The RNA sequencing data used and analyzed in the current study are available from sequence read archive (SRA), NCBI BioProject under the accession number PRJNA882638 (https://www.ncbi.nlm.nih.gov/sra).

Competing interests

The authors declare that we have no competing interests.

_____

Authors' contributions

JMF conceived the study and was the principal investigator of the project, and collected the animal samples. LCM, DCP and AC performed the total RNA isolation. JMF and YRC designed the RNA-Seq experiment. YRC proposed the statistical methodology. JVH performed the bioinformatic and statistical analyses including parameter tuning of rCCA, all RIF step, manuscript scripts in R language, functional analysis, and classification of FA-interconnected genes in the network. DCP, MP, and CS participated in data processing and bioinformatic methods. YRC reviewed the statistical methodology and manuscript scripts. JVH, JMF, DCP and YRC conceived the structure of the paper. JVH, YRC, DCP, and JMF wrote the paper. All authors participated in the analysis and interpretation of the data, as well as critically revised and approved the final manuscript.

Acknowledgements

_____

References

1. Zhang J, Cui L, Ma J, Chen C, Yang B, Huang L. Transcriptome analyses reveal genes and pathways associated with fatty acid composition traits in pigs. Anim Genet. 2017;48:645–52.

2. Wood JD, Richardson RI, Nute GR, Fisher A V., Campo MM, Kasapidou E, et al. Effects of fatty acids on meat quality: a review. Meat Sci. 2003;66:21–2.

3. Puig-Oliveras A, Revilla M, Castelló A, Fernández AI, Folch JM, Ballester M. Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. Sci Rep. 2016;6:31803.

4. Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castelló A, Ballester M, et al. Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. BMC Genomics. 2012;13:1.

5. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Estellé J, Castelló A, Alves E, et al. Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. BMC Genomics. 2013;14:1.

6. Puig-Oliveras A, Ramayo-Caldas Y, Corominas J, Estellé J, Pérez-Montarelo D, Hudson NJ, et al. Differences in Muscle Transcriptome among Pigs Phenotypically Extreme for Fatty Acid Composition. PLoS One. 2014;9:e99720.

7. Valdés-Hernández J, Ramayo-Caldas Y, Passols M, Sebastià C, Criado-Mesas L, Crespo-Piazuelo D, et al. Global analysis of the association between pig muscle fatty acid composition and gene expression using RNA-Seq. Sci Rep. 2023;13:535.

8. Rohart F, Gautier B, Singh A, Lê Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. PLOS Comput Biol. 2017;13:e1005752.

9. Cesar ASM, Regitano LCA, Koltes JE, Fritz-Waters ER, Lanna DPD, Gasparin G, et al. Putative Regulatory Factors Associated with Intramuscular Fat Content. PLoS One. 2015;10:e0128350.

10. Reverter A, Hudson NJ, Nagaraj SH, Pérez-Enciso M, Dalrymple BP. Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. Bioinformatics. 2010;26:896–4.

_____

11. Martínez-Montes ÁM, Fernández A, Muñoz M, Noguera JL, Folch JM, Fernández AI. Using genome wide association studies to identify common QTL regions in three different genetic backgrounds based on Iberian pig breed. PLoS One. 2018;13:e0190184.

12. Mach N, Devant M, Díaz I, Font-Furnols M, Oliver MA, García JA, et al. Increasing the amount of n-3 fatty acid in meat from young Holstein bulls through nutrition. J Anim Sci. 2006;84:3039–48.

13. Crespo-Piazuelo D, Criado-Mesas L, Revilla M, Castelló A, Noguera JL, Fernández AI, et al. Identification of strong candidate genes for backfat and intramuscular fatty acid composition in three crosses based on the Iberian pig. Sci Rep. 2020;10:1.

14. Andrews S. Babraham Bioinformatics-FastQC A Quality Control tool for High Throughput Sequence Data. 2010.

15. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–48.

16. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15.

17. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:1.

18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;26:139–40.

19. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4:1184–91.

20. Lê Cao KA, González I, Déjean S. integrOmics: an R package to unravel relationships between two omics datasets. Bioinformatics. 2009;25:2855–56.

21. Ramayo-Caldas Y, Mármol-Sánchez E, Ballester M, Sánchez JP, González-Prendes R, Amills M, et al. Integrating genome-wide co-association and gene expression to identify putative regulators and predictors of feed efficiency in pigs. Genet Sel Evol. 2019;51:1.

22. Csardi, Gabor and Nepusz T. The igraph software package for complex network research. InterJournal. 2006.

23. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–04.

24. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32:2847–49.

25. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25:1091–93.

26. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995;57:289.

27. Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag; 2016.

28. Oliveira de Biagi CA, Nociti RP, Brotto DB, Funicheli BO, Cássia Ruy P de, Bianchi Ximenez JP, et al. CeTF: an R/Bioconductor package for transcription factor co-expression networks using regulatory impact factors (RIF) and partial correlation and information (PCIT) analysis. BMC Genomics. 2021;22:1.

29. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2020.

30. Guney E, Garcia-Garcia J, Oliva B. GUILDify: a web server for phenotypic characterization of genes through biological data integration and network-based prioritization algorithms. Bioinformatics. 2014;30:1789–90.

31. Neitzel J. Fatty Acid Molecules: A Role in Cell Signaling. Nat Educ. 2010;3:57.

32. Spector AA. Essentiality of fatty acids. Lipids. 1999;34:S1–3.

33. Cockcroft S. Mammalian lipids: structure, synthesis and function. Essays Biochem. 2021;65:813–45.

34. Wall R, Ross RP, Fitzgerald GF, Stanton C. Fatty acids from fish: the anti-inflammatory potential of long-chain omega-3 fatty acids. Nutr Rev. 2010;68:280–9.

35. O'Neill LM, Miyazaki M, Bond LM, Lewis SA, Ding F, Liu Z, et al. Fatty acid desaturation and elongation in mammals. Biochem Lipids, Lipoproteins Membr; 2021. p. 201–226.

36. Cameron ND, Enser M, Nute GR, Whittington FM, Penman JC, Fisken AC, et al. Genotype with nutrition interaction on fatty acid composition of intramuscular fat and the relationship with flavour of pig meat. Meat Sci. 2000;55:187–95.

37. Pena RN, Ros-Freixedes R, Tor M, Estany J. Molecular Sciences Genetic Marker Discovery in Complex Traits: A Field Example on Fat Content and Composition in Pigs. J Mol Sci. 2016;17: 2100.

38. Cao H, Gerhold K, Mayers JR, Wiest MM, Watkins SM, Hotamisligil GS. Identification of a Lipokine, a Lipid Hormone Linking Adipose Tissue to Systemic Metabolism. Cell. Cell Press; 2008;134:933–44.

39. Bickel PE, Tansey JT, Welte MA. PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. Biochim. Biophys. Acta. 2009. p. 419.

40. Zhang F, Chen Y, Heiman M, DiMarchi R. Leptin: Structure, Function and Biology. Vitam Horm. 2005;71:345–72.

41. Barb CR, Hausman GJ, Houseknecht KL. Biology of leptin in the pig. Domest Anim Endocrinol. 2001. p. 297–317.

42. Goodman HM. Hormonal Regulation of Fuel Metabolism. Basic Med Endocrinol. Academic Press; 2009. p. 151–174.

43. Pérez-Montarelo D, Fernández A, Folch JM, Pena RN, Ovilo C, Rodríguez C, et al. Joint effects of porcine leptin and leptin receptor polymorphisms on productivity and quality traits. Anim Genet. 2012;43:805–9.

44. Liu Y, Long H, Feng S, Ma T, Wang M, Niu L, et al. Trait correlated expression combined with eQTL and ASE analyses identified novel candidate genes affecting intramuscular fat. BMC Genomics. 2021;22:1.

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

46. Kang X, Zhao Y, Van Arsdell G, Nelson SF, Touma M. Ppp1r1b-lncRNA inhibits PRC2 at myogenic regulatory genes to promote cardiac and skeletal muscle development in mouse and human. RNA. 2020;26:481–1.

47. Ceder MM, Lekholm E, Hellsten S V., Perland E, Fredriksson R. The Neuronal and Peripheral Expressed Membrane-Bound UNC93A Respond to Nutrient Availability in Mice. Front Mol Neurosci. 2017;10:351.

48. Oyelami FO, Zhao Q, Xu Z, Zhang Z, Sun H, Zhang Z, et al. Haplotype Block Analysis Reveals Candidate Genes and QTLs for Meat Quality and Disease Resistance in Chinese Jiangquhai Pig Breed. Front Genet. 2020;11:752.

49. Karbowska J, Z K. Role of adiponectin in the regulation of carbohydrate and lipid metabolism. J Physiol Pharmacol. 2006;57:103.

50. Jakobsson A, Westerberg R, Jacobsson A. Fatty acid elongases in mammals: Their regulation and roles in metabolism. Prog Lipid Res. Pergamon; 2006;45:237–49.

51. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, et al. Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. PLoS One. 2013;8:e53687.

52. Jung Y, Kwon S, Ham S, Lee D, Park HH, Yamaoka Y, et al. Caenorhabditis elegans Lipin 1 moderates the lifespan-shortening effects of dietary glucose by maintaining ω-6 polyunsaturated fatty acids. Aging Cell. 2020;19:e13150.

53. Dong C, Yuan T, Wu Y, Wang Y, Fan TWM, Miriyala S, et al. Loss of FBP1 by Snail-Mediated Repression Provides Metabolic Advantages in Basal-like Breast Cancer. Cancer Cell. 2013;23:316–31.

54. Ahn J, Oh SA, Suh Y, Moeller SJ, Lee K. Porcine G0/G1 switch gene 2 (G0S2) expression is regulated during adipogenesis and short-term in-vivo nutritional interventions. Lipids. 2013;48:209–18.

55. Cui R, Gao M, Qu S, Liu D. Overexpression of superoxide dismutase 3 gene blocks high-fat diet-induced obesity, fatty liver and insulin resistance. Gene Ther. 2014;21:840–8.

56. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinforma. 2016;54:30.

57. Long JZ, Cravatt BF. The metabolic serine hydrolases and their functions in mammalian physiology and disease. Chem Rev. 2011;111:6022–63.

_____

58. Zhou X, Curbo S, Li F, Krishnan S, Karlsson A. Inhibition of glutamate oxaloacetate transaminase 1 in cancer cell lines results in altered metabolism with increased dependency of glucose. BMC Cancer. 2018;18:1.

59. Inoue M, Akama T, Jiang Y, Chun TH. The Exocyst Complex Regulates Free Fatty Acid Uptake by Adipocytes. PLoS One. 2015;10:e0120289.

60. Bennett JA, Mastrangelo MA, Ture SK, Smith CO, Loelius SG, Berg RA, et al. The choline transporter Slc44a2 controls platelet activation and thrombosis by regulating mitochondrial function. Nat Commun. 2020;11:3479.

61. Nair TS, Kommareddi PK, Galano MM, Miller DM, Kakaraparthi BN, Telian SA, et al. SLC44A2 single nucleotide polymorphisms, isoforms, and expression: Association with severity of Meniere's disease? Genomics. Academic Press; 2016;108:201–8.

62. Wang YL, Faiola F, Xu M, Pan S, Martinez E. Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. J Biol Chem. 2008;283:33808–15.

63. Desert C, Baéza E, Aite M, Boutin M, Le Cam A, Montfort J, et al. Multi-tissue transcriptomic study reveals the main role of liver in the chicken adaptive response to a switch in dietary energy source through the transcriptional regulation of lipogenesis. BMC Genomics. 2018;19:1.

64. Fan Y, Guo Y, Hamblin M, Chang L, Zhang J, Chen YE. Inhibition of gluconeogenic genes by calcium-regulated heat-stable protein 1 via repression of peroxisome proliferator-activated receptor α. J Biol Chem. 2011;286:40584–94.

65. Fritzen AM, Lundsgaard AM, Kiens B. Tuning fatty acid oxidation in skeletal muscle with dietary fat and exercise. Nat Rev Endocrinol. 2020;16:683–96.

66. Akram M. Citric Acid Cycle and Role of its Intermediates in Metabolism. Cell Biochem Biophys. 2014;68:475–8.

67. Consitt LA, Bell JA, Houmard JA. Intramuscular lipid metabolism, insulin action, and obesity. IUBMB Life. 2009;61:47.

68. MacPherson REK, Peters SJ. Piecing together the puzzle of perilipin proteins and skeletal muscle lipolysis. Appl Physiol Nutr Metab. 2015;40:641–51.

_____

69. Li B, Weng Q, Dong C, Zhang Z, Li R, Liu J, et al. A Key Gene, PLIN1, Can Affect Porcine Intramuscular Fat Content Based on Transcriptome Analysis. Genes. 2018;9:194.

70. Bolsoni-Lopes A, Festuccia WT, Farias TSM, Chimin P, Torres-Leal FL, Derogis PBM, et al. Palmitoleic acid (n-7) increases white adipocyte lipolysis and lipase content in a PPARα-dependent manner. Am J Physiol - Endocrinol Metab. 2013;305:E1093–102.

Figures

Figure 1 Correlation circle plot from the PCA applied to of the FA phenotypes and gene expression in muscle of BC1_DU pigs displayed from the first two rCCA dimensions (15 FAs and 365 genes selected in total).

Figure 2 Network plot for the _longissimus dorsi_ muscle study in BC1_DU pigs. Green and red edges indicate positive and negative correlation.

Legend: Output obtained for the first three rCCA dimensions (13 FAs and 365 genes were selected), showing the correlation structure for all bipartite relationships with a correlation cutoff of 0.29. Color legend: FAs: magenta = SFA members; royal blue = MUFA members; orange red = PUFA members; and genes: darkorange = enzyme; aquamarine = adipokine; chartreuse = TF; turquoise = TF cofactors; yellow = lipid metabolism-related genes; navajowhite = carbohydrate metabolism; crimson = glycolysis; gold = transporter; lightpink = fatty acid beta-oxidation; coral = amino acid metabolism; cornsilk = receptor family; deepskyblue = nucleic acid metabolism. Out of a total of 365 genes, the 176 colored genes were referring to a functional or gene family classification (while 189 genes were unclassified).

Figure 3 Functional analysis with the correlated genes from the rCCA approach that were significantly enriched in GO terms according to metabolic pathway delimitation.

Legend: This output is a representation of the original results table generated with the ClueGO plugin in Cytoscape software.

_____

Figure 4 Principal component analysis (PCA) summarizing the separation and similarities among FA conditions with the extreme values.

Legend: PC1 and PC2 explained 77.40% of the total variance. The visualization of the eigenvectors was implemented to delimit the FAs belonging to their respective group (SFA, MUFA, PUFA).

Additional files

Additional file 1 Table S1

Format: Excel

Title: Manual list with functional annotation and plausible function in different aspects of FA, lipid, and carbohydrate metabolism.

Description: This list was compiled based in ClueGO results for selected genes from rCCA.

Additional file 2 Table S2

Format: Excel

Title: Trained list containing genes associated with lipid and/or carbohydrate metabolism in lipogenic tissues (adipose, liver and muscle-skeletal) was obtained from GUILDify tool.

Description: This procedure uses BIANA knowledge base to create a species-specific interaction network for each gene detected. Here, the netcombo prioritization algorithm based on network topology, and the highest guild score for the top 100 gene products (with only seed) were considered to compose the trained list.

Additional file 3 Table S3

Format: Excel

Title: List of all transcription factors (TFs) and TF cofactors for sus scrofa specie obtained from AnimalTFDB v3.0 database.

Description: This resource is also available at: http://bioinfo.life.hust.edu.cn/AnimalTFDB/#!/.

_____

Additional file 4 Table S4

Format: Excel

Title: The correlation matrix for all bipartite relationships (13 FAs and 365 genes) obtained via rCCA approach with a correlation above 0.29 using the network function in the mixOmics package.

Description: NAs values in this table mean that the correlation between both variables did not exceed the pre-established cut-off point.

Additional file 5 Figure S1

Format: png

Title: Correlation circle plot from the PCA applied to the FA phenotypes and gene expression in muscle of BC1_DU pigs displayed the first versus third rCCA dimensions.

Description: This output was obtained using the plotVar function of the mixOmics package.

Additional file 6 Figure S2

Format: png

Title: Heatmap displaying the correlation structure of the rCCA in the *longissimus dorsi* muscle from BC1_DU pigs.

Description: This output was obtained using the functions and dependencies of the ComplexHeatmap package. All bipartite relationships between FAs and gene expression variables (cutoff of 0.29) are shown, including hierarchical clusters for both variables. Heatmap with the rCCA variables (365 genes and 13 FAs selected in total) was computed. To complement the network plot, heatmap was used. The color key indicates positive (green) and negative (red) correlation.

Additional file 7 Figure S3

Format: png

Title: Density distribution of correlation between FA and gene expression profiles, including five quantile levels and mean value in the *longissimus dorsi* muscle from BC1_DU pigs.

Description: Distribution of correlations (FA vs. gene expression) as Density heatmap using *densityHeatmap* function. Here, the density was calculated by columns from input data passed as a list item.

Additional file 8 Figure S4

Format: png

Title: Functional analysis with correlated genes from rCCA that were significantly enriched in GO terms according to biological processes delimitation.

Description: This output is a representation of the original results table generated with the ClueGO plugin in Cytoscape software.

Additional file 9 Table S5

Format: Excel

Title: ClueGO Gene Ontology (GO) results of selected genes in the rCCA approach.

Description: Functional analysis included significant GO terms with and without "GO Term Fusion" (biological process, pathways and molecular functions, BH-corrected P-value < 0.05) and was performed with the background of all genes expressed in muscle.

Additional file 10 Table S6

Format: Excel

Title: Mean comparison between the two FA conditions with the extreme values obtained through classification using PCA (see details in Figure 4).

Description: This analysis included the FAs profile (n = 60 animals) with the 13 phenotypes selected in the rCCA approach. Significant differences (P-value < 0.05) between FA conditions were determined using a t-test, and the SEM = standard error of the mean was calculated.

Additional file 11 Table S7

_____

Format: Excel

Title: Results of the co-expression correlation analysis between the TF (n = 19) and rCCA-derived genes (n = 343) in each experimental condition.

Description: This output was obtained using the *runAnalysis* function of the CeTF package.

Additional file 12 Table S8

Format: Excel

Title: Results of the correlation differential (alias DW) written as the correlation of the TF with all possible target genes in conditions 1 and 2, respectively.

Description: This output was obtained using the functions of the CeTF package.

Additional file 13 Table S9

Format: Excel

Title: Results of the distribution of DE genes and specific TF in both conditions, as well as the relationship between log(baseMean) and expression difference.

Description: This output was obtained using the functions of the CeTF package.

Chapter 4. General Discussion

The links between fatty acid (FA) profile and gene expression in muscle has received special attention in recent years, mainly due to its effects on meat quality parameters and human health. Skeletal muscle is an extremely complex organ, which consists of approximately 90% muscle fibers and 10% of connective and fat tissues (Listrat et al., 2016). Muscle is also a highly metabolically active tissue, that both stores (e.g., of metabolites) and consumes energy (Liu et al., 2015). Moreover, muscle is also controlling many physiological processes and structural and metabolic features. Since skeletal muscle is one of the primary regulators of lipid and glucose metabolisms, it is highly susceptible to changes in glucose and FA availability (Morales et al., 2017). Muscle has an intrinsic ability to change its mass and phenotype in response to activity, and these processes may involve both quantitative and qualitative changes in gene expression (Goldspink, 2002).

Based on the intramuscular FA content in the *longissimus dorsi* (LD) muscle of pigs from an Iberian $\times$ Duroc backcross (called BC1_DU), the present discussion emphasizes the integration of the results from the three studies: (1) outlining the rationale summary for the relationships between gene expression levels and FA composition traits, and (2) highlighting the important FA phenotypes, key regulators, candidate genes, biological processes and metabolic pathways modulating the intramuscular FA profile of the LD muscle.

4.1. Identification of differentially expressed genes and polymorphisms related to oleic-to-stearic acid ratio

We present results evaluating the muscle transcriptomic profiles of divergent animals (16 high and 16 low, H and L groups) for the intramuscular oleic-to-stearic acid (or C18:1n-9/C18:0) ratio. The overall objective was to identify DE genes and polymorphisms related to C18:1n-9/C18:0 phenotype. Results from DESeq2 method (Love et al., 2014), reports a total of 81 DE genes (57 down-regulated and 24 up-regulated DE genes). The functional analyses of DE genes indicate that mainly PPAR signaling pathway was enriched among the DE genes.

In general, the literature search of the 81 DE genes revealed that ten genes of them (*ACAD10, ACADVL, ECHDC3, FABP3, FADS3, ILVBL, MMAA, PLIN1, PPARG,* and *SCD*) are well-

145

known for their role in energy, lipid and carbohydrate metabolism, and mitochondrial fatty acid beta-oxidation; which are important pathways to modulate the FA composition. Three of these genes (*ACADVL, PLIN1*, and *SCD*) and four other DEGs (*DZANK1, GOT1, PANK1,* and *PPP1R1B*) were previously detected as DE genes in the LD muscle transcriptome of phenotypically extreme animals (H vs. L groups, only females) for intramuscular FA composition from an Iberian × Landrace backcross population (BC1_LD) (Puig-Oliveras et al., 2014). In addition, we found that the most significant DE gene (i.e., *HGFAC*, P-adj = 4.34E-07) has been reported as a ChREBP-regulated hepatokine in mice and humans. *HGFAC* codes for a protease that activates the pleiotropic hormone hepatocyte growth factor (Sargsyan et al., 2023). Likewise, this gene enhances lipid and glucose homeostasis, mediated in part by the activation of hepatic peroxisome proliferator-activated receptor gamma (*PPARG*).

4.1.1. Differentially expressed genes between high and low groups

Among the up-regulated genes in high group animals we detected the stearoyl-CoA desaturase (*SCD*), a regulatory enzyme catalyzing the rate-limiting step in the overall *de novo* synthesis of C18:1n-9 and C16:1n-7 acids (Ntambi & Miyazaki, 2004). Both FAs are produced by the Δ9-desaturation of their respective precursors (i.e., C18:0 or stearic acid and C16:0 or palmitic acid, respectively) (Guillou et al., 2010). A higher *SCD* gene expression will likely increase the desaturation ratio values (C18:1n-9/C18:0, C16:1n-7/C16:0, C18:1n-9+C16:1n-7/C18:0+C16:0, and MUFA/SFA) in muscle in high group animals. Regarding these traits, we also investigated the correlations with the expression of *SCD* in muscle. Moderate correlation was observed with values of r = 0.65 for C18:1n-9/C18:0, r = 0.68 for C16:1n-7/C16:0, and r = 0.54 for C18:1n-9+C16:1n-7/C18:0+C16:0.

Several TF binding sites were found in the pig *SCD* gene promoter (Estany et al., 2014), including the PUFA response element (PUFA-RE), the sterol regulatory element (SRE) where the SREBP-1C transcription factor binds, the *PPARG* motif, and the retinoic receptor and the retinoic acid receptor α (RXR:RARα) binding elements, suggesting a fine regulation of its expression. In our RNA-Seq study, the *PPARG* gene expression was increased in the high group. *PPARG* is a master regulator of adipogenesis (Lee & Ge, 2014) and its expression was correlated with the expression of lipid metabolism genes in backfat (Revilla et al., 2018). As a member of peroxisome proliferator-activated receptors (PPARs), it controls the mRNA

_____

expression of genes involved in lipid metabolism, including the induction of the *SCD* gene expression (Miller & Ntambi, 1996). Hence, a higher expression of *PPARG* in high group animals was associated with higher *SCD* levels in pig muscle, but we cannot conclude that there is a direct effect. Moreover, PUFA has also been reported to repress the induction of lipogenic genes (including *SCD*) (Liang et al., 2017; Mater et al., 1999; Nakamura et al., 2004). This modulation occurs, for example, by binding to the PUFA-RE element and suppression of SREBP-1c gene transcription and proteolytic processing. In fact, we hypothesized that the lower PUFA content in the muscle of high group animals ($P = 7.07E-05$) might repress the *PPARG* gene expression (r $_{\text{PUFA-PPARG}}$ = -0.52, $P = 2.01E-03$), and consequently modulate the *SCD* gene expression in LD muscle.

Another attractive DE gene between the two groups of animals was the perilipin 1 gene (*PLIN1*), which showed overexpression in the high group. *PLIN1* is a member of the PAT family of proteins that are able to bind lipid droplets (Bickel et al., 2009), and plays a role in regulating intracellular lipid storage and mobilization, controlling lipid homeostasis. *PLIN1* has been identified among the LD muscle overexpressed genes in pigs with a high intramuscular fat (IMF) content, both in Iberian purebred pigs (Muñoz et al., 2018) and in commercial hybrid pigs (Li et al., 2018). Similarly, high group animals, with higher *PLIN1* expression, showed a higher IMF content in LD muscle ($P = 2.59E-03$).

Furthermore, we found that the fatty acid desaturase 3 gene (*FADS3*) was overexpressed in the high group animals. *FADS3* is a member of the FADS family of fatty acid desaturases, including *FADS1* and *FADS2* genes, clustered together on porcine chromosome 2 (Taniguchi et al., 2015). *FADS3* function is unknown, but its high sequence homology to functional *FADS1* and *FADS2* suggest that it is involved in long chain PUFA biosynthesis (Park et al., 2009). In addition, the expression of *FADS3* has been shown to be up-regulated markedly in mice lacking *FADS2* (Stroud et al., 2009), which suggested that *FADS3* is a homologous gene to *FADS2*.

As another example, our results also showed an overexpression of the nerve growth factor receptor (*NGFR*) in high group animals. *NGFR* gene plays a role in the regulation of the translocation of *GLUT4* to the cell surface in adipocytes and muscle cells in response to insulin (Liang et al., 2017).

_____

Among the down-regulated genes in high vs. low groups, the fatty acid binding protein 3 (*FABP3*) showed a lower expression. The *FABP3*, also known as H-FABP (heart-type cytoplasmic fatty acid-binding protein), belongs to a family of FA binding proteins which are involved in intracellular transport of long-chain FAs and the regulation of FA uptake (Chmurzyńska, 2006; Glatz et al., 2003). As a member, *FABP3* is expressed in a large number of tissues, including heart, skeletal muscle and adipose tissue among others, but also it has been reported to be involved in FAs uptake and transport to mitochondria for β-oxidation system in muscle (Furuhashi & Hotamisligil, 2008). Furthermore, all FABPs bind long-chain FAs with differences in ligand selectivity. Thus, FABPs binding FAs can potentially enter the nucleus to target members of the PPARs family (including *PPARG*) of transcription factors (Furuhashi & Hotamisligil, 2008).

4.1.2. Verification of RNA-Seq results with RT-qPCR

For three protein coding genes with an important role in lipid metabolism (*FABP3*, *PPARG*, and *SCD*), a comparison of the RNA-Seq results with the RT-qPCR via Fluidigm system (n = 32 extreme animals) (Criado-Mesas et al., 2020) was carried out. There was a good concordance between the expression values measured by RNA-Seq (log_2CPM) and the RT-qPCR (log_2NQ) methods, showing significant correlations (P < 0.05) of 0.776 for *FABP3*, *0.693* for *PPARG*, and 0.834 for *SCD* genes. Since *PPARG* is a transcription factor binding to the *SCD* promoter (Estany et al., 2014), the correlation between gene expression of both genes was screened. To this particular, our findings confirmed by another methodology (i.e., RT-qPCR) suggested a transcriptional regulation of *SCD* by the *PPARG* (r = 0.75, $P = 9.295$E-07). A similar correlation value (r = 0.78, $P < 1.00 \times 10^{-16}$) was reported in a larger study by real-time qPCR in an Iberian × Landrace backcross (Puig-Oliveras et al., 2016).

4.1.3. Analysis of single nucleotide polymorphisms (SNPs) located in the *SCD* gene

To identify putative SNPs, ten polymorphisms were genotyped and analyzed in depth for the *SCD* gene. The SNP rs80912566 (AY487830:g.2228T>C SNP) is located in the promoter region of the *SCD* gene in purebred Duroc pigs (Estany et al., 2014) was included in our list. We considered other nine additional polymorphisms located in the promoter, 5'UTR and 3'UTR regions of the *SCD* gene. The main results indicated that animals of the high group had

_____

a higher allelic frequency of the T allele of the rs80912566 SNP (0.91) in comparison to low group animals (0.59). Interestingly, the haplotype reconstruction reveals that seven genetic variants within the *SCD* were distributed in two haplotypes. Haplotype H1 (C-C-C-T-A-C-G) contains the T allele of the rs80912566 SNP, while haplotype H2 (A-G-T-C-G-T-C) has the alternative C allele. Among the high animals 13 were homozygous for haplotype H1 (H1H1) and three heterozygous (H1H2). Conversely, among low animals four were homozygous for haplotype (H1H1), 11 heterozygous (H1H2) and one homozygous for haplotype H2 (H2H2). Hence, high animals showed a higher frequency of the H1H1 genotype than low animals. However, our results also showed the haplotype is not explaining completely the classification in high and low animals, nor the *SCD* expression or the C18:1*n*-9/C18:0 ratio. These results suggested that other additional factors or other causal variants in linkage disequilibrium with rs80912566 may be explaining the IMF C18:1*n*-9/C18:0 ratio and *SCD* expression differences. Hence, further studies will be required to validate the causal mutation in the BC1_DU population, as well as in other pig populations.

4.1.4. Gene ontology and pathways analysis

Functional analysis with the 81 DEGs indicated that two gene ontology (GO) term and one KEGG pathway were significantly overrepresented. Overall, the main GO term detected was the PPAR signaling pathway containing four genes related to lipid metabolism (*FABP3, PLIN1, PPARG,* and *SCD*). In concordance, the peroxisome proliferator-activated receptor (PPAR) signaling pathway was significantly overrepresented in the muscle transcriptome analysis of pigs (Iberian x Landrace backcross) with divergent muscle FA composition (Puig-Oliveras et al., 2014).

4.2. Global and integrative analysis of gene expression and fatty acid traits

Overall, our results illustrate an assessment of the gene expression and fatty acid (FA) profiles of pig muscle in the context of multifactorial design using a combination of univariate and multivariate analyses. Collectively, we showed that expression was significantly associated with FA composition (second paper) using the univariate ELMSeq approach (Liu et al., 2018). Notwithstanding, only 53.52% of the gene expression values (547/1022 FA-associated genes) significantly correlated (i.e., Pearson's correlation coefficient (r) ranged from low to moderate

_____

(r [-0.19 to 0.51]) with the phenotypic values of representative FAs (i.e., 21 of 36 FA traits). In fact, these correlation values were similar to those reported in independent studies on muscle, such as: gene expression vs. FA composition (Wang et al., 2013; Puig-Oliveras et al., 2016; Zhang et al., 2017), or gene expression vs. IMF content (Wang et al., 2020; Liu et al., 2021). Subsequently, in the third paper we performed a multivariate integrative analysis via rCCA approach (Lê et al., 2009), and with them we showed a bipartite relationship between the intramuscular FA and gene expression profiles. We highlight a subset of canonical variables (13 FAs and 365 genes) that maximize the correlation between both datasets (low to moderate range, r [-0.39 to 0.41]) was identified by applying a pre-established cutoff (above 0.29) (Ramayo-Caldas et al., 2019). As an intriguing result, FAs C20:2n-6, and C20:3n-3 showed only negatively correlated genes.

4.2.1. Overlap between ELMSeq and rCCA results

Broadly, the intersection between the set of genes identified with the univariate (second study) and multivariate approach (third paper) is evaluated below.

6.2.1.1. Summary of metabolic enzymes

Taken together, both approaches yielded a set of candidate genes coding for metabolic enzymes (ELMSeq= 118 and rCCA=30), but only nine out of theses enzymes (*GOT1*, *PDHA1*, *MDH1*, *IDH3A*, *ALAD*, *HMGCR*, *ATP6V1C1*, *G0S2*, and *FBP1*) were commonly detected in the two studies. Moreover, specific enzyme activities with lipolytic activity such as *LPL* and *ACSL1* (ELMSeq), and lipogenic activity such as *ELOVL6* and *SC5D* (ELMSeq), and *SOD3* (rCCA) were detected only by one approach.

6.2.1.2. Genes related to carbohydrate and lipid metabolism

We also observed overlapping of 30 lipid and/or carbohydrate metabolism genes (out of a total of ELMSeq=146 and rCCA=49). Among them, we highlight *SLC44A2, LPIN1, PNPLA8, HMGCR, ARSA, MED1, PLIN1, PIK3C2G, INPPL1, G0S2, PPP1CB, ATP5F1A, IDH3A, PDHA1, ATP5MC3, CYCS, SDHD, NDUFS1, RPN2,* and *POMGNT1*. Of note, adipokines

_____

(*LEP*, *ADIPOQ*, *NAMPT, and SFRP5*) were found, being *ADIPOQ* and *NAMPT* specifically detected by the rCCA method. *LEP*, *ADIPOQ* and *NAMPT* genes participate in the adipogenesis pathway, while *SFRP5* gene is also known as an anti-inflammatory adipokine (Wang et al., 2020). In addition, the *CHUK* cytokine gene was commonly reported by the two methods, which plays an essential role in the negative feedback of NF-kappa-B (NFKB) canonical signaling to limit inflammatory gene activation.

6.2.1.3. Fatty acid and glucose transport genes

Regarding genes involved in transport of molecules, three members of the solute carrier (SLC) gene superfamily (*SLC31A1*, *SLC44A2*, and *SLC35G1*) were highlighted. Members of SLC superfamily encode membrane-bound transporters, which play essential roles in transporting a variety of substrates across cellular membranes, including amino acids, glucose, inorganic cations and anions, FAs and lipids, acetyl coenzyme A, and vitamins, among others (He et al., 2009). Meanwhile, for glucose transport, the *TRARG1* [alias trafficking regulator of GLUT4 (SLC2A4) 1] gene was only detected by ELMSeq approach presenting numerous associations with MUFA and PUFA traits, and MUFA/PUFA, MUFA/SFA ratios. *TRARG1* gene has been predicted to be involved in endosome to plasma membrane protein transport and glucose import in response to insulin stimulus (NCBI Gene ID: 286753). Moreover, a component of the exocyst complex (like *EXOC7*) was detected by the two methods. *EXOC7* gene was found to be associated with C18:0 and/or C20:1n-9, respectively. The exocyst complex is known to regulate the free FA (FFA) uptake by adipocytes, and the trafficking/docking of insulin-sensitive glucose transporter type 4 (*GLUT4*) (Inoue et al., 2003, 2015).

6.2.1.4. Genes involved in meat quality

On the other hand, overlap for genes involved in meat quality was observed including six candidate genes (*MDH1*, *GOT1*, *NEU3*, *GALNT15*, *GALNT17*, and *PDHA1*). Three of them, *MDH1*, *GOT1* and *PDHA1* were detected by both methods, while *NEU3* and *GALNT15*, was detected only by the ELMSeq approach and *GALNT17* only with rCCA.

_____

6.2.1.5. Association of fatty acid with long non-coding RNAs

Unlike RNAs that are translated into proteins (mRNAs), there are other types of RNAs such as the long non-coding RNAs (lncRNAs). lncRNAs play important roles in biological processes such as chromatin remodeling, transcriptional activation, transcriptional interference, RNA processing, and mRNA translation (Zhang et al., 2019). They also regulate gene expression at the epigenetic, transcriptional, post-transcriptional, levels in a variety of ways. Results based on ELMSeq indicated the presence of three lncRNAs (*ENSSSCG00000015889* alias *TANK*, *ENSSSCG00000011196* or previously known as *DPH3*, and *ENSSSCG00000038429*) significantly associated with FA composition. For instance, *TANK* was negatively correlated with C18:1n-9 (r = -0.26, *P*-value = 3.024E-03) but positively with C18:2n-6/C18:3n-3 (r = 0.30, *P*-value = 6.517E-04). Moreover, C18:2n-6/C18:3n-3 showed a positive correlation with *ENSSSCG00000011196* (r = 0.28, *P*-value = 1.286E-03) and *ENSSSCG00000038429* (r = 0.31, *P*-value = 3.817E-04). Meanwhile, ω6/ω3 ratio was positively correlated (r = 0.28, *P*-value = 1.562E-03) with *ENSSSCG00000038429*. However, only *TANK* is annotated in the current version of Sscrofa11.1 (i.e., Pig Ensembl Genes 109). The other members need to be categorized or named in terms of gene annotation. To the best of our knowledge, the effects of these lncRNAs on FA phenotypes have not yet been established in the literature. *TANK* in human cells has been reported as a negative regulator of NF-kappaB signaling activation upon DNA damage, as well as on cell survival and, consequently on increased inflammatory cytokine production (Wang et al., 2015). Therefore, it appears that our identification of both *TANK* lncRNA and *CHUK* cytokine may suggest a parallel activity negatively regulating the NFKB signaling cascade.

6.2.1.6. Putative regulators affecting fatty acid composition, including transcription factor and co-factors

When focused on putative regulators, a cross-referencing of ELMSeq list (n = 62 TFs) against the rCCA detection (n = 22 TFs) indicated 11 commonly reported TFs (*HBP1, HOXB6, LBX1, MAFA, SIX5, TADA2A, TFAM, YBX2, ZNF146, ZNF326,* and *ZNF407*). Of note, the detection of TFs by ELMSeq predominated in PUFA traits such as ω6/ω3, C18:2n-6/C18:3n-3, C20:2n-6 and C20:1n-9 MUFA, whereas by rCCA the highest number of TFs were detected in C18:2n-

_____

6, C20:4n-6, and C20:3n-6 PUFA and C18.1n-9 MUFA. We complemented this step using the RIF approach (Oliveira de Biagi et al., 2021; Reverter et al., 2010) on rCCA outputs, identifying six DE regulators (TF genes) between FA conditions, including TFs (*LBX1, MAFA, SIX5,* and *TADA2A*) shared with ELMSeq approach as well as the remaining TFs (*CARHSP1* and *PAX7*) not detected by the two methods. RIF suggested a relevant role of *TADA2A* and *CARHSP1* with extreme value for the RIF metrics, being scored as the first and second most relevant TF according to RIF1 and RIF2, respectively. The information from RIF1 and RIF2 is complementary. RIF1 score classified the TFs most differentially co-expressed with the highly abundant and highly DE genes, whereas RIF2 score classified those TF which are the best predictors of the abundance of DE genes. Notwithstanding, our data also provide additional evidence of *TADA2A* and *CARHSP1* as novel regulators in pigs with a pivotal role in gene expression and, consequently, in the modulation of FA traits. In fact, both regulators were linked with the six most interconnected FAs. For instance, *TADA2A* was positively associated with C20:4n-6, C18:2n-6, and C20:3n-6 but negatively related with C18:1n-9; while *CARHSP1* was positively linked with the C16:1n-7 lipokine.

Furthermore, the *in silico* prediction of TF-binding sites suggest some shared target genes related to lipid metabolism (e.g., *PLIN1*, *TFRC*, and *GOT1*). In line with biological relevance, the protein encoded by *TADA2A* is part of the ATAC (Ada-Two-A-Containing) complex that interacts with the TATA-binding protein for transcriptional activation (Wang et al., 2008). In addition, *TADA2A* was suggested to be involved in *de novo* hepatic lipogenesis in chicken fed with different diets (Desert et al., 2018). On the other hand, *CARHSP1* is regulated by the nutrient status in the liver and was suggested to inhibit hepatic gluconeogenic gene expression via repression of transcriptional activity of the PPARα transcription factor in mice (Fan et al., 2011). Taken together, these findings suggest a putative transcriptional circuits through which key regulatory genes may exert their impact on their targets and FAs.

Considering the collection of TF co-factors, the identification of the two different approaches (ELMSeq: 77 co-factors versus rCCA: 22 co-factors) revealed a total of 14 overlapping. These 14 co-factors are involved in the regulation of transcription (*CCDC85B*, *CHUK*, *EDF1*, *GNL3*, *GTF2B*, *MED1*, *PCGF2*, *SNW1*, *TCEA1*, and *VGLL2*), in adipogenesis (*LGALS12* and *LPIN1*), in gluconeogenesis and glycolysis (*FBP1*), and also in epigenetic regulation of myogenesis (*HDGFL2*). Our data indicated that lipin 1 (*LPIN1*) was positively correlated with the

abundance of C20:4n-6 and ω6/ω3 ratio. This gene participates in the metabolism of the arachidonic acid in *Caenorhabditis elegans* organism (Jung et al., 2020). Likewise, *LPIN1* was identified as a potential candidate gene in a study conducted by our group on the intramuscular FA composition of Iberian × Landrace backcross population (Puig-Oliveras et al., 2016). Notably, although the galectin-related inhibitor of proliferation (*LGALS12*) was exclusively associated with C16:1n-7 via rCCA (positively correlated), this gene simultaneously showed positive associations with six FAs via ELMSeq (including e.g., with C16:1n-7), while it was negatively correlated with C20:2n-6. In addition, in our GO analysis the *LGALS12* gene was assigned to the functional categories related to signal transduction and regulation of cellular processes. In fact, Wu et al. (2018) have reported that *LGALS12* knockdown could inhibit adipogenesis in porcine adipocytes by downregulating lipogenic genes and activating the PKA–Erk1/2 signaling pathway.

Another interesting co-factor, detected by ELMSeq method, was the nuclear receptor coactivator 2 (*NCOA2*), which showed a positive correlation with the ω6/ω3 ratio. *NCOA2* encodes a protein that acts as a transcriptional coactivator for nuclear hormone receptors, including steroid, thyroid, retinoid, and vitamin D. In fact, in BC1_LD pigs a key role of *NCOA2* as modulator of the intramuscular FA composition and control energy homeostasis has been reported (Ramayo-Caldas et al., 2014), as well as it was reported among lipid-related genes that exhibited differential gene expression between sexes (Puig-Oliveras et al., 2016).

6.2.1.7. Association of PUFA ratios with gene expression

ELMSeq reports links between PUFA ratios including ω6/ω3 and C18:2n-6/C18:3n-3 traits, which share 57.32% of genes between them, including also a sizeable fraction of enzymes and TFs. Arguably, this considerable degree of overlap could be explained because the ω6/ω3 ratio contains C18:2n-6 and C18:3n-3 as major FAs in its numerator and denominator. In both ratios we identified 10 common TFs (of 24 and 18 found). However, with the exception of *ESRRA, IRF2, HBP1, TSC22D3,* and *MECP2* TFs; not much is known about the functions and the transcriptional regulatory relationships of *ZNF407, ZSCAN20, HOXB5, YBX2,* and *NFE2L3* TFs. We highlight the significant association of estrogen related receptor alpha (*ESRRA*) gene with PUFA ratios. *ESRRA* was one of the candidate genes for backfat found in a region on SSC2 related to the oxidation of different FAs in different backcrossed pigs (including BC1_DU

_____

animals) (Crespo-Piazuelo et al., 2020). ERR alpha is an orphan member of the superfamily of nuclear hormone receptors, which has been reported as a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene (*ACADM*) (Sladek et al., 1997). On the other hand, concerning the enzyme activity underlying the metabolism of specific PUFA ratios, 41 enzymes (of 77 and 55 found) were detected in the ω6/ω3 and C18:2*n*-6/C18:3*n*-3 traits. Among these enzymes, we highlight (*ATP5F1A, PIK3C3, PRPF19, AACS, ACSL1, SC5D, GALC,* and *HMOX1*) for which their relevance in lipid metabolism has been documented.

4.2.2. Methodological assessment of the assumptions and features of ELMSeq and rCCA approaches

We have already described the degree of overlap and the differences in the results obtained by the ELMSeq and rCCA approaches. However, it is important to highlight some methodological differences between them. The algorithm known as ELMSeq computes a single gene-by-gene regression analysis, the assumptions of our model include the normality of the response variable, but also that the gene expression data is log-normally distributed. Then, it automatically performs inter-sample normalization, and thus can handle e.g., batch effects and variation in sequencing depths of different samples. In addition to features of the model: (1) sample-specific normalization are modeled as in the gene-wise linear models and jointly estimated with the regression coefficients, and (2) by imposing sparsity-inducing L1 penalty on the regression coefficients. As an advantage, ELMSeq allows to test the association of gene expression with a particular variable, while adjusting for all other covariates.

On the other hand, the rCCA is an unsupervised multivariate exploratory approach that integrates two datasets (X and Y, respectively of sizes n × p and n × q). This method is known as the regularized variant of the primary canonical-correlation analysis (rCCA). rCCA performs a dimension reduction in each dataset with the final goal of selecting the variables that maximize the correlation between the input matrices (FA and gene-expression). In our case we used as input data FA composition (measured by gas chromatography) and muscle transcriptomic data (estimated by RNA-Seq) (X and Y, respectively). It is important to note that prior to the analysis, it is assumed that the researcher has appropriately normalized and pre-processed the data sets when applicable. We introduce both matrices in continuous scale, the X matrix in relative percentages of FAs and the Y matrix in log_2CPM. Consequently, the issue of

_____

high dimensionality can be by-passed by introducing regularization into the CCA method. Meanwhile, the optimal regularization parameters ($\lambda 1$and $\lambda 2$) were estimated and added to the diagonal of X and Y respectively to make them invertible. Among the regularization methods, we used the shrinkage approach, which allows an analytical calculation of the regularization parameters for large-scale correlation matrices. Although rCCA has the advantage of a reasonable computation time (about 15-20 minutes), a shortcoming of the shrinkage approach is that $\lambda 1$and $\lambda 2$ values are calculated independently, i.e., they are not manually tuned as in the case of the cross-correlation approach. However, we opted for the shrinkage approach because of the experience that the cross-validation approach does not converge with data containing more than 10K predictor variables. Also as positive control, we have tested with a smaller data set "Case Study: rCCA Nutrimouse" that both methods provide similar results (http://mixomics.org/case-studies/rcca-nutrimouse-case-study/).

_____

Chapter 5. Conclusions

1. We identified 81 differentially expressed genes and with an over-representation of the PPAR signaling pathway from the LD muscle transcriptome analysis of pigs with extreme values of the intramuscular C18:1n-9/C18:0 ratio.

2. Two haplotypes of seven SNPs located in the *SCD* gene were associated with the C18:1n-9/C18:0 ratio variation and the *SCD* gene expression in muscle. However, additional factors or other genetic variants are likely involved in the observed differences in the desaturation ratio.

3. The univariate association analysis identified several candidate genes linked to intramuscular FA composition in muscle, including well-known (e.g., *ACSL1, ELOVL6, ESRRA, FBP1, GOT1, LEP, LGALS12, LPL, MDH1, NCOA2, PLIN1, SC5D,* and *TFRC*) and novel candidate genes (e.g., *TRARG1, TANK, ENSSSCG00000011196,* and *ENSSSCG00000038429*).

4. A multivariate analysis pinpointed a subset of 378 canonical variables (13 FAs and 365 genes) maximizing the correlation between FAs and expression datasets in pig muscle. The FA-correlated genes were mainly related to lipid and/or carbohydrate metabolism, meat quality, and transport of molecules.

5. Our results derived from integrative and regulatory impact factor analyses showed the relevance of six transcription factors (*CARHSP1, LBX1, MAFA, PAX7, SIX5*, and *TADA2A*) as putative regulators of gene expression and intramuscular FA composition.

Chapter 6. References

Abdellaoui, A., Conor V.D., Karin J. H.V., and Michel G.N. (2022). "Gene–environment correlations across geographic regions affect genome-wide association studies". *Nature Genetics*, 54(9):1345–1354.

Albuquerque, A., Óvilo, C., Núñez, Y., Benítez, R., López-Garcia, A., García, F., Félix, M. do R., Laranjo, M., Charneca, R., and Martins, J. M. (2021). "Transcriptomic profiling of skeletal muscle reveals candidate genes influencing muscle growth and associated lipid composition in portuguese local pig breeds". *Animals*, 11(5), 1423.

Ameer, F., Scandiuzzi, L., Hasnain, S., Kalbacher, H., and Zaidi, N. (2014). "De novo lipogenesis in health and disease". *Metabolism*, 63(7):895–902.

Amills, M., Clop, A., and Óvilo, C. (2020). "Nutrigenomics of lipid supplementation in ruminants and pigs". *In Lipids and Edible Oils* (pp. 93-131). Academic Press.

Andersson, L., Haley, C.S., Ellegren, H., Knott, S. A., Johansson, M., Andersson, K., Andersson-Eklund, L., Edfors-Lilja, I., Fredholm, M., Hansson, I., Håkansson, J., and Lundström, K. (1994). "Genetic mapping of quantitative trait loci for growth and fatness in pigs". *Science*, 263(5154):1771–1774.

Archibald, A.L., Bolund, L., Churcher, C., Fredholm, M., Groenen, M.A.M., Harlizius, B., Lee, K.T., Milan, D., Rogers, J., Rothschild, M.F., Uenishi, H., Wang, J., Schook, L.B., and Swine Genome Sequencing Consortium. (2010). "Pig genome sequence - analysis and publication strategy". *BMC Genomics*, 11(1):438.

Astrup, A., Teicholz, N., Magkos, F., Bier, D.M., Brenna, J. T., King, J. C., Mente, A., Ordovas, J. M., Volek, J. S., Yusuf, S., and Krauss, R. M. (2021). "Dietary saturated fats and health: are the U.S. guidelines evidence-based?". *Nutrients*, 13(10):3305.

Bacle, A., and Ferreira, T. (2019). "Strategies to counter saturated fatty acid (SFA)-mediated lipointoxication". Pp. 347–363 in *The Molecular Nutrition of Fats*. Elsevier.

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

Ballester, M., Puig-Oliveras, A., Castelló, A., Revilla, M., Fernández, A. I., and Folch, J. M. (2017). "Association of genetic variants and expression levels of porcine FABP4 and FABP5 genes". *Animal Genetics*, 48(6), 660–668.

Ballester, M., Ramayo-Caldas, Y., Revilla, M., Corominas, J., Castelló, A., Estellé, J., Fernández, A.I., and Folch, J.M. (2017a). "Integration of liver gene co-expression networks and EGWAs analyses highlighted candidate regulators implicated in lipid metabolism in pigs". *Scientific Reports*, 7(1):46539.

Bartlett, K., and Simon, E. (2004). "Mitochondrial beta-oxidation". *European Journal of Biochemistry*, 271(3):462–469.

Barton-Gade, P.A. (1990). "Pork quality in genetic improvement programmes –the danish experience". in Record of Proceedings National Swine Improvement Federation and Annual Meeting. Des Moines, Iowa,USA.

Benítez, R., Núñez, Y., Fernández, A., Isabel, B., Fernández, A. I., Rodríguez, C., Barragán, C., Martín- Palomino, P., López-Bote, C., Silió, L., and Óvilo, C. (2015). "Effects of dietary fat saturation on fatty acid composition and gene transcription in different tissues of Iberian pigs". *Meat Science*, 102:59–68.

Benítez, R., Fernández, A., Isabel, B., Núñez, Y., De Mercado, E., Gómez-Izquierdo, E., García-Casco, J, López-Bote, C, and Óvilo, C. (2017). "Modulatory effects of breed, feeding status, and diet on adipogenic, lipogenic, and lipolytic gene expression in growing Iberian and Duroc pigs". *International Journal of Molecular Sciences*, 19(1):22.

Bergen, W. G., and Mersmann, H. J. (2005). "Comparative aspects of lipid metabolism: impact on contemporary research and use of animal models". *The Journal of Nutrition*, 135(11):2499–2502.

Bickel, P. E., Tansey, J. T., and Welte, M. A. (2009). "PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores". *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1791*(6), 419–440.

Black, M. B., Parks, B. B., Pluta, L., Chu, T. M., Allen, B. C., Wolfinger, R. D., and Thomas, R. S. (2014). "Comparison of microarrays and RNA-seq for gene expression analyses of dose-response experiments". *Toxicological Sciences*, 137(2), 385-403.

Blasco, A., and Pena, R. N. (2018). "Current status of genomic maps: genomic selection/GBV in livestock". Pp. 61–80 in *Animal Biotechnology 2*. Cham: Springer International Publishing.

Blasco, A. (2017). "Bayesian data analysis for animal scientists". *Journal of Chemical Information and Modeling*, 53(9).

Bond, L. M., Miyazaki, M., O'Neill, L. M., Ding, F., and Ntambi, J. M. (2016). "Fatty Acid Desaturation and Elongation in Mammals". In *Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 185–208). Elsevier.

Brondz, I. (2005). "Lipids | fatty acids." Pp. 76–88 in *Encyclopedia of Analytical Science*. Elsevier.

Cameron, N. D. (1990). "Genetic and phenotypic parameters for carcass traits, meat and eating quality traits in pigs". *Livestock Production Science,* 26(2):119–135.

Cameron, N.D. (1993). "Selection for meat quality: objectives and criteria". *Pig News and Information*, 14:161N-168N.

Cameron, N. D., Enser, M., Nute, G. R., Whittington, F. M., Penman, J. C., Fisken, A. C., Perry, A. M., and Wood, J. D. (2000). "Genotype with nutrition interaction on fatty acid composition of intramuscular fat and the relationship with flavour of pig meat". *Meat Science*, 55(2):187–195.

Cánovas, A., Quintanilla, R., Amills, M., and Pena, R. N. (2010). "Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits". *BMC genomics*, 11, 1-15.

Cao, H., Gerhold, K., Mayers, J. R., Wiest, M. M., Watkins, S. M. and Hotamisligil, G. S. (2008). "Identification of a lipokine, a lipid hormone linking adipose tissue to systemic metabolism". *Cell*, 134(6):933–944.

Cardoso, T. F., Cánovas, Á., Canela-Xandri, O., González-Prendes, R., Amills, M., and Quintanilla, R. (2017). "RNA-seq based detection of differentially expressed genes in the skeletal muscle of duroc pigs with distinct lipid profiles". *Scientific Reports*, 7(1), 40005.

Cardoso, T. F., Quintanilla, R., Castelló, A., González-Prendes, R., Amills, M., and Cánovas, Á. (2018). "Differential expression of mRNA isoforms in the skeletal muscle of pigs with distinct growth and fatness profiles". *BMC Genomics*, 19(1), 145.

Cenedella, R.J., and Allen, A. (1969). "Differences between the metabolism of linoleic and palmitic acid: utilization for cholesterol synthesis and oxidation to respiratory CO2". *Lipids*, 4(2):155–158.

Chin, C.S., and Nyam, K.L. (2020). "Refining of edible oils". Pp. 213–241 in *Lipids and Edible Oils*. Elsevier.

Chmurzyńska, A. (2006). "The multigene family of fatty acid-binding proteins (FABPs): Function, structure and polymorphism". *Journal of Applied Genetics*, *47*(1), 39–48.

Chriki, S., Gardner, G. E., Jurie, C., Picard, B., Micol, D., Brun, J. P., ... and Hocquette, J. F. (2012). "Cluster analysis application identifies muscle characteristics of importance for beef tenderness". *BMC biochemistry*, 13(1), 1-11.

_____

Ciobanu, D., Bastiaansen, J., Malek, M., Helm, J., Woollard, J., Plastow, G., and Rothschild, M. (2001). "Evidence for new alleles in the protein kinase adenosine monophosphate-activated γ3-subunit gene associated with low glycogen content in pig skeletal muscle and improved meat quality". *Genetics*, 159(3), 1151–1162.

Clop, A., Ovilo, C., Perez-Enciso, M., Cercos, A., Tomas, A., Fernandez, A, Coll, A., Folch, J.M., Barragan, C., Diaz, I., Oliver, M. A., Varona, L., Silio, L., Sanchez, A., and Noguera, J. L. (2003). "Detection of QTL affecting fatty acid composition in the pig". *Mammalian Genome*, 14(9):650–656.

Cockcroft, S. (2021). "Mammalian lipids: structure, synthesis and function". *Essays in Biochemistry*, 65(5):813–845.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D. J., Elo, L.L., Zhang, X., and Mortazavi, A. (2016). "A survey of best practices for RNA-Seq data analysis". *Genome Biology*, 17(1):13.

Cook, H. W., and McMaster, C. R. (2002). "Fatty acid desaturation and chain elongation in eukaryotes". Pp. 181–204 in *New Comprehensive Biochemistry*, 36.

Corominas, J., Ramayo-Caldas, Y., Castelló, A., Muñoz, M., Ibáñez-Escriche, N., Folch, J. M., and Ballester, M. (2012). "Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs". *Animal Genetics*, 43(6):714–720.

Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Estellé, J., Castelló, A, Alves, E., Pena, R.N., Ballester, M., and Folch, J.M. (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(1):843.

Corominas, J., Marchesi, J. A., Puig-Oliveras, A., Revilla, M., Estellé, J., Alves, E., ... and Ballester, M. (2015). "Epigenetic regulation of the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs". *Genetics Selection Evolution*, 47(1), 1-11.

Crespo-Piazuelo, D., Criado-Mesas, L., Revilla, M., Castelló, A., Noguera, J. L., Fernández, A. I., Ballester, M., and Folch, J. M. (2020). "Identification of strong candidate genes for backfat and intramuscular fatty acid composition in three crosses based on the Iberian pig". *Scientific Reports*, *10*(1), 13962.

Criado-Mesas, L., Ballester, M., Crespo-Piazuelo, D., Castelló, A., Benítez, R., Fernández, A. I., and Folch, J. M. (2019). "Analysis of porcine IGF2 gene expression in adipose tissue and its effect on fatty acid composition". *PLOS ONE*, 14(8), e0220708.

Criado-Mesas, L., Ballester, M., Crespo-Piazuelo, D., Castelló, A., Fernández, A. I., and Folch, J. M. (2020). "Identification of eQTLs associated with lipid metabolism in Longissimus dorsi muscle of pigs with different genetic backgrounds". *Scientific Reports*, 10(1).

Cui, J. X., Zeng, Y. Q., Wang, H., Chen, W., Du, J. F., Chen, Q. M., Hu, Y. X., and Yang, L. (2011). "The effects of DGAT1 and DGAT2 mRNA expression on fat deposition in fatty and lean breeds of pig". *Livestock Science*, 140(1–3), 292–296.

Damon, M., Wyszynska-Koko, J., Vincent, A., Hérault, F., and Lebret, B. (2012). "Comparison of muscle transcriptome between pigs with divergent meat quality phenotypes identifies genes related to muscle metabolism and structure". *PLoS ONE*, 7(3), e33763.

Davies, A.S., and Pryor, W. J. (1977). "Growth changes in the distribution of dissectable and intramuscular fat in pigs". *The Journal of Agricultural Science*, 89(2):257–266.

Day, D.A., and Tuite, M.F. (1998). "Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview". *Journal of Endocrinology*, 157(3):361–371.

DeLany, J. P., Windhauser, M. M., Champagne, C. M., and Bray, G. A. (2000). "Differential oxidation of individual dietary fatty acids in humans". *The American journal of clinical nutrition*, 72(4), 905-911.

Désert, C., Baéza, E., Aite, M., Boutin, M., Le Cam, A., Montfort, J., ... and Lagarrigue, S. (2018). "Multi-tissue transcriptomic study reveals the main role of liver in the chicken adaptive response to a switch in dietary energy source through the transcriptional regulation of lipogenesis". *BMC genomics*, 19, 1-18.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., ... and Gingeras, T.R. (2013)." STAR: ultrafast universal RNA-seq aligner". *Bioinformatics*, 29(1), 15-21.

Domaradzki, P., Nowakowicz-Dębek, B., Wlazło, Ł., Ossowski, M., Dmoch, M., and Florek, M. (2022). "Fatty Acid Composition of Muscle and Adipose Tissue in Pigs Fed with Addition of Natural Sorbents". *Animals*, 12(13), 1681.

Eggen, A. and Hocquette, J.F. (2004). "Genomic approaches to economic trait loci and tissue expression profiling: application to muscle biochemistry and beef quality". *Meat Science*, 66(1):1–9.

Estany, J., Ros-Freixedes, R., Tor, M., and Pena, R. N. (2014). "A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork". *PLoS ONE*, *9*(1), e86177.

Esteve-Codina, A., Kofler, R., Palmieri, N., Bussotti, G., Notredame, C., and Pérez-Enciso, M. (2011). "Exploring the gonad transcriptome of two extreme male pigs with RNA-seq". *BMC genomics*, 12(1), 1-14.

Fan, Y., Guo, Y., Hamblin, M., Chang, L., Zhang, J., and Chen, Y. E. (2011). "Inhibition of Gluconeogenic Genes by Calcium-regulated Heat-stable Protein 1 via Repression of Peroxisome Proliferator-activated Receptor α". *Journal of Biological Chemistry*, *286*(47), 40584–40594.

FAOSTAT. (2022). "FAOSTAT." Retrieved January 11, 2023 (https://www.fao.org/faostat/en/#data/QCL/visualize).

Fiego, D.P. Lo, P. Santoro, P. Macchioni, and E. De Leonibus. (2005). "Influence of genetic type, live weight at slaughter and carcass fatness on fatty acid composition of subcutaneous adipose tissue of raw ham in the heavy pig". *Meat Science*, 69(1):107–114.

Fillmore, N., Mori, J., and Lopaschuk, G.D. (2014). "Mitochondrial fatty acid oxidation alterations in heart failure, ischaemic heart disease and diabetic cardiomyopathy". *British Journal of Pharmacology*, 171(8):2080–2090.

Folch, J., Lees, M., and Sloane Stanley, G.H. (1957). "A simple method for the isolation and purification of total lipides from animal tissues". *J Biol Chem.*, 226(1):497–509.

Fontanesi, L., Scotti, E., Buttazzoni, L., Davoli, R., and Russo, V. (2009). "The porcine fat mass and obesity associated (FTO) gene is associated with fat deposition in Italian Duroc pigs". *Animal Genetics*, 40(1), 90–93.

Fontanesi, L., Scotti, E., Buttazzoni, L., Dall'Olio, S., Bagnato, A., Lo Fiego, D. Pietro, Davoli, R., and Russo, V. (2010). "Confirmed association between a single nucleotide polymorphism in the (FTO) gene and obesity-related traits in heavy pigs". *Molecular Biology Reports*, 37(1), 461–466.

Frazee, A. C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., and Leek, J.T. (2015). "Ballgown bridges the gap between transcriptome assembly and expression analysis". *Nature biotechnology*, 33(3), 243-246.

Fritzen, A.M, Lundsgaard, A.M., and Kiens, B. (2020). "Tuning fatty acid oxidation in skeletal muscle with dietary fat and exercise". *Nature Reviews Endocrinology*, 16(12):683–696.

Frühbeck, G., Méndez-Giménez, L., Fernández-Formoso, J.A., Fernández, S., and Rodríguez, A. (2014). "Regulation of adipocyte lipolysis". *Nutrition Research Reviews*, 27(1):63–93.

Furuhashi, M., and Hotamisligil, G.S. (2008). "Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets". *Nature Reviews Drug Discovery*, *7*(6), 489–503.

Gagaoua, M., and Picard, B. (2020). Current advances in meat nutritional, sensory and physical quality improvement. *Foods*, 9(3), 321.

Gandemer, G. (2002). "Lipids in muscles and adipose tissues, changes during processing and sensory properties of meat products". *Meat Science*, 62(3):309–321.

Gandolfi, G., Cinar, M. U., Ponsuksili, S., Wimmers, K., Tesfaye, D., Looft, C., Jüngst, H., Tholen, E., Phatsara, C., Schellander, K., and Davoli, R. (2011). "Association of PPARGC1A and CAPNS1 gene polymorphisms and expression with meat quality traits in pigs". *Meat Science*, 89(4), 478–485.

Gao, P., Cheng, Z., Li, M., Zhang, N., Le, B., Zhang, W., Song, P., Guo, X., Li, B., and Cao, G. (2019). "Selection of candidate genes affecting meat quality and preliminary exploration of related molecular mechanisms in the Mashen pig". *Asian-Australasian Journal of Animal Sciences*, 32(8), 1084–1094.

García-Olmo, J., E. De Pedro, A. Garrido, A. Paredes, C. Sanabria, M. Santolalla, J. Salas, J. R. García-Hierro, I. Gonzalez, M. D. García-Cachan, and J. Guirao. (2002). "Determination of the precision of the fatty acid analysis of Iberian pig fat by gas chromatography. Results of a mini collaborative study". *Meat Science*, 60(1):103–109.

Ghosh, M., Sodhi, S. S., Song, K.-D., Kim, J. H., Mongre, R. K., Sharma, N., Singh, N. K., Kim, S. W., Lee, H. K., and Jeong, D. K. (2015). "Evaluation of body growth and immunity-related differentially expressed genes through deep RNA sequencing in the piglets of Jeju native pig and Berkshire". *Animal Genetics*, 46(3), 255–264.

Glatz, J. F. C., Schaap, F. G., Binas, B., Bonen, A., Van Der Vusse, G. J., and Luiken, J. J. F. P. (2003). "Cytoplasmic fatty acid-binding protein facilitates fatty acid utilization by skeletal muscle". *Acta Physiologica Scandinavica*, *178*(4), 367–371.

Godfray, H. C. J., Aveyard, P., Garnett, T., Hall, J. W., Key, T. J., Lorimer, J., ... and Jebb, S. A. (2018). "Meat consumption, health, and the environment". *Science*, 361(6399), eaam5324.

Gol, S., Pena, R. N., Rothschild, M. F., Tor, M., and Estany, J. (2018). "A polymorphism in the fatty acid desaturase-2 gene is associated with the arachidonic acid metabolism in pigs". *Scientific Reports*, 8(1), 14336.

Goldspink, G. (2002). "Gene expression in skeletal muscle". *Biochemical Society Transactions*, *30*(2), 285–290.

González-Prendes, R., Mármol-Sánchez, E., Quintanilla, R., Castelló, A., Zidi, A., Ramayo-Caldas, Y., ... and Amills, M. (2019). "About the existence of common determinants of gene expression in the porcine liver and skeletal muscle". *BMC genomics*, 20(1), 1-15.

González-Prendes, R., Quintanilla, R., Mármol-Sánchez, E., Pena, R. N., Ballester, M., Cardoso, T. F., ... and Amills, M. (2019a). "Comparing the mRNA expression profile and the genetic determinism of intramuscular fat traits in the porcine gluteus medius and longissimus dorsi muscles". *BMC genomics*, 20, 1-18.

Gozalo-Marcilla, M., Buntjer, J., Johnsson, M., Batista, L., Diez, F., Werner, C. R., Chen, C.-Y., Gorjanc, G., Mellanby, R. J., Hickey, J. M., and Ros-Freixedes, R. (2021). "Genetic architecture and major genes for backfat thickness in pig lines of diverse genetic backgrounds". *Genetics Selection Evolution*, 53(1), 76.

Groenen, M. (2015). "Development of a high-density Axiom® porcine genotyping array to meet research and commercial needs". in *Plant and Animal Genome XXIII*. San Diego, CA, USA.

Groenen, M. A. M., Schook, L. B., and Archibald, A. L. (2011). "Pig genomics". Pp. 179–99 in *the genetics of the pig*. UK: CABI.

Groenen, M. A., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., ... and Schook, L. B. (2012). "Analyses of pig genomes provide insight into porcine demography and evolution". *Nature*, 491(7424), 393-398.

Guillou, H., Zadravec, D., Martin, P.G., and Jacobsson, A. (2010). "The key roles of elongases and desaturases in mammalian fatty acid metabolism: insights from transgenic mice". *Progress in Lipid Research*, 49(2):186–199.

Guo, B., and Dalrymple, B. P. (2017). "Transcriptomics of meat quality". *New Aspects of Meat Quality*, 259–320.

Haley, C. S., Archibald, A., Andersson, L., Bosma, A. A., Davies, W., Fredholm, M., Geldermann, H., Groenen, M., Gustavsson, I., Ollivier, L., Tucker, E. M., and van de Weghe, A. (1990). "The pig gene mapping project - PiGMap". *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, Edinburgh 23-27 July 1990. XIII. Plenary Lectures, Molecular Genetics and Mapping, Selection, Prediction and Estimation. 67–70.

Hamill, R. M., McBryan, J., McGee, C., Mullen, A. M., Sweeney, T., Talbot, A., ... and Davey, G. C. (2012). "Functional analysis of muscle gene expression profiles associated with tenderness and intramuscular fat content in pork". *Meat Science*, 92(4), 440-450.

He, L., Vasiliou, K., and Nebert, D. W. (2009). "Analysis and update of the human solute carrier (SLC) gene superfamily". *Human Genomics*, *3*(2), 195.

Hu, Z. L., Park, C. A., and Reecy, J. M. (2022). "Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services.". *Nucleic acids research*, 50(D1), D956-D961.

Humphray, S. J., Scott, C. E., Clark, R., Marron, B., Bender, C., Camm, N., ... and Rogers, J. (2007). "A high utility integrated map of the pig genome". *Genome Biology*, 8(7):1–11.

IDESCAT. (2020). "Idescat. anuario estadístico de Cataluña. producción de carne. por especies". *Departamento de Acción Climática, Alimentación y Agenda Rural*. Retrieved January 11, 2023. https://www.idescat.cat/indicadors/?id=aec&n=15435&lang=es.

Inoue, M., Chang, L., Hwang, J., Chiang, S. H., and Saltiel, A. R. (2003). "The exocyst complex is required for targeting of Glut4 to the plasma membrane by insulin". *Nature*, 422(6932), 629-633.

Inoue, M., Akama, T., Jiang, Y., and Chun, T. H. (2015). "The exocyst complex regulates free fatty acid uptake by adipocytes". *PLoS One*, 10(3), e0120289.

Iqbal, J., and Hussain, M. M. (2009). "Intestinal lipid absorption". *American Journal of Physiology-Endocrinology and Metabolism*, 296(6), E1183–E1194.

Jiang, S., Wei, H., Song, T., Yang, Y., Peng, J., and Jiang, S. (2013). "Transcriptome comparison between porcine subcutaneous and intramuscular stromal vascular cells during adipogenic differentiation". *PloS one*, 8(10), e77094.

Johnston, M.R., and Sobhi, H.F. (2018). "Advances in fatty acid analysis for clinical investigation and diagnosis using GC/MS methodology". *Journal of Biochemistry and Analytical Studies*, 3(1).

Jung, Y., Kwon, S., Ham, S., Lee, D., Park, H. H., Yamaoka, Y., Jeong, D., Artan, M., Altintas, O., Park, S., Hwang, W., Lee, Y., Son, H. G., An, S. W. A., Kim, E. J. E., Seo, M., and Lee, S. V. (2020). "Caenorhabditis elegans Lipin 1 moderates the lifespan-shortening effects of dietary glucose by maintaining ω-6 polyunsaturated fatty acids". *Aging Cell*, *19*(6).

Kapke, P., L. Wang, J. Helm, and M. F. Rothschild. (1996). "Integration of the PiGMaP and USD a maps for porcine chromosome 14", *Animal Genetics*, 27(3):187–190.

Kennes, Y. M., Murphy, B. D., Pothier, F., and Palin, M.-F. (2001). "Characterization of swine leptin (LEP) polymorphisms and their association with production traits". *Animal Genetics*, 32(4), 215–218.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). "HISAT: a fast spliced aligner with low memory requirements". *Nature methods*, 12(4), 357-360.

Kim, K. S., Larsen, N., Short, T., Plastow, G., and Rothschild, M. F. (2000). "A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits". *Mammalian Genome*, 11(2), 131–135.

Kim, N.K., Park, H.R., Lee, H.C., Yoon, D., Son, E.S., Kim, Y.S., Kim, S.R., Kim, O.H., and Lee, C.S. (2010). "Comparative studies of skeletal muscle proteome and transcriptome profilings between pig breeds". *Mammalian Genome*, 21(5–6), 307–319.

Kloareg, M., Noblet, J., and van Milgen, J. (2007). "Deposition of dietary fatty acids, *de novo* synthesis and anatomical partitioning of fatty acids in finishing pigs". *British Journal of Nutrition*, 97(1):35–44.

Knol, E. F., Nielsen, B., & Knap, P. W. (2016). "Genomic selection in commercial pig breeding". *Animal Frontiers*, 6(1):15–22.

Kojima, M., Nakajima, I., Arakawa, A., Mikawa, S., Matsumoto, T., Uenishi, H., ... and Taniguchi, M. (2018). "Differences in gene expression profiles for subcutaneous adipose, liver, and skeletal muscle tissues between Meishan and Landrace pigs with different backfat thicknesses". *PloS one*, 13(9), e0204135.

yriazakis, I., and Whittemore, C.T. (2008). "Whittemore's science and practice of pig production." John Wiley and Sons. Oxford, UK: Blackwell Publishing Ltd.

Latorre, P., Burgos, C., Hidalgo, J., Varona, L., Carrodeguas, J. A., and López-Buesa, P. (2016). "c.A2456C-substitution in Pck1 changes the enzyme kinetic and functional properties modifying fat distribution in pigs". *Scientific Reports*, 6(1), 19617.

Lê Cao, K.-A., González, I., and Déjean, S. (2009). "IntegrOmics: an R package to unravel relationships between two omics datasets". *Bioinformatics*, *25*(21), 2855–2856.

Lebret, B. (2008). "Effects of feeding and rearing systems on growth, carcass composition and meat quality in pigs". *Animal*, 2(10):1548–1558.

Lee, J.-E., and Ge, K. (2014). "Transcriptional and epigenetic regulation of PPARγ expression during adipogenesis". *Cell & Bioscience*, *4*(1), 29.

Li, B., Weng, Q., Dong, C., Zhang, Z., Li, R., Liu, J., Jiang, A., Li, Q., Jia, C., Wu, W., and Liu, H. (2018). "A key gene, PLIN1, can affect porcine intramuscular fat content based on transcriptome analysis". *Genes*, *9*(4), 194.

Liang, R., Han, B., Li, Q., Yuan, Y., Li, J., and Sun, D. (2017). "Using RNA sequencing to identify putative competing endogenous RNAs (ceRNAs) potentially regulating fat metabolism in bovine liver". *Scientific Reports*, *7*(1), 6396.

Liao, Y., Smyth, G. K., and Shi, W. (2019). "The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads". *Nucleic acids research*, 47(8), e47-e47.

Lim, K. S., Lee, K. T., Park, J. E., Chung, W. H., Jang, G. W., Choi, B. H., ... and Kim, T. H. (2017). "Identification of differentially expressed genes in longissimus muscle of pigs with high and low intramuscular fat content using RNA sequencing". *Animal Genetics*, 48(2), 166-174.

Listrat, A., Lebret, B., Louveau, I., Astruc, T., Bonnet, M., Lefaucheur, L., ... & Bugeon, J. (2016). "How muscle structure and composition influence meat and flesh quality". *The Scientific World Journal*, 2016.

Liu, J., Damon, M., Guitton, N., Guisle, I., Ecolan, P., Vincent, A., ... and Gondret, F. (2009). "Differentially-expressed genes in pig longissimus muscles with contrasting levels of fat, as identified by combined transcriptomic, reverse transcription PCR, and proteomic analyses". *Journal of agricultural and food chemistry*, 57(9), 3808-3817.

Liu, K., Ye, J., Yang, Y., Shen, L., and Jiang, H. (2018). "A unified model for joint normalization and differential gene expression detection in RNA-Seq data". *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *16*(2), 442–454.

Liu, Xianxian, Zhang, J., Xiong, X., Chen, C., Xing, Y., Duan, Y., Xiao, S., Yang, B., and Ma, J. (2021). "An integrative analysis of transcriptome and GWAS data to identify potential candidate genes influencing meat quality traits in pigs". *Frontiers in Genetics*, 12.

Liu, Xuan, Yang Du, Nares Trakooljul, Bodo Brand, Eduard Muráni, Carsten Krischek, Michael Wicke, Manfred Schwerin, Klaus Wimmers, and Siriluck Ponsuksili. (2015). "Muscle transcriptional profile based on muscle fiber, mitochondrial respiratory activity, and metabolic enzymes". *International Journal of Biological Sciences*, 11(12):1348–1362.

Lopaschuk, Gary D., John R. Ussher, Clifford D. L. Folmes, Jagdip S. Jaswal, and William C. Stanley. (2010). "Myocardial fatty acid metabolism in health and disease". *Physiological Reviews*, 90(1):207–258.

López, M. E., Neira, R., and Yáñez, J. M. (2015). "Applications in the search for genomic selection signatures in fish". *Frontiers in Genetics*, 5.

Lopez-Bote, C. J. (1998). "Sustained utilization of the Iberian pig breed". *Meat science*, 49, S17-S27.

Love, M. I., Huber, W., and Anders, S. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". *Genome Biology*, *15*(12), 550.

MAPA. (2021). "Indicadores económicos e informes trimestrales." Ministerio de Agricultura, Pesca y Alimentación. Retrieved January 11, 2023 (https://www.mapa.gob.es/es/ganaderia/temas/produccion-y-mercados-ganaderos/sectores-ganaderos/porcino/indicadoreseconomicos.aspx).

Martínez-Montes, A. M., Fernández, A., Pérez-Montarelo, D., Alves, E., Benítez, R. M., Nuñez, Y., Óvilo, C., Ibañez-Escriche, N., Folch, J. M., and Fernández, A. I. (2017). "Using

RNA-Seq SNP data to reveal potential causal mutations related to pig production traits and RNA editing". *Animal Genetics*, 48(2), 151–165.

Martínez-Montes, A. M., Fernández, A., Muñoz, M., Noguera, J.L., Folch, J.M., and Fernández, A.I. (2018). "Using genome wide association studies to identify common QTL regions in three different genetic backgrounds based on Iberian pig breed". *PLOS ONE*, 13(3), e0190184.

Mater, M. K., Thelen, A. P., Pan, D. A., and Jump, D.B. (1999). "Sterol response element-binding protein 1c (SREBP1c) is involved in the polyunsaturated fatty acid suppression of hepatic S14 gene transcription". *Journal of Biological Chemistry*, *274*(46), 32725–32732.

McGarry, J. D., Mannaerts, G. P., and Foster, D. W. (1977). "A possible role for malonyl-CoA in the regulation of hepatic fatty acid oxidation and ketogenesis". *The Journal of Clinical Investigation*, 60(1):265–270.

McLean, R., Skeaff, M., Mann, J., and Morenga, L.T. (2015). "Are some diets 'mass murder'? dietary guidelines worldwide advise limiting saturated fat in favour of monounsaturated and polyunsaturated fats". *BMJ*, 350(feb11 10):h625–h625.

Melmed, S.M.B. (2020). "Disorders of lipid metabolism". In Williams Textbook of Endocrinology.

Mensink, R. P. (2013). "Fatty acids: health effects of saturated fatty acids". Pp. 215–219. In Encyclopedia of Human Nutrition. Elsevier.

Mercadé, A., Estellé, J., Noguera, J. L., Folch, J. M., Varona, L., Silió, L., Sánchez, A., and Pérez-Enciso, M. (2005). "On growth, fatness, and form: a further look at porcine chromosome 4 in an Iberian × Landrace cross". *Mammalian Genome*, 16(5), 374–382.

Mercade, A., Estelle, J., Perez-Enciso, M., Varona, L., Silio, L., Noguera, J. L., Sanchez, A., and Folch, J. M. (2006). "Characterization of the porcine acyl-CoA synthetase long-chain 4

gene and its association with growth and meat quality traits". *Animal Genetics*, 37(3), 219–224.

Messad, F., Louveau, I., Koffi, B., Gilbert, H., and Gondret, F. (2019). "Investigation of muscle transcriptomes using gradient boosting machine learning identifies molecular predictors of feed efficiency in growing pigs". *BMC genomics*, 20, 1-14.

Milan, D., Jeon, J.-T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannuccelli, N., Rask, L., Ronne, H., Lundström, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P. Le, Chardon, P., and Andersson, L. (2000). "A Mutation in PRKAG3 Associated with Excess Glycogen Content in Pig Skeletal Muscle". *Science*, 288(5469), 1248–1251.

Miller, C. W., and Ntambi, J. M. (1996). "Peroxisome proliferators induce mouse liver stearoyl-CoA desaturase 1 gene expression". *Proceedings of the National Academy of Sciences*, *93*(18), 9443–9448.

Miller, R. K. (2014). "Chemical and physical characteristics of meat | palatability". Pp. 252–61 in Encyclopedia of Meat Sciences. Elsevier.

Molinero, E., Pena, R. N., Estany, J., and Ros-Freixedes, R. (2022). "Identification of a missense variant in the porcine AGPAT gene family associated with intramuscular fat content through whole-genome sequencing". *Animal Genetics*, 53(6), 782–793.

Mondal, S., and Singh, R. L. (2020). "Advances in animal genomics". 1st ed. Academic Press. Singh. Elsevier.

Morales, P. E., Bucarey, J. L., and Espinosa, A. (2017). "Muscle lipid metabolism: role of lipid droplets and perilipins". *Journal of Diabetes Research*, 1–10.

Mote, B.E., and Rothschild, M.F. (2020). "Modern genetic and genomic improvement of the pig". Pp. 249–262 in Animal Agriculture. Elsevier.

Mottram, D.S. (1998). "Flavour formation in meat and meat products: a review". *Food Chemistry*, 62(4):415–424.

Muñoz, G., Alves, E., Fernández, A., Óvilo, C., Barragán, C., Estellé, J., Quintanilla, R., Folch, J. M., Silió, L., Rodríguez, M. C., and Fernández, A. I. (2007). "QTL detection on porcine chromosome 12 for fatty-acid composition and association analyses of the fatty acid synthase, gastric inhibitory polypeptide and acetyl-coenzyme A carboxylase alpha genes". *Animal Genetics*, 38(6), 639–646.

Muñoz, M., Rodríguez, M.C., Alves, E., Folch, J.M., Ibañez-Escriche, N., Silió, L., Fernández, A.I. (2013). "Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data". *BMC Genomics*, 14(1):845.

Muñoz, M., García-Casco, J.M., Caraballo, C., Fernández-Barroso, M.Á., Sánchez-Esquiliche, F., Gómez, F., Rodríguez, M. del C., and Silió, L. (2018). "Identification of candidate genes and regulatory factors underlying intramuscular fat content through longissimus dorsi transcriptome analyses in heavy Iberian pigs". *Frontiers in Genetics*, 9.

Nakamura, M.T., Cheon, Y., Li, Y., and Nara, T.Y. (2004). "Mechanisms of regulation of gene expression by fatty acids". *Lipids*, *39*(11), 1077–1083.

Ramona, N.P., Ros-Freixedes, R., Tor, M., and Estany, J. (2016). "Genetic marker discovery in complex traits: a field example on fat content and composition in pigs." *International Journal of Molecular Sciences*, 17(12):2100.

Nelson, D. L., Lehninger, A. L., and Cox, M. M. (2008). "Lehninger principles of biochemistry." 5th ed. Macmillan, New York.

Ngapo, Tania M., and Claude Gariépy. (2008). "Factors affecting the eating quality of pork". *Critical Reviews in Food Science and Nutrition*, 48(7):599–633.

Nold, R. (2006). "Meat quality." Pork Information Gateway. https://porkgateway.org/resource/meat-quality/.

Nookaew, I., Papini, M., Pornputtapong, N., Scalcinati, G., Fagerberg, L., Uhlén, M., and Nielsen, J. (2012). "A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae". *Nucleic acids research*, 40(20), 10084-10097.

Ntambi, J. M., and Miyazaki, M. (2004). "Regulation of stearoyl-CoA desaturases and role in metabolism". *Progress in lipid research*, 43(2), 91-104.

Ntawubizi, M., Colman, E., Janssens, S., Raes, K., Buys, N., and De Smet, S. (2010). "Genetic parameters for intramuscular fatty acid composition and metabolism in pigs". *Journal of Animal Science*, 88(4):1286–1294.

Nürnberg, K., Wegner, J., and Ender, K. (1998). "Factors influencing fat composition in muscle and adipose tissue of farm animals". *Livestock Production Science*, 56(2):145–156.

O'Fallon, J. V., Busboom, J. R., Nelson, M. L., and Gaskins, C. T. (2007). "A direct method for fatty acid methyl ester synthesis: application to wet meat tissues, oils, and feedstuffs". *Journal of Animal Science*, 85(6):1511–1521.

O'Hea, Eugene Kevin, and Gilbert A. Leveille. (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". *The Journal of Nutrition*, 99(3):338–344.

Oliveira de Biagi, C. A., Nociti, R. P., Brotto, D. B., Funicheli, B. O., Cássia Ruy, P. de, Bianchi Ximenez, J. P., Alves Figueiredo, D. L., and Araújo Silva, W. (2021). "CeTF: an R/Bioconductor package for transcription factor co-expression networks using regulatory impact factors (RIF) and partial correlation and information (PCIT) analysis". *BMC Genomics*, 22(1), 624.

Óvilo, C., Pérez-Enciso, M., Barragán, C., Clop, A., Rodríguez, C., Oliver, M. A., Toro, M. A., and Noguera, J. L. (2000). "A QTL for intramuscular fat and backfat thickness is located on porcine Chromosome 6". *Mammalian Genome*, 11(4), 344–346.

Óvilo, C., Clop, A., Noguera, J. L., Oliver, M. A., Barragán, C., Rodríguez, C., Silió, L., Toro, M. A., Coll, A., Folch, J. M., Sánchez, A., Babot, D., Varona, L., and Pérez-Enciso, M. (2002). "Quantitative trait locus mapping for meat quality traits in an Iberian × Landrace F2 pig population1". *Journal of Animal Science*, 80(11), 2801–2808.

Óvilo, C., Fernández, A., Noguera, J. L., Barragán, C., Letón, R., Rodríguez, C., Mercadé, A., Alves, E., Folch, J. M., Varona, L., and Toro, M. (2005). "Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross". *Genetical Research*, 85(1), 57–67.

Óvilo, C., Fernández, A., Fernández, A. I., Folch, J. M., Varona, L., Benítez, R., Nuñez, Y., Rodríguez, C., and Silió, L. (2010). "Hypothalamic expression of porcine leptin receptor (LEPR), neuropeptide Y (NPY), and cocaine- and amphetamine-regulated transcript (CART) genes is influenced by LEPR genotype". *Mammalian Genome*, 21(11–12), 583–591.

Óvilo, C., Benítez, R., Fernández, A., Núñez, Y., Ayuso, M., Fernández, A., Rodríguez, C., Isabel, B., Rey, A., López-Bote, C, and Silió, L. (2014). "Longissimus dorsi transcriptome analysis of purebred and crossbred Iberian pigs differing in muscle characteristics". *BMC Genomics*, 15(1):413.

Oyelami, F. O., Zhao, Q., Xu, Z., Zhang, Z., Sun, H., Zhang, Z., ... and Pan, Y. (2020). "Haplotype block analysis reveals candidate genes and QTLs for meat quality and disease resistance in Chinese Jiangquhai Pig Breed". *Frontiers in Genetics*, 11, 752.

Park, W. J., Kothapalli, K. S. D., Reardon, H. T., Kim, L. Y., and Brenna, J. T. (2009). "Novel fatty acid desaturase 3 (FADS3) transcripts generated by alternative splicing". *Gene*, 446(1), 28–34.

Pena, R. N., Noguera, J. L., Casellas, J., Díaz, I., Fernández, A. I., Folch, J. M., and Ibáñez-Escriche, N. (2013). "Transcriptional analysis of intramuscular fatty acid composition in the longissimus thoracis muscle of Iberian × Landrace back-crossed pigs". Animal Genetics, 44(6):648–660.

Pena, R. N., R. Quintanilla, A. Manunza, D. Gallardo, J. Casellas, and M. Amills. (2014). "Application of the microarray technology to the transcriptional analysis of muscle phenotypes in pigs". *Animal Genetics*, 45(3):311–321.

Pérez-Enciso, M., A. Clop, J. L. Noguera, C. Ovilo, A. Coll, J. M. Folch, D. Babot, J. Estany, M. A. Oliver, I. Díaz, and A. Sánchez. (2000). "A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross". *Journal of Animal Science*, 78(10):2525.

Pérez-Enciso, M., Mercadé, A., Bidanel, J. P., Geldermann, H., Cepica, S., Bartenschlager, H., Varona, L., Milan, D., and Folch, J. M. (2005). "Large-scale, multibreed, multitrait analyses of quantitative trait loci experiments: The case of porcine X chromosome1". *Journal of Animal Science*, 83(10), 2289–2296.

Pérez-Enciso, M., Clop, A., Folch, J. M., Sánchez, A., Oliver, M. A., Óvilo, C., Barragán, C., Varona, L., and Noguera, J. L. (2002). "Exploring alternative models for sex-linked quantitative trait loci in outbred populations: application to an Iberian × Landrace pig intercross". *Genetics*, 161(4), 1625–1632.

Perez-Montarelo, D., Hudson, N. J., Fernandez, A. I., Ramayo-Caldas, Y., Dalrymple, B. P., and Reverter, A. (2012). "Porcine tissue-specific regulatory networks derived from meta-analysis of the transcriptome". *PLoS ONE*, 7(9), e46159.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). "StringTie enables improved reconstruction of a transcriptome from RNA-seq reads". *Nature biotechnology*, 33(3), 290-295.

Pertea, Mihaela, Daehwan Kim, Geo M. Pertea, Jeffrey T. Leek, and Steven L. Salzberg. (2016). "Transcript-level expression analysis of RNA-Seq experiments with HISAT, stringtie and ballgown". *Nature Protocols*, 11(9):1650–1667.

Piles, M., Fernandez-Lozano, C., Velasco-Galilea, M., González-Rodríguez, O., Sánchez, J. P., Torrallardona, D., Ballester, M., and Quintanilla, R. (2019). "Machine learning applied to transcriptomic data to identify genes associated with feed efficiency in pigs". *Genetics Selection Evolution*, 51(1), 10.

Piórkowska, K., Małopolska, M., Ropka-Molik, K., Szyndler-Nędza, M., Wiechniak, A., Żukowski, K., Lambert, B., and Tyra, M. (2020). "Evaluation of SCD, ACACA and FASN Mutations: effects on pork quality and other production traits in pigs selected based on RNA-Seq results". *Animals*, 10(1), 123.

Poirier, Yves, Vasily D. Antonenkov, Tuomo Glumoff, and J. Kalervo Hiltunen. (2006). "Peroxisomal β-oxidation—a metabolic pathway with multiple functions". *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1763(12):1413–1426.

Poklukar, K., Čandek-Potokar, M., Batorek Lukač, N., Tomažin, U., and Škrlep, M. (2020). Lipid deposition and metabolism in local and modern pig breeds: A review. *Animals*, 10(3), 424.

Puig-Oliveras, A., Ramayo-Caldas, Y., Corominas, J., Estellé, J., Pérez-Montarelo, D., Hudson, N. J., Casellas, J., Folch, J.M., and Ballester, M. (2014). "Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition". *PLoS ONE*, 9(6):e99720.

Puig-Oliveras, A., Revilla, M, Castelló, A., Fernández, A. I., Folch, J.M., and Ballester, M. (2016). "Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat". *Scientific Reports*, 6(1):31803.

Ramayo-Caldas, Y., Mercadé, A., Castelló, A., Yang, B., Rodríguez, C., Alves, E., Díaz, I., Ibáñez-Escriche, N., Noguera, J. L., Pérez-Enciso, M., Fernández, A. I., and Folch, J.M. (2012). "Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace Cross". *Journal of Animal Science*, 90(9):2883–2893.

Ramayo-Caldas, Y., Mach, N., Esteve-Codina, A., Corominas, J., Castelló, A., Ballester, M., Estellé, J., Ibáñez-Escriche, N., Fernández, A. I., Pérez-Enciso, M., and Folch, J.M. (2012a). "Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition". *BMC Genomics*, 13(1):547.

Ramayo-Caldas, Y., Ballester, M., Fortes, M. R., Esteve-Codina, A., Castelló, A., Noguera, J. L., Fernández, A. I., Pérez-Enciso, M., Reverter, A., and Folch, J.M. (2014). "From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine". *BMC Genomics*, *15*(1), 232.

Ramayo-Caldas, Y., Mármol-Sánchez, E., Ballester, M., Sánchez, J. P., González-Prendes, R., Amills, M., and Quintanilla, R. (2019). "Integrating genome-wide co-association and gene expression to identify putative regulators and predictors of feed efficiency in pigs". *Genetics Selection Evolution*, 51(1), 48.

Ramos, A. M., Crooijmans, R. P., Affara, N. A., Amaral, A. J., Archibald, A. L., Beever, J. E., ... and Groenen, M. A. (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(8), e6524.

Raudsepp, T., and Chowdhary, B. P. (2011). "Cytogenetics and chromosome maps". Pp. 134–78 in The genetics of the pig. UK: CABI.

Renaville, B., Prandi, A., Fan, B., Sepulcri, A., Rothschild, M. F., and Piasentier, E. (2013). "Candidate gene marker associations with fatty acid profiles in heavy pigs". *Meat Science*, 93(3), 495–500.

Reverter, A., Hudson, N. J., Nagaraj, S. H., Pérez-Enciso, M., and Dalrymple, B. P. (2010). "Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data". *Bioinformatics*, 26(7), 896–904.

Revilla, M., Puig-Oliveras, A., Crespo-Piazuelo, D., Criado-Mesas, L., Castelló, A., Fernández, A. I., Ballester, M., and Folch, J. M. (2018). "Expression analysis of candidate genes for fatty acid composition in adipose tissue and identification of regulatory regions". *Scientific Reports*, 8(1), 1–13.

Roberts, M. C., Mickelson, J. R., Patterson, E. E., Nelson, T. E., Jane Armstrong, P., Brunson, D. B., and Hogan, K. (2001). "Autosomal dominant canine malignant hyperthermia is caused by a mutation in the gene encoding the skeletal muscle calcium release channel (RYR1)". *Anesthesiology*, 95(3), 716–725.

Rosenvold, K., and Andersen, H. J. (2003). "Factors of significance for pork quality—a review". *Meat Science*, 64(3):219–237.

Ros-Freixedes, R., Gol, S., Pena, R. N., Tor, M., Ibáñez-Escriche, N., Dekkers, J. C. M., and Estany, J. (2016). "Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in duroc pigs". *PLOS ONE*, 11(3), e0152496.

Ruiz-Carrascal, J., J. Ventanas, R. Cava, A. I. Andrés, and C. García. (2000). "Texture and appearance of dry cured ham as affected by fat content and fatty acid composition". *Food Research International*, 33(2):91–95.

Santos, O.O., Carbonera, F., Vagula, J. M., Maruyama, S. A., and Visentainer, J. V. (2017). "Incorporation and Bioconversion of Omega-3 Fatty Acids for Obtention of Enriched Fish". In Food Bioconversion (pp. 385-409). Academic Press.

Sargsyan, A., Doridot, L., Hannou, S. A., Tong, W., Srinivasan, H., Ivison, R., Monn, R., Kou, H. H., Haldeman, J. M., Arlotto, M., White, P. J., Grimsrud, P. A., Astapova, I., Tsai,

L. T., and Herman, M. A. (2023). "HGFAC is a ChREBP-regulated hepatokine that enhances glucose and lipid homeostasis". *JCI Insight*, 8(1).

Schook, L. B., Beever, J. E., Rogers, J., Humphray, S., Archibald, A., Chardon, P., ... and Eversole, K. (2005). "Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome". *Comparative and functional genomics*, 6(4), 251-255.

Segura, J., and Lopez-Bote, C. J. (2014). "A laboratory efficient method for intramuscular fat analysis". *Food Chemistry*, 145:821–825.

Sellier, P. (1994). "The future role of molecular genetics in the control of meat production and meat quality." Meat Science 36(1–2):29–44.

Sellier, P. (1998). "Genetic improvement of the pig". Pp. 511–40. UK: CABI.

Seppänen-Laakso, T., Laakso, I., and Hiltunen, R. (2002). "Analysis of fatty acids by gas chromatography, and its relevance to research on health and nutrition". *Analytica Chimica Acta*, 465(1–2):39–62.

Sharpe, A., and McKenzie, M. (2018). "Mitochondrial fatty acid oxidation disorders associated with short-chain enoyl-CoA hydratase (ECHS1) deficiency". *Cells*, 7(6):46.

Sladek, R., Bader, J. A., and Giguère, V. (1997). "The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene". *Molecular and Cellular Biology*, 17(9), 5400–5409.

Sodhi, S. S., Song, K.-D., Ghosh, M., Sharma, N., Lee, S. J., Kim, J. H., Kim, N., Mongre, R. K., Adhikari, P., Kim, J. Y., Hong, S. P., Oh, S. J., and Jeong, D. K. (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(2), 233–242.

Solé, E., Ros-Freixedes, R., Tor, M., Pena, R. N., and Estany, J. (2021). "A sequence variant in the diacylglycerol O-acyltransferase 2 gene influences palmitoleic acid content in pig muscle". *Scientific Reports*, 11(1), 14797.

Spector, A. A. (1999). "Essentiality of fatty acids". *Lipids*, 34(1):S1–S3.

Stachowiak, M., Nowacka-Woszuk, J., Szydlowski, M., and Switonski, M. (2013). "The ACACA and SREBF1 genes are promising markers for pig carcass and performance traits, but not for fatty acid content in the longissimus dorsi muscle and adipose tissue". *Meat Science*, 95(1), 64–71.

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., and Lancet, D. (2016). "The genecards suite: from gene data mining to disease genome sequence analyses". *Current Protocols in Bioinformatics*, 54(1).

Stroud, C. K., Nara, T. Y., Roqueta-Rivera, M., Radlowski, E. C., Lawrence, P., Zhang, Y., Cho, B. H., Segre, M., Hess, R. A., Brenna, J. T., Haschek, W. M., and Nakamura, M. T. (2009). "Disruption of FADS2 gene in mice impairs male reproduction and causes dermal and intestinal ulceration". *Journal of Lipid Research*, 50(9), 1870–1880.

Taniguchi, M., Arakawa, A., Motoyama, M., Nakajima, I., Nii, M., and Mikawa, S. (2015). "Genomic structural analysis of porcine fatty acid desaturase cluster on chromosome 2". *Animal Science Journal*, 86(4), 369–377.

Tikk, Kaja, Meelis Tikk, Margit D. Aaslyng, Anders H. Karlsson, Gunilla Lindahl, and Henrik J. Andersen. (2007). "Significance of fat supplemented diets on pork quality – connections between specific fatty acids and sensory attributes of pork". *Meat Science*. 77(2):275–286.

Timperio, A. M., Gevi, F., Longo, V., Pallotta, V., and Zolla, L. (2013). "Proteomics and transcriptomics investigation on longissimus muscles in Large White and Casertana pig breeds". In Farm animal proteomics (pp. 298–301). Wageningen Academic Publishers.

Toorop, Alfons I., Dale R. Romsos, and Gilbert A. Leveille. (2016). "The metabolic fate of dietary 1-[14C] linoleate and 1-[14C] palmitate in meal-fed rats". *Proceedings of the Society for Experimental Biology and Medicine*, 160(3):312–316.

Torres-Pitarch, A., V. Moset, P. Ferrer, M. Cambra-López, P. Hernández, J. Coma, M. Pascual, P. Serrano, and A. Cerisuelo. (2014). "The inclusion of rapeseed meal in fattening pig diets, as a partial replacer of soybean meal, alters nutrient digestion, faecal composition and biochemical methane potential from faeces". *Animal Feed Science and Technology*, 198:215–223.

TRARG1 trafficking regulator of GLUT4 (SLC2A4) 1 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=286753.

Uemoto, Y., Nakano, H., Kikuchi, T., Sato, S., Ishida, M., Shibata, T., Kadowaki, H., Kobayashi, E., and Suzuki, K. (2012). "Fine mapping of porcine SSC14 QTL and SCD gene effects on fatty acid composition and melting point of fat in a Duroc purebred population". *Animal Genetics*, 43(2), 225–228.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., and Posthuma, D. (2021). "Genome-wide association studies", *Nature Reviews Methods Primers*, 2021 1:1, 1(1), 1–21.

Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C.S., Buys, N., Tally, M., Andersson, G., Georges, M., and Andersson, L. (2003). "A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig". *Nature*, 425(6960), 832–836.

van Son, M., Enger, E. G., Grove, H., Ros-Freixedes, R., Kent, M. P., Lien, S., and Grindflek, E. (2017). "Genome-wide association study confirm major QTL for backfat fatty acid composition on SSC14 in Duroc pigs". *BMC Genomics*, 18(1):369.

Varona, L., Óvilo, C., Clop, A., Noguera, J. L., Pérez-Enciso, M., Coll, A., Folch, J. M., Barragán, C., Toro, M. A., Babot, D., and Sánchez, A. (2002). "QTL mapping for growth and carcass traits in an Iberian by Landrace pig intercross: additive, dominant and epistatic effects". *Genetical Research*, 80(2), 145–154.

Viterbo, V. S., Lopez, B. I. M., Kang, H., Kim, H., Song, C., and Seo, K.S. (2018). "Genome wide association study of fatty acid composition in Duroc swine". *Asian-Australasian Journal of Animal Sciences*, 31(8), 1127–1133.

Wajner, M., and Amaral, A.U. (2016). "Mitochondrial dysfunction in fatty acid oxidation disorders: insights from human and animal studies". *Bioscience Reports*, 36(1).

Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983). "Fatty acid synthesis and its regulation". *Annual review of biochemistry*, 52(1), 537-579.

Wang, D., Zhang, Y., and Shen, C. (2020). "Research update on the association between SFRP5, an anti-inflammatory adipokine, with obesity, type 2 diabetes mellitus and coronary heart disease". *Journal of Cellular and Molecular Medicine*, 24(5), 2730–2735.

Wang, H., Wang, X., Yan, D., Sun, H., Chen, Q., Li, M., Dong, X., Pan, Y., and Lu, S. (2022). "Genome-wide association study identifying genetic variants associated with carcass backfat thickness, lean percentage and fat percentage in a four-way crossbred pig population using SLAF-seq technology". *BMC Genomics*, 23(1), 594.

Wang, W., Huang, X., Xin, H.-B., Fu, M., Xue, A., and Wu, Z.-H. (2015). "TRAF family member-associated NF-κB activator (TANK) inhibits genotoxic nuclear factor κB activation by facilitating deubiquitinase USP10-dependent deubiquitination of TRAF6 ligase". *Journal of Biological Chemistry*, 290(21), 13372–13385.

Wang, W., Xue, W., Jin, B., Zhang, X., Ma, F., and Xu, X. (2013). "Candidate gene expression affects intramuscular fat content and fatty acid composition in pigs". *Journal of Applied Genetics*, 54(1), 113–118.

Wang, Y.L., Faiola, F., Xu, M., Pan, S., and Martinez, E. (2008). "Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein". *Journal of Biological Chemistry*, 283(49), 33808–33815.

WHO. (2019). "Coronavirus disease (COVID-19) pandemic." Retrieved January 11, 2023 (https://www.who.int/europe/emergencies/situations/covid-19).

Wimmers, K., Murani, E., Te Pas, M. F. W., Chang, K. C., Davoli, R., Merks, J. W. M., Henne, H., Muraniova, M., Da Costa, N., Harlizius, B., Schellander, K., Böll, I., Braglia, S., De Wit, A. A. C., Cagnazzo, M., Fontanesi, L., Prins, D., and Ponsuksili, S. (2007). "Associations of functional candidate genes derived from gene-expression profiles of prenatal porcine muscle tissue with meat quality and muscle deposition". *Animal Genetics*, 38(5), 474–484.

Wood, J. D., M. Enser, A. V. Fisher, G. R. Nute, P. R. Sheard, R. I. Richardson, S. I. Hughes, and F. M. Whittington. (2008). "Fat deposition, fatty acid composition and meat quality: a review". *Meat Science*, 78(4):343–358.

Wood, J. R., Richardson, G., Nute, A., Fisher, M., Campo, E., Kasapidou, P., Sheard, and M. Enser. (2003). "Effects of fatty acids on meat quality: a review". *Meat Science*, 66(1):21–32.

Wu, Jiayan, Jingfa Xiao, Zhang Zhang, Xumin Wang, Songnian Hu, and Jun Yu. (2014). "Ribogenomics: the science and knowledge of RNA". *Genomics, Proteomics and Bioinformatics,* 12(2):57–63.

Wu, W., Yin, Y., Xu, K., Peng, Y., and Zhang, J. (2018). "Knockdown of LGALS12 inhibits porcine adipocyte adipogenesis via PKA&ndash;Erk1/2 signaling pathway". *Acta Biochimica et Biophysica Sinica*, 50(10), 960–967.

Xing, Kai, Feng Zhu, Liwei Zhai, Shaokang Chen, Zhen Tan, Yangyang Sun, Zhuocheng Hou, and Chuduan Wang. (2016). "Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data". *Scientific Reports*, 6(1):1–10.

Xu, P., Ni, L., Tao, Y., Ma, Z., Hu, T., Zhao, X., Yu, Z., Lu, C., Zhao, X., and Ren, J. (2020). "Genome-wide association study for growth and fatness traits in Chinese Sujiang pigs". *Animal Genetics*, 51(2), 314–318.

Yu, B. T. P., Wang, L., Tuggle, C. K., and Rothschild, M. F. (1999). "Mapping genes for fatness and growth on pig chromosome 13: a search in the region close to the pig PIT1 gene". *Journal of Animal Breeding and Genetics*, 116(4), 269–280.

Yu, K., Shu, G., Yuan, F., Zhu, X., Gao, P., Wang, S., Wang, L., Xi, Q., Zhang, S., Zhang, Y., Li, Y., Wu, T., Yuan, L., and Jiang, Q. (2013). "Fatty Acid and Transcriptome Profiling of Longissimus Dorsi Muscles between Pig Breeds Differing in Meat Quality". *International Journal of Biological Sciences*, 9(1), 108–118.

Zappaterra, M., Luise, D., Zambonelli, P., Mele, M., Serra, A., Costa, L. N., and Davoli, R. (2019). "Association study between backfat fatty acid composition and SNPs in candidate genes highlights the effect of FASN polymorphism in large white pigs". *Meat Science*, 156, 75–84.

Zhang, D., Wang, L., Ma, S., Ma, H., and Liu, D. (2023). "Characterization of pig skeletal muscle transcriptomes in response to low temperature". *Veterinary Medicine and Science*, 9(1), 181–190.

Zhang, J., Cui, L., Ma, J., Chen, C., Yang, B., and Huang, L. (2017). "Transcriptome analyses reveal genes and pathways associated with fatty acid composition traits in pigs". *Animal Genetics*, 48(6), 645–652.

Zhang, J., Zhang, Y., Gong, H., Cui, L., Ma, J., Chen, C., ... and Yang, B. (2019). "Landscape of loci and candidate genes for muscle fatty acid composition in pigs revealed by multiple population association analysis". *Frontiers in genetics*, 10, 1067.

Zhang, P., Li, Q., Wu, Y., Zhang, Y., Zhang, B., and Zhang, H. (2022). "Identification of candidate genes that specifically regulate subcutaneous and intramuscular fat deposition using transcriptomic and proteomic profiles in Dingyuan pigs". *Scientific Reports*, 12(1), 2844.

Zhang, Wanchang, Bin Yang, Junjie Zhang, Leilei Cui, Junwu Ma, Congying Chen, Huashui Ai, Shijun Xiao, Jun Ren, and Lusheng Huang. (2016). "Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations". *Scientific Reports*, 6(1):24718.

Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., and Shen, F. (2019). "Mechanisms and functions of long non-coding RNAs at multiple regulatory levels". *International Journal of Molecular Sciences*, *20*(22), 5573.

Zhang, Y., Manjunath, M., Kim, Y., Heintz, J., and Song, J.S. (2019). "SequencEnG: an interactive knowledge base of sequencing techniques". *Bioinformatics*, 35(8), 1438-1440.

Zhao, X., Mo, D., Li, A., Gong, W., Xiao, S., Zhang, Y., ... and Chen, Y. (2011). "Comparative analyses by sequencing of transcriptomes during skeletal muscle development between pig breeds differing in muscle growth rate and fatness". *PloS one*, 6(5), e19774.

_____

Chapter 7. Annexes

All supplementary tables, figures, and related documents included and referred in the present Ph.D. thesis and papers and studies section are available at their corresponding online versions. Additionally, all supplementary material is publicly available and can be downloaded from the following link:

VALDES, JESUS (2023), "Annexes. Paper II. Supplementary material [Sci. Rep. 13, 535 (2023)]", Mendeley Data, V1, doi: 10.17632/yx4ykgrdfg.1

https://data.mendeley.com/drafts/yx4ykgrdfg

VALDES, JESUS (2023), "Annexes. Paper III. Supplementary material [GSEV-D-23-00032 (2023)]", Mendeley Data, V1, doi: 10.17632/myr7vxwh5t.1

https://data.mendeley.com/drafts/myr7vxwh5t

_____

Aknowledgments

El tiempo de reflexionar y agradecer es grato y me conecta con la vida real. Las cosas que nos suceden tienen un porque, aunque a veces inexplicable. Miramos hacia atrás y vemos donde estamos ahora, hemos sobrevivido ¡qué dicha!, muchas personas e instituciones que convergen y colaboran para que uno transite en busca de los sueños y metas. Hoy te hablo de uno, la tesis doctoral resumida en un transcurso de cuatro años.

Quisiera agradecer:

Al Dr. Josep M. Folch Albareda (profe Folch) por brindarme la oportunidad y respaldo para realizar la tesis doctoral en su grupo de investigación. Lograr una ayuda FI-AGAUR no es fácil, ello no hubiera sido posible sin la mentoría y experiencia profesional del profe Folch. En verdad, muchas gracias por tu tiempo y paciencia para conmigo, así como por las enseñanzas brindadas en genómica animal. Usted es un CRAG profe. ¡La tesis tiene tu sello!. Así pues, donde sea que me lleve el destino siempre estaré hay para lo que necesite.

A la "Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR)" de la Generalitat de Catalunya por concederme la ayuda FI de tres años, y a la "Universitat Autònoma de Barcelona" por darme respaldo para solicitud de visa de España desde el exterior, así como por concederme un año adicional de beca. Asimismo, al Centro de Investigación en Agrigenómica (CRAG) por acogerme y brindarme los recursos e instalaciones necesarios para la estadía y realización de la tesis.

Al Dr. Yuliaxis Ramayo Caldas por aceptarme como su doctorando y por presentarme a tu familia. Tu hijo es mi sobrino también. En Pinar del Río tienes una casa donde puedes visitar. Como guajiro que soy (natural de allá del campo macurijeño), he tenido la dicha de formarme

_____

y aprender múltiples cosas contigo, me has inculcado la importancia de buscar y entender el fondo de los algoritmos, y que el hecho de tunear los parámetros lleva su tiempo (un paso muy importante de los análisis bioinformáticos). ¡Gracias por tu tiempo y paciencia para conmigo, la tesis tiene también tu sello!.

A todos investigadores del grupo de Animal Genomics del CRAG. A los dres. Anna Castelló, Miguel Pérez Enciso, Sebastián Ramos, Alex Clop y Marcel Amills. Gracias por vuestros conocimientos científicos, los aportes en los seminarios y por preguntar cómo le va a uno cuando nos cruzamos en los pasillos.

A todos los colegas ya graduados que me han ayudado y compartido su sabiduría durante la realización del PhD: Lino, Laura, Yordi, Marta, María G. Luigi, Emilio y Lourdes "la Jefa". También a los compañeros de oficina CRAG-310, Natalia, Ari, Lorena y Carlos. Por su puesto, a los amigos del grupo IBMAP: Yron, Magí, Cristina, Liu y Ferraán, gracias a ustedes por la ayuda y hacer ameno el paso del tiempo durante la tesis, aún más por los códigos y por compartir en actividades recreativas (días de pizza, calçotadas, disfraces, cumpleaños, disco, etc.), "pafff" "ya tu sabes cómo es eso". A Joan por su amistad y la buena atención en la tierra de Granada. A Uriol por su amistad y siempre tener la iniciativa de organizar las actividades culturales e internacionales. A mis amigos de Italia, hermano Giacomo, Alice, Antonia y Elena, gracias por dejarme probar las delicias de la gastronomía italiana, también por los códigos e informaciones cruzadas. A los más nuevos, Leonardo, Alex, Manuel sepan que os habéis ganado mi confianza.

A Júlia Lacuesta, por su excelente gestión administrativa como secretaria del Doctorado en Producción Animal (UAB).

A mi amigo el pana Bolivar Samuel Sosa Madrid por su amistad, ayuda y conocimientos compartidos con migo. También, al amigo Ricardo por dejarme su bicicleta plegable, todos estos años me ha servido de maravilla para ir y venir a la universidad.

_____

Quiero agradecer especialmente a toda mi familia cubana y boliviana, que siempre me apoyan espiritualmente. Ambas son mi fuente de motivación para salir hacia adelante, esforzarme cada día más y poder hacer frente a la realidad que nos toca.

A los que en paz descansen "mis padres": Jesús Valdés Berra (Boruco) y María Magdalena Hernández (Magdi). Gracias a ellos tengo este ser y sé que a pesar de ya no están, siempre me acompañan con su imagen ejemplarizante. A mis queridos abuelos por parte de madre: "papi Pedro" y "mima María" y por parte de padre: "Edelmiro" y "Hilaria". Ustedes me vieron crecer y cuidaron de mí hasta mi adolescencia. A mí admirado hermano Jasiel Valdés por acompañarme en los momentos más difíciles, por salir adelante, volverse independiente y siempre estar hay luchándola.

A mi otra familia boliviana: Sr. Benedicto Miranda y Sra. Sonia Alejo por aceptarme en su familia, brindándome todo su apoyo en los momentos más necesitados, así como a sus demás hijos.

Finalmente, a Judith Carmen Miranda Alejo, por su amor, cariño, comprensión, compañía, paciencia para conmigo. Mil gracias por la confianza depositada en mí durante todos estos años, también por dedicar tu tiempo en la búsqueda y sugerencia de becas para continuar la formación post-universitaria. Asimismo, muchas gracias por considerarme como co-autor en las publicaciones derivadas de tu tesis doctoral.