UAB
Universitat Autònoma de Barcelona

INTEGRATIVE ANALYSIS OF GUT MICROBIOME, GENE EXPRESSION, AND LIPID METABOLISM IN PIGS

Cristina Sebastià Navarro

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: https://creativecommons.org/licenses/?lang=ca

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: https://creativecommons.org/licenses/?lang=es

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Ciència Animal i dels Aliments

Facultat de Veterinària

CENTRE DE RECERCA EN AGRIGENÒMICA

Grup de Recerca de Genòmica Animal

INTEGRATIVE ANALYSIS OF GUT MICROBIOME, GENE EXPRESSION, AND LIPID METABOLISM IN PIGS

Cristina Sebastià Navarro

Doctoral thesis to obtain the PhD degree in Animal Production of the Universitat Autònoma de Barcelona, June 2024

Supervisors

Dr. Josep Maria Folch Albareda

Dr. Daniel Crespo Piazuelo

Dr. Josep Maria Folch Albareda, Professor of the Department of Animal and Food Science at the Universitat Autònoma de Barcelona (UAB), and Dr. Daniel Crespo Piazuelo, researcher at the R&D Department of Cuarte S.L. (Grupo Jorge),

state

that the research work and the writing of the doctoral thesis titled '*Integrative analysis of gut microbiome, gene expression, and lipid metabolism in pigs*' have been carried out under their supervision by

CRISTINA SEBASTIÀ NAVARRO

and certify

that this work has been conducted at the Department of Animal and Food Science of the Faculty of Veterinary Medicine at the Universitat Autònoma de Barcelona and at the Animal Genomics unit of the Center for Research in Agricultural Genomics,

considering

that the resulting thesis is suitable for the degree of Doctor in Animal Production from the Universitat Autònoma de Barcelona,

and to provide proof of this, they sign this document in Bellaterra,

on June 25, 2024.

Dr. Josep Maria Folch Albareda

Dr. Daniel Crespo Piazuelo

Cristina Sebastià Navarro

This thesis was funded by the *Ministerio de Ciencia, Innovación y Universidades* (MCIN: MCIN/AEI/10.13039/501100011033) and by "ERDF A way of making Europe", projects AGL2017-82641-R and PID2020-12677RB-C22. We acknowledge the Spanish Ministry of Science, Innovation and Universities for the "*Severo Ochoa Programme for Centres of Excellence in R&D*" (SEV-2015-0533 and CEX2019-000902-S) awarded to the Centre for Reseach in Agricultural Genomics (CRAG) and by the CERCA Programme, Generalitat de Catalunya.

Cristina Sebastià Navarro was funded with a "*Formació i Contractació de Personal Investigador Novell* (FI-DGR)" fellowship provided by the "Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR)" from the *Generalitat de Catalunya* (2020FI_B_00225).

Cuando sea de noche miraré las estrellas
Y allá donde brillen Elnath y Aldebarán sonreiré
Sabiendo que allí brilla también el pilar de mi tierra
Y solo podré decir “Bona nit”
Hasta que nos volvamos a ver

Table of contents

Summary **13**
List of tables **19**
List of figures **21**
List of publications **27**
1. General introduction **29**
1.1. Pig: origins, production, physiology 31
1.1.1. Pig domestication 31
1.1.2. Pig meat production 32
1.1.3. Pig meat quality 33
1.1.3.1. Fatty acid composition 34
1.1.4. Fatty acid metabolism 35
1.1.4.1. Fatty acid biosynthesis 35
1.1.4.2. Fatty acid β-oxidation 37
1.2. Gut microbiota 39
1.2.1. Micro-concepts 39
1.2.2. Gut microbiota and the host 39
1.2.3. SCFA as microbiota metabolites 41
1.2.3.1. Acetate 41
1.2.3.2. Propionate 42
1.2.3.3. Butyrate 42
1.3. Omes and omics 43
1.3.1. Sequencing technologies 43
1.3.2. Transcriptomics 48
1.3.3. Microbiomics 49
1.3.3.1. 16S rRNA gene sequencing 50
1.3.3.2. Whole-metagenome shotgun sequencing 52
1.3.3.3. Metatranscriptome shotgun sequencing 53
1.3.4. Data integration 54
2. Objectives **57**
3. Articles and studies **61**
Paper I. Gene co-expression network analysis for porcine intramuscular fatty acid composition 63
Paper II. Interrelation between gut microbiota, SCFA and fatty acid composition in pigs 97
Paper III. Metagenomic and metatranscriptomic interplay between the pig gut microbiota and short-chain fatty acids 133
Paper IV. Gut microbial genera as predictors of liver gene expression in pigs 161
4. General discussion **189**
4.1. Association between gene expression and fatty acid composition 193
4.2. Association between microbiota and liver gene expression 196
4.3. Gut microbial composition comparison 200

4.4. Comparison between 16S, metagenome and metatranscriptome..... 203
4.5. Association between the gut microbiota and the fatty acid composition..... 206
4.6. Microbiota - gene expression - lipid metabolism..... 208
4.7. Future perspectives.....210
5. Conclusions.....213
6. References.....217
7. Annexes..... 253
7.1. Supplementary material Paper I: “Gene co-expression network analysis for porcine intramuscular fatty acid composition”.....255
7.2. Supplementary material Paper II: “Interrelation between gut microbiota, SCFA and fatty acid composition in pigs”..... 261
7.3. Supplementary material Paper III: “Metagenomic and metatranscriptomic interplay between the pig gut microbiota and short-chain fatty acids..... 262
7.4. Supplementary material Paper IV.....265
Acknowledgements..... 287

Summary

The pig is an important species in production and research, as pork is one of the most demanded meats worldwide and pig animal models are key for different research fields such as gut health, cardiovascular diseases or xenotransplants. Therefore, it is important to deepen the knowledge of the complex interactions between the gene expression in different tissues, the gut microbiota and the fatty acid (FA) composition in pigs.

We performed a *Weighted Gene Co-expression Network Analysis* (WGCNA) and identified co-expressed gene modules in muscle tissue of a backcrossed Iberian × Duroc pig population, revealing significant correlations with FA profiles. Notably, two gene modules were positively correlated with saturated and monounsaturated FAs, and negatively correlated with polyunsaturated FAs and the omega-6/omega-3 ratio, highlighting pathways related to insulin and FA biosynthesis. Conversely, another two modules showed positive correlations with polyunsaturated FAs and omega-6/omega-3 ratio, and negative correlations with saturated and monounsaturated FAs, linked to pathways involved in adipocytokine signalling and FA β-oxidation in the mitochondria.

Further, we explored the relationship between the gut microbiota and host metabolism, particularly the production of short-chain fatty acids (SCFAs) from non-digestible compounds, and their subsequent accumulation in tissues. Significant associations were found between bacterial and archaeal genera with FA profiles in rectal content, backfat, and muscle of Duroc × Iberian pigs. Specifically, the abundance of butyric acid was positively correlated with *Prevotella* and negatively correlated with *Akkermansia*, while exhibiting the inverse relationship between these two genera and the relative abundance of acetic acid. Additionally, *Rikenellaceae RC9 gut group* was positively associated with palmitic acid in muscle and negatively correlated with stearic acid in backfat and oleic acid in muscle, suggesting the indirect role of the microbiota in modulating FA composition.

Then, we studied the differences between the metagenome and metatranscriptome of Duroc × Iberian pigs with extreme differences in SCFAs content in their rectal content. At genus level, *Clostridium* was the most active, despite being third in abundance. The most abundant genera were *Bacteroides* and *Prevotella*, although *Treponema* was the second most active genus. In addition, some genera such as *Dorea*, *Dysgonomonas* and *Butyriciccocus* were more active than abundant. At phylum level, Spirochaetota showed higher expression activity than expected relative to its abundance. Differential metatranscriptomic analysis revealed that *Clostridium* and *Lactobacillus* species were more active in samples with high

acetic acid and low propionic acid content. Additionally, *Akkermansia glycaniphila* was identified as an acetate producer, while *Prevotella copri* was more active in high butyric acid samples, supporting its role in butyrate production via the butyryl-CoA pathway. Functional analysis highlighted significant expression of spore-related proteins and bacteriophage activity, such as antB, which can be explained by the harsh conditions of the environment.

Lastly, the liver's crucial role in FA metabolism and its interplay with gut microbiota and immune function was investigated. We identified several bacterial genera, such as *Agathobaculum*, *Collinsella*, *Floccifex*, *Phascolarctobacterium*, *SFLA01* and *Streptococcus*, which potentially influence the expression of key liver genes related to energy metabolism. Genes such as *ABCD3*, *CXCL8*, *CYP8B1*, *DEFB1*, *ELOVL2*, *FADS2*, *HMGCS2*, *MZB1* and *SCD* were identified as being modulated by these microbial predictors, suggesting potential avenues for microbiota-based modulation of gene expression.

In summary, this work provides comprehensive insights into the connections between gene networks, gut microbiome, and lipid metabolism, offering potential strategies for enhancing meat quality through targeted genetic and microbial interventions.

Resumen

El cerdo es una especie importante tanto en producción como en investigación, ya que la carne de cerdo es una de las más demandadas en todo el mundo y los modelos animales porcinos son clave en diversos campos de investigación como la salud intestinal, las enfermedades cardiovasculares o los xenotrasplantes. Por lo tanto, es importante profundizar en el conocimiento de las complejas interacciones entre la expresión génica en diferentes tejidos, la microbiota intestinal y la composición de ácidos grasos (AGs) en los cerdos.

Realizamos un análisis de redes de coexpresión génica ponderada (WGCNA) e identificamos módulos de genes coexpresados en el tejido muscular de una población de cerdos Ibérico × Duroc retrocruzada, revelando correlaciones significativas con el perfil de AGs. Notablemente, dos módulos génicos se correlacionaron positivamente con ácidos grasos saturados y monoinsaturados, y negativamente con ácidos grasos poliinsaturados y la proporción omega-6/omega-3, destacando vías relacionadas con la insulina y la biosíntesis de AGs. Por el contrario, otros dos módulos mostraron correlaciones positivas con ácidos grasos poliinsaturados y la proporción omega-6/omega-3, y correlaciones negativas con ácidos grasos saturados y monoinsaturados, vinculados a vías involucradas en la señalización de adipocitoquinas y la β-oxidación de ácidos grasos en las mitocondrias.

Además, exploramos la relación entre la microbiota intestinal y el metabolismo del huésped, particularmente la producción de ácidos grasos de cadena corta (AGCCs) a partir de compuestos no digestibles, y su posterior acumulación en los tejidos. Se encontraron asociaciones significativas entre géneros bacterianos y arqueales con los perfiles de AGs en el contenido rectal, la grasa dorsal y el músculo de los cerdos Duroc × Ibérico. Específicamente, la abundancia de ácido butírico se correlacionó positivamente con *Prevotella* y negativamente con *Akkermansia*, mientras que mostraba la relación inversa entre estos dos géneros y la abundancia relativa de ácido acético. Además, *Rikenellaceae RC9* se asoció positivamente con el ácido palmítico en el músculo y se correlacionó negativamente con el ácido esteárico en la grasa dorsal y el ácido oleico en el músculo, sugiriendo el papel indirecto de la microbiota en la modulación de la composición de AGs.

Luego, estudiamos las diferencias entre el metagenoma y el metatranscriptoma de cerdos Duroc × Ibérico con diferencias extremas en el contenido de AGCCs en su contenido rectal. A nivel de género, *Clostridium* fue el más activo, a pesar de ser el tercero en abundancia. Los géneros más abundantes fueron *Bacteroides* y *Prevotella*, aunque *Treponema* fue el

segundo género más activo. Además, algunos géneros como *Dorea*, *Dysgonomonas* y *Butyricicoccus* fueron más activos que abundantes. A nivel de filo, Spirochaetota mostró una actividad de expresión mayor de la esperada en relación con su abundancia. El análisis metatranscriptómico diferencial reveló que las especies de *Clostridium* y *Lactobacillus* fueron más activas en muestras con alto contenido de ácido acético y bajo contenido de ácido propiónico. Además, se identificó a *Akkermansia glycaniphila* como un productor de acetato, mientras que *Prevotella copri* fue más activa en muestras con alto contenido de ácido butírico, respaldando su papel en la producción de butirato a través de la vía de butiril-CoA. El análisis funcional destacó una expresión significativa de proteínas relacionadas con esporas y actividad de bacteriófagos, como antB, lo que puede explicarse por las condiciones adversas del entorno.

Por último, se investigó el papel crucial del hígado en el metabolismo de AGs y su interacción con la microbiota intestinal y la función inmune. Identificamos varios géneros bacterianos, como *Agathobaculum*, *Collinsella*, *Floccifex*, *Phascolarctobacterium*, *SFLA01* y *Streptococcus*, que potencialmente influyen en la expresión de genes clave del hígado relacionados con el metabolismo energético. Genes como *ABCD3*, *CXCL8*, *CYP8B1*, *DEFB1*, *ELOVL2*, *FADS2*, *HMGCS2*, *MZB1* y *SCD* fueron identificados como modulados por estos predictores microbianos, sugiriendo posibles vías para la modulación de la expresión génica basada en la microbiota.

En resumen, este trabajo proporciona una visión integral de las conexiones entre las redes génicas, la microbiota intestinal y el metabolismo de los lípidos, ofreciendo estrategias potenciales para mejorar la calidad de la carne a través de intervenciones genéticas y microbianas dirigidas.

Resum

El porc és una espècie important tant en producció com en investigació, ja que la carn de porc és una de les més demandades a tot el món i els models animals porcí són clau en diversos camps de recerca com la salut intestinal, les malalties cardiovasculars o els xenotrasplantaments. Per tant, és important aprofundir en el coneixement de les complexes interaccions entre l'expressió gènica en diferents teixits, la microbiota intestinal i la composició d'àcids grassos (AGs) en els porcs.

Vam realitzar una anàlisi de xarxes de coexpressió gènica ponderada (WGCNA) i vam identificar mòduls de gens coexpressats en el teixit muscular d'una població de porcs Ibèric × Duroc retrocreuada, revelant correlacions significatives amb el perfil d'AGs. Notablement, dos mòduls gènics es van correlacionar positivament amb àcids grassos saturats i monoinsaturats, i negativament amb àcids grassos poliinsaturats i la proporció omega-6/omega-3, destacant vies relacionades amb la insulina i la biosíntesi d'AGs. Per contra, altres dos mòduls van mostrar correlacions positives amb àcids grassos poliinsaturats i la proporció omega-6/omega-3, i correlacions negatives amb àcids grassos saturats i monoinsaturats, vinculats a vies involucrades en la senyalització d'adipocitoquines i la β-oxidació d'àcids grassos en les mitocòndries.

A més, vam explorar la relació entre la microbiota intestinal i el metabolisme de l'hoste, particularment la producció d'àcids grassos de cadena curta (AGCCs) a partir de compostos no digeribles, i la seva posterior acumulació en els teixits. Es van trobar associacions significatives entre gèneres bacterians i arqueals amb els perfils d'AGs en el contingut rectal, la greix dorsal i el múscul dels porcs Duroc × Ibèric. Específicament, l'abundància d'àcid butíric es va correlacionar positivament amb *Prevotella* i negativament amb *Akkermansia*, mentre que mostrava la relació inversa entre aquests dos gèneres i l'abundància relativa d'àcid acètic. A més, *Rikenellaceae RC9* es va associar positivament amb l'àcid palmític en el múscul i es va correlacionar negativament amb l'àcid esteàric en greix dorsal i l'àcid oleic en el múscul, suggerint el paper indirecte de la microbiota en la modulació de la composició d'AGs.

Després, vam estudiar les diferències entre el metagenoma i el metatranscriptoma dels porcs Duroc × Ibèric amb diferències extremes en el contingut de AGCCs en el seu contingut rectal. A nivell de gènere, *Clostridium* va ser el més actiu, malgrat ser el tercer en abundància. Els gèneres més abundants van ser *Bacteroides* i *Prevotella*, tot i que *Treponema* va ser el segon gènere més actiu. A més, alguns gèneres com *Dorea*,

Dysgonomonas i *Butyricicoccus* van ser més actius que abundants. A nivell de fílum, Spirochaetota va mostrar una activitat d'expressió més gran de l'esperada en relació amb la seva abundància. L'anàlisi metatranscriptòmic diferencial va revelar que les espècies de Clostridium i Lactobacillus van ser més actives en mostres amb alt contingut d'àcid acètic i baix contingut d'àcid propiònic. A més, es va identificar *Akkermansia glycaniphila* com un productor d'acetat, mentre que *Prevotella copri* va ser més activa en mostres amb alt contingut d'àcid butíric, donant suport al seu paper en la producció de butirat a través de la via de butiril-CoA. L'anàlisi funcional va destacar una expressió significativa de proteïnes relacionades amb espores i activitat de bacteriòfags, com antB, que es pot explicar per les condicions adverses de l'entorn.

Finalment, es va investigar el paper crucial del fetge en el metabolisme d'AGs i la seva interacció amb la microbiota intestinal i la funció immune. Vam identificar diversos gèneres bacterians, com *Agathobaculum*, *Collinsella*, *Floccifex*, *Phascolarctobacterium*, *SFLA01* i *Streptococcus*, que potencialment influeixen en l'expressió de gens clau del fetge relacionats amb el metabolisme energètic. Gens com *ABCD3*, *CXCL8*, *CYP8B1*, *DEFB1*, *ELOVL2*, *FADS2*, *HMGCS2*, *MZB1* i *SCD* van ser identificats com a modulats per aquests predictors microbians, suggerint possibles vies per a la modulació de l'expressió gènica basada en la microbiota.

En resum, aquest treball proporciona una visió integral de les connexions entre les xarxes gèniques, la microbiota intestinal i el metabolisme dels lípids, oferint estratègies potencials per a millorar la qualitat de la carn a través d'intervencions genètiques i microbianes dirigides.

List of tables

General introduction

Table 1.1. Advantages and drawbacks of the main sequencing platforms.

Table 1.2. Omics definition and sequencing technologies.

Table 1.3. Main candidate genes found related with meat quality and involved in lipid metabolism.

Table 1.4. Main advantages of the most used bioinformatic pipelines for 16S rRNA amplicon data analysis.

Paper I

Paper I. Table 1. Mean relative abundance (%) and standard deviation (SD) values of FA composition in the *Longissimus dorsi* muscle of the 129 BC1_DU pigs.

Paper I. Table 2. Over-represented pathways for each significant module. Only the pathways with a corrected (by Benjamini and Hochberg) *p*-value of ≤ 0.05 are considered.

Paper II

Paper II. Table 1. Mean relative abundances (%) and standard deviations (SD) of the analysed fatty acids from backfat and *longissimus dorsi* muscle.

Paper II. Table 2. Mean relative abundances (%) and standard deviations (SD) of the analysed short-chain fatty acids from rectal content.

Paper III

Paper III. Table 1. Average content (%) of high and low groups for the three studied SCFAs.

Paper IV

Paper IV. Table 1. Relative abundance (%) of the main genera found and the phylum to which they belong. Genera with a lower relative abundance than 1% were grouped as “Others”.

Paper IV. Table 2. List of rCCA relevant genes and the genera with which they are connected.

General discussion

Table 4.1. Mean relative abundance (percentage) ± SD of the main fatty acids of the three studied tissues in the three different populations. Backfat and one set of the muscle samples belong to the Duroc × Iberian population (u × Ib, 288 animals), while another set of the muscle samples belong to the backcross BC1_DU (129 animals), and the liver samples belong to Duroc pigs (310 animals).

Table 4.2. List of genes whose expression was significantly correlated with the abundance of *SFLA01*.

Annexes: Paper I

Supplementary Table S1. Gene pathways and source.

Supplementary Table S2. LogCPM expression of lipid metabolism genes.

Supplementary Table S3. Gene - module membership.

Supplementary Table S4. Significantly enriched pathways.

Annexes: Paper III

Supplementary Table S1. Main programs and functions of the bioinformatic tools used for the metagenomic and metatranscriptomic analyses. a) Programs included in metaWRAP for metagenomic analysis. b) Programs included in SAMSA2 for metatranscriptomic analysis.

Supplementary Table S2. Results of the differential abundance and expression analysis for the content in acetic, propionic and butyric acids.

Annexes: Paper IV

Supplementary Table S1. Genes involved in energy metabolism, lipid metabolism or immunity based on GO keywords retrieved with DAVID database.

Supplementary Table S2. rCCA connections between genes and genera filtered for an absolute weight > 0.15. A positive weight indicates a positive relationship and a negative weight indicates a negative relationship.

List of figures

General introduction

Figure 1.1. Worldwide meat production of four of the main livestock species for the last thirty years (FAOSTAT, 2024).

Figure 1.2. Fatty acid biosynthesis in animals. Adapted from (Guillou et al., 2010; Naudí et al., 2013).

Figure 1.3. Fatty acid β-oxidation in the mitochondria. Adapted from (Abegaz et al., 2021; Sharpe & McKenzie, 2018).

Figure 1.4. Relationships between the different omes that end up defining the phenotype. Adapted from (Varona, 2017).

Figure 1.5. Structure of the 16S rRNA. Hypervariable regions are highlighted and numbered V1-V9.

Paper I

Paper I. Figure 1. MDS plot of the gene distribution. The colours represent the modules assigned for each gene.

Paper I. Figure 2. Relationships between the four significant modules identified in WGCNA and the FA phenotypes in the *Longissimus dorsi* muscle of (Iberian × Duroc) × Duroc pigs. Red colour points out the positive correlations, while blue represents the negative ones. The *p*-value of the correlation is shown between parenthesis.

Paper I. Figure 3. Phenotypic correlations between the analysed FAs in *Longissimus dorsi* muscle.

Paper I. Figure 4. Magenta module gene network of the most connected genes (genes with six or more neighbours are highlighted in magenta colour and, for visualisation purposes, genes without direct connection with these most connected items are hidden). The size of the node is based on the number of direct neighbours of the gene. The edges are coloured based on the correlation between the genes, green for positive and red for negative. In this network, all the correlations are positive.

Paper II

Paper II. Figure 1. Population diversity in the 285 analysed samples from rectal content. A. Stacked barplot for the relative abundance of genera. Genera with a total relative abundance lower than 2% are grouped as “Other”. B. Shannon’s α-diversity in the 285 samples. C. Bray-Curtis dissimilarities between samples.

Paper II. Figure 2. Heatmap of Pearson correlations between fatty acids in backfat, muscle and rectal content. “BF” extension means that the FA belongs to backfat and “LD” to longissimus dorsi. Acetic, propionic, iso- and n-butyric and iso- and n-valeric acids come from rectal content. Significant correlations were considered with a p-value < 0.05. Non-significant correlations were set as 0.

Paper II. Figure 3. Correlation circular plot between the relative abundance of short-chain fatty acids and bacterial and archaeal genera. For clarity purposes, only variables with a component value greater than |0.3| are represented.

Paper II. Figure 4. Heatmap for the significant correlations between genera and fatty acid composition in backfat (BF) and muscle (LD) and short-chain fatty acid composition in rectum. Significant correlations were considered at ρ > |0.15| and p-value < 0.05. Non-significant correlations were set as 0.

Paper II. Figure 5. NMDS plots for acetic acid (left) and n-butyric acid (right) with the fatty acid abundances and R^2 from PERMANOVA results. Colour saturation represents the SCFA relative abundance on each sample. Black lines show the estimated distribution of the SCFA relative abundances regarding the ordination of the samples based on their β-diversity values (Bray-Curtis dissimilarities). The less curved and the more parallel straight lines are represented, the higher relationship between the FA relative abundance and the microbiota diversity can be assumed.

Figure 6. Correlation circular plot between the relative abundance of the short-chain fatty acids and the KEGG orthologs related with lipid metabolism.

Paper III

Paper III. Figure 1. Stacked area plots representing the relative abundance of top 10 classified taxa of the metagenomic and metatranscriptomic analyses. a) Metagenomic results at phylum level. The rest of the 49 annotated phyla were grouped as “Other”, which accounted for 2.74% of the total abundance. b) Metagenomic results at genus level. The rest

of the 1,456 annotated genera were grouped as “Other”, which accounted for 54.76% of the total abundance. c) Metatranscriptomic results at phylum level. The rest of the 28 annotated phyla were grouped as “Other”, which accounted for 0.76% of the total expression. d) Metatranscriptomic results at genus level. The rest of the 1,884 annotated taxa were grouped as “Other”, which accounted for 37.07% of the total expression.

Paper III. Figure 2. Stacked bar plots representing the mean of the relative abundance of (A) phyla and (B) genera found in the metagenome (n = 20) and metatranscriptome (n = 24) of pigs.

Paper III. Figure 3. Scatter plot showing the abundance of genera in metagenome and metatranscriptome results. Blue dots represent the 20 most abundant genera, while red dots highlight the 20 genera with the greatest differences between their metagenomic and metatranscriptomic abundances.

Paper III. Figure 4. Venn diagram representing the shared species between the metagenome and metatranscriptome datasets. Species with a relative abundance lower than 0.005% were filtered in both datasets.

Paper III. Figure 5. PCA biplot with the sample distribution based on the three main SCFAs content. Samples are coloured depending on their batch, which was taken into account to choose the high and low animals for each SCFA.

Paper III. Figure 6. Differential abundance analysis between samples with high and low acetic acid content. a) MA plot showing the differences in bacterial activity. Species with $\log_2FC \leq -1.5$ were more active in high-acetic gut samples. Significantly active species (adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-acetic gut samples.

Paper III. Figure 7. Differential abundance analysis between samples with high and low propionic acid content. a) MA plot showing the differences in bacterial activity. Species with $\log_2FC \leq -1.5$ were more active in high-propionic gut samples. Significantly active species (adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-propionic gut samples.

Paper III. Figure 8. Differential abundance analysis between samples with high and low butyric acid content. a) MA plot showing the differences in bacterial activity. Species with $\log_2FC \leq -1.5$ were more active in high-butyric gut samples. Significantly active species

(adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-butyric gut samples.

Paper IV

Paper IV. Figure 1. Pearson correlation results. Significance threshold was set at an adjusted p-value < 0.05. For visualisation purposes, we only included the gene expressions and genera abundances that had at least one significant correlation.

Paper IV. Figure 2. Circular plot of the rCCA results. For visualisation purposes, we established a 0.2 cut-off to keep the more relevant genes (n = 203, represented in orange) and genera (n = 105, represented in blue).

Paper IV. Figure 3. Correlation network of the rCCA results. We filtered the expressed genes by lipid metabolism, energy metabolism and immunity based on the functional annotation analysis with DAVID and established a 0.15 cut-off to keep the more relevant genes (n = 32, represented as orange squared nodes) and genera (n = 31, represented as blue round nodes with a different size depending on their relative abundance). Green edges indicate positive relationships, while negative relationships are marked by magenta edges.

General discussion

Figure 4.1. Relative abundance (%) barplot of the main bacterial and archaeal phyla in the two studied populations. Phyla with a relative abundance lower than 1% have been grouped as “Other”. Light pink colour represents the samples taken from Duroc piglets and analysed with Greengenes2 (GG2), while the dark red colour represents the Duroc × Iberian crossbred pigs analysed with SILVA138.

Figure 4.2. Relative abundance (%) barplot of the main bacterial and archaeal genera in the two studied populations. Genera with a relative abundance lower than 1% have been grouped as “Other”. Light pink colour represents the samples taken from Duroc piglets and analysed with Greengenes2 (GG2), while the dark red colour represents the Duroc × Iberian crossbred pigs analysed with SILVA138.

Figure 4.3. Relative abundance (%) barplot of the main bacterial and archaeal phyla based on the analysis in the crossbred Duroc × Iberian population. Phyla with a relative abundance/activity lower than 1% were grouped as “Other”.

Annexes: Paper I

Supplementary Figure S1. Co-expression networks of relevant modules. Edges in green show positive relationships and edges in red show negative relationships. a) Yellow module. b) Green module. c) Brown module.

List of publications

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

- Paper I. **Sebastià, C.**, Gallopin, M., Ramayo-Caldas, Y., Estellé, J., Valdés-Hernández, J., Castelló, A., Sánchez, A., Crespo-Piazuelo, D., and Folch, J.M. (2023) 'Gene co-expression network analysis for porcine intramuscular fatty acid composition'. *animal: The international journal of animal biosciences* (submitted)

- Paper II. **Sebastià, C.**, Folch, J.M., Ballester, M., Estellé, J., Passols, M., Muñoz, M., García-Casco, J.M., Fernández, A.I., Castelló, A., Sánchez, A., and Crespo-Piazuelo, D. (2024) 'Interrelation between gut microbiota, SCFA, and fatty acid composition in pigs'. *mSystems*, 9(1), p. e01049-23.

- Paper III. **Sebastià, C.**, Folch, J.M., Estellé, J., … Castelló, A., Sánchez, A., and Crespo-Piazuelo, D. (2024) 'Metagenomic and metatranscriptomic comparison of the pig gut microbiota'. (in preparation)

- Paper IV. **Sebastià, C.**, Crespo-Piazuelo, D., Ramayo-Caldas, Y., Ballester, M., Liu, J., Llobet-Cabau, F., Castelló, A., and Folch, J.M. (2024) 'Gut microbial genera as predictors of liver gene expression in pigs'. (in preparation)

Related publications

(Not included in the thesis.) The author also appears in the following publications:

- 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, J.M. (2023) 'Global analysis of the association between pig muscle fatty acid composition and gene expression using RNA-Seq'. *Scientific Reports*, 13(1), p. 535.
- Ramayo-Caldas, Y., Crespo-Piazuelo, D., Morata, J., González, O., **Sebastià, C.**, Castelló, A., Dalmau, A., Ramos-Onsins, S., Alexiou, K., Folch, J.M., Quintanilla, R., Ballester, M., and Gil, R. (2023). 'Copy number variation on ABCC2-DNMBP loci affects the diversity and composition of the fecal microbiota in pigs'. *Microbiology Spectrum*, 11(4), p. e05271-22.
- Passols, M., Llobet-Cabau, F., **Sebastià, C.**, Castelló, A., Valdés-Hernández, J., Criado-Mesas, L., Sánchez, A., and Folch, J.M. (2023). 'Identification of genomic regions, genetic variants and gene networks regulating candidate genes for lipid metabolism in pig muscle'. *animal*, 17(12), p. 101033.
- Valdés-Hernández, J., Folch, J.M., Crespo-Piazuelo, D., Passols, M., **Sebastià, C.**, Criado-Mesas, L., Castelló, A., Sánchez, A., and Ramayo-Caldas, Y. (2024). Identification of candidate regulatory genes for intramuscular fatty acid composition in pigs by transcriptome analysis. *Genet Sel Evol,* 56(12).

1. General introduction

1.1. Pig: origins, production, physiology

1.1.1. Pig domestication

Nowadays, the domestic pig species is highly interrelated with human daily life, particularly as an important source of meat, but also as a model species for research in human physiology, diseases or xenotransplants (Lunney et al., 2021). The predator-prey relation between wild boars and humans shifted ten thousand years ago, when the domestication of pigs began mainly in the Crescent Fertile and in China. Around three thousand years later, domestic pigs were initially introduced to Europe from the Near East through the Danubian and Mediterranean corridors, but they were replaced by their European relatives in a short period of time (Amills et al., 2010; Ramos-Onsins et al., 2014).

The development of different breeds, however, started millennia after the initial stages of domestication. During the Industrial Revolution, in the 18th century in Europe, there was an increase in meat consumption and demand, which caused the start of phenotypic selection for pig breeds, markedly in Britain (White, 2011). Even after improving the desired traits by crossbreeding with other breeds, including those from distant origins such as China (Giuffra et al., 2000), the European breeds remained well differentiated. This specialisation intensified with the development of modern genetics in the 20th century, which enabled a strong selection of the animals for different production traits including carcass weight, lean content or prolificacy (Muñoz et al., 2018; Poklukar et al., 2023). Consequently, modern pig breeds became highly specialised and its production increased considerably. Genetic analyses also revealed different signatures of selection of the domestic pig, providing molecular evidence of the selection process that the species has undergone (Rubin et al., 2012).

The animal material used on this thesis belonged to two different European breeds with very different levels of selection and specialisation: the Iberian and Duroc pig breeds.

The Iberian pig is a rustic breed from the Iberian Peninsula, characterised by a slow growth and high fat deposition, both subcutaneous and intramuscular. The fat composition of this breed has a characteristic higher content in oleic acid (C18:1n-9) than the rest of the commercial breeds, a trait that is influenced by its own genetics and the extensive rearing conditions, which usually includes an acorn-rich diet. This characteristic fat composition and the rest of the organoleptic qualities of the Iberian pork are very appreciated by consumers and the meat industry, where it is used for high-quality dry-cured products such as ham (López-Bote, 1998; Nieto et al., 2019).

Conversely, Duroc pigs are a relatively modern breed that has been intensely selected for over a century and has produced different lines depending on the country of breeding. They are characterised by a high resilience and an excellent growth rate, which makes them a good choice to be used as parental lines in crossbreeding. Furthermore, Duroc pigs have a higher rate of intramuscular fat deposition compared to other commercial breeds (e.g. Large White, Landrace, and Piétrain), making them the ideal parental line for the Iberian sows to enhance the growth performance of their offspring, while still obtaining high quality fresh and dry-cured crossbred products.

1.1.2. Pig meat production

Pork is one of the most produced and consumed meats in the world, along with chicken and, to a lesser extent, beef. As of the most recent data available from FAOSTAT (2022), worldwide pork production reached more than 120 million tons. Pork was the most produced meat in the world until 2018, when it was surpassed by chicken, as illustrated in Figure 1.1.

China is the biggest pork producer with more than 55 million tons in 2022, followed by the United States and Brazil. Spain is the fourth producer in the world and the first one in Europe. Germany was also one of the biggest producers in the last three decades, but it has decreased its production since 2016 (FAOSTAT, 2024).

Among the regions within Spain, Catalonia stands as one of the regions that has a major production and herd heads. Aragon and Catalonia account for more than 50% of the total herd heads in Spain, mostly commercial lines of lean breeds reared in intensive conditions (Ministerio de Agricultura, Pesca y Alimentación (MAPA), 2022).

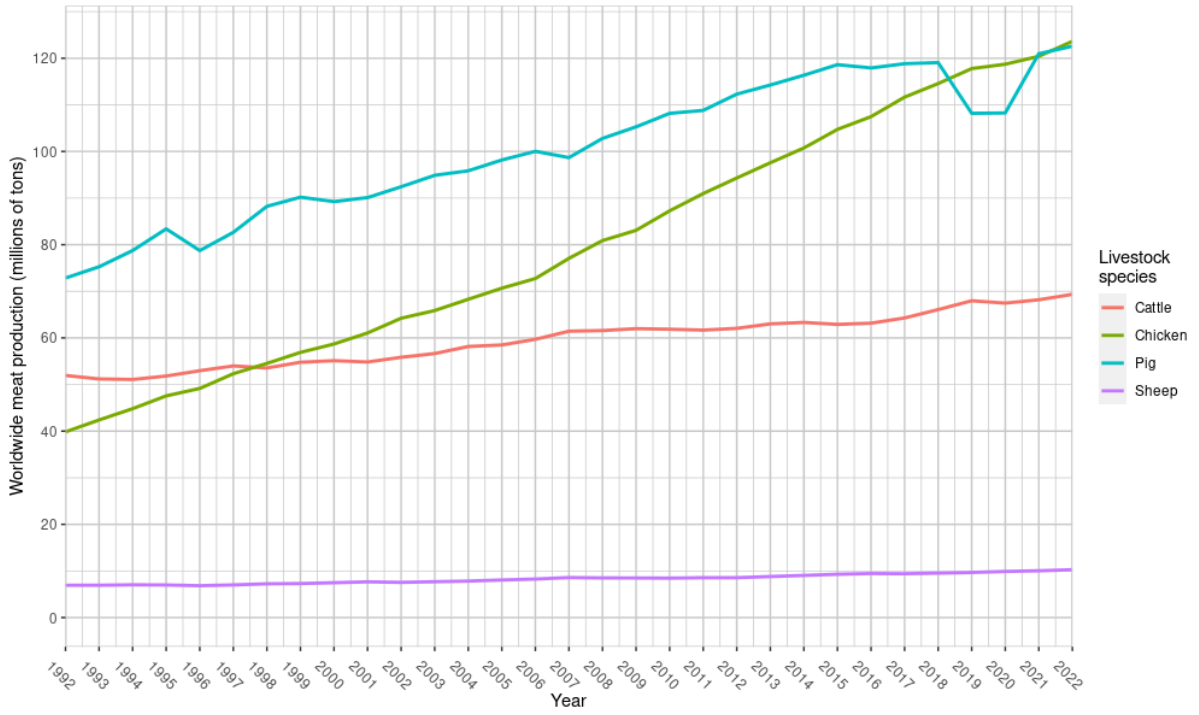


Figure 1.1. Worldwide meat production of four of the main livestock species for the last thirty years (FAOSTAT, 2024).

In the last few years, the production of the Iberian breed has increased since the start of its conservation program in the 1980s. This program was created to preserve the genetic diversity and reevaluate the breed as a high-quality meat producer (Nieto et al., 2019). The majority of Iberian pigs are bred in the West of Spain, as it is an autochthonous breed of the Iberian Peninsula. Regions such as Extremadura have extensions of oak holms known as “dehesas”, where Iberian pigs can be reared in extensive conditions and feed on acorns, a special aspect of production that gives a high-quality final product based on fatty acid profile (Nieto et al., 2019).

1.1.3. Pig meat quality

Meat quality is a complex trait influenced by numerous factors. While these factors depend on consumer perception, some of them can be objectively quantified, whereas others are subjective and rely on the consumer preferences. Therefore, these factors can be categorised into:

- Nutritional value and healthiness: determined by the composition of lean and fat content in the meat. Intramuscular fat content and its fatty acid profile are a key point for both factors (Scollan et al., 2017). Over the past decades, consumer preference has shifted from lean pork with low intramuscular fat content towards pork with higher

fat content that contains the minimum amount of detrimental fatty acids for human health.

- Sensory aspects: such as juiciness, texture, taste and aroma. As these aspects are subjectively perceived by the consumer, preferences vary depending on each individual as well as the country, due to cultural and regional preferences, availability, and product marketing. In recent years, there has been a noticeable shift in preferences towards pork that not only reflects health and nutrition awareness, but also takes into account its sensory aspects. Consumer preference has switched from lean pork with minimal fat content, often perceived as less flavourful and juicy, towards a more palatable meat (Aaslyng et al., 2007; Font-i-Furnols & Guerrero, 2014).

- Food safety: ensuring the absence of pathogens and pharmaceutical products such as hormones or antibiotics is an important aspect for food safety, which is highly regulated in Europe. In 2006, the European Union banned the use of antibiotics as growth promoters and restricted their usage. Nowadays, antibiotics can only be administered with a veterinary prescription to lower the risk of developing antimicrobial resistances (Maron et al., 2013; Schmerold et al., 2023). Although the two previous points, nutritional and organoleptic properties, have greatly influenced pig breeding programs, there has been an increasing focus on food safety, developing new strategies to breed resilient pigs (Ballester et al., 2020).

1.1.3.1. Fatty acid composition

The fat content, especially intramuscular fat, and its composition are critical factors in defining meat quality. Variations in the length and degree of unsaturation of fatty acids influence their melting points, which affects some sensory aspects of meat such as tenderness or colour (Maw et al., 2003; J. D. Wood et al., 2008). Based on the carbon chain length, fatty acids can be divided in:

- Short-chain fatty acids (SCFAs): consisting of up to five carbons. SCFAs are also highly volatile, which earns them the alternative name of volatile fatty acids (VFAs).

- Medium-chain fatty acids (MCFAs): from six to 12 carbons.

- Long-chain fatty acids (LCFAs): from 13 to 21 carbons.

- Very long-chain fatty acids (VLCFAs): 22 or more carbons.

As previously mentioned, the level of saturation (i.e., the number of double bonds in the carbon chain) also influences the melting point, which increases with a higher number of double bonds. Additionally, the saturation level confers different properties to the meat:

- Saturated fatty acids (SFAs): do not present carbon-carbon double bonds in the carbon chain. SFAs have been related with cardiovascular diseases in the consumer due to their role in elevating cholesterol and low-density lipoprotein levels in blood (Hammad et al., 2016).

- Monounsaturated fatty acids (MUFAs): only present one carbon-carbon double bond that can be located at different positions along the chain, resulting in varying properties of the fatty acid. MUFAs, mainly represented by the oleic acid (C18:1n-9), are associated with positive sensory traits of the meat (Benet et al., 2015). In addition, these fatty acids are related with positive health effects as they reduce the metabolic syndrome and the cardiovascular disease risk (Gillingham et al., 2011). Iberian pork has a particularly high content in oleic acid.

- Polyunsaturated fatty acids (PUFAs): contain two or more carbon-carbon double bonds. PUFAs are named based on the position of the first double bond, which imparts distinct properties to the fatty acid. Omega-3 and omega-6 PUFAs are noteworthy examples, with omega-3 having anti-inflammatory properties and omega-6 showing pro-inflammatory effects (Marion-Letellier et al., 2015; J. D. Wood et al., 2008). In addition, PUFAs are more susceptible to oxidation, leading to rancidity in meat and affecting its sensory attributes (J. D. Wood et al., 2008).

1.1.4. Fatty acid metabolism

Fatty acids belong to the category of lipids, which are classified as biomolecules along with carbohydrates, proteins and nucleic acids. In addition to fatty acids, the lipids category is composed of phospholipids, glycerides, sterols, waxes, and liposoluble vitamins (A, D, E and K). Lipids have various and essential functions within living organisms. They can be found as structural components of cell membranes or as energy storage, but also act as regulators of different metabolic pathways and physiological processes. The bioactive lipids called lipokines are able to establish cellular communications (Kumar et al., 2019). Thus, lipids play important roles in maintaining homeostasis and cellular function.

1.1.4.1. Fatty acid biosynthesis

Animals obtain their fatty acids through dietary intake or *de novo* synthesis, mainly in the liver and adipose tissue. In pigs, the main organ for *de novo* biosynthesis is the adipose

tissue (O'Hea & Leveille, 1969), while the liver is key for PUFAs and cholesterol production (Smink et al., 2012). However, in other animals, the liver is the primary site for *de novo* biosynthesis. The biosynthesis of fatty acids occurs when the animal is in the fed state, which means it has an energy supply, and the excess can be converted and stored as fatty acids.

Short-chain fatty acids may be obtained via the transformation of acetyl-CoA and malonyl-CoA into acetyl-ACP and malonyl-ACP, respectively. The *FASN* gene encodes for the fatty acid synthase (FAS), which is a key element in *de novo* fatty acid biosynthesis as it elongates acetyl- and malonyl-ACP by two carbons and the subsequent saturated fatty acids until the production of palmitate (C16:0). Additionally, the gut microbiota produces SCFAs as metabolites, which can either be utilised by other microorganisms or absorbed and integrated into the host metabolism (Zeng et al., 2023).

Different elongases (ELOVLs) and desaturases (SCDs and FADSs) convert saturated fatty acids into longer chain or unsaturated fatty acids, respectively. ELOVLs catalyse the first step of the elongation reactions, with ELOVL1, 3, 6 and 7 using saturated and monounsaturated fatty acids as substrate, while ELOVL2, 4 and 5 elongate polyunsaturated fatty acids. Desaturases introduce double bonds between carbons, thereby changing the level of saturation of the fatty acids and their properties. Stearoyl-CoA desaturases (SCDs) introduce a double bond at the 9th position, while the fatty acid desaturases (FADSs) work at the 5th or 6th positions, counting from the carboxylic acid end (Guillou et al., 2010). Depending on the position of the first unsaturation, fatty acids are categorised into omega-3, omega-6, omega-7, or omega-9 families (Guillou et al., 2010). In animals, linoleic acid (C18:2n-6) and α-linolenic acid (C18:3n-3) are the initial molecules of the omega-6 and omega-3 families, respectively, and are considered essential FAs, as they cannot be endogenously synthesised and must be obtained through dietary sources (Das, 2006). A diagram of the *de novo* biosynthesis with the involved enzymes can be seen in Figure 1.2.

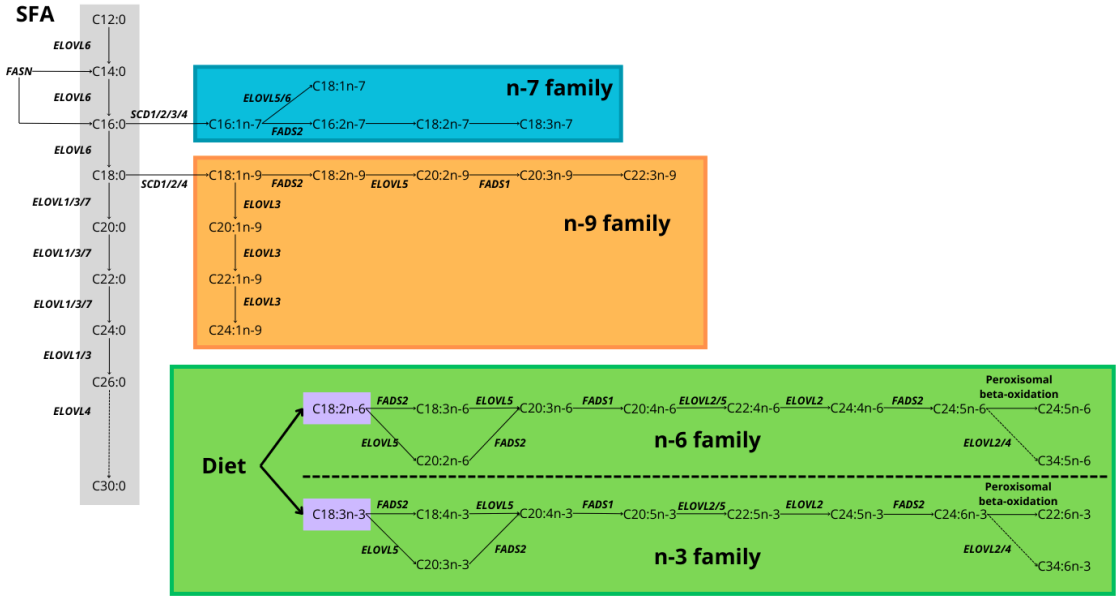


Figure 1.2. Fatty acid biosynthesis in animals. Adapted from (Guillou et al., 2010; Naudí et al., 2013).

1.1.4.2. Fatty acid β-oxidation

When the animal is in a fasting period, the energy supply is obtained from the stored fatty acids. The breakdown of FAs occurs predominantly through β-oxidation in the mitochondria, where fatty acyl-CoA molecules undergo a cycle of four steps that sequentially cleaves off two carbons, as shown in Figure 1.3. The acetyl-CoA molecules obtained enter the tricarboxylic acid (TCA) cycle, also known as Krebs cycle, where they undergo further oxidation to produce ATP, the primary energy source of the cell.

The SCFAs play a regulatory role in this metabolic pathway, as they can stimulate the FA oxidation and are also used as substrates in ATP generation. Specifically, butyrate can be converted into acetyl-CoA and enter the Krebs cycle. Acetyl-CoA, in turn, can also generate acetate during fatty acid oxidation in peroxisomes (Beauvieux et al., 2001; Yamashita et al., 2001).

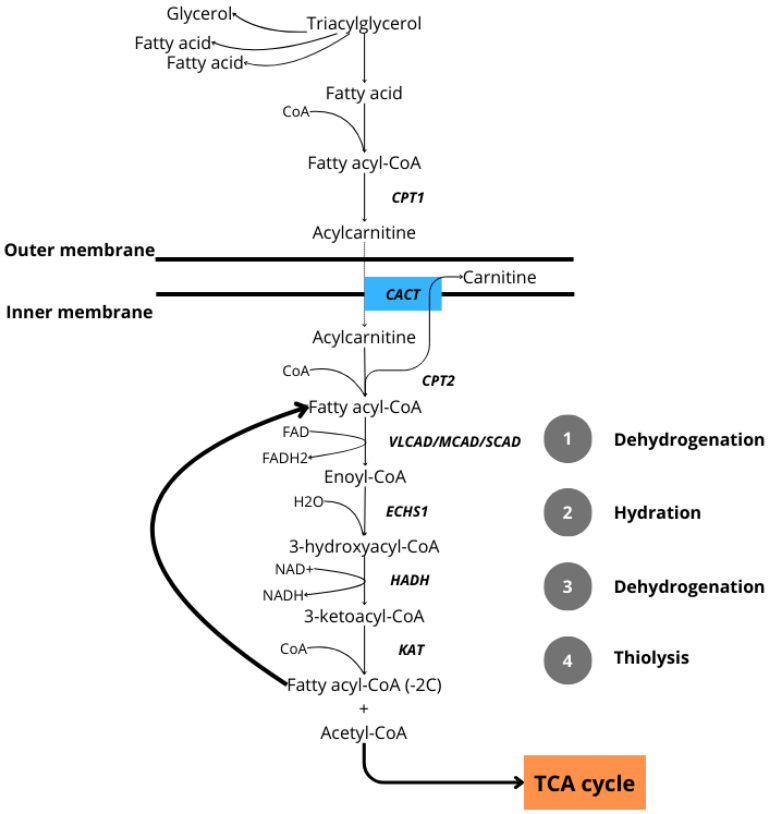


Figure 1.3. Fatty acid β-oxidation in the mitochondria. Adapted from (Abegaz et al., 2021; Sharpe & McKenzie, 2018).

1.2. Gut microbiota

1.2.1. Micro-concepts

According to Marchesi and Ravel, the microbiota can be defined as a population of microorganisms, including bacteria, archaea, viruses and lower and higher eukaryotes. The collection of all genomes of a microbial population is then termed as metagenome, wherein the concept of microbiome, in turn, encompasses not only all the microorganisms, but also their genomes and the environmental conditions (Marchesi & Ravel, 2015).

Microbial populations can be found everywhere, in the environment or in living beings, where they colonise the skin, the mucosal cavities and the gut. These microbial populations, in a healthy status, co-exist in symbiosis with the host and establish deep and complex relationships, even though some of them can be pathogenic or opportunistic species and harm the host (Afzaal et al., 2022; Fouhse et al., 2016; Patil et al., 2020). This profound level of interaction has recently spread the use of the term ‘holobiont’ to refer to both the host organism and its associated microbial community (Simon et al., 2019).

1.2.2. Gut microbiota and the host

The gut microbiota produces a diverse variety of metabolites that act as a source of energy or molecules, which can be incorporated into the host metabolism, consequently influencing growth, productivity and overall performance in production animals such as dairy cattle (Wallace et al., 2019), chicken (Crisol-Martínez et al., 2017) or pigs (Han et al., 2017; H. Yang et al., 2016). Moreover, the gut microbiome plays a vital role in maintaining host health by modulating immune function, as any perturbation in microbial composition (dysbiosis) can be detrimental for the animal (Patil et al., 2020). In addition to the health issues, potentially pathogenic microorganisms reduce the productivity of the host (Mach et al., 2015; H. Yang et al., 2016).

The composition of the gut microbiota, however, is dynamic and changes depending on different factors:

- Diet: among other traits, the amount and source of fibre in the diet modulate microbial composition. A diet rich in fibre promotes the abundance of beneficial bacteria, such as *Lactobacillus* or Bifidobacteria (R. Hu et al., 2023).
- Age: the microbial diversity and composition of the gut microbiota differ throughout the ageing stages. Before weaning, the diversity is lower, and the bacterial profile is oriented towards a milk-based diet. This profile changes drastically during weaning.

In pigs, this process occurs at 3-4 weeks of life, and piglets become more susceptible to dysbiosis and infections. Even though some bacterial genera such as *Treponema* have a relatively stable abundance, other genera like *Bacillus*, *Lactobacillus*, *Bacteroides* and *Escherichia* are more abundant in the first 40 days of age, while *Clostridium* decreases its abundance after weaning and increases in the finishing period, following an opposite trend to *Prevotella* (Luo et al., 2022).

- Intestinal region: the abundance and diversity of microorganisms are lower in the small intestine, where Firmicutes and Proteobacteria are the predominant phyla. In the colon, both the abundance and diversity of the microbiota increase, with Proteobacteria reducing their presence and Bacteroidota (formerly known as Bacteroidetes) showing a higher prevalence (Crespo-Piazuelo et al., 2018; W. Zhao et al., 2015).
- Host genome: the presence and abundance of some microbial species are influenced by the host genome, with a wide range of heritability values. In the literature, the species that show higher heritability values usually belong to the Firmicutes phylum (Bergamaschi et al., 2020; Camarinha-Silva et al., 2017).

Despite the different factors that affect the microbial composition of the gut, this composition can be used as a predictor of growth traits (Camarinha-Silva et al., 2017; Khanal et al., 2020; Maltecca et al., 2019). Specifically, in pigs, some bacterial genera such as *Bifidobacterium*, *Lactobacillus*, *Roseburia* or *Prevotella* have been positively related with production traits like body weight, average daily gain or feed conversion ratio (Mach et al., 2015; McCormack et al., 2017; Ramayo-Caldas et al., 2016; C. Wang et al., 2018). These genera produce lactate or SCFAs (acetate, propionate and/or butyrate), which have an important role in gut health and metabolism.

Some of the previous genera can be considered part of the “core” microbiota, a term that is used when the taxon is found in over 90% of the samples. However, defining a “core” microbiota is challenging due to its dynamic composition throughout the lifespan of the animal, and the presence of a taxon is not guaranteed in every sample (Dong et al., 2023; Holman et al., 2017).

The localization within the gastrointestinal tract plays a key role, with different microbial phyla dominating the small intestine versus the colon. The small intestine is primarily dominated by Firmicutes, represented by abundant genera such as *Lactobacillus* and *Clostridium*, and Proteobacteria. In colon, Firmicutes and Proteobacteria reduce their abundance while Bacteroidota increases, predominantly represented by *Prevotella*. Overall, Firmicutes and Bacteroidota appear as the most abundant phyla, followed by Proteobacteria

(Crespo-Piazuelo et al., 2018; Holman et al., 2017). The meta-analysis from Holman and colleagues found the genera *Lactobacillus*, *Clostridium*, *Prevotella*, *Ruminococcus*, *Blautia*, *Roseburia*, *Subdoligranulum,* and the RC9 gut group from the Rikenellaceae family, present in more than the 90% of the samples, thus defining them as the core microbiota, with *Prevotella* being found as the most abundant genus. Notably, a further meta-analysis performed by Dong and colleagues (Dong et al., 2023) in pigs of a wider variety of stages of life, identified twenty-four genera appearing in more than the 90% of the samples, but *Prevotella* was notably absent. This difference could be attributed to the inclusion of samples from suckling animals, where *Prevotella* is not part of the core microbiota in the early days of life. This genus increases its abundance and reaches a peak between weaning and the early growth stages before declining. However, other studies such as from Luo and colleagues (Luo et al., 2022) report *Prevotella* as part of the core microbiota despite its lesser prevalence in lactating piglets.

1.2.3. SCFA as microbiota metabolites

SCFAs are mainly produced by microbial fermentation of non-digestible polysaccharides within the gut (G. T. Macfarlane & Macfarlane, 2012). Acetate, propionate, and butyrate represent the most prevalent SCFAs in the gut milieu. These volatile fatty acids serve as an energy source for other microorganisms, or are incorporated into the host metabolism, where they play diverse roles that contribute to enhance the feed efficiency, boost the immune response, and regulate the lipid metabolism, among other functions (Liu et al., 2022; Vasquez et al., 2022).

1.2.3.1. Acetate

Acetate is the most abundant SCFA in the hindgut, originating from the fermentation of dietary fibre or through cross-feeding of metabolites like lactate by different bacteria (Brestenský et al., 2017). When the acetate is obtained from the degradation of polysaccharides, these molecules undergo enzymatic transformations within either the glycolytic or the pentose phosphate pathways, where phosphoenolpyruvate is produced, which subsequently undergoes conversion into acetyl-CoA. Acetyl-CoA can be further metabolised into acetate or butyrate, while pyruvate may derive into propionate or lactate (S. Macfarlane & Macfarlane, 2003).

In addition to becoming an energy source for other microorganisms, acetate assumes roles within the host metabolism, particularly as a precursor for cholesterol and fatty acid

synthesis. Acetate enters the host bloodstream, where it is integrated into the animal metabolic pathways and contributes to different physiological processes.

1.2.3.2. Propionate

Propionate is produced primarily by the fermentation of dietary fibres. It is formed by the decarboxylation of succinate, which in turn is produced by the pyruvate resulting from the glycolytic or the pentose phosphate pathways (S. Macfarlane & Macfarlane, 2003). This SCFA is predominantly synthesised by bacteria from the phylum Bacteroidota, but also *Akkermansia muciniphila*, which is a Verrucomicrobiota that also produces acetate (Parada Venegas et al., 2019).

Propionate improves lipid metabolism (K. Yu et al., 2019) and enhances immune response (Y. Zhang et al., 2019).

1.2.3.3. Butyrate

Butyrate can be synthesised by acetate and lactate, but also by amino acids or carbohydrates. This SCFA is produced by certain families from the Clostridiales order, including Lachnospiraceae (*Coprococcus*, *Eubacterium*, *Anaerostipes* and *Roseburia*), Ruminococcaceae (*Faecalibacterium* and *Subdoligranulum*) and Erysipelotrichaceae (*Holdemanella*) (Louis et al., 2014; Louis & Flint, 2009), as well as by different genera such as *Clostridium*, *Blautia*, *Lachnoclostridium*, *Oscillibacter* and others (Martin-Gallausiaux et al., 2021).

Functionally, butyrate serves as an energy source for colonocytes, enhances its proliferation and reduces apoptosis (Bartholome et al., 2004; Kien et al., 2007). In addition, it plays a pivotal role in enhancing immunity and reducing inflammation, contributing to the maintenance of gut health (C. L. Fang et al., 2014; Zeng et al., 2013).

1.3. Omes and omics

The suffix “-omes” denotes a collective of constituents, whereas “-omics” represents the scientific discipline dedicated to the study of these “-omes”. For example, the complete set of genes of an individual comprises its genome, which is studied through genomics.

The genes and other regulatory elements of the genome play an important role in shaping the phenotype, yet there are multiple steps between the genetic information and the final phenotype. The genome and the phenotype have an indirect relationship and the phenotype is affected by other biological processes that are connected between them. In addition, the environment affects the phenotype and previous steps such as the epigenetic modifications or the microbial composition, which is heavily determined by environmental conditions. A diagram of these relationships is shown in Figure 1.4.

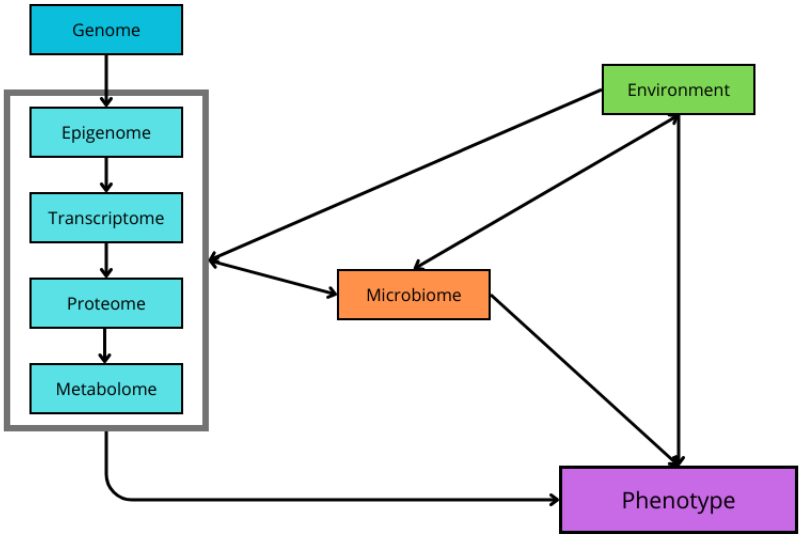


Figure 1.4. Relationships between the different omes that end up defining the phenotype. Adapted from (Varona, 2017).

1.3.1. Sequencing technologies

In the seventies, the first sequencing methods were developed. Sanger technique was published in 1977 and it became the preferred method to sequence DNA for almost three decades (Sanger et al., 1977). During the following decades, the process was automated and the sequencing of genomes of different species became a priority. The Human Genome Project was created in 1990, but the first draft of the human genome was published after

over a decade of research, in 2001. The final human genome was published in 2004 (International Human Genome Sequencing Consortium, 2004). In the meantime, many important species had their genomes sequenced and published, starting with the yeast *Saccharomyces cerevisiae* in 1996 (Goffeau et al., 1996). In 1998, the first genome of an animal, *Caenorhabditis elegans*, was published (The C. elegans Sequencing Consortium, 1998), while the mouse genome was published in 2002 (Mouse Genome Sequencing Consortium et al., 2002).

The advent of next-generation sequencing (NGS) marked a milestone in the field of genomics, being faster and having a better performance than the traditional Sanger sequencing methods. The first NGS method was commercialised in 2005 by 454 Life Sciences, nowadays known as Roche (Margulies et al., 2005), shortly after the publication of the first human genome. This technology allowed the launching of the 1000 Genomes Project, which was able to sequence thousands of genomes and characterise human variability (Siva, 2008). In pigs, genome sequencing started with the creation of the Swine Genome Sequencing Consortium (SGSC) in 2003 (Schook et al., 2005), which used first Sanger sequencing and was completed with NGS data obtained with Illumina, allowing the publication of the first draft genome in 2012 (Groenen et al., 2012).

NGS technologies addressed several limitations of Sanger sequencing. By operating in a cell-free system and removing the need for bacterial cloning of the DNA, NGS platforms such as Illumina or Ion Torrent enabled high-throughput sequencing with parallel processing of numerous reactions simultaneously. Moreover, the direct detection of sequencing output eliminated the need for electrophoresis previously required for base interrogation (T. Hu et al., 2021).

Despite the advantages of NGS, one of their main drawbacks at the time was the generation of short reads, which needed new alignment methods for genome assembly. The release of the PacBio RS sequencer in 2011 marked a significant advancement (Eid et al., 2009). Using Single-Molecule Real-Time (SMRT) sequencing, PacBio produced longer reads. Shortly thereafter, Oxford Nanopore Technologies (ONT) released the MinION, and these two events marked the beginning of the third-generation sequencing (TGS) technologies or long-read sequencing technologies (Jain et al., 2015, 2016).

TGS platforms, characterised by real-time single-molecule sequencing (SMS), have emerged as invaluable tools for transcriptomics analyses, particularly for whole-transcriptomic analysis as they do not need specialised assembly bioinformatic tools (Athanasopoulou et al., 2021). Despite lower accuracy and higher costs compared to NGS, TGS technologies have seen improvements in error rates with the release of newer

instruments such as the PacBio Sequel IIe System or the PromethION from ONT, which recently has released the second version. Table 1.1 shows the main sequencing platforms and their advantages and drawbacks.

Table 1.1. Advantages and drawbacks of the main sequencing platforms.

Sequencing generation	Sequencers	Platforms	How it works	Advantages	Drawbacks
NGS	Roche 454	454 GS Junior, 454 GS FLX+...	Pyrosequencing (pyrophosphate luminescence marks when a nucleotide is incorporated)	First NGS: high throughput compared to Sanger, fast run time	High error rates (homopolymers), high cost
	Illumina	MiSeq™, NextSeq, Hiseq…	Sequencing by synthesis, fluorescently labelled deoxynucleoside triphosphates	High accuracy	Short reads, expensive, long turnaround time
	Ion Torrent (by Thermo Fisher)	Ion PGM™ Torrent, Ion 5 S™…	Sequencing by synthesis, detection of hydrogen ions	Fast run time	Short reads, homopolymer errors
TGS	Pacific Biosciences (PacBio)	Sequel, RSII…	Sequencing by synthesis (SMRTbell replication)	Long reads	Expensive, variable accuracy, long turnaround time
	Oxford Nanopore	MinION™, GridION, PromethION, VolTRAX…	Changes in current as the molecule passes through the nanopore	Long reads, cheaper, short turnaround time	Low accuracy

Adapted from (PacBio, 2020; Petersen et al., 2019).

The choice of technology in omics studies depends on the specific level. Broadly, genomics and metagenomics will need DNA sequencing, while transcriptomics and metatranscriptomics use RNA sequencing. Other omics work in a different way, as listed in Table 1.2.

Table 1.2. Omics definition and sequencing technologies.

Omics	Definition	Technologies
Genomics	Investigates the structure and function of the genome, encompassing the entire complement of an organism's genetic material	DNA sequencing techniques such as whole-genome or whole-exome sequencing are utilised to study the genome. Whole-genome sequencing analyses the complete genome of the individual, while whole-exome sequencing focuses specifically on the exonic regions. High-density genotyping is another technique used in genomics
Epigenomics	Explores the chemical modifications undergone by the genome, the chromatin structure, the conformational changes, and the interaction with associated proteins	Chromatin immunoprecipitation (ChIP-Seq) is a key technology in epigenomics, allowing the analysis of chromatin structure and DNA accessibility. This technique allows the identification of DNA binding sites for transcription factors and other proteins, as well as the analysis of the histone methylation and acetylation. Bisulfite sequencing is another highly used method in epigenomics. Other methods of analysis are DNase-Seq, 3C and 4C, HiC and ATAC-Seq
Transcriptomics	Focuses on the dynamic expression levels of the different genes. The gene expression depends on the time and tissue	RNA sequencing (RNA-Seq) is the primary method for studying the transcriptome, over microarrays, high-throughput RT-qPCR or single-cell transcriptome analysis
Proteomics	Evaluates the produced proteins in a qualitative and quantitative way	Tandem mass spectrometry is commonly used to identify and characterise the proteins of an organism, considering their structure, functions, interactions and modifications

Lipidomics	Studies the produced lipids in a qualitative and quantitative way	Mass spectrometry is used frequently combined with other technologies such as gas chromatography or HPLC
Metabolomics	Analyses the qualitative or quantitative variations in metabolite production	Gas chromatography, mass spectrometry, or nuclear magnetic resonance allow the analysis of the metabolites. These techniques inform about the biochemical activity of the individual and are related with phenotypic traits
Microbiomics	Studies the microbiota, their genomes and the environmental conditions. Microorganisms are affecting host phenotype	A widely used method to study the microbiome is the 16S rRNA gene analysis, which analyses the hyper-variable regions of this gene, allowing the identification of the microorganisms. Whole-metagenome or whole-metatranscriptome analysis offer deeper and more accurate results, being more expensive options
Phenomics	Studies the observable traits of an organism	Phenotypic traits are analysed using diverse methods depending on the trait of interest, with the goal of analysing the highest number of traits easily and in a reduced time. In production animals, image or video analysis-based systems are used for traits such as water and feed consumption, growth or environmental conditions. Ultrasound are used for lean content and carcass quality, and digital imaging techniques can be employed for pig leg conformation

This thesis is focused on the study of pig transcriptomics and microbiomics, so the following sections explain in further detail about these disciplines.

1.3.2. Transcriptomics

The transcriptome can be defined as the set of all transcribed RNAs of a cell or a tissue. RNA molecules derive from DNA through transcription, which is regulated by different mechanisms such as epigenetics. Therefore, the transcriptome is different depending on the tissue, stage of life, and other conditions such as environment or disease. There are different types of RNA molecules, which have different functions within the cell, but the messenger RNA (mRNA) is the one that encodes the translation for amino acids that will constitute the proteins.

Transcriptome analysis has been widely employed in farm animals to evaluate the relationship between gene expression and productive traits such as growth and fat composition. In pigs, the fatty acid profile of intramuscular fat is crucial as it confers various properties related with meat quality. Therefore, elucidating the genes involved in this composition enables better selection of animals (T. Yu et al., 2023). Our research group has studied the porcine muscle transcriptome over the years in populations with different genetic backgrounds. The expression genome-wide association studies (eGWASs), which relate gene expression quantification with the genotype of single nucleotide polymorphisms (SNPs) distributed along the genome, have been utilised to locate eQTLs for traits of interest and to detect variants that may influence gene expression. Additionally, this method helps identify candidate genes that regulate different traits, listed in Table 1.3 (Corominas, Ramayo-Caldas, Puig-Oliveras, Estellé, et al., 2013; Criado-Mesas et al., 2020; Passols et al., 2023; Puig-Oliveras et al., 2016; Ramayo-Caldas et al., 2012).

Table 1.3. Main candidate genes found related with meat quality and involved in lipid metabolism.

Gene	Trait	Reference
IGF2	Muscle growth, fat deposition and carcass weight	(Criado-Mesas et al., 2019; Estellé et al., 2005; Laere et al., 2003)
ELOVL6	Percentage of palmitic and palmitoleic acids in muscle and backfat	(Corominas, Ramayo-Caldas, Puig-Oliveras, Pérez-Montarelo, et al., 2013)
FABP4 and *FABP5*	FA deposition and growth	(Estellé et al., 2006; Ojeda et al., 2006)
SCD	MUFA content	(Estany et al., 2014)

Nowadays, transcriptome analysis is mainly carried out by RNA-Seq, which can be divided into Whole Transcriptome Sequencing, mRNA sequencing or small RNA sequencing. As it analyses the transcribed RNA, it is useful for functional genomic studies (e.g. differential gene expression, alternative splicing, and variant discovery) (Mutz et al., 2013).

In order to analyse the transcriptome with RNA-Seq, RNA-Seq libraries must be prepared. The tissue to be analysed is lysed, the RNA is extracted, and the quality and quantity of that RNA are then assessed. RNA-Seq libraries are generated by selecting mRNA and converting it to cDNA through reverse transcription. The cDNA is fragmented, and adapters are added for sequencing. This material is then amplified by PCR and purified to be sequenced (Hrdlickova et al., 2017).

After the sequencing process, the bioinformatic analysis of RNA-Seq data can be divided in three main stages:

1. Base calling. This step is carried out by the sequencing instrument, which also conducts preliminary quality control. At this point, quality control is based on read filtering and trimming, and all the information is stored in FASTQ files.
2. Read mapping and transcriptome reconstruction. Reads are aligned to a reference genome and stored in BAM/SAM files using software tools such as STAR (Spliced Transcripts Alignment to a Reference) (Dobin et al., 2013) or Hisat2 (D. Kim et al., 2019). Typically, the reference genome comes from the ENSEMBL project (Martin et al., 2023), which has the current reference genome *Sscrofa11.1* for pigs, published in 2017. The quality control of this step can be performed by other tools such as Picard (Broad Institute, 2019) or SAMTools (H. Li et al., 2009).
3. Expression quantification and differential expression analysis. In this phase, the reads mapped to individual transcripts are quantified, and differentially expressed genes are identified. Bioinformatic tools such as HTseq (Anders et al., 2015) and R packages such as DESeq2, edgeR or limma are commonly employed for this step (Ji & Sadreyev, 2018; Love et al., 2014, 2015; Ritchie et al., 2015).

1.3.3. Microbiomics

The microbiome, as defined before, is the set of all the microorganisms, their genomes and the environmental conditions (Marchesi & Ravel, 2015). These microbial populations can be analysed through 16S rRNA gene sequencing or through whole-metagenome shotgun sequencing. While the former approach is cost-effective, enabling the characterization of a larger number of samples, the latter provides a more in-depth and accurate classification. In

addition, whole-metatranscriptome shotgun sequencing has also emerged recently as a method to assess the activity of the microbiota.

While microbiome analysis can be conducted on any microbial population, its application in production animals is particularly interesting because the microbiome of animals, especially the gut microbiome, has a close relationship with the host's metabolism and overall health. In ruminants, research has primarily focused on characterising the intestinal (Mao et al., 2015; J. Wang et al., 2016) and rumen microbial communities (Jami & Mizrahi, 2012; Y. K. Zhang et al., 2021) using the 16S rRNA gene sequencing method. Conversely, in monogastric animals like pigs, studies have employed a combination of methodologies, including 16S rRNA gene sequencing, shotgun metagenomics, and culturomics, providing a broader and more accurate view of the microbial population (Fenske et al., 2020; Holman et al., 2021).

In previous works carried out by our research group, 16S rRNA gene sequencing has been used to characterise the microbiome in different gut sections of the pig (Crespo-Piazuelo et al., 2018) and also to associate the microbiota of the rectal content with the pig genome (Crespo-Piazuelo et al., 2019). The use of 16S rRNA gene profiling allowed the sequencing of a high number of samples and a reliable taxonomic classification at genus level. This approach confirmed that the bacterial genera *Prevotella* and *Treponema* were the most abundant in rectal content, consistent with findings from previous studies (Ramayo-Caldas et al., 2016; H. Yang et al., 2018). Furthermore, it allowed the identification of associations between certain bacterial genera and genes that could potentially regulate the microbiota, as explored in other works (Camarinha-Silva et al., 2017; Chen et al., 2018).

In the following sections, the main sequencing methods used in this thesis will be discussed in detail.

1.3.3.1. 16S rRNA gene sequencing

The 16S rRNA is a component of the 30S small subunit of prokaryotic ribosomes. The 16S rRNA gene is formed by nine hypervariable regions that are located between well-conserved regions as shown in Figure 1.5, with a total length of approximately 1,500 base pairs. This structure allows the taxonomic classification of prokaryotes through the sequencing of this gene, as the hypervariable regions exhibit species-specific differences, while the conserved regions serve as primer binding sites for sequencing. The first primers were designed in the early 1990s using 16S rRNA gene consensus sequences of different microorganisms, and it has become the preferred method for prokaryote analysis and classification since then (Giovannoni et al., 1990; Ward et al., 1990; Weisburg et al., 1991).

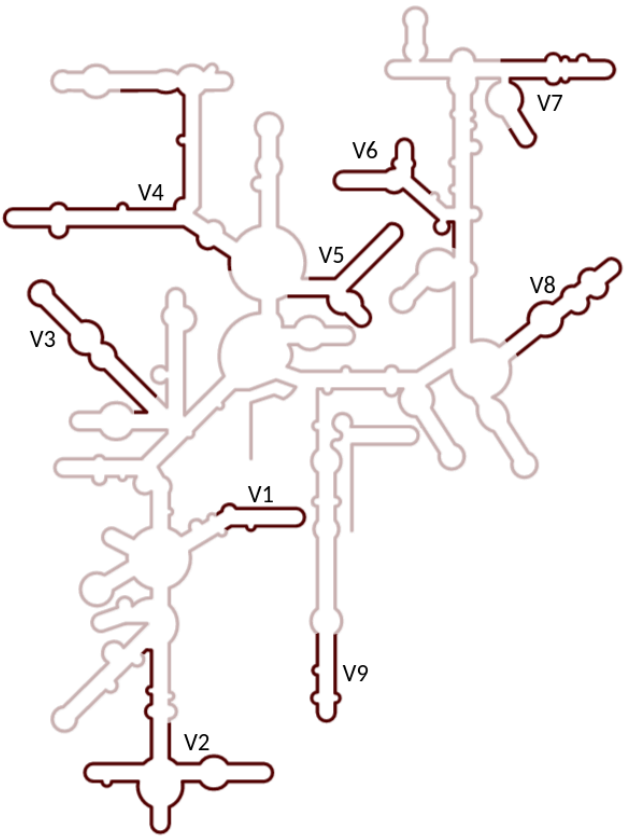


Figure 1.5. Structure of the 16S rRNA. Hypervariable regions are highlighted and numbered V1-V9.

Woese and colleagues defined the main conserved regions of the 16S rRNA in 1975 (C. R. Woese et al., 1975). Two years later, the 16S rRNA sequencing method was used to classify the cellular life forms into the three-domain system that divided them into archaebacteria, eubacteria and eurkaryotes, corresponding to the current Archaea, Bacteria, and Eukarya domains (C. Woese, 1998; C. R. Woese et al., 1990; C. R. Woese & Fox, 1977).

Primers for 16S rRNA gene sequencing can be designed to target both bacteria and archaea or be specific to one domain. Primers must maximise the detection coverage and the length of the sequenced amplicons to identify lower taxonomy levels, as the different amplified regions have varying levels of accuracy and affect diversity estimates and taxonomic classification (B. Yang et al., 2016). The sequencing of the 16S rRNA gene has followed the evolution of the sequencing technologies, transitioning from Sanger to NGS and TGS methods. Nowadays, the Illumina MiSeq sequencer is one of the most employed technologies for 16S rRNA gene sequencing (Regueira-Iglesias et al., 2023).

The bioinformatic analysis of this gene involves several steps. Initial quality control typically involves assessing a phred score of 20 (base call error probability of 1%). Subsequently,

reads are merged into unique sequences, which are then clustered or denoised and binned into Operational Taxonomic Units (OTUs) or, more recently, into Amplicon Sequence Variants (ASVs). OTUs represent clusters of organisms with a higher similarity than 97% or 99%, depending on the taxonomic level (genus or species, respectively) (Edgar, 2013; Morgan & Huttenhower, 2012). In contrast, ASVs are generated on a 100% of sequence similarity using single-nucleotide resolution, which improves the taxonomic determination (Nearing et al., 2018). The amplicon sequence data is processed with tools such as the ones listed in Table 1.4.

Taxonomic assignment requires a reference database. The most used ones are RDP (Ribosomal Database Project) (Cole et al., 2009), SILVA (Quast et al., 2012) and Greengenes (DeSantis et al., 2006), which recently released a second version integrating genomic and 16S rRNA data (McDonald et al., 2023).

Table 1.4. Main advantages of the most used bioinformatic pipelines for 16S rRNA amplicon data analysis.

Software	**Language**	**Advantages**	**Reference**
mothur	shell	More efficient (faster), good for high number of samples	(Schloss et al., 2009)
USEARCH			(Edgar, 2010)
dada2	R	High accuracy	(Callahan et al., 2016)
QIIME2	Python	Different plugins with more functionalities, visualisation tools	(Bolyen et al., 2019)

1.3.3.2. Whole-metagenome shotgun sequencing

Whole-metagenome sequencing addresses the 16S rRNA gene limitations by enabling deeper and more accurate classification, and also characterising other organisms besides bacteria and archaea. In 2004, NGS techniques were first applied in metagenomic samples (Tyson et al., 2004; Venter et al., 2004), and the first study of the human gut metagenome was published the next year (Eckburg et al., 2005). In pigs, whole-metagenome sequencing allowed the creation of the reference gene catalogue of the gut microbiome, which identified 7.7 million microbial genes that belonged to 719 species in 287 pig faecal samples from different breeds (Xiao et al., 2016). Recently there has been published an expanded catalogue of the pig gut microbiome, which sequenced 500 samples from different gut

sections of different pig breeds or crossbreeds. This study also included the samples of the previous catalogue, and identified more than 17 million complete genes that were clustered into 2673 species (Chen et al., 2021).

The sequencing of metagenomic samples has been conducted by the main sequencing companies such as Illumina or Ion Torrent (Van Dijk et al., 2018). Metagenomic assembly is challenging compared to genomic assembly due to reads originating from many different genomes, which may include related species or strains, and coverage variability depending on the abundance of each genome. In addition, the genome of the host is also sequenced and needs to be removed by mapping it against the host species reference genome (Pereira-Marques et al., 2019). Commonly used assemblers include metaSPAdes (Nurk et al., 2017) and MEGAHIT (D. Li et al., 2015), following an initial quality control performed with software such as FastQC (Andrews, 2010). The assembly is highly fragmented in thousands of contigs, which are subsequently binned or grouped into species via supervised binning methods, which use databases of already sequenced genomes to label the contigs into taxa, or unsupervised methods, that look for natural groups or clusters in the data.

Taxonomic profiling of metagenomic data can be performed without an assembly using a publicly available reference genome, which is faster than a *de novo* assembly, but loses information about the uncharacterized microbes. Alternatively, taxonomic profiling can be achieved by selecting markers or representative genes from available reference sequences. Several bioinformatic tools execute the full pipeline for metagenomic analysis, starting from the quality control and obtaining the taxonomic classification or putative functions as the final output. Examples include SqueezeMeta (Tamames & Puente-Sánchez, 2019) and metaWRAP (Uritskiy et al., 2018).

Metagenomic genes can be identified using adapted genomic characterization tools. Databases such as KEGG (Kanehisa & Goto, 2000) or UniProt (The UniProt Consortium et al., 2023) can be used to assess the putative functions of the metagenome. Furthermore, specific functions of interest can be identified, such as the resistome, which is the complete set of antibiotic resistance genes (Wright, 2007).

1.3.3.3. Metatranscriptome shotgun sequencing

In recent years, sequencing the whole metatranscriptome has become an increasingly popular technology for microbiome elucidation. Whole-transcriptome analysis sequences the expressed genes of the microbiome at a specific time point, which is an advantage compared to the metagenomic analysis that sequences all the genes regardless of their expression. These genes can be annotated and mapped to metabolic pathways, allowing the

characterization of the active microbial species. Similarly to metagenome sequencing, the bioinformatic analysis of metatranscriptomic data starts with a quality control of the sequenced reads using programs like FastQC (Andrews, 2010) or Trimmomatic (Bolger et al., 2014), which are used for short reads sequenced with Illumina. In the case of metatranscriptome, it is important to remove the ribosomal RNA transcripts using programs like SortMeRNA (Kopylova et al., 2012), as they constitute a significant part of the data (up to 90%) and do not contribute in subsequent analyses such as differential expression or pathway characterization. If specific target organisms need to be removed, such as host reads of the samples, mapping methods that tag and remove this genome (P.-E. Li et al., 2017) or alignment-free methods like BMTagger (Rotmistrovsky & Agarwala, 2011) can be applied as a previous step to the assembly.

De novo assembly can be performed using transcriptomic tools such as Oases (Schulz et al., 2012), rnaSPAdes (Bushmanova et al., 2019), or Trinity (Grabherr et al., 2011). Alternatively, specific options for metatranscriptome such as IDBA-MT (Leung et al., 2013), IDBA-MTP (Leung et al., 2014) or TAG (Transcript Assembly Graph) (Ye & Tang, 2016) are available. The taxonomy of the active species can be assessed using classification tools like MetaPhlan2 (Truong et al., 2015) or Kraken2 (D. E. Wood et al., 2019), which include eukaryotes in their databases. Functional annotation can be performed with read-based functional profilers such as MetaCLADE (Ugarte et al., 2018), or with assembled contigs and tools like DIAMOND (Buchfink et al., 2015), which searches against functional databases like KEGG (Kanehisa & Goto, 2000) or NCBI RefSeq (O'Leary et al., 2016), among others. Some bioinformatic tools like MetaTrans (Martinez et al., 2016), SAMSA2 (Westreich et al., 2018) or HUMAnN2 (Franzosa et al., 2018) perform the different steps of metatranscriptomic analysis.

1.3.4. Data integration

In organisms, processes are interconnected, with existent and transcribed genes influencing protein production and, therefore, metabolite levels. Similarly, the microbiota is both influencing and being influenced by the host metabolism, as well as being shaped by environmental factors. Systems biology approaches offer a means to characterise these complex interactions of the different components in study, aiming to reduce and integrate the different parts independently of their structural and functional complexity (Kohl & Noble, 2009). There are several approaches to designing a systems biology study:

- Bottom-up: starts with elements placed at the bottom of any multiscale representation, such as genes and proteins, and using these components and their interactions tries to reconstruct the system.
- Top-down: studies first the system at high level and tries to infer the parts and their functionality in a form of reverse engineering.
- Middle-out: combines the two previous approaches, as it starts at an intermediate level and goes into higher and lower levels in the exploration of the relationships.

Given the multidimensional interactions in which the parameters may vary due to perturbations in one or more of the dimensions, the landscape concept has emerged. The landscape approach tries to identify trajectories for dynamic parameter variation when the responses are kept within a certain range (Ao, 2009; Kohl et al., 2010).

Different bioinformatic tools have been developed for systems biology, including MATLAB (Schmidt & Jirstrand, 2006), PCIT (Reverter & Chan, 2008) and Pathway Tools (Karp et al., 2021). In R, different packages enable the integration of two or more datasets and the interpretation of their connections. Improving basic statistics like correlations, packages such as WGCNA (Weighted Gene Co-expression Network Analysis) reduces the dimensionality of the gene data to find hidden relationships with other traits (Langfelder & Horvath, 2008). Additionally, packages like mixOmics employ various methods to find relationships between two or more datasets (Lê Cao et al., 2016; Rohart et al., 2017).

In production animals, the systems biology approach was also used to relate two or more data sets that contain information on complex traits, such as the genome or transcriptome of the animal with productivity traits like meat quality or feed efficiency (Puig-Oliveras et al., 2014; Schettini et al., 2022; H. Yang et al., 2017). While in cattle it is important dairy production or the rumen microbiome (Fan et al., 2021; Neves et al., 2020), in pigs there is also interest in the immunity of the animal or regarding the gut microbiota (Crespo-Piazuelo et al., 2021; Reverter et al., 2021).

In our research group, we used mixOmics previously to find relevant correlations between the RNA-Seq data and the fatty acid composition in muscle. In the study of Valdés-Hernández and colleagues, the regularised canonical correlation analysis (rCCA) from mixOmics was applied to identify candidate regulatory genes influencing the IMF fatty acid profile (Valdés-Hernández et al., 2024).

2. Objectives

This PhD thesis was done under the framework of the AGL2017-82641-R and the PID2020-12677RB-C22 projects funded by the Ministerio de Ciencia e Innovación (MCIN).

The main aim of this thesis was to study the interactions among the pig gut microbiome, rectal SCFA composition, gene expression and FA composition in different tissues, and to assess their impact in meat quality traits.

The main goal was divided in four objectives that correspond to the four works that we have developed:

- Objective 1: find modules of co-expressed genes in muscle related with intramuscular FA composition.
- Objective 2: find links among gut microbial composition, rectal SCFA content and the FA composition in muscle and adipose tissue.
- Objective 3: find differentially active species of the gut microbiota and their expressed genes in pigs with divergent content in rectal SCFA composition.
- Objective 4: find connections between the gut microbial composition and liver gene expression, focusing on energy metabolism.

3. Articles and studies

Paper I

Gene co-expression network analysis for porcine intramuscular fatty acid composition

Sebastià, C., Gallopin, M., Ramayo-Caldas, Y., Estellé, J., Valdés-Hernández, J., Castelló, A., Sánchez, A., Crespo-Piazuelo, D., and Folch, J.M.

animal: The international journal of animal biosciences (under review)

Gene co-expression network analysis for porcine intramuscular fatty acid composition

C. Sebastià [a,b,*], M. Gallopin [c,*], Y. Ramayo-Caldas [d], J. Estellé [e], J. Valdés-Hernández [a,b], A. Castelló [a,b], A. Sánchez [a,b], D. Crespo-Piazuelo [d,f], J.M. Folch [a,b]

[a] Plant and Animal Genomics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB Consortium, C. de la Vall Moronta, 08193, Bellaterra, Spain.

[b] Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), Edifici V, Travessera dels Turons, 08193, Bellaterra, Spain.

[c] Institute for Integrative Biology of the Cell (I2BC), Université Paris-Saclay, CEA, CNRS, 1, Avenue de la Terrasse, Bâtiment 21, 91190, Gif-sur-Yvette, France.

[d] Departament de Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Torre Marimon, 08140, Caldes de Montbui, Spain.

[e] Université Paris-Saclay, INRAE, AgroParisTech, GABI, Domaine de Vilvert, 78350, Jouy-en-Josas, France.

[f] R&D Department, Cuarte S.L., Grupo Jorge, Autov. Zaragoza-Logroño, km.9, 50120, Monzalbarba, Spain.

[*] These authors contributed equally.

Corresponding author: Cristina Sebastià. Email: cristina.sebastia@cragenomica.es

Abstract

The aim of this work was to identify the link between muscle gene expression and its fatty acid (FA) composition. In an (Iberian × Duroc) × Duroc backcrossed pig population, we identified modules of co-expressed genes, and correlation analyses were performed for each of them versus the phenotypes, finding four relevant modules. Two of the modules were positively correlated with saturated FAs (SFAs) and monounsaturated FAs (MUFAs), while negatively correlated with polyunsaturated FAs (PUFAs) and the omega-6/omega-3 ratio. The gene-enrichment analysis showed that these modules had over-representation of pathways related with the biosynthesis of unsaturated FAs, the PPAR signalling pathway and FA elongation. The two other relevant modules were positively correlated with PUFA and the n-6/n-3 ratio, but negatively correlated with SFA and MUFA. In this case, they had an over-representation of pathways related with fatty and amino acid degradation, and with oxidative phosphorylation. Using a graphical Gaussian model, we inferred a network of connections between the genes within each module. The first module had 52 genes with 87 connections, and the most connected genes were *ADIPOQ*, which is related with FA oxidation, and *ELOVL6* and *FABP4*, both involved in FA metabolism. The second module showed 196 genes connected by 263 edges, being *FN1* and *MAP3K11* the most connected genes. On the other hand, the third module had 161 genes connected by 251 edges and *ATG13* was the top neighbouring gene, while the fourth module had 224 genes and 655 connections, and its most connected genes were related with mitochondrial pathways. Overall, this work successfully identified relevant muscle gene networks and modules linked with FA composition, providing further insights on how the physiology of the pigs influences FA composition.

Keywords: gene network, pig muscle, lipid metabolism, WGCNA, graphical Gaussian model

Implications

Pig meat quality is influenced by the fatty acid composition in intramuscular fat, which, in turn, is determined by different factors such as the diet, the animal age or the breed. Aiming to deepen the knowledge of the links between muscle gene expression and fatty acid composition, our study identified groups of genes that play a role in controlling fatty acid metabolism. Groups with positive relationships with saturated and monounsaturated fatty acids had genes involved in the synthesis of fatty acids and groups with positive relationships with polyunsaturated fatty acids had genes involved in the degradation of fatty acids.

Introduction

Pig is a main livestock species, being pork one of the most produced meats worldwide only surpassed by poultry, with 1.35 billion swine heads produced in 2019 and 110 million tonnes of meat (FAOSTAT, 2023). This high productivity is sustained by the efficiency of the commercial pig breeds, which are strongly selected based on their prolificity, growth capacity, and muscle and fat deposition, among other traits (Latorre et al., 2008). Nowadays, other traits are also added to the selective breeding schemes, as consumers demand healthier meats, but also keeping their good organoleptic qualities such as flavour or juiciness (Aaslyng et al., 2007), being determinant the level of intramuscular fat (Fernandez et al., 1999; van Laack et al., 2001). Lipid metabolism plays an important role in energy production and storage among other processes (Bergen & Mersmann, 2005), and it is a key factor in pork production, as it affects the composition of the fat and, therefore, consumer acceptance of the meat (Fernandez et al., 1999). In addition, pig is used in biomedicine as a model species for obesity and cardiovascular diseases, therefore, increasing the importance of deepening the knowledge on its lipid metabolism (Spurlock & Gabler, 2008).

Iberian pig is a Spanish rustic breed, less efficient than the highly efficient lean breeds, but also endowed with a high fat deposition, both intramuscular and subcutaneous (Nieto et al., 2019). Iberian pork fatty acid (FA) profile, when compared with leaner breeds, has higher content in monounsaturated FAs (MUFAs), but also in saturated FAs (SFAs) (Serra et al., 1998). MUFAs are involved in meat quality and taste, but SFAs have been associated with obesity and cardiovascular diseases in humans (Briggs et al., 2017; Livingstone et al., 2022). Iberian sows can be crossed with Duroc boars to enhance the growth performance while keeping good organoleptic properties of the meat (Ayuso et al., 2016). This crossbreed and the backcross (Duroc × Iberian) × Iberian are also used in the dry-cured products industry, producing more affordable pork while keeping an Iberian background (BOE, 2014).

Recent studies have been performed by our research group to assess relationships between the genes and the FA composition in muscle (Valdés-Hernández et al., 2023)using different approaches. This analysis found candidate genes for intramuscular FA profile. However, these results can be further explored using co-expression networks that highlight the relationships between the genes.

The analysis of gene expression data is a challenging task and extensive efforts have been made to provide a large range of methods to extract relevant interaction networks from transcriptome data. Among these approaches, the Weighted Gene Co-expression Network Analysis (WGCNA) detects co-expressed genes associated with complex target traits

(Horvath & Dong, 2008; Langfelder & Horvath, 2008; B. Zhang & Horvath, 2005). WGCNA has been used to investigate co-expression modules associated with obesity (Kogelman et al., 2014) or feed efficiency in pigs (Ramayo-Caldas et al., 2018). However, WGCNA and other network inference approaches lack the ability to distinguish between direct and indirect relations among the expressed genes in a statistically sound way. A feasible alternative that can overcome this limitation is the combination of the WGCNA approach with graphical Gaussian models (GGMs) (Schäfer & Strimmer, 2005). Under this statistical framework, gene regulatory networks are modelled as sparse graphs where genes are nodes, i.e. Gaussian random variables, and edges between nodes are the relationships between genes, i.e. conditional dependencies between random variables. GGMs distinguish direct dependencies (true biological genes regulation network) from spurious correlations (undirected relationships between genes). Furthermore, this approach has been already used in plants (Ma et al., 2007), on other omics (Benedetti et al., 2017) or in cancer research (Svoboda et al., 2018).

Overall, WGCNA and GGM are complementary approaches to unravel the complexity of biological mechanisms (Le Novère, 2015) and in this article, we used a two-step approach similar to L. Zhang et al. (2018) to apply them to a RNA-Seq dataset that targets the pig muscle transcriptome in animals with intramuscular FA composition.

The aim of this work was to study the relationships between the muscle gene expression and the FA profile, in order to identify genes that may be influencing the FA composition in muscle and therefore, the meat quality. For that purpose, we first used the WGCNA method to identify muscle co-expressed genes modules associated with FA composition. Subsequently, on each module, we used the statistical framework of GGMs to infer regulatory networks.

Material and methods

Animal material

The 129 pigs used in this study belonged to an experimental backcrossed population (BC1_DU), which was obtained by mating five Iberian × Duroc F1 boars with 22 Duroc sows. The population is fully described in Crespo-Piazuelo et al. (2020). In brief, the backcrossed animals (59 females and 70 males) were raised in intensive conditions and fed a cereal-based diet. Pigs were slaughtered at an age of 190 ± 15 days in a commercial abattoir, where samples from the *longissimus dorsi* muscle were collected and snap-frozen in liquid nitrogen. Muscle samples were later stored at −80ºC until further use.

The FA profile was determined by a gas chromatography protocol for methyl esters, as explained in Crespo-Piazuelo et al. (2020). Then, the results of the FA composition were represented as relative abundances in percentages (Table 1). In addition, other indices such as the sum of SFA, MUFA and PUFA, and the ratio between the omega-6 and omega-3 (n-6/n-3) PUFA were also calculated.

Transcriptome sequencing and estimation of gene expression

The RiboPureTM Isolation kit for High Quality Total RNA (Ambion; Austin, TX, USA) was used, following the manufacturer's recommendations, to isolate the RNA from the Longissimus dorsi muscle of the 129 animals. A NanoDrop ND-1000 spectrophotometer (NanoDrop products; Wilmington, DE, USA) was used to quantify the RNA, and its integrity was checked with an Agilent Bioanalyzer-2100 equipment (Agilent Technologies, Inc.; Santa Clara, CA, USA). The CNAG institute (Centre Nacional d'Anàlisi Genòmica; Barcelona, Spain) performed the library preparation and sequencing. The analysis was run in an Illumina HiSeq 3000/4000 machine (Illumina, Inc.; San Diego, CA, USA) using the paired-end library prepared with TruSeq Stranded mRNA kit also from Illumina. Full details are available in Crespo-Piazuelo et al. (2020).

Bioinformatic analyses for RNA alignment and quantification can be found in Valdés-Hernández et al. (2023). In brief, initial quality control was performed with FastQC v0.11.9 (Andrews, 2010). STAR v2.7.9a (Dobin et al., 2013) was used to map the reads against the Sscrofa11.1 pig reference genome and gene expression was assessed with RSEM v1.2.28 (B. Li & Dewey, 2011). In summary, a mean of 45.09 million of 2 × 75 bp paired-end reads per sample were obtained, resulting in an average of 90.06% of uniquely mapped reads. The raw count matrix consisted of 19,263 genes.

Weighted Gene Co-expression Network Analysis (WGCNA) and gene enrichment analysis

The FA composition and ratios indicated in Table 1 were used in the WGCNA analysis.The list of expressed genes obtained from the RNA-Seq expression data was filtered, keeping 3,538 genes related to energy metabolism. The filter was created using a list of gene functional annotation based on gene ontology and pathway terms, which included the following gene ontology databases: KEGG, Reactome, WikiPathways, STRING-db, AmiGO 2, MGI, and BioSystems from NCBI. This list was manually curated by our research group, as further explained in Valdés-Hernández et al. (2023). The list with the pathways and the source can be found in Supplementary Table S1. The counts were corrected by sex and batch, normalised using the counts per million (CPM) calculation and log-transformed with the edgeR package v3.38.4 (Robinson et al., 2010). A complete list of the lipid-related genes can be found in Supplementary Table S2. To identify co-expressed and highly interconnected genes associated with the intramuscular profile of 17 FAs (Table 1), the WGCNA package v.1.72-1 (Langfelder & Horvath, 2008) was used in R as follows:

(1) The network construction as well as the identification of the modules of co-expressed genes was done on gene expression data. The soft-thresholding power as function of the scale-free topology index was defined at seven, which represent a power-law model of $R2 = 0.88$.

(2) Pearson's correlation between the module eigengene and the FA phenotype information was estimated. The eigengene is defined as the first principal component of a given module and can be considered a representative of the gene expression profiles in a module (Langfelder & Horvath, 2008). A module was chosen for downstream analysis if it presented module-trait relationship $\geq |0.1|$ and p-value ≤ 0.05 in at least six of the analysed traits.

Gene function classification and pathway enrichment analyses were performed using the ClueGO Cytoscape plug-in (Bindea et al., 2009). The cut-off for considering a significant over-representation was established with a p-value ≤ 0.05 after Benjamini and Hochberg multiple-test correction (Benjamini & Hochberg, 1995).

Table 1. Mean relative abundance (%) and standard deviation (SD) values of FA composition in the *longissimus dorsi* muscle of the 129 BC1_DU pigs.

Phenotypes	Name	Mean	SD
Saturated FA			
SFA_LD	Total SFA	39.98	3.107
C140_LD	Myristic acid	1.27	0.231
C160_LD	Palmitic acid	23.90	1.663
C180_LD	Stearic acid	14.35	1.727
Monounsaturated FA			
MUFA_LD	Total MUFA	43.37	6.233
C161n7_LD	Palmitoleic acid	2.79	0.524
C181n9_LD	Oleic acid	35.83	5.800
C181n7_LD	Vaccenic acid	3.83	0.303
C201n9_LD	Gondoic acid	0.73	0.166
Polyunsaturated FA			
PUFA_LD	Total PUFA	16.16	8.306
C182n6_LD	Linoleic acid	12.25	5.962
C183n3_LD	α-linolenic acid	0.40	0.131
C202n6_LD	Eicosadienoic acid	0.43	0.122
C203n6_LD	Dihomo-γ-linolenic acid	0.46	0.294
C203n3_LD	Eicosatrienoic acid	0.18	0.098
C204n6_LD	Arachidonic acid	2.62	1.983
W6W3_LD	Ratio n-6/n-3	26.60	8.159

Network inference using graphical Gaussian model

Within each module of genes, we inferred a network of “direct” regulation between genes (i.e., conditional dependencies) using the graphical Gaussian model in high dimension (Friedman et al., 2008). Analyses were implemented in R by using the glasso (Friedman et al., 2008) and huge (T. Zhao et al., 2012) packages.

For consistent estimation of the network, it was inferred 30 times on 30 subsets from the original dataset (sampling with replacement). Based on the Bolasso variable selection algorithm (Bach, 2008), only edges appearing all the times in the inferred networks were kept in the final model. Inferred networks were subsequently uploaded to Cytoscape program for visualisation.

Results

Co-expressed gene modules and correlation with FA traits

The co-variation between gene-expression and FA phenotypic information was estimated by using the WGCNA approach. First, genes were clustered according to their connectivity, resulting in 12 modules. Supplementary Table S3 provides the complete list of genes with their module membership. The WGCNA procedure assigned a specific colour to each gene-module that will henceforth be used to refer to each module, as seen in the summary plot of Fig. 1, which also shows the colour assigned to each significant module. Fig. 2 pictures the correlation coefficient between the eigengene values of these modules and the 17 FA-related phenotypes. From the 12 modules, we only kept for further analyses the ones that had six or more significant correlations (p-value ≤ 0.05) of their eigengene values with the phenotypes, which resulted in four relevant modules (Magenta, Yellow, Green and Brown), as shown in Fig. 2. SFA and MUFA followed similar relationships with the module eigengene values, while PUFA showed opposite correlation patterns (Fig. 2). This is an expected result taking into consideration the negative phenotypic correlations between the different FA. As Fig. 3 shows, SFA and MUFA have positive correlations between them and negative with PUFA.

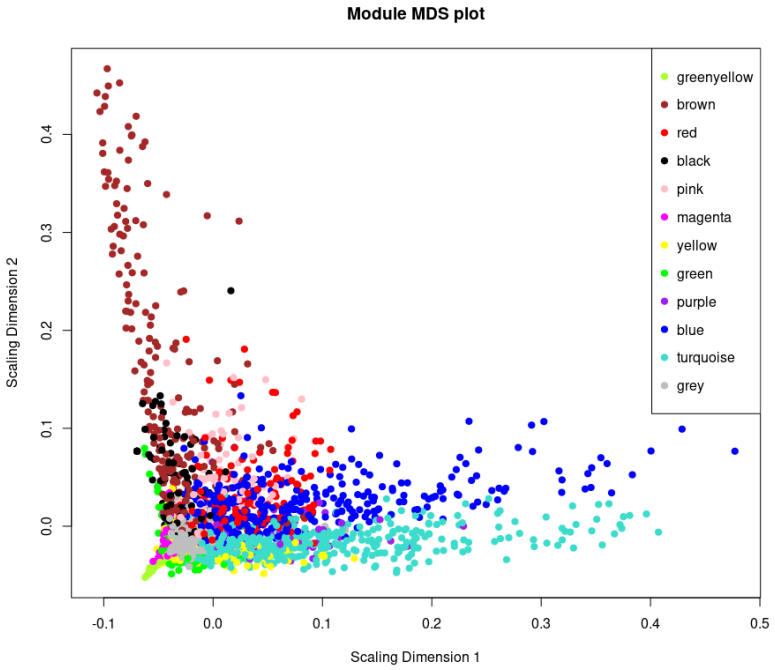


Fig. 1. MDS plot of the gene distribution. The colours represent the modules assigned for each gene.

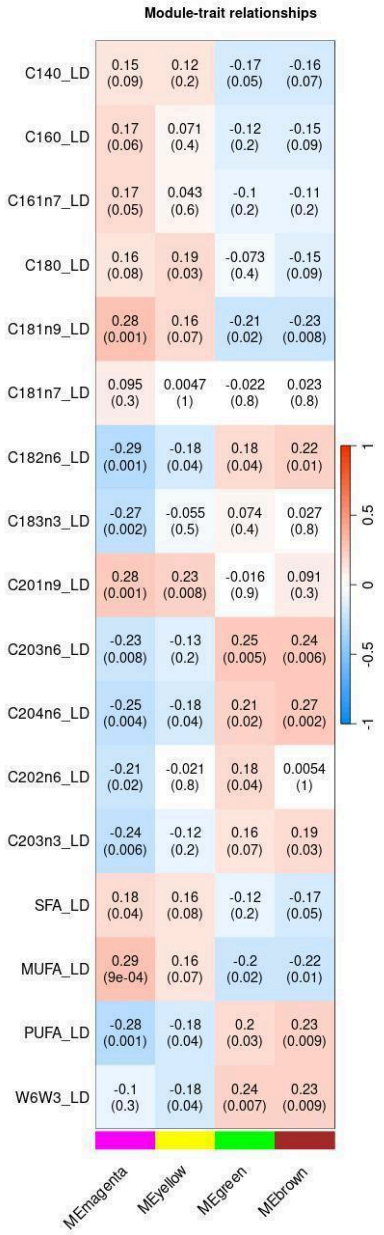


Fig. 2. Relationships between the four significant modules identified in WGCNA and the FA phenotypes in the *longissimus dorsi* muscle of (Iberian × Duroc) × Duroc pigs. Red colour points out the positive correlations, while blue represents the negative ones. The *p*-value of the correlation is shown between parenthesis.

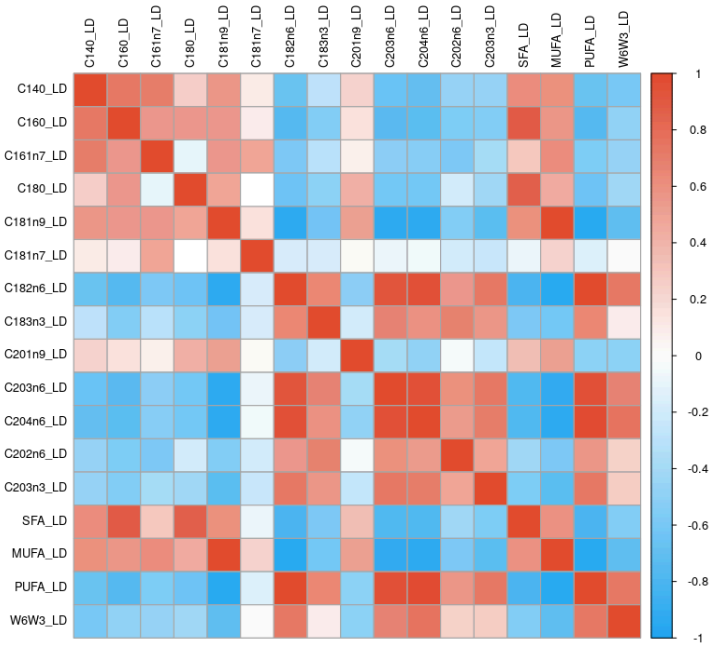


Fig. 3. Phenotypic correlations between the analysed FAs in *longissimus dorsi* muscle.

The eigengene values of Magenta and Yellow modules showed a similar correlation trend, which was opposite to the patterns observed in the Green and Brown modules. Magenta and Yellow modules included 57 and 221 genes, respectively; which resulted to be positively correlated with palmitoleic (C16:1n-7), oleic (C18:1n-9) and the sum of the MUFA in Magenta module, and with stearic acid (C18:0) in Yellow module, but both modules were negatively correlated with linoleic acid (C18:2n-6), arachidonic (C20:4n-6), the ratio n-6/n-3 and the sum of the PUFA. The Magenta module was also negatively correlated with other PUFA such as α-linolenic acid (C18:3n-3), eicosadienoic acid (C20:2n-6), eicosatrienoic acid (C20:3n-3) and dihomo-γ-linolenic acid (C20:3n-6). Meanwhile, the Green and Brown modules included 174 and 234 co-expressed genes, respectively; both of them following an opposite relationship than Magenta and Yellow modules, being positively correlated with C18:2n-6, C20:4n-6, C20:3n-6, C20:2n-6, ratio n-6/n-3 or PUFA, but negatively correlated with C14:0, C18:1n-9 and MUFA.

Biological processes and pathways enriched within gene modules

The gene-enrichment analysis indicated that the genes gathered in the four modules strongly associated with FA are involved in a wide variety of physiological and biological events (Supplementary Table S4). Among significantly enriched pathways, it is worth highlighting the pathways listed in Table 2.

Table 2. Over-represented pathways for each significant module. Only the pathways with a corrected (by Benjamini and Hochberg) *p*-value of ≤ 0.05 are considered.

Over-represented pathways[1]	**Corrected *p*-value**
Magenta module	
Biosynthesis of unsaturated fatty acids (KEGG:01040)	0.0013
PPAR signalling pathway (KEGG:03320)	0.0013
Fatty acid elongation (KEGG:00062)	0.0076
Cholesterol metabolism (KEGG:04979)	0.0294
Yellow module	
Chemokine signalling pathway (KEGG:04062)	0.0000
Platelet activation (KEGG:04611)	0.0000
Glutamatergic synapse (KEGG:04724)	0.0009
Calcium signalling pathway (KEGG:04020)	0.0013
Phospholipase D signalling pathway (KEGG:04072)	0.0037
PI3K-Akt signalling pathway (KEGG:04151)	0.0041
Regulation of lipolysis in adipocytes (KEGG:04923)	0.0298
Green module	
Adipocytokine signalling pathway (KEGG:04920)	0.0225
Brown module	
Oxidative phosphorilation (KEGG:00190)	0.0000
Citrate cycle (TCA cycle) (KEGG:00020)	0.0000
Fatty acid degradation (KEGG:00071)	0.0001
Amino acid degradation	
Valine, leucine and isoleucine degradation (KEGG:00280)	0.0004
Tryptophan metabolism (KEGG:00380)	0.0004
Pyruvate metabolism (KEGG:00620)	0.0044
Lysine degradation (KEGG:00310)	0.0058

[1]Numbers between parenthesis are the KEGG entry for the pathway.

To be noted, module gene-enrichment analysis agreed with the observed module and FA correlations patterns. As previously commented, modules Magenta and Yellow were positively correlated with MUFA and SFA, showing over-representation of pathways related to Biosynthesis of unsaturated fatty acids (KEGG:01040), PPAR signalling pathway (KEGG:03320) and Fatty acid elongation (KEGG:00062). Meanwhile, genes belonging to Green and the Brown modules were negatively correlated with those FAs, but positively linked to PUFA, showing over-representation of pathways associated with Fatty acid (KEGG:00071) and amino acid degradation (KEGG:00280; KEGG:00380; KEGG:00620; KEGG:00310) or Oxidative phosphorylation (KEGG:00190).

Network inference in each module

For each of the four WGCNA modules associated with FA composition we performed network inference by using the graphical Gaussian model. Fig. 4 shows the connections between the most connected genes and their neighbours from the Magenta module, the remaining networks of the significant modules (Yellow, Green and Brown) are available in Supplementary Figure S1. The Magenta network had 52 nodes with 87 edges. The most connected gene was *ADIPOQ* with 10 neighbours; *TRARG1* and *ELOVL6* with eight neighbours; *FABP4* with seven neighbours, and *GYS2*, *PLIN1*, *CIDEC*, and *AK4* with six. The Yellow network inferred using the graphical Gaussian model had 196 nodes with 263 edges. The most connected genes were *FN1* and *MAP3K11* with eight neighbours, and *PLAU*, *CSF1R* and *EHD2* with seven neighbours. The Green network had 161 nodes with 351 edges. The most connected gene was *ATG13* with 11 neighbours. *MEF2C*, *GOT1*, *IRS2*, *NCEH1* and *MAP2K3* had nine neighbours, and *PBX1*, ENSSSCG00000002433 (*PSMC1*), *ILVBL*, *GPAT4* and *EIF4A1* had eight neighbours. The Brown network had 224 and 655 edges. The most connected gene was ENSSSCG00000034019 (*ATP5MF*) with 16 neighbours; whereas *COX6A2*, *ATP5F1C*, *ATP5PB*, *COX7B* and ENSSSCG00000002907 (*COX6B*) had 13 neighbours.

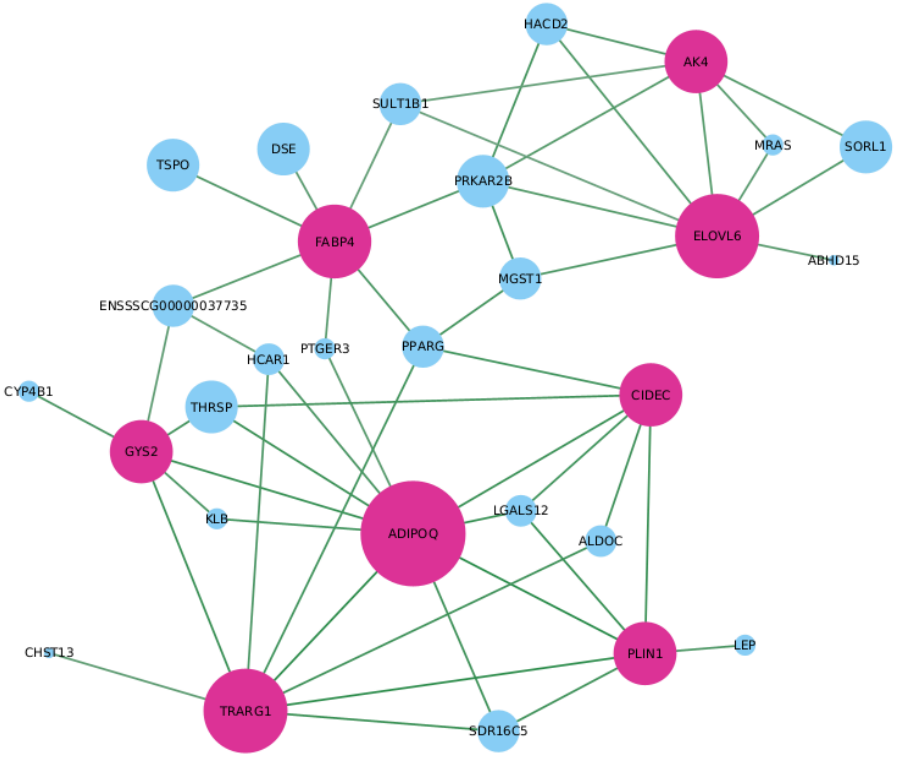


Fig. 4. Magenta module gene network of the most connected genes (genes with six or more neighbours are highlighted in magenta colour and, for visualisation purposes, genes without direct connection with these most connected items are hidden). The size of the node is based on the number of direct neighbours of the gene. The edges are coloured based on the correlation between the genes, green for positive and red for negative. In this network, all the correlations are positive.

Discussion

Co-expression networks enable the clustering of different genes that have similar behaviours. This is especially useful when working with a high number of data such as tables of expressed genes analysed in a whole-transcriptome approach, but it is still often needed to define subsets of genes of particular interest. In our study, these groups were defined as modules of co-expressed genes in muscle that showed correlations with FA composition. Among the four significant modules identified, Magenta and Yellow modules showed positive correlations with SFA and MUFA, while Green and Brown modules had positive correlations with PUFA and its omega-6/omega-3 ratio, which is in agreement with previous results of our research group (Valdés-Hernández et al., 2023). Out of the four modules, the Magenta module clustered the least number of co-expressed genes, while the Brown module had the highest number of co-expressed genes. The most connected genes in the Magenta module were related to the insulin and FA biosynthesis pathways and also aligned with the significant phenotypic correlations found in this module, which showed positive results with three of the MUFAs and their sum. On the other hand, the most connected genes of the Brown module were related to the mitochondrial FA oxidation. In this case, the phenotypic correlations showed positive results with four of the PUFAs, their sum, and the omega-6/omega-3 ratio.

In the following sections, we provide an in-depth analysis of the interaction network defining the gene-gene connections within each of the four modules.

Insulin pathway and FA biosynthesis (Magenta module)

In the Magenta module we found *ADIPOQ* (adiponectin) as the most connected gene. *ADIPOQ* is an adipokine involved in meat quality, and this module was found positively correlated with the palmitoleic acid (C16:1n-7). Besides, the palmitoleic acid has been described as a lipokine and has been related with metabolism homeostasis (Frigolet & Gutiérrez-Aguilar, 2017). In addition to being mainly synthesised by the adipocytes, *ADIPOQ* is involved in FA oxidation and inhibits lipogenesis, but it also regulates insulin sensitivity (Kadowaki & Yamauchi, 2005) and has been associated with average daily weight gain in pigs (Dall'Olio et al., 2009). The insulin sensitivity regulation is thought to be performed also by *TRARG1* (trafficking regulator of GLUT4-1) (Duan et al., 2018), which was found in this module with a positive correlation with *ADIPOQ*. *ADIPOQ* is associated with one of the enriched pathways of this module, the PPAR (peroxisome proliferator-activated receptor) signalling pathway, which has associated other highly connected genes such as *GYS2* (glycogen synthase 2) and *PPARG* (peroxisome proliferator activated receptor gamma).

GYS2 plays a role in obesity through abnormal carbohydrate and lipid metabolism (Morton et al., 2011). PPARs are regulators of the lipid metabolism genes expression (Varga et al., 2011), and *PPARG* can be found, in addition to adipose cells, in other tissues such as breast and colon (Walczak & Tontonoz, 2002). The *PPARG* gene is a key regulator of adipogenesis (Tontonoz et al., 1995), but also plays a role in insulin sensitivity and glucose homeostasis (Walczak & Tontonoz, 2002). It interacts with other of the highly neighbouring genes of this module, such as *FABP4* (fatty acid binding protein 4), whose encoded protein promotes insulin resistance (Garin-Shkolnik et al., 2014), and *CIDEC* (cell death inducing DFFA like effector c), which has been found directly activated by *PPARG* in liver (Matsusue et al., 2008) and up-regulated by *PPARG2* in bovine adipose tissue (Zhou et al., 2018).

PLIN1 (perilipin 1), which belongs to the perilipin family, is found in the lipid droplets of the adipocytes and is involved in intramuscular fat deposition (Gol et al., 2016; Kimmel et al., 2010). Our research group found in a previous work (Valdés-Hernández et al., 2023) that this gene is positively correlated to palmitoleic acid (C16:1n-7), oleic acid (C18:1n-9) and total MUFA content, and negatively correlated to the PUFAs eicosadienoic acid (C20:2n-6) and α-linolenic acid (C18:3n-3). In addition, other studies found that *PLIN1* interacts with *CIDEC* to promote the formation of larger lipid droplets (B. Li et al., 2018; Moreno-Navarrete et al., 2014), which is in agreement with the connections found in our analysis (Fig. 4).

Among the other highly connected genes in this module, *ELOVL6* (elongation of long-chain fatty acids family number 6) has also an important role in insulin sensitivity, and it acts in the elongation of saturated and monounsaturated long-chain FAs (Matsuzaka & Shimano, 2009). This enzyme is involved in the elongation of FAs, e.g., from myristic (C14:0) to palmitic acid (C16:0) or from palmitic to stearic acid (C18:0). Although another gene found in this module, *SCD* (stearoyl-CoA desaturase), did not have a direct connection in the network analysis with *ELOVL6*, both participate in FA metabolism. *SCD* converts palmitic to palmitoleic acid (C16:1n-7) and stearic acid to oleic acid (C18:1n-9), being both of them MUFA, which were positively correlated with this module. Genetic polymorphisms in *ELOVL6* have been previously reported by our research group (Corominas et al., 2013, 2015; Crespo-Piazuelo et al., 2020) as associated with the abundance of myristic, palmitic and palmitoleic acids (C14:0, C16:0 and C16:1n-7). In a previous study we also found this elongase positively associated with MUFA content (Valdés-Hernández et al., 2023). In addition, the analysis suggested (corrected p-value < 0.1) a positive correlation between the Magenta module and the saturated FA: myristic (C14:0), palmitic (C16:0) and stearic (C18:0) acids (Fig. 2). *ELOVL6* is also related to *PPARA*, which, interestingly, has been found in the Green module. A deficiency in *ELOVL6* implies a reduction of *PPARA* and a suppression of

FA synthesis and degradation, and increases insulin sensitivity (Matsuzaka & Shimano, 2009).

FA synthesis genes (Yellow module)

Even though the Yellow module showed many enriched pathways related with cancer and other diseases, there were two significant pathways related with lipid metabolism: PI3K-Akt signalling pathway and the regulation of lipolysis in adipocytes. The PI3K-Akt axis plays a central role in the insulin pathway and acts in liver increasing gluconeogenesis and glycogen synthesis, but also lipid synthesis (Khorami et al., 2015; Petersen & Shulman, 2018), which agrees with the positive correlations found with the SFA and MUFA. Interestingly, one of the most connected genes of this module, *FN1* (fibronectin 1), an activator of the PI3K-Akt signalling pathway (Ji et al., 2020), was found enriched in this pathway. This pathway also had *PIK3CG* as an associated gene, which is one of the four genes that constitute the catalytic subunit of the class I PI3K and works in a wide variety of processes (Engelman et al., 2006). Other enriched genes found, such as *PDGFRA* and *PDGFRB*, were positive regulators of the PI3K-Akt signalling pathway (Cunningham et al., 2022).

The other enriched pathway related with lipid metabolism is the one that regulates lipolysis in adipocytes, which breaks the stored triacylglycerol into glycerol and FAs (Duncan et al., 2007). Three genes of the same family of adenylate cyclases (*ADCY3*, *ADCY4* and *ADCY5*), which regulate lipolysis in adipocytes (Kuo et al., 2018), were enriched in this pathway. *MGLL*, another gene enriched in this pathway, is involved in the final step of lipolysis, as it breaks monoglycerides into glycerol and FAs (Fredrikson et al., 1986).

Another of the most connected genes, *EHD2* (EH domain-containing 2), can be linked to obesity, as it controls the pathway related with caveolate-dependent FA uptake (Matthaeus et al., 2020). Its over-expression suppresses the activity of *PPARG*, a gene that was enriched in the previous module, and its silencing affects insulin sensitivity, lipid storage capacity and lipolysis (Matthaeus et al., 2020; Morén et al., 2019). *EHD2* and caveoline-1 are up-regulated during adipocyte differentiation, while caveoline-1 also regulates the adenylate cyclase pathway (Kuo et al., 2018; Morén et al., 2019).

Adipocytokine signalling (Green module)

The Green module only had one significantly enriched pathway, the adipocytokine signalling pathway, which plays a central role in metabolic homeostasis and is an important component in obesity (Cao, 2014). The gene with the most neighbouring genes that was associated with the adipocytokine signalling pathway was *IRS2* (insulin receptor substrate 2). *IRS2* is

involved in insulin signalling and may be related with human life span and nutrient homeostasis (Taguchi et al., 2007). *PPARA* and *PPARGC1A* (peroxisome proliferator-activated receptor gamma-coactivator 1A) genes were also associated with this pathway. As previously explained for the Magenta module results, *PPARA* interacts with *ELOVL6*. *PPARA* is a nuclear receptor related to FA oxidation, and its decrease can promote obesity (Matsuzaka & Shimano, 2009). *PPARGC1A* is involved in lipid and energy metabolism, and it has been found as a candidate gene for pig meat quality traits (Gandolfi et al., 2011).

Regarding the highly neighbouring genes, some were also found in previous works of our research group, such as *GOT1*, *ILVBL* or *ACSL1*. *GOT1* (glutamate oxaloacetate transaminase 1) is involved in different catalytic processes and also in FA homeostasis (Cunningham et al., 2022), while *ILVBL* (IlvB acetolactate synthase like) is involved in the α-oxidation of FAs (Kitamura et al., 2017). *ACSL1* (acyl-CoA synthetase long chain family number 1) plays a key role in lipid metabolism, as seen for humans and cattle (Brasaemle et al., 2004; Z. Zhao et al., 2020), but also for different pig breeds (Q. Li et al., 2012) and it may promote lipogenic-related genes such as *FABP4*, *APOE* and *FASN* (Shan et al., 2010). These three genes were found in the Magenta and Yellow modules, which have enriched pathways related with FA biosynthesis.

Mitochondrial FA oxidation (Brown module)

The FA degradation pathway was found enriched in the Brown module. One of the highly neighbouring genes found was related with this pathway, *ACADVL* (acyl-CoA dehydrogenase very long chain). *ACADVL* is a target gene of *PPARA* and is involved in mitochondrial FA β-oxidation, and has already been found differentially expressed in the muscle of pigs with extreme FA composition (Puig-Oliveras et al., 2014). Another gene related to this pathway is *HADH* (hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex), whose subunits alpha and beta, are also involved in FA β-oxidation.

The FA degradation pathway produces metabolites that can be used in other enriched pathways of this module related with the mitochondria, such as the citrate pathway (Raimundo et al., 2011; Scheffler, 2007). The most neighbouring genes of this module were composed of three different ATP synthase subunits (*ATP5MF*, *ATP5F1C* and *ATP5PB*) and three cytochrome c oxidases (*COX6A2*, *COX6B* and *COX7B*). ATP synthase and the COX proteins are related to mitochondrial energy metabolism. The ATP synthase is part of an enzyme complex that stimulates the production of ATP more efficiently (Jonckheere et al., 2012). *COX6A2*, which can be found in striated muscle (Taanaman et al., 1993), is a

regulator of respiratory uncoupling in muscle, and it is involved in obesity and insulin resistance (Quintens et al., 2013). *COX6B* is involved in mitochondrial respiration (Lazarou et al., 2009), whereas *COX7B* plays a role in apoptosis (Indrieri et al., 2012). The genes of other ATP synthase subunits were highly neighbouring genes, and they were also found in previous studies of our research group (Valdés-Hernández et al., 2023). However, other highly neighbouring genes, such as some genes from the NADH:ubiquinone reductases family (*NDUFAF7*, *NDUFS1*, and *NDUFS2*) were newly reported.

Another member of this family, *NDUFB8* (NADH:ubiquinone reductase subunit B8), is also another of the top neighbouring genes. *NDUFB8* is a subunit of the mammalian mitochondrial complex I and it is involved in the stability and activity of the mitochondrial complex (Arumugam & Manickam, 2022). *GOT1* and *SCD* were located close to the *NDUFB8* gene. *GOT1* was one of the top neighbouring genes in the Green module, while *SCD* was found in the Magenta module. Furthermore, *ELOVL3* (elongation of long-chain fatty acids family number 3) and *HIF1AN* (hypoxia inducible factor 1 subunit alpha inhibitor) were also found within this SSC14 region. *ELOVL3* is involved in SFA elongation, whereas *HIF1AN* is related to the PI3K signalling pathway, which was enriched in the Yellow module (Puig-Oliveras et al., 2016).

Conclusions

Here, we present an analysis of related gene regulatory networks in the porcine muscle that are potentially related to FA composition in the pig meat. We found four modules that contain relevant muscle gene networks linked with FA composition of pig meat, providing further insights on how the physiology of the pigs influences this relevant phenotype for both meat quality and human health.

Ethics approval

Animal handling and sample collection were performed based on the regulations from the Spanish Policy for Animal Protection RD1201/05, in accordance with the European Union Directive 86/609 about the protection of animals used in experimentation. In addition, this study was carried out meeting the relevant guidelines and regulations of the animal care and use committee of the Institut de Recerca i Tecnologia Agroalimentàries (IRTA). Therefore, the protocol was approved by the Ethical Committee of the IRTA. These regulations adopt the European Code of Conduct for Research Integrity and Good Experimental Practices. In addition, the study is also reported in full compliance with ARRIVE guidelines (https://arriveguidelines.org/).

Data and model availability statement

All data used in this study are present in the paper or its supplementary material. The RNA sequencing data used for these analyses are available in Sequence Read Archive (SRA) from the NCBI BioProject, under the accession number PRJNA882638 (https://www.ncbi.nlm.nih.gov/sra/).

Declaration of generative AI and AI-assisted technologies in the writing process

We declare that no AI technologies were used in the writing process of this article.

Author ORCIDs

C. Sebastià: https://orcid.org/0000-0003-2250-6451

M. Gallopin: https://orcid.org/0000-0002-2431-7825

Y. Ramayo-Caldas: https://orcid.org/0000-0002-8142-0159

J. Estellé: https://orcid.org/0000-0002-6241-1732

J. Valdés-Hernández: https://orcid.org/0000-0001-5314-0988

A. Castelló: https://orcid.org/0000-0001-8497-6251

A. Sánchez: https://orcid.org/0000-0001-9160-1124

D. Crespo-Piazuelo: https://orcid.org/0000-0001-7896-2507

J.M. Folch: https://orcid.org/0000-0003-3689-1303

Declaration of interest

The authors declare no competing interests.

Acknowledgements

We thank the INIA, IRTA and UAB institutions, and their members that contributed to the generation of the animal material used in this study.

Financial support statement

This research received funding from projects AGL2017-82641-R funded by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”; and PID2020-12677RB-C22 funded by MCIN/ AEI /10.13039/501100011033. Research at CRAG was also supported by grants SEV-2015-0533 and CEX2019-000902-S funded by MCIN/AEI/10.13039/501100011033, and by the CERCA Programme, Generalitat de Catalunya. Cristina Sebastià was funded with a FI-AGAUR PhD grant from the Generalitat de Catalunya (2020FI_B_00225) and Jesús Valdés-Hernández with another FI-AGAUR PhD grant from the Generalitat de Catalunya (2019FI_B_00787).

References

Aaslyng, M. D., Oksama, M., Olsen, E. V., Bejerholm, C., Baltzer, M., Andersen, G., Bredie, W. L. P., Byrne, D. V., & Gabrielsen, G. (2007). The impact of sensory quality of pork on consumer preference. *Meat Science*, 76(1), 61–73. https://doi.org/10.1016/j.meatsci.2006.10.014

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Arumugam, S. K., & Manickam, V. (2022). NDUFB8 mutation in a neonate with Leigh's disease. *Journal of Pediatric and Neonatal Individualized Medicine*, 11(2), e110209. https://doi.org/10.7363/110209

Ayuso, M., Fernández, A., Núñez, Y., Benítez, R., Isabel, B., Fernández, A. I., Rey, A. I., González-Bulnes, A., Medrano, J. F., Cánovas, Á., López-Bote, C. J., & Óvilo, C. (2016). Developmental Stage, Muscle and Genetic Type Modify Muscle Transcriptome in Pigs: Effects on Gene Expression and Regulatory Factors Involved in Growth and Metabolism. *PLOS ONE*, 11(12), e0167858. https://doi.org/10.1371/journal.pone.0167858

Bach, F. R. (2008). Bolasso: Model consistent Lasso estimation through the bootstrap. *Proceedings of the 25th International Conference on Machine Learning - ICML '08, 33–40*. https://doi.org/10.1145/1390156.1390161

Benedetti, E., Pučić-Baković, M., Keser, T., Wahl, A., Hassinen, A., Yang, J.-Y., Liu, L., Trbojević-Akmačić, I., Razdorov, G., Štambuk, J., Klarić, L., Ugrina, I., Selman, M. H. J., Wuhrer, M., Rudan, I., Polasek, O., Hayward, C., Grallert, H., Strauch, K., … Krumsiek, J. (2017). Network inference from glycoproteomics data reveals new reactions in the IgG glycosylation pathway. *Nature Communications*, 8(1), 1483. https://doi.org/10.1038/s41467-017-01525-0

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bergen, W. G., & 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. https://doi.org/10.1093/jn/135.11.2499

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8), 1091–1093. https://doi.org/10.1093/bioinformatics/btp101

BOE. (2014). Real Decreto 4/2014, de 10 de enero, por el que se aprueba la norma de calidad para la carne, el jamón, la paleta y la caña de lomo ibérico. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2014-318

Brasaemle, D. L., Dolios, G., Shapiro, L., & Wang, R. (2004). Proteomic Analysis of Proteins Associated with Lipid Droplets of Basal and Lipolytically Stimulated 3T3-L1 Adipocytes. *Journal of Biological Chemistry*, 279(45), 46835–46842. https://doi.org/10.1074/jbc.M409340200

Briggs, M., Petersen, K., & Kris-Etherton, P. (2017). Saturated Fatty Acids and Cardiovascular Disease: Replacements for Saturated Fat to Reduce Cardiovascular Risk. *Healthcare*, 5(2), 29. https://doi.org/10.3390/healthcare5020029

Cao, H. (2014). Adipocytokines in obesity and metabolic disease. *Journal of Endocrinology*, 220(2), T47–T59. https://doi.org/10.1530/JOE-13-0339

Corominas, J., Marchesi, J. A., Puig-Oliveras, A., Revilla, M., Estellé, J., Alves, E., Folch, J. M., & 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), 20. https://doi.org/10.1186/s12711-015-0111-y

Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Pérez-Montarelo, D., Noguera, J. L., Folch, J. M., & Ballester, M. (2013). Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE*, 8(1), e53687. https://doi.org/10.1371/journal.pone.0053687

Crespo-Piazuelo, D., Criado-Mesas, L., Revilla, M., Castelló, A., Noguera, J. L., Fernández, A. I., Ballester, M., & 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. https://doi.org/10.1038/s41598-020-70894-2

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin

Fioretto, L., … Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, 50(D1), D988–D995. https://doi.org/10.1093/nar/gkab1049

Dall'Olio, S., Davoli, R., Buttazzoni, L., Zambonelli, P., & Russo, V. (2009). Study of porcine adiponectin (ADIPOQ) gene and association of a missense mutation with EBVs for production and carcass traits in Italian Duroc heavy pigs. *Livestock Science*, 125(1), 101–104. https://doi.org/10.1016/j.livsci.2009.03.003

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Duan, X., Krycer, J. R., Cooke, K. C., Yang, G., James, D. E., & Fazakerley, D. J. (2018). Membrane Topology of Trafficking Regulator of GLUT4 1 (TRARG1). *Biochemistry*, 57(26), 3606–3615. https://doi.org/10.1021/acs.biochem.8b00361

Duncan, R. E., Ahmadian, M., Jaworski, K., Sarkadi-Nagy, E., & Sul, H. S. (2007). Regulation of Lipolysis in Adipocytes. *Annual Review of Nutrition*, 27(1), 79–101. https://doi.org/10.1146/annurev.nutr.27.061406.093734

Engelman, J. A., Luo, J., & Cantley, L. C. (2006). The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nature Reviews Genetics*, 7(8), 606–619. https://doi.org/10.1038/nrg1879

FAOSTAT. (2023). FAOSTAT. FAOSTAT. https://www.fao.org/faostat/en/#home

Fernandez, X., Monin, G., Talmant, A., Mourot, J., & Lebret, B. (1999). Influence of intramuscular fat content on the quality of pig meat. Consumer acceptability of m. Longissimus lumborum. *Meat Science*.

Fredrikson, G., Tornqvist, H., & Belfrage, P. (1986). Hormone-sensitive lipase and monoacylglycerol lipase are both required for complete degradation of adipocyte triacylglycerol. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, 876(2), 288–293. https://doi.org/10.1016/0005-2760(86)90286-9

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Frigolet, M. E., & Gutiérrez-Aguilar, R. (2017). The Role of the Novel Lipokine Palmitoleic Acid in Health and Disease. *Advances in Nutrition*, 8(1), 173S-181S. https://doi.org/10.3945/an.115.011130

Gandolfi, G., Cinar, M. U., Ponsuksili, S., Wimmers, K., Tesfaye, D., Looft, C., Jüngst, H., Tholen, E., Phatsara, C., Schellander, K., & Davoli, R. (2011). Association of PPARGC1A and CAPNS1 gene polymorphisms and expression with meat quality traits in pigs. *Meat Science*, 89(4), 478–485. https://doi.org/10.1016/j.meatsci.2011.05.015

Garin-Shkolnik, T., Rudich, A., Hotamisligil, G. S., & Rubinstein, M. (2014). FABP4 Attenuates PPARγ and Adipogenesis and Is Inversely Correlated With PPARγ in Adipose Tissues. *Diabetes*, 63(3), 900–911. https://doi.org/10.2337/db13-0436

Gol, S., Ros-Freixedes, R., Zambonelli, P., Tor, M., Pena, R. N., Braglia, S., Zappaterra, M., Estany, J., & Davoli, R. (2016). Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs. *Journal of Animal Breeding and Genetics*, 133(1), 24–30. https://doi.org/10.1111/jbg.12159

Horvath, S., & Dong, J. (2008). Geometric Interpretation of Gene Coexpression Network Analysis. *PLoS Computational Biology*, 4(8), e1000117. https://doi.org/10.1371/journal.pcbi.1000117

Indrieri, A., van Rahden, V. A., Tiranti, V., Morleo, M., Iaconis, D., Tammaro, R., D'Amato, I., Conte, I., Maystadt, I., Demuth, S., Zvulunov, A., Kutsche, K., Zeviani, M., & Franco, B. (2012). Mutations in COX7B Cause Microphthalmia with Linear Skin Lesions, an Unconventional Mitochondrial Disease. *The American Journal of Human Genetics*, 91(5), 942–949. https://doi.org/10.1016/j.ajhg.2012.09.016

Ji, J., Chen, L., Zhuang, Y., Han, Y., Tang, W., & Xia, F. (2020). Fibronectin 1 inhibits the apoptosis of human trophoblasts by activating the PI3K/Akt signaling pathway. *International Journal of Molecular Medicine*. https://doi.org/10.3892/ijmm.2020.4735

Jonckheere, A. I., Smeitink, J. A. M., & Rodenburg, R. J. T. (2012). Mitochondrial ATP synthase: Architecture, function and pathology. *Journal of Inherited Metabolic Disease*, 35(2), 211–225. https://doi.org/10.1007/s10545-011-9382-9

Kadowaki, T., & Yamauchi, T. (2005). Adiponectin and Adiponectin Receptors. *Endocrine Reviews*, 26(3), 439–451. https://doi.org/10.1210/er.2005-0005

Khorami, S. A. H., Huzwah, A., Mutalib, A., & Sokhini, M. (2015). PI3K/AKT pathway in modulating glucose homeostasis and its alteration in Diabetes. *Annals of Medical and Biomedical Sciences*, 1(2), 46–55.

Kimmel, A. R., Brasaemle, D. L., McAndrews-Hill, M., Sztalryd, C., & Londos, C. (2010). Adoption of PERILIPIN as a unifying nomenclature for the mammalian PAT-family of intracellular lipid storage droplet proteins. *Journal of Lipid Research*, 51(3), 468–471. https://doi.org/10.1194/jlr.R000034

Kitamura, T., Seki, N., & Kihara, A. (2017). Phytosphingosine degradation pathway includes fatty acid α-oxidation reactions in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences*, 114(13). https://doi.org/10.1073/pnas.1700138114

Kogelman, L. J. A., Cirera, S., Zhernakova, D. V., Fredholm, M., Franke, L., & Kadarmideen, H. N. (2014). Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. *BMC Medical Genomics*, 7(1), 57. https://doi.org/10.1186/1755-8794-7-57

Kuo, A., Lee, M. Y., Yang, K., Gross, R. W., & Sessa, W. C. (2018). Caveolin-1 regulates lipid droplet metabolism in endothelial cells via autocrine prostacyclin–stimulated, cAMP-mediated lipolysis. *Journal of Biological Chemistry*, 293(3), 973–983. https://doi.org/10.1074/jbc.RA117.000980

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1), 559. https://doi.org/10.1186/1471-2105-9-559

Latorre, M. A., Pomar, C., Faucitano, L., Gariépy, C., & Méthot, S. (2008). The relationship within and between production performance and meat quality characteristics in pigs from three different genetic lines. *Livestock Science*, 115(2–3), 258–267. https://doi.org/10.1016/j.livsci.2007.08.013

Lazarou, M., Smith, S. M., Thorburn, D. R., Ryan, M. T., & McKenzie, M. (2009). Assembly of nuclear DNA-encoded subunits into mitochondrial complex IV, and their preferential integration into supercomplex forms in patient mitochondria: Mitochondrial complex IV assembly. *FEBS Journal*, 276(22), 6701–6713. https://doi.org/10.1111/j.1742-4658.2009.07384.x

Le Novère, N. (2015). Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics*, 16(3), 146–158. https://doi.org/10.1038/nrg3885

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.

Li, B., Weng, Q., Dong, C., Zhang, Z., Li, R., Liu, J., Jiang, A., Li, Q., Jia, C., Wu, W., & Liu, H. (2018). A Key Gene, PLIN1, Can Affect Porcine Intramuscular Fat Content Based on Transcriptome Analysis. *Genes*, 9(4), 194. https://doi.org/10.3390/genes9040194

Li, Q., Tao, Z., Shi, L., Ban, D., Zhang, B., Yang, Y., Zhang, H., & Wu, C. (2012). Expression and genome polymorphism of ACSL1 gene in different pig breeds. *Molecular Biology Reports*, 39(9), 8787–8792. https://doi.org/10.1007/s11033-012-1741-6

Livingstone, K. M., Brayner, B., Celis-Morales, C., Ward, J., Mathers, J. C., & Bowe, S. J. (2022). Dietary patterns, genetic risk, and incidence of obesity: Application of reduced rank regression in 11,735 adults from the UK Biobank study. *Preventive Medicine*, 158, 107035. https://doi.org/10.1016/j.ypmed.2022.107035

Ma, S., Gong, Q., & Bohnert, H. J. (2007). An Arabidopsis gene network based on the graphical Gaussian model. *Genome Research*, 17(11), 1614–1625. https://doi.org/10.1101/gr.6911207

Matsusue, K., Kusakabe, T., Noguchi, T., Takiguchi, S., Suzuki, T., Yamano, S., & Gonzalez, F. J. (2008). Hepatic Steatosis in Leptin-Deficient Mice Is Promoted by the PPARγ Target Gene Fsp27. *Cell Metabolism*, 7(4), 302–311. https://doi.org/10.1016/j.cmet.2008.03.003

Matsuzaka, T., & Shimano, H. (2009). Elovl6: A new player in fatty acid metabolism and insulin sensitivity. *Journal of Molecular Medicine*, 87(4), 379–384. https://doi.org/10.1007/s00109-009-0449-0

Matthaeus, C., Lahmann, I., Kunz, S., Jonas, W., Melo, A. A., Lehmann, M., Larsson, E., Lundmark, R., Kern, M., Blüher, M., Olschowski, H., Kompa, J., Brügger, B., Müller, D. N., Haucke, V., Schürmann, A., Birchmeier, C., & Daumke, O. (2020). EHD2-mediated restriction of caveolar dynamics regulates cellular fatty acid uptake. *Proceedings of the National Academy of Sciences*, 117(13), 7471–7481. https://doi.org/10.1073/pnas.1918415117

Morén, B., Hansson, B., Negoita, F., Fryklund, C., Lundmark, R., Göransson, O., & Stenkula, K. G. (2019). EHD2 regulates adipocyte function and is enriched at cell surface–associated lipid droplets in primary human adipocytes. *Molecular Biology of the Cell*, 30(10), 1147–1159. https://doi.org/10.1091/mbc.E18-10-0680

Moreno-Navarrete, J. M., Ortega, F., Serrano, M., Rodriguez-Hermosa, J. I., Ricart, W., Mingrone, G., & Fernández-Real, J. M. (2014). CIDEC/FSP27 and PLIN1 gene expression run in parallel to mitochondrial genes in human adipose tissue, both increasing after weight loss. *International Journal of Obesity*, 38(6), 865–872. https://doi.org/10.1038/ijo.2013.171

Morton, N. M., Nelson, Y. B., Michailidou, Z., Di Rollo, E. M., Ramage, L., Hadoke, P. W. F., Seckl, J. R., Bunger, L., Horvat, S., Kenyon, C. J., & Dunbar, D. R. (2011). A Stratified Transcriptomics Analysis of Polygenic Fat and Lean Mouse Adipose Tissues Identifies Novel Candidate Obesity Genes. *PLoS ONE*, 6(9), e23944. https://doi.org/10.1371/journal.pone.0023944

Nieto, R., García-Casco, J., Lara, L., Palma-Granados, P., Izquierdo, M., Hernandez, F., Dieguez, E., Luis Duarte, J., & Batorek-Lukač, N. (2019). Ibérico (Iberian) Pig. In M. Candek-Potokar & R. M. Nieto Linan (Eds.), European Local Pig Breeds—Diversity and Performance. A study of project TREASURE. IntechOpen. https://doi.org/10.5772/intechopen.83765

Petersen, M. C., & Shulman, G. I. (2018). Mechanisms of Insulin Action and Insulin Resistance. *Physiological Reviews*, 98(4), 2133–2223. https://doi.org/10.1152/physrev.00063.2017

Puig-Oliveras, A., Ramayo-Caldas, Y., Corominas, J., Estellé, J., Pérez-Montarelo, D., Hudson, N. J., Casellas, J., Folch, J. M., & Ballester, M. (2014). Differences in Muscle Transcriptome among Pigs Phenotypically Extreme for Fatty Acid Composition. *PLoS ONE*, 9(6), e99720. https://doi.org/10.1371/journal.pone.0099720

Puig-Oliveras, A., Revilla, M., Castelló, A., Fernández, A. I., Folch, J. M., & 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. https://doi.org/10.1038/srep31803

Quintens, R., Singh, S., Lemaire, K., De Bock, K., Granvik, M., Schraenen, A., Vroegrijk, I. O. C. M., Costa, V., Van Noten, P., Lambrechts, D., Lehnert, S., Van Lommel, L., Thorrez, L., De Faudeur, G., Romijn, J. A., Shelton, J. M., Scorrano, L., Lijnen, H. R., Voshol, P. J., … Schuit, F. (2013). Mice Deficient in the Respiratory Chain Gene Cox6a2 Are Protected against High-Fat Diet-Induced Obesity and Insulin Resistance. *PLoS ONE*, 8(2), e56719. https://doi.org/10.1371/journal.pone.0056719

Raimundo, N., Baysal, B. E., & Shadel, G. S. (2011). Revisiting the TCA cycle: Signaling to tumor formation. *Trends in Molecular Medicine*, 17(11), 641–649. https://doi.org/10.1016/j.molmed.2011.06.001

Ramayo-Caldas, Y., Ballester, M., Sánchez, J. P., González-Rodríguez, O., Revilla, M., Reyer, H., Wimmers, K., Torrallardona, D., & Quintanilla, R. (2018). Integrative approach using liver and duodenum RNA-Seq data identifies candidate genes and pathways associated with feed efficiency in pigs. *Scientific Reports*, 8(1), 558. https://doi.org/10.1038/s41598-017-19072-5

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Schäfer, J., & Strimmer, K. (2005). An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics*, 21(6), 754–764. https://doi.org/10.1093/bioinformatics/bti062

Scheffler, I. E. (2007). Mitochondria (2nd ed.).

Serra, X., Gil, F., Pérez-Enciso, M., Oliver, M. A., Vázquez, J. M., Gispert, M., Díaz, I., Moreno, F., Latorre, R., & Noguera, J. L. (1998). A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livestock Production Science*, 56(3), 215–223. https://doi.org/10.1016/S0301-6226(98)00151-1

Shan, D., Li, J., Wu, L., Li, D., Hurov, J., Tobin, J. F., Gimeno, R. E., & Cao, J. (2010). GPAT3 and GPAT4 are regulated by insulin-stimulated phosphorylation and play distinct roles in adipogenesis. *Journal of Lipid Research*, 51(7), 1971–1981. https://doi.org/10.1194/jlr.M006304

Spurlock, M. E., & Gabler, N. K. (2008). The Development of Porcine Models of Obesity and the Metabolic Syndrome. *The Journal of Nutrition*, 138(2), 397–402. https://doi.org/10.1093/jn/138.2.397

Svoboda, M., Mungenast, F., Gleiss, A., Vergote, I., Vanderstichele, A., Sehouli, J., Braicu, E., Mahner, S., Jäger, W., Mechtcheriakova, D., Cacsire-Tong, D., Zeillinger, R., Thalhammer, T., & Pils, D. (2018). Clinical Significance of Organic Anion Transporting Polypeptide Gene Expression in High-Grade Serous Ovarian Cancer. *Frontiers in Pharmacology*, 9, 842. https://doi.org/10.3389/fphar.2018.00842

Taanaman, J.-W., Hall, R. E., Tang, C., Marusich, M. F., Kennaway, N. G., & Capaldi, R. A. (1993). Tissue distribution of cytochrome c oxidase isoforms in mammals. Characterization with monoclonal and polyclonal antibodies. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1225(1), 95–100. https://doi.org/10.1016/0925-4439(93)90128-N

Taguchi, A., Wartschow, L. M., & White, M. F. (2007). Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis. *Science*, 317(5836), 369–372. https://doi.org/10.1126/science.1142179

Tontonoz, P., Hu, E., & Spiegelman, B. M. (1995). Regulation of adipocyte gene expression and differentiation by peroxisome proliferator activated receptor y. *Current Opinion in Genetics & Development*, 5, 571–576.

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. (2023). Global analysis of the association between pig muscle fatty acid composition and gene expression using RNA-Seq. *Scientific Reports*, 13(1), 535. https://doi.org/10.1038/s41598-022-27016-x

van Laack, R. L., Stevens, S. G., & Stalder, K. J. (2001). The influence of ultimate pH and intramuscular fat content on pork tenderness and tenderization. *Journal of Animal Science*, 79(2), 392. https://doi.org/10.2527/2001.792392x

Varga, T., Czimmerer, Z., & Nagy, L. (2011). PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1812(8), 1007–1022. https://doi.org/10.1016/j.bbadis.2011.02.014

Walczak, R., & Tontonoz, P. (2002). PPARadigms and PPARadoxes: Expanding roles for PPARγ in the control of lipid metabolism. *Journal of Lipid Research*, 43(2), 177–186. https://doi.org/10.1016/S0022-2275(20)30159-0

Zhang, B., & Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Statistical Applications in Genetics and Molecular Biology*, 4(1). https://doi.org/10.2202/1544-6115.1128

Zhang, L., Wu, W., Lee, Y.-K., Xie, J., & Zhang, H. (2018). Spatial Heterogeneity and Co-occurrence of Mucosal and Luminal Microbiome across Swine Intestinal Tract. *Frontiers in Microbiology*, 9, 48. https://doi.org/10.3389/fmicb.2018.00048

Zhao, T., Liu, H., Roeder, K., Lafferty, J., & Wasserman, L. (2012). The huge Package for High-dimensional Undirected Graph Estimation in R. *Journal of Machine Learning Research*, 13, 1059–1062.

Zhao, Z., Abbas Raza, S. H., Tian, H., Shi, B., Luo, Y., Wang, J., Liu, X., Li, S., Bai, Y., & Hu, J. (2020). Effects of overexpression of ACSL1 gene on the synthesis of unsaturated fatty acids in adipocytes of bovine. *Archives of Biochemistry and Biophysics*, 695, 108648. https://doi.org/10.1016/j.abb.2020.108648

Zhou, Y., Li, M., Hu, X., Cai, H., Hua, L., Wang, J., Huang, Y., Lan, X., Lei, C., Liu, G. E., Li, C., Plath, M., & Chen, H. (2018). Characterization of candidate genes for bovine adipogenesis reveals differences of TUSC5 isoforms caused by novel alternative splicing. *Oncotarget*, 0(0). https://doi.org/10.18632/oncotarget.23482

Paper II

Interrelation between gut microbiota, SCFA and fatty acid composition in pigs

Sebastià, C., Folch, J.M., Ballester, M., Estellé, J., Passols, M., Muñoz, M., García-Casco, J.M., Fernández, A.I., Castelló, A., Sánchez, A., Crespo-Piazuelo, D.

mSystems (2024), 9(1), p. e01049-23.

Interrelation between gut microbiota, SCFA and fatty acid composition in pigs

Sebastià, C.[a,b,#], Folch, J.M.[a,b], Ballester, M.[c], Estellé, J.[d], Passols, M.[a], Muñoz, M.[e], García-Casco, J.M.[f], Fernández, A.I.[e], Castelló, A.[a,b], Sánchez, A.[a,b], Crespo-Piazuelo, D.[c]

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

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

[c]Departament de Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Caldes de Montbui, Spain.

[d]Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, France.

[e]Departamento de Mejora Genética Animal, INIA-CSIC, Madrid, Spain.

[f]Centro I+D en Cerdo Ibérico INIA-Zafra, Zafra, Spain.

Address correspondence to Cristina Sebastià, cristina.sebastia@cragenomica.es

Abstract

The gut microbiota is a key player in the host metabolism. Some bacteria are able to ferment non-digestible compounds and produce short-chain fatty acids that the host can later transform and accumulate in tissue. In this study, we aimed to better understand the relationships between the microorganisms and the short-chain fatty acid composition of the rectal content, including the possible linkage with the fatty acid composition in backfat and muscle of the pig. We studied a Duroc × Iberian crossbred population and we found significant correlations between different bacterial and archaeal genera and the fatty acid profile. The abundance of n-butyric acid in the rectal content was positively associated with *Prevotella* spp. and negatively associated with *Akkermansia* spp., while conversely, the abundance of acetic acid was negatively and positively associated with the levels of *Prevotella* spp. and *Akkermansia* spp., respectively. The most abundant genus, *Rikenellaceae RC9 gut group*, had a positive correlation with palmitic acid in muscle and negative correlations with stearic acid in backfat and oleic acid in muscle. These results suggest the possible role of *Prevotella* spp. and *Akkermansia* spp. as biomarkers for acetic and n-butyric acids, and the relationship of *Rikenellaceae RC9 gut group* with the lipid metabolism, building up the potential, although indirect, role of the microbiota in the modification of the backfat and muscle fatty acid composition of the host.

Importance: The vital role of the gut microbiota on its host metabolism makes it essential to know how its modulation is mirrored on the fatty acid composition of the host. Our findings suggest *Prevotella* spp. and *Akkermansia* spp. as potential biomarkers for the levels of beneficial short-chain fatty acids, and the possible influence of *Rikenellaceae RC9 gut group* in the backfat and muscle fatty acid composition of the pig.

Introduction

Pig is one of the most important livestock species, being pork the second most produced meat worldwide after chicken (1). In Spain, the Iberian pig is an autochthonous breed, usually raised in “dehesas” (holm oaks) and fed with acorns as part of its diet (2). The Iberian pig has lower growth efficiency and higher fat deposition than conventional breeds, but it excels in meat quality due to its different fatty acid (FA) profile. Iberian pork has higher percentage of oleic acid and saturated FA (SFA), and less abundance of polyunsaturated FAs (PUFA) (3, 4), which makes this meat ideal for cured products and appeals to the consumer due to its organoleptic properties. In production farms, Iberian pigs are crossbred with Duroc pigs, producing more efficient animals without overly impairing the production of ham and other cured products (2, 5).

Aside from its interest in meat production, the pig, as a monogastric omnivorous animal, can be used as model species for different human syndromes and conditions, including gut microbiota-related diseases and dietary modulation of the gut population (6, 7). Thus, gut microbiota composition influences host metabolism and its general welfare (8), and is deeply modulated by the diet (9, 10), albeit being also affected by environmental factors and by the host genetics (11–13). During the last decades, the distinction and classification of gut microbial communities has been performed through 16S rRNA gene sequencing among other methods (14, 15). In addition, the metabolites produced by the gut microbiota have also been studied. Gut microbiota is able to ferment non-digestible compounds such as dietary fibres and produce, amongst other metabolites, short-chain FAs (SCFAs) (16). In the gut, SCFAs are an energy source available for the host (17) and can be incorporated in different metabolic pathways (18, 19). SCFAs that cross the intestinal barrier and are incorporated into the host metabolism can influence the FA profile across tissues, as they act as a substrate for FA biosynthesis and are also involved in the regulation of the lipid metabolism (20). Hence, identifying the main microbial organisms involved in SCFA production is key to imagine strategies capable of influencing the FA composition of the host through any mechanism able to modify the gut microbiota, such as the diet. Therefore, this study aimed to assess the relationship between gut microbiota, SCFA abundance and FA profile in backfat and muscle of the pig, which to our knowledge, has not yet been explored.

Materials and Methods

Ethics approval

Animal handling and sample collection were performed based on the regulations from the Spanish Policy for Animal Protection RD53/2013 in accordance with the European Union Directive 2010/63/EU about the protection of animals used in experimentation. Pigs were slaughtered in a commercial abattoir following national and institutional guidelines for Good Experimental Practices.

Animal material

The animal data for this study consisted of 288 healthy pigs from an F1 Duroc × Iberian crossbred population. All the animals were housed in the same farm under a controlled environment in intensive conditions and were fed *ad libitum* with a standard commercial diet based on barley and wheat. Pigs were slaughtered in a commercial abattoir with an average weight of 138.8 kg (SD = 11.46 kg) in four different days. Samples of backfat, *longissimus dorsi* muscle and rectal content were collected, snap-frozen in liquid nitrogen and preserved at −80 ºC until further used.

Microbial DNA extraction and sequencing

Microbial DNA was extracted from 0.2 g of rectal content using the PowerFecal kit (MoBio Laboratories, Carlsbad, CA, USA) and following the manufacturer's guidelines. DNA purity and concentration was assessed with a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

The V3-V4 region of the 16S rRNA gene was amplified following the recommendations on the *16S Metagenomic Sequencing Library Preparation* guide (Illumina, San Diego, CA, USA). The complete description of the procedure can be accessed at Crespo-Piazuelo et al. (2019). In summary, all the amplicon pooled libraries from the 288 animals were sequenced in three runs by a MiSeq System (Illumina, San Diego, CA, USA) using the MiSeq Reagent Kit v3 (600-cycle format, 2×300 bp paired-end reads) in the Sequencing Service of the FISABIO (*Fundació per al Foment de la Investigació Sanitària i Biomèdica de la Comunitat Valenciana*). A total of 17.991 Gb was obtained with an average of 104,115 reads per sample. An outlier was discarded because it had 1,758,983 reads while the rest of the samples were in a 34,186 to 218,360 range. Raw sequencing data can be accessed under

the BioProject accession number PRJNA540380 in the NCBI Sequence Read Archive (SRA).

Taxonomy classification

The remaining 287 samples were analysed with QIIME2 v2021.11.0 (21). Sequences with a Phred score lower than 33 were filtered out, and were processed in QIIME2 with the DADA2 algorithm, denoising and trimming 280 nucleotides in the forward read sequences and 230 in the reverse read sequences. Two of the samples were discarded because they did not match the quality criteria. The three technical replicates were merged per individual using the sum overlap method. For the 285 remaining samples, phylogeny was calculated and they were taxonomically classified at 99% of similarity using the SILVA database release 138 (22) specifying the V3-V4 region forward and reverse primers. After the QIIME2 pipeline, 46,284 ASVs (Amplicon Sequence Variants) were obtained and filtered with the following procedure to avoid artefacts and singletons. Using R v4.1.3 software (23), we kept only those ASVs with a greater abundance than 0.005% of the total number of counts:

$$\frac{sample\ ASVs \cdot 100}{\Sigma ASVs} > 0.005$$

Hence, 2,671 ASVs were obtained and grouped in 119 genera, 67 families and 17 phyla. For diversity analyses, filtered samples were rarefied with phyloseq v1.38.0 (24) to an equal number of ASVs per individual (n = 8,864), which was the minimum number of counts. Then, α- and β-diversities were calculated with phyloseq and vegan v2.5-7 R packages (25). Thus, α-diversity was estimated by the Shannon index, while β-diversities were determined by Bray-Curtis' dissimilarities (26).

Fatty acid composition

Fatty acid composition in backfat and *longissimus dorsi* muscle was measured by gas chromatography. For backfat, the analysis was performed in ASICI (*Asociación Interprofesional del Cerdo Ibérico*) at Zafra (Spain) and following the official method explained in BOE (27). Two Perkin Elmer chromatographs equipped with a flame ionization detector (FID) were used, with autosamplers and a fused silica capillary column (20 m × 0.32 mm internal diameter and 0.25 μm film thickness). Intramuscular fatty acid composition was performed in STAB (*Servicio de Técnicas Aplicadas a la Biociencia*) from Extremadura University, where lipids were extracted from muscle according to Bligh & Dyer (28) and analysed by gas chromatography using a Bruker Scion 456 GC equipped with a FID and a DB-225ms (Agilent Technologies, Santa Clara, CA, USA) capillary column (30 m × 0.25 mm

internal diameter and 0.25 µm film thickness). Results were expressed as the percentage of total fatty acids identified. Out of the total 285 samples, 14 samples were discarded because there was not enough material for measuring FA composition in one of the two tissues. Thus, downstream analyses were carried out with 271 samples. Information on the 16 measured FAs and the three calculated indices is shown in Table 1.

Table 1. Mean relative abundances (%) and standard deviations (SD) of the analysed fatty acids from backfat and *longissimus dorsi* muscle.

Group	**Trait**	**Backfat**		***Longissimus dorsi***	
		Mean	**SD**	**Mean**	**SD**
SFA	C14:0 (myristic acid)	1.57	0.140	1.52	0.147
	C15:0 (pentadecanoic acid)	0.03	0.006	0.05	0.034
	C16:0 (palmitic acid)	24.94	0.888	27.67	1.986
	C17:0 (margaric acid)	0.20	0.037	0.14	0.102
	C18:0 (stearic acid)	10.71	0.834	12.46	1.276
	C20:0 (arachidic acid)	0.20	0.023	0.18	0.108
	SFA (saturated fatty acids)	37.67	1.419	42.02	3.067
MUFA	C16:1 (palmitoleic acid)	2.85	0.482	4.62	0.628
	C17:1 (heptadecenoic acid)	0.25	0.038	0.18	0.088
	C18:1 (oleic acid)	49.11	1.419	47.75	3.276
	C20:1 (gondoic acid)	1.28	0.158	0.67	0.231
	MUFA (monounsaturated fatty acids)	53.49	1.266	53.23	3.196
PUFA	C18:2 (linoleic acid)	7.64	0.691	3.81	0.758
	C18:3 (α-linolenic acid)	0.43	0.043	0.14	0.037
	C20:2 (eicosadienoic acid)	0.46	0.047	0.10	0.097
	C20:3(n-3) eicosatrienoic acid)	0.06	0.018	0.07	0.054
	C20:3(n-6) (dihomo-γ-linolenic acid)	0.11	0.041	0.08	0.116
	C20:4 (arachidonic acid)	0.14	0.025	0.55	0.291
	PUFA (polyunsaturated fatty acids)	8.84	0.751	4.75	1.045

Short chain fatty acid composition in the rectal content

The abundance of SCFAs in the rectal content was assessed through the analytical method SM 5560-D, using a Varian CP-3800 gas chromatograph with an FID detector and a TRB-FFAP (free fatty acids phase) chromatographic column (15 m × 0.53 mm × 0.5 mm). The FFAP stationary phase was a nitroterephthalic-acid-modified polyethylene glycol (PEG) column. SCFA composition of the 271 individuals is available in Table 2. In addition, principal component analysis (PCA) of the SCFA composition was performed with the prcomp function in R.

Table 2. Mean relative abundances (%) and standard deviations (SD) of the analysed short-chain fatty acids from rectal content.

SCFA	Mean	SD
Acetic acid	59.23	2.991
Propionic acid	19.88	1.723
n-Butyric acid	10.13	2.025
iso-Butyric acid	3.11	0.531
n-Valeric acid	2.84	0.347
iso-Valeric acid	4.76	0.682

Correlations and associations analyses between microbiota, SCFA and FA compositions

Pearson correlations were calculated within and between the traits from the three different datasets, i.e., microbiota using bacterial genera abundances, SCFA and FA compositions, using the corr.test function from psych v2.1.9 R package (29). Thereafter, Benjamini and Hochberg method (30) was applied for multiple testing correction. The significance threshold was established at $p.adj < 0.05$. Correlation results were plotted as heatmaps with the Pheatmap R package (v1.0.12) (31).

In addition, the relationship between the taxonomic data, FA and SCFA compositions in the 271 samples was studied with the mixOmics v6.18.1 R package (32, 33). The two datasets used as input were (1) the 38 FA phenotypes (16 backfat FA, 16 muscle FA, and 6 SCFA)

pre-corrected by sex (2 categories, male and female) and batch (4 categories, based on the slaughter day); and (2) the filtered taxonomic data at genus level, which was normalized using the centred log-ratio transformation (CLR) option from mixOmics. Then, data was analysed by two different methods in mixOmics: the regularized Canonical Correlation Analysis (rCCA), using the shrinkage method (λ_1 = 0.28; λ_2 = 0.15), and the PLS (Partial Least Squares) method.

Furthermore, the adonis2 function from the vegan R package was used to implement a PERMANOVA analysis for each phenotype, related with the ASV relative abundances table. A non-metric multidimensional scaling (NMDS) using Bray-Curtis dissimilarities was carried out to plot the results together with an ordination diagram, using vegan's package different functions.

Functional prediction analysis

PICRUSt2 (34) plug-in for QIIME2 v2021.11 was used to carry out the functional prediction analysis, using the *full-pipeline* option and the default parameters to obtain the results. The KEGG Orthology abundances were then processed with R, we calculated the Pearson correlations with Benjamini and Hochberg multiple testing correction, and we performed an association with the FA composition and with the microbial genera through the rCCA method from mixOmics.

Results

Taxonomic classification and diversity analysis of gut microbiota

We found 2,671 ASVs in the rectal content of 285 pigs, which were classified in 119 genera, 67 families and 17 phyla. At phylum level, two out of the 17 phyla belonged to the Archaea domain, Euryarchaeota and Thermoplasmatota. There was only one genus within each of these two archaeal phyla, *Methanobrevibacter* and an uncultured Methanomethylophilaceae genus, respectively.

The two most abundant phyla were Bacteroidota (44.08%) and Firmicutes (38.17%) (Figure 1A). Other remarkable phyla in terms of abundance were Spirochaetota (8.18%) and Verrucomicrobiota (3.51%). At genus level, *Rikenellaceae RC9 gut group* (8.98%) was the most abundant, closely followed by *Treponema* (7.56%), *Prevotella* (5.48%) and an *unspecified Muribaculaceae* genus (5.27%). The total counts and the relative abundances at phylum and genus level are available in Supplemental Table S1.

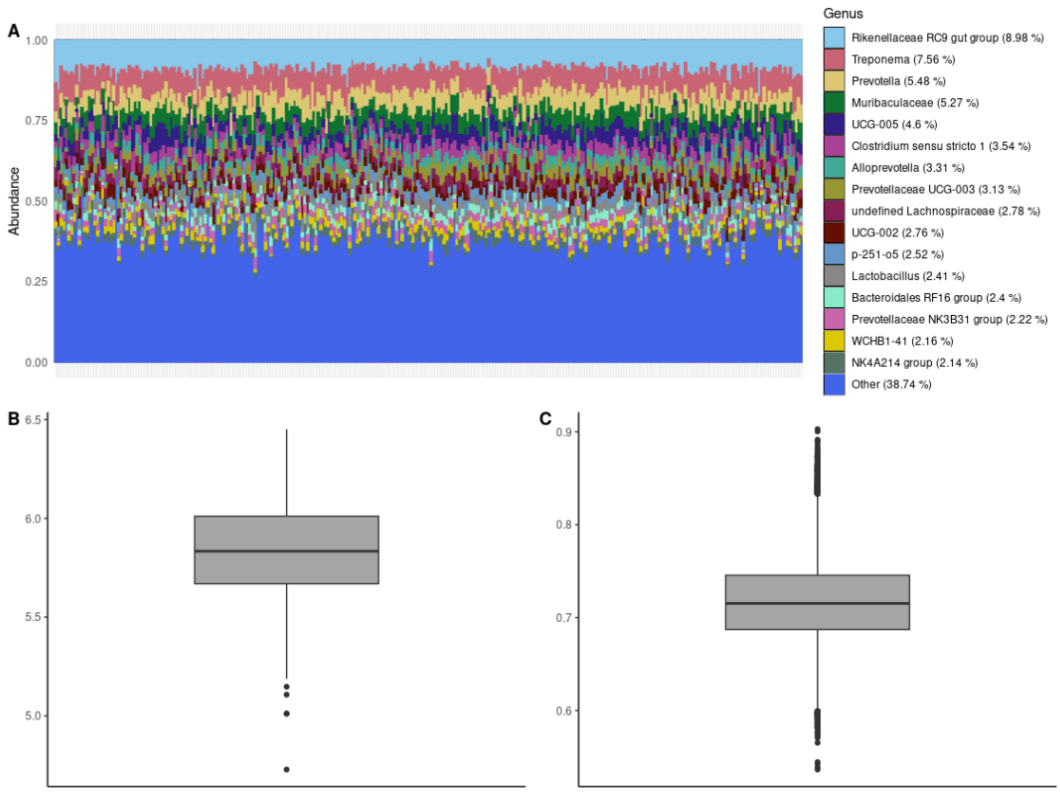


Figure 1. Population diversity in the 285 analysed samples from rectal content. A. Stacked barplot for the relative abundance of genera. Genera with a total relative abundance lower than 2% are grouped as “Other”. B. Shannon’s α-diversity in the 285 samples. C. Bray-Curtis dissimilarities between samples.

To evaluate the richness and evenness of the gut microbiota, α-diversity was calculated using the Shannon index. The average value of the α-diversity was 5.82, ranging from 4.73 to 6.45 (Figure 1B). The differences between samples were computed with the Bray-Curtis dissimilarities. On average, the β-diversity value was 0.72, ranging from 0.54 to 0.9 (Figure 1C). After the rarefaction, the diversity values barely changed; the mean α-diversity was 5.8 [4.72, 6.39] and the mean β-diversity was 0.73 [0.54, 0.91].

Phenotypic analysis

Pearson correlations between tissular FA and rectal SCFA compositions were analysed in 271 samples. The results of such correlations can be found in Supplemental Table S2 and were represented as a heatmap in Figure 2. In general, the abundance of SFA was strongly correlated with the abundance of MUFA in both backfat and muscle (ρ = −0.85 and ρ = −0.94, respectively), while MUFA and PUFA had only a slight negative correlation in muscle (ρ = −0.27). In backfat, PUFA had a strong negative correlation (ρ = −0.45) with SFA. However, the unsaturated FAs (C16:1, C17:1, C20:1, C20:2, C20:3(n-3), C20:3(n-6) and C20:4) had a positive correlation with their respective SFA (C16:0, C17:0 and C20:0). The only exceptions were the stearic acid (C18:0) and its unsaturated forms (C18:1, C18:2 and C18:3), which showed negative correlations.

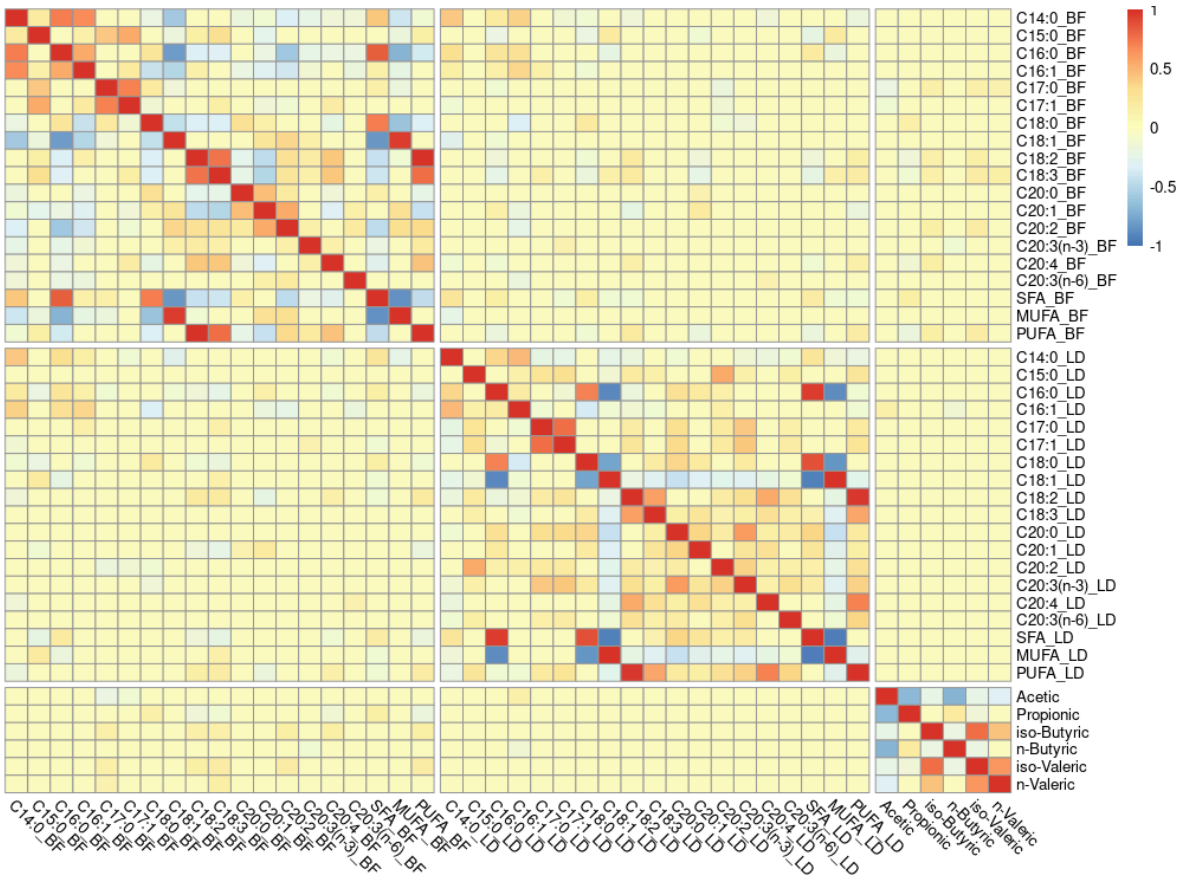


Figure 2. Heatmap of Pearson correlations between fatty acids in backfat, muscle and rectal content. “BF” extension means that the FA belongs to backfat and “LD” to longissimus dorsi. Acetic, propionic, iso- and n-butyric and iso- and n-valeric acids come from rectal content. Significant correlations were considered with a p-value < 0.05. Non-significant correlations were set as 0.

Acetic acid was found negatively correlated with all the other SCFAs, especially with propionic and n-butyric acids (ρ = −0.67 and ρ = −0.70 respectively), while propionic and n-butyric acids had a slight positive correlation between them (ρ = 0.20).

Phenotype and taxonomy association

The rCCA method from mixOmics was used to assess the relationship between the rectal microbiota at genus level from the 271 samples with complete data and their rectal SCFA and FA compositions in backfat and muscle. Figure 3 shows a circular plot of the microbial genera and the FA with the strongest correlations (regarding the components), showing only the variables with a component correlation > |0.3|. Globally, the tissular FAs had weaker correlations with the microbiota than the SCFAs, which showed stronger correlations. The n-butyric acid had a clear positive correlation with the bacterial genera *Prevotellaceae*

NK3B31 group, *Anaerovibrio*, *Oscillospiraceae UCG-005*, an *uncultured Coriobacteriales*, *Subdoligranulum*, *Prevotella*, *Lactobacillus* and *Solobacterium*, while the acetic acid was negatively correlated with the same genera. The two most abundant SCFAs, acetic and propionic, had lower correlations with the microbiota, and the acetic acid was placed on the opposing side to propionic and n-butyric acids, showing a negative correlation that can be also appreciated in Figure 2. After Pearson correlation tests and Benjamini & Hochberg correction, the strongest positive correlations were found between the levels of n-butyric acid and *Prevotella* (ρ = 0.47), the iso-valeric acid and *Oscillospiraceae UCG-002* (ρ = 0.38), and the n-valeric acid and archaea of the *Methanobrevibacter* (ρ = 0.35) genus. On the other hand, the n-butyric acid also had the strongest negative correlations with *Clostridia vadin BB60 group* (ρ = −0.36), *Akkermansia* (ρ = −0.36), and *Oscillospiraceae NK4A214 group* (ρ = −0.35). Albeit weaker than those reported for the previous SCFAs, the strongest positive correlations with the acetic acid were with *Akkermansia* (ρ = 0.24), *Christensenellaceae R-7 group* (ρ = 0.23), and *Escherichia-Shigella* (ρ = 0.23), while the strongest negative correlations were found with *Prevotella* (ρ = −0.24) and *Prevotellaceae NK3B31 group* (ρ = −0.24). It is also worth noting that the strongest positive correlations with propionic acid were with *Prevotellaceae UCG-001* (ρ = 0.26) and *Alloprevotella* (ρ = 0.22), and the strongest negative correlations were with *Christensenellaceae R-7 group* (ρ = −0.28), *Oscillospirales UCG-010* (ρ = −0.24) and *Clostridium sensu stricto 1* (ρ = −0.23). All the correlations between the microbial genera and the FA composition can be found in Supplemental Table S3.

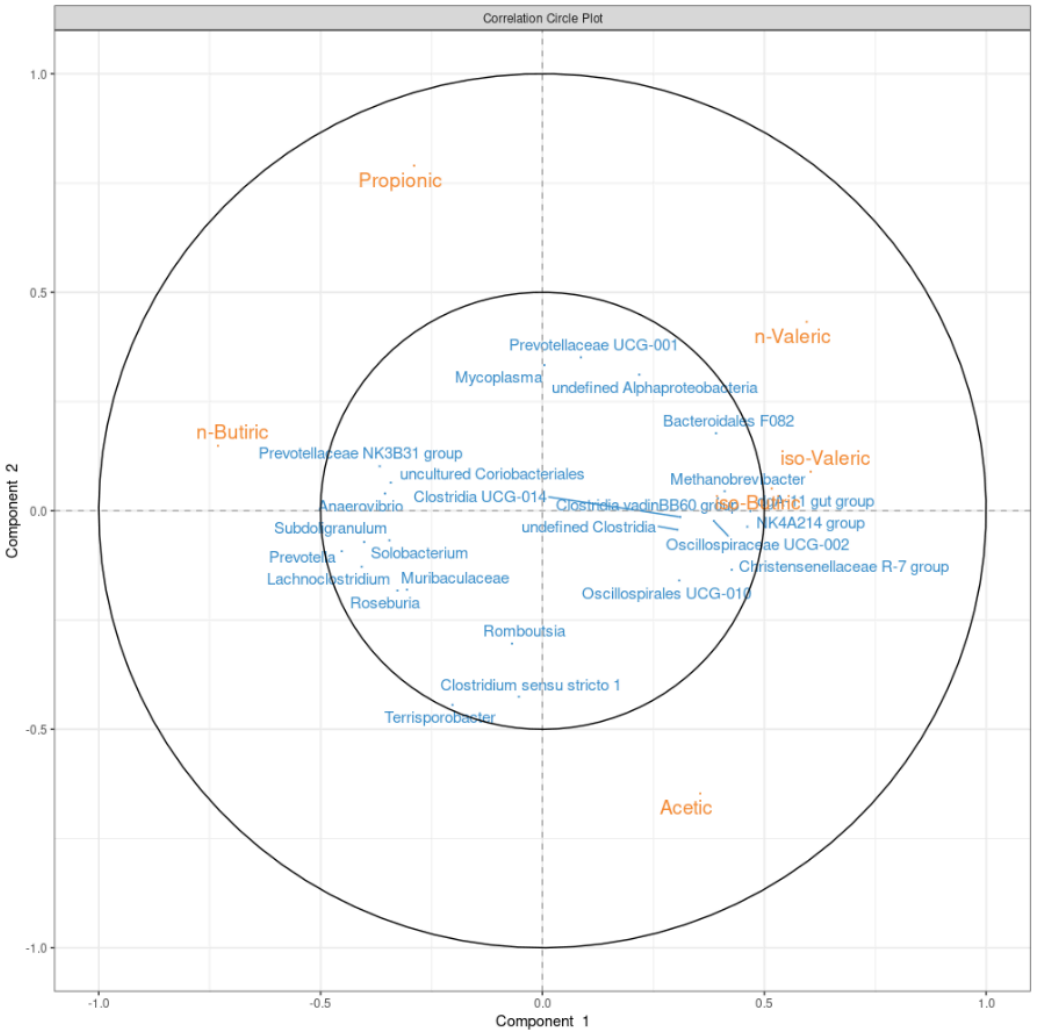


Figure 3. Correlation circular plot between the relative abundance of short-chain fatty acids and bacterial and archaeal genera. For clarity purposes, only variables with a component value greater than |0.3| are represented.

There were genera correlated with the six SCFAs: *Prevotellaceae NK3B31 group* and *uncultured Coriobacteriales* were positively correlated with propionic and n-butyric acid, and negatively correlated with acetic, iso-butyric, iso-valeric and n-valeric acids. On the other hand, *Christensenellaceae R-7 group*, *Oscillospiraceae NK4A214 group* and *Oscillospirales UCG-010* were positively correlated with acetic, iso-butyric, iso-valeric and n-valeric acids, and negatively correlated with propionic and n-butyric acids.

Regarding the FA composition in tissue, *Rikenellaceae RC9 gut group* had the strongest correlations with some of the most abundant FAs. In backfat, *Rikenellaceae RC9 gut group* had a negative correlation with stearic acid (ρ = −0.22), while it showed a positive correlation (ρ = 0.17) with palmitic acid a negative correlation with oleic acid (ρ = −0.18) in muscle.

Figure 4 shows the significant correlations (ρ > |0.15| and p-value < 0.05) between the FA phenotypes and the microbial genera. These microbial genera were also correlated with each other, finding the strongest negative correlation between *Prevotella* and *Oscillospiraceae NK4A214 group* (ρ = −0.54) and the most positive one between *Ureibacillus* and *Hydrogenophilus* (ρ = 0.92). A heatmap with the correlations between all the genera can be found in Supplemental Material S4.

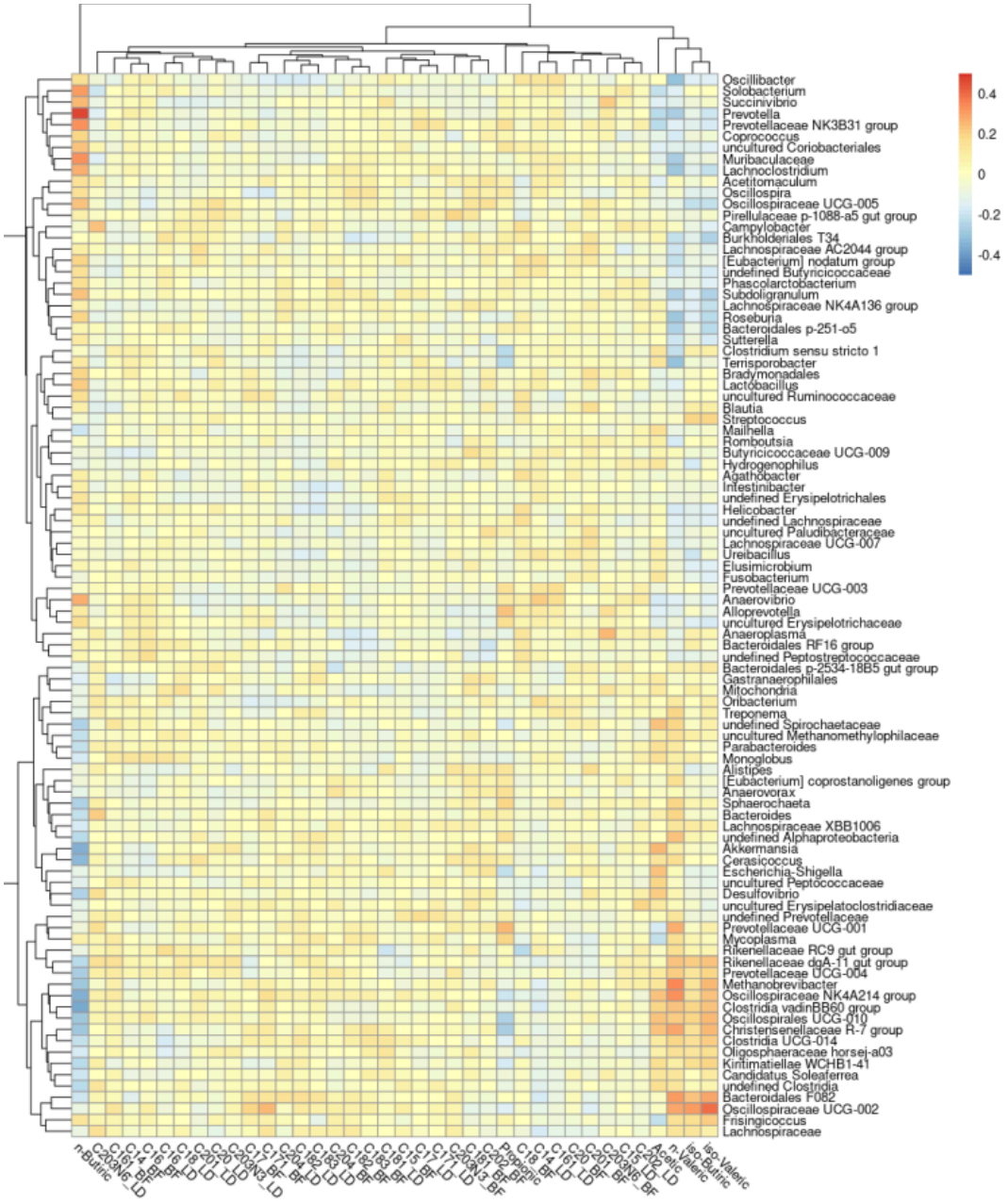


Figure 4. Heatmap for the significant correlations between genera and fatty acid composition in backfat (BF) and muscle (LD) and short-chain fatty acid composition in rectum. Significant correlations were considered at ρ > |0.15| and p-value < 0.05. Non-significant correlations were set as 0.

PLS was conducted to find putative predictors for each FA abundance, its results indicate the relationship between genera and each specific FA by assigning a positive or negative weight to each genus to represent their contribution to each phenotype. The results of the PLS analysis carried out independently for each FA phenotype were in agreement with Pearson's correlations and those reported by the rCCA method. Supplemental Table S5 shows the top 10 most contributing microbial genera for each FA. *Prevotella* was the second main negative contributor to acetic acid and the main positive contributor to n-butyric acid, while *Akkermansia* was the main positive contributor to acetic acid and the second main negative contributor to n-butyric acid.

The NMDS analysis with the ASV relative abundances reflected a clear relationship between the ordination of the samples based on their β-diversity values and the relative abundances of acetic and n-butyric acids (Figure 5). The PERMANOVA results found out that a small percentage of the β-diversity was explained by the abundance of these SCFAs, 0.85% for acetic and 1.58% for n-butyric acids. Regarding the FA composition in backfat and muscle, the results reported a clear relationship between the arachidonic acid in both tissues and the ordination of the samples. NMDS plots with suggestive PERMANOVA results (*p*-value < 0.1) can be accessed at Supplemental Material S6.

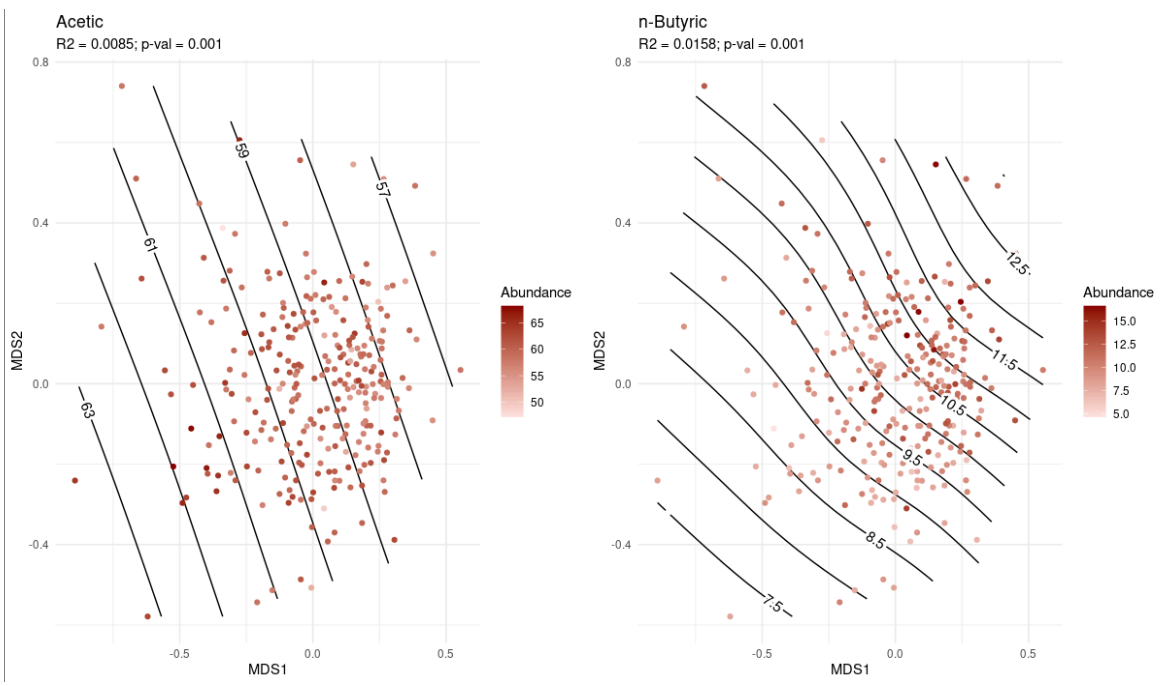


Figure 5. NMDS plots for acetic acid (left) and n-butyric acid (right) with the fatty acid abundances and R^2 from PERMANOVA results. Colour saturation represents the SCFA relative abundance on each sample. Black lines show the estimated distribution of the SCFA relative abundances regarding the ordination of the samples based on their β-diversity

values (Bray-Curtis dissimilarities). The less curved and the more parallel straight lines are represented, the higher relationship between the FA relative abundance and the microbiota diversity can be assumed.

Functional prediction

From the 2,025 KEGG Orthologs (KOs) found with PICRUSt2, 60 were related with lipid metabolism, according to the KEGG Orthology hierarchy. Propionic acid had the highest number of significant correlations with the lipid metabolism KOs, 44 out of the total 60 were significant, while acetic acid and n-butyric acid had 32 and 28 significant correlations respectively. The strongest positive correlation of acetic acid was with K00111 (ρ = 0.43), while the strongest negative one was with K06131 (ρ = −0.3). For propionic acid, the strongest correlation was with K16363 (ρ = 0.38), while the most negative one was with the K00111 (ρ = −0.42), the same KO as the most positively correlated with acetic acid. In n-butyric acid the highest positive correlation was with K00655 (ρ = 0.34), while the strongest negative one was with K07406 (ρ = −0.40). In n-valeric acid, the highest correlation was with K00648 (ρ = 0.34) and the most negative one was with K01048 (ρ = −0.37). Iso-butyric and iso-valeric acids had weaker correlations than |0.3|. In general, the negative correlations were stronger than the positive ones. Figure 6 shows the rCCA results between the SCFAs and the KOs related with lipid metabolism. This plot highlights the pattern of the correlations between the KOs and the SCFAs, which follows opposite directions for acetic versus propionic and n-butyric acids, similarly to the taxonomic results. In this plot there are two KOs that surpass the 0.5 radius threshold, K00111 and K03429, which are two of the KOs with stronger correlations with the three most abundant SCFAs, positive with acetic (ρ = 0.43 and ρ = 0.39 respectively), and negative with propionic (ρ = −0.42 and ρ = −0.37 respectively) and n-butyric acids (ρ = −0.25 and ρ = −0.36 respectively).

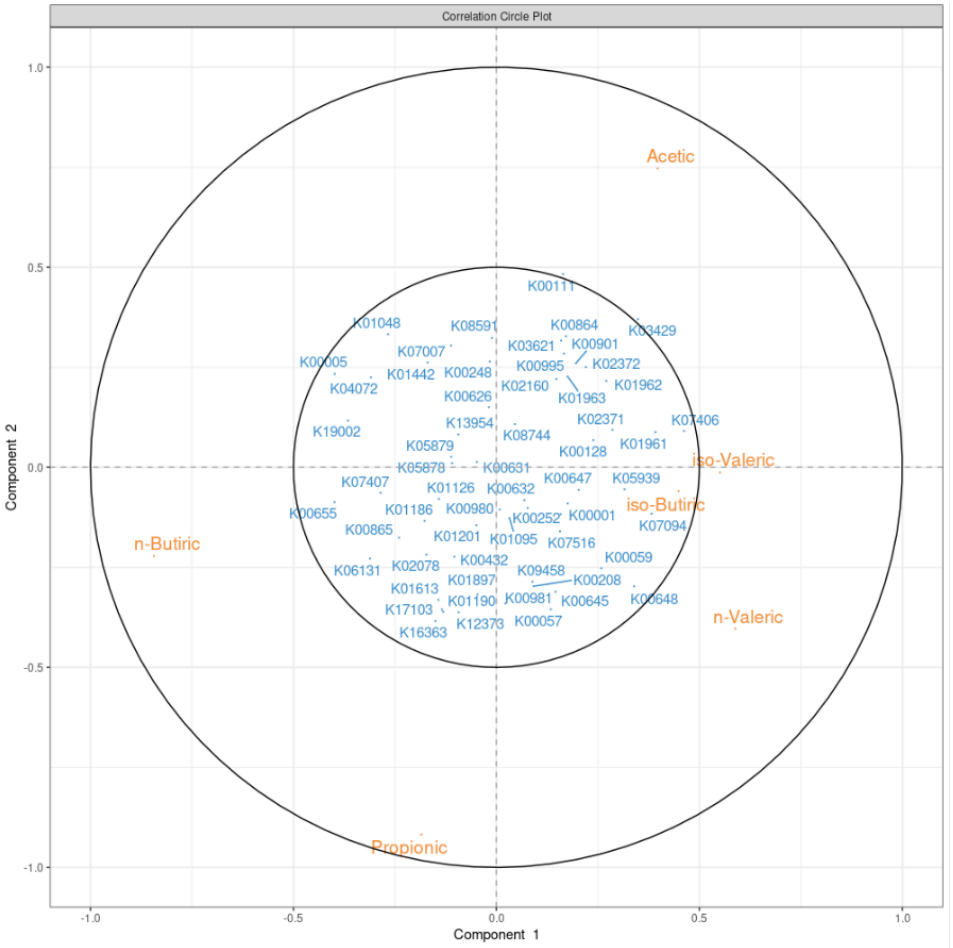


Figure 6. Correlation circular plot between the relative abundance of the short-chain fatty acids and the KEGG orthologs related with lipid metabolism.

Similarly to the previous results, tissular FAs had weaker correlations with KOs and less significant results. Only nine FAs (out of 32), eight of backfat (pentadecanoic, margaric, heptadecenoic, stearic, eicosadienoic, eicosatrienoic, dihomo-γ-linolenic and arachidonic acids) and palmitoleic acid from muscle had stronger correlations than |0.2|, most of them unclassified, poorly characterised or KOs belonging to different protein families. Supplemental Table S7 shows the significant correlations between the FA composition and the KEGG Orthology results.

KOs correlations with microbial genera had higher and more significant values. The highest correlation was found between *Treponema* and K00209 (ρ = 0.71), while the strongest negative correlation was found between the genus *Bacteroidales F082* and K00655 (ρ = −0.58). Supplemental Material S8 shows a heatmap of all the significant correlations between the KEGG orthologs and the microbial genera.

Discussion

In this study, we analysed the correlation between the microbiota and SCFA composition of the rectal content and backfat and muscle FA composition of Duroc × Iberian finishing pigs. Even though this is not a traditional backcross of Landrace, Pietrain or Large White breeds, it is a hybrid population used for dry-cured products and the fact that it is a F1 cross minimises the genetic effect of a F2 three-way cross or backcross. In addition, the samples were taken at the moment of slaughter and immediately frozen, which gives results that resemble more the reality in the pig gut than taking faeces samples in the farm.

In accordance with previous studies about the pig gut microbiota in the distal colon (35–38), Bacteroidota and Firmicutes were the two most abundant phyla in our study, representing 82.25% of the total number of counts. Although Firmicutes were less abundant than Bacteroidota in our study, Firmicutes usually is the most prevalent phylum along the swine gastrointestinal tract, while Bacteroidota only has an increased presence in the colon (39–41). In a previous study of the same population conducted by our research group (12), Firmicutes and Bacteroidota also represented 82.83% of the total number of counts, albeit Firmicutes was reported as the most abundant phylum, probably due to the differences between the databases used for the taxonomic classification (GreenGenes 13.8 and SILVA138) and the different QIIME versions used (QIIME 1.9.1 and QIIME2 v2021.11.0). Other studies, conducted with pigs at different age stages and supplemented with probiotics, also found a higher relative abundance of Bacteroidota versus Firmicutes in faeces (42, 43). These differences between studies are difficult to address, as the pigs were not raised in the same conditions, and neither the breeds nor the diets were similar. In addition, the composition of the gut microbiota is dynamic and changes through age and sections (39, 41, 44). A possible explanation for the higher presence of Bacteroidota than Firmicutes in our study may be due to the adaptation of the microbiota to the diet, supported by the fact that three out of the four most abundant genera belonged to the Bacteroidota phylum (*Rikenellaceae RC9 gut group*, *Prevotella* and an *unspecified Muribaculaceae* genus), whereas the remaining one (*Treponema*) belonged to the Spirochaetota phylum. *Prevotella* is usually reported as one of the most abundant genera in the porcine colon and it is favoured in the gut of pigs and human populations with high-fibre diets (45, 46). In the same manner, the levels of Muribaculaceae decrease in high-fat-fed and trans-fatty-acid-fed obese rats (47), and together with *Rikenellaceae RC9 gut group*, pigs supplemented with a fruit and vegetable diet also had an increase of Muribaculaceae (48).

The strongest correlation found between all the FA phenotypes and the microbiota

composition was between the relative abundances of *Prevotella* and n-butyric acid. In agreement with the PLS results, *Prevotella* was the highest positive contributor to the levels of n-butyric acid in the rectal content and the second highest negative contributor to acetic acid levels after *Prevotellaceae NK3B31 group*. Remarkably, *Prevotella* was found negatively correlated with acetic acid and with butyrate-producing clostridia. Thus, unsurprisingly, some of these butyrate-producing bacteria were positively correlated with the acetic acid and negatively correlated with the n-butyric acid. Gut butyrate can be produced through four pathways, having one of them the acetate as the main intermediate product, whereas the other ones use succinate, glutarate or lysine (49). Although our 16S rRNA gene sequencing method was not designed for the taxonomic classification at species level, the results suggest that *Prevotella* genus is involved in the succinate pathway, hinting that the species present in the samples may be *P. copri* or *P. ruminicola* (50, 51). *P. copri* is the most abundant species of the *Prevotella* genus in the gut of adult pigs (52) and produces succinate as a metabolite, which can be transformed to butyrate through its conversion to butyryl-CoA (53).

Spirochaetota was the third most abundant phylum, mainly due to the abundance of *Treponema*. *Treponema* has been associated with crude fibre digestibility (44) and feed efficiency (54). This genus plays a role in dietary fibre degradation in finishing pigs (44), which provides energy to the host catabolizing indigestible components (55). The negative correlation found between *Treponema* and *Prevotella* has been already reported in pigs (56, 57), possibly due to the fact that both genera may compete for the degradation of the dietary fibre. However, *Prevotella* genus can break down dietary polysaccharides, such as arabinoxylans from wheat and barley, and produce SCFAs (58), while *Treponema* has an essential role in catabolizing dietary non-digestible components, such as cellulose and lignin (44).

The most abundant taxon at genus level, *Rikenellaceae RC9 gut group*, was negatively correlated with the stearic acid levels in backfat and the oleic acid levels in muscle, while it was also positively correlated with the content of palmitic acid in muscle. Supporting this finding was the strong negative correlation observed between palmitic and oleic in both tissues. As previously mentioned, *Rikenellaceae RC9 gut group* was negatively correlated with *Prevotella*, whereas it was positively correlated with another member of the same family, *Rikenellaceae dgA-11 gut group*. *Rikenellaceae RC9 gut group* may be involved in fatty acid metabolism, as suggested in (59). It has been found in higher abundance in high-fat fed mice than in control animals (60, 61), and it is a dominant genus in the ileum of pigs with low food conversion ratio (62). These results may suggest the negative role of *Rikenellaceae*

RC9 gut group in meat quality.

Despite *Akkermansia* not being as abundant as other genera, it presented a positive correlation with acetic acid and a negative correlation with n-butyric acid. The positive correlation with acetic acid can be explained because *A. muciniphila*, the main species of this genus, breaks down the mucin of the colon, producing acetic and propionic acids (63). The PLS results showed *Akkermansia* as the first positive contributor to acetic acid and the second negative contributor to n-butyric acid after *Clostridia vadin BB60 group*. These results for *Akkermansia* were in contraposition to the relationship with acetic and n-butyric acid reported for *Prevotella*; thus, as expected, a negative relationship between *Akkermansia* and *Prevotella* was observed in our rCCA and PLS results. This antagonistic relationship has also been found in certain situations in the human gut microbiome (64, 65). Regarding its potential role, *Akkermansia* belongs to the Verrucomicrobiota phylum and has been related with host health, having a reduced presence in mouse models with metabolic disorders (66). It also has been reported its negative relationship with obesity development in humans and other species (67). In summary, the *Akkermansia* genus could be a biomarker of gut health, as it has been reported in humans (68).

Methanobrevibacter (phylum Euryarchaeota) and the *uncultured Methanomethylophilaceae* (phylum Thermoplasmatota) were the only archaeal genera found, but they were not significantly correlated between them. These two archaea are methanogens, and although methane increases SCFA production (69), their correlation with all the studied SCFAs was not always positive. Both genera had negative correlations with n-butyric acid and positive correlations with n-valeric acid. In addition, *Methanobrevibacter* was positively correlated with acetic, iso-butyric and iso-valeric acids. Methanogens from genus *Methanobrevibacter*, such as *M. smithii*, need acetate to grow (70), which means that in an environment rich in acetate they will be found in higher abundance.

Regarding the functional analysis, the two KOs related with lipid metabolism with the strongest correlations were K00111 and K03429, both with positive correlations with acetic acid and negatives with propionic and n-butyric acids. These two KOs play a role in lipid metabolism, K00111 is the glycerol-3-phosphate dehydrogenase [EC:1.1.5.3], which is involved in glycerophospholipid metabolism. This enzyme can be found in the bacterial cytoplasmic membrane and catalyzes the change from glycerol 3-phosphate to glycerone phosphate and vice versa (71, 72). Glycerol is a constitutive molecule of lipids and glycerophospholipids, and is key to phospholipid biosynthesis and carbon metabolism, while glycerone phosphate is involved in glycolysis (73). Glycerol and glycerone phosphate can be fermented and transformed in simpler molecules such as acetate, propionate or butyrate,

which will depend on the environmental conditions and bacteria found (74).

K03429 is the processive 1,2-diacylglycerol beta-glucosyltransferase [EC:2.4.1.315] and is involved in different pathways such as glycerolipid metabolism or teichoic acid biosynthesis. This enzyme catalyzes the reversible transformation from diacylglycerol to monoglucosyldiacylglycerol, from this one to diglucosyldiacylglycerol, which can be transformed into lipoteichoic acid, and from diglucosyldiacylglicerol to triglucosyldiacylglycerol (71, 75). Lipoteichoic acid can be found in the cell wall of gram positive bacteria that belong to Firmicutes phylum, such as clostridia (76). Our results showed positive correlations of different clostridia (*Clostridia UCG-014*, *undefined Clostridia*, *Clostridia vadin BB60 group* and *Clostridium sensu stricto*) with acetic acid and negative correlations with propionic and n-butyric acid, most of them significant, which agrees with this finding in the functional prediction.

On the other hand, K00655 and K16363 showed positive correlations with n-butyric and propionic acids (they were the strongest correlation for the respective SCFA) and negative with acetic acid. K00655 is the 1-acyl-sn-glycerol-3-phosphate acyltransferase [EC:2.3.1.51], which is involved in phosphatidic acid production and, therefore, in cell membrane composition. To our knowledge, there is no previous report of a relationship found between K00655 and butyrate, but interestingly this KO was the one with the strongest negative correlation with a microbial genus, *F082*, which is a member of the order Bacteroidales. This genus also had a slight negative correlation with n-butyric acid and no significant correlation with propionic acid, although *Bacteroidales F082* is considered a propionate producer in rumen (77), but its role in intestinal fermentation is not well defined yet.

K16363 encodes for UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase / 3-hydroxyacyl-[acyl-carrier-protein] dehydratase [EC:3.5.1.108 4.2.1.59]. The first enzyme is involved in the first step of the lipopolysaccharide production and, therefore, in the cell membrane of gram-negative bacteria (78), such as *Phascolarctobacterium*, *Sphaerochaeta*, *Anaerovibrio*, *Bacteroidales RF16 group* or all the members of the Prevotellaceae family, with which the correlations were positive. This KO had negative correlations with gram-positive bacteria such as clostridia.

Treponema and K00209 shared the strongest Pearson correlation. K00209 encodes for enoyl-[acyl-carrier protein] reductase / trans-2-enoyl-CoA reductase (NAD+) [EC:1.3.1.9 1.3.1.44], which is involved in fatty acid biosynthesis and elongation. Trans-2-enoyl-CoA reductase (NAD+) was described for the first time in the microalga *Euglena gracilis* (79), but the first enzymatic characterization of a trans-2-enoyl-CoA reductase was done in

Treponema denticola (80). This family of proteins are involved in fatty acid biosynthesis, but they work with CoA esters instead of ACP-esters (79).

Conclusions

In summary, the most abundant phyla found in rectal content were Bacteroidota and Firmicutes, while the most abundant genera, *Rikenellaceae RC9 gut group*, *Treponema*, *Prevotella* and an *unspecified Muribaculaceae* belonged to the Bacteroidota or Spirochaetota phyla. While some bacteria may be involved in the lipid metabolism of the host, as indicated by the correlations observed between their abundances and the FA composition, specifically the SCFA content in the gut, these relationships were weak, maybe due to the indirect and long paths between both phenotypes. Of note, the levels of *Akkermansia* and *Prevotella* were negatively correlated in our study, and their main positive or negative contribution to acetic and n-butyric acid in the rectal content strengthen their role as potential biomarkers for these two SCFAs. The functional prediction analysis showed coherent results regarding the relationships between the KEGG orthologs and the microbial genera or the SCFA. The most abundant genus, *Rikenellaceae RC9 gut group*, had the highest correlations with stearic, palmitic and oleic acid levels in backfat and muscle, further indicating the potential, although indirect, role of the microbiota in the modification of the FA composition in tissue. Nonetheless, FA composition is a cumulative trait that is developed and changed throughout the life of the animal, same as the modification of the microbiota that growing animals experiment. Thus, further analyses are warranted to measure the microbiota and the FA composition in different time-points, to better determine the implications of the microbiota over the FA composition in tissue.

Acknowledgements

We would like to thank all the members of the institutions that contributed to the generation of the animal material used in this study. We acknowledge the technical and human support provided by the Facility of Bioscience Applied Techniques of SAIUEx (financed by UEX, Junta de Extremadura, MICINN, FEDER and FSE) and from the Asociación Interprofesional del Cerdo Ibérico (ASICI). This work was funded by the Ministerio de Ciencia e Innovación (MICIIN) and the Fondo Europeo de Desarrollo Regional (FEDER), with project numbers AGL2017-82641-R and PID2020-12677RB-C22, and supported by the Ministerio de Economía y Competitividad (MINECO), with the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000902-S) awarded to the Centre for Research in Agricultural Genomics (CRAG) and to the programmes of Centres de Recerca de Catalunya (CERCA). Cristina Sebastià was funded with a FI-AGAUR PhD grant from the Generalitat de Catalunya (2020FI_B_00225). Magí Passols was funded by an FPI grant from the AGL2017-82641-R project.

Supplemental Material

Supplemental Table S1. Total counts and relative abundance of taxa at phylum and genus level. The number of positive samples, mean, standard deviation (SD), minimum and maximum counts are also calculated.

Supplemental Table S2. Correlations between all the fatty acids and their *p*-values.

Supplemental Table S3. Correlations between all the fatty acids and the microbial genera. *P*-values are included.

Supplemental Material S4. Heatmap with the correlations between the microbial genera. Non-significant relationships (*p*-value > 0.05) were set to 0.

Supplemental Table S5. Top 10 significant correlations between each fatty acid and the microbial genera.

Supplemental Material S6. Significant (*p*-value < 0.05) and suggestive (*p*-value < 0.1) PERMANOVA results for fatty acids represented in NMDS plots.

Supplemental Table S7. Correlations between all the KEGG orthologs and the fatty acid composition.

Supplemental Material S8. Heatmap of the significant correlations between the KEGG orthologs involved in lipid metabolism and the microbial genera. Non-significant relationships (*p*-value > 0.05) were set to 0.

References

1. FAOSTAT. 2022. FAOSTAT. FAOSTAT. https://www.fao.org/faostat/en/#home.

2. López-Bote CJ. 1998. Sustained utilization of the Iberian pig breed. Meat Sci 49:S17–S27.

3. Barea R, Isabel B, Nieto R, López-Bote C, Aguilera JF. 2013. Evolution of the fatty acid profile of subcutaneous back-fat adipose tissue in growing Iberian and Landrace × Large White pigs. Animal 7:688–698.

4. Cava R, Manuel Ferrer J, Estévez M, Morcuende D, Toldrá F. 2004. Composition and proteolytic and lipolytic enzyme activities in muscle *Longissimus dorsi* from Iberian pigs and industrial genotype pigs. Food Chem 88:25–33.

5. Ramírez MR, Cava R. 2008. Effect of Iberian × Duroc genotype on composition and sensory properties of dry-cured ham: Composition, instrumental and sensory analysis of dry-cured hams. J Sci Food Agric 88:667–675.

6. Heinritz SN, Mosenthin R, Weiss E. 2013. Use of pigs as a potential model for research into dietary modulation of the human gut microbiota. Nutr Res Rev 26:191–209.

7. Hou N, Du X, Wu S. 2022. Advances in pig models of human diseases. Anim Models Exp Med 5:141–152.

8. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S. 2012. Host-gut microbiota metabolic interactions. Science 336:1262–1267.

9. Chen H, Mao XB, Che LQ, Yu B, He J, Yu J, Han GQ, Huang ZQ, Zheng P, Chen DW. 2014. Impact of fiber types on gut microbiota, gut environment and gut function in fattening pigs. Anim Feed Sci Technol 195:101–111.

10. González-Prendes R, Pena RN, Solé E, Seradj AR, Estany J, Ramayo-Caldas Y. 2019. Modulatory effect of protein and carotene dietary levels on pig gut microbiota. Sci Rep 9:14582.

11. Carmody RN, Gerber GK, Luevano JM, Gatti DM, Somes L, Svenson KL, Turnbaugh PJ. 2015. Diet dominates host genotype in shaping the murine gut microbiota. Cell Host Microbe 17:72–84.

12. Crespo-Piazuelo D, Migura-Garcia L, Estellé J, Criado-Mesas L, Revilla M, Castelló

A, Muñoz M, García-Casco JM, Fernández AI, Ballester M, Folch JM. 2019. Association between the pig genome and its gut microbiota composition. Sci Rep 9:8791.

13. Spor A, Koren O, Ley R. 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol 9:279–290.

14. Rapin A, Pattaroni C, Marsland BJ, Harris NL. 2017. Microbiota Analysis Using an Illumina MiSeq Platform to Sequence 16S rRNA Genes: Microbiota Analysis by Sequencing 16 S rRNA. Curr Protoc Mouse Biol 7:100–129.

15. Tringe SG, Hugenholtz P. 2008. A renaissance for the pioneering 16S rRNA gene. Curr Opin Microbiol 11:442–446.

16. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM. 2013. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. J Lipid Res 54:2325–2340.

17. Hu J, Lin S, Zheng B, Cheung PCK. 2018. Short-chain fatty acids in control of energy metabolism. Crit Rev Food Sci Nutr 58:1243–1249.

18. Morrison DJ, Preston T. 2016. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes 7:189–200.

19. Tian M, Chen J, Liu J, Chen F, Guan W, Zhang S. 2020. Dietary fiber and microbiota interaction regulates sow metabolism and reproductive performance. Anim Nutr 6:397–403.

20. He J, Zhang P, Shen L, Niu L, Tan Y, Chen L, Zhao Y, Bai L, Hao X, Li X, Zhang S, Zhu L. 2020. Short-Chain Fatty Acids and Their Association with Signalling Pathways in Inflammation, Glucose and Lipid Metabolism. Int J Mol Sci 21:6356.

21. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D,

McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857.

22. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.

23. R Core Team. 2022. R: A language and environment for statistical computing (4.1.3). R Foundation for Statistical Computing, Vienna, Austria.

24. McMurdie PJ, Holmes S. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217.

25. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2020. vegan: Community Ecology Package.

26. Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of Southern Wisconsin. Ecol Monogr 27:325–349.

27. BOE. 2004. ORDEN PRE/3844/2004, de 18 de noviembre, por la que se establecen los métodos oficiales de toma de muestras en canales de cerdos ibéricos y el método de análisis para la determinación de la composición de ácidos grasos de los lípidos totales del tejido adiposo subcutáneo de cerdos ibéricos.

28. Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. Natl Res Counc Can 37:7.

29. Revelle W. 2018. psych: Procedures for Personality and Psychological Research (2.1.9).

30. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300.

31. Kolde R. 2019. pheatmap: Pretty Heatmaps (1.0.12).

32. Lê Cao K-A, Rohart F, González I, Déjean S, Gautier B, Bartolo F, Monget P, Coquery J, Yao F, Liquet B. 2016. mixOmics: Omics Data Integration Project (6.1.1).

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

34. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 38:685–688.

35. Holman DB, Brunelle BW, Trachsel J, Allen HK. 2017. Meta-analysis to define a core microbiota in the swine gut. mSystems 2:e00004-17.

36. Wang C, Wei S, Chen N, Xiang Y, Wang Y, Jin M. 2022. Characteristics of gut microbiota in pigs with different breeds, growth periods and genders. Microb Biotechnol 15:793–804.

37. Wang X, Tsai T, Deng F, Wei X, Chai J, Knapp J, Apple J, Maxwell CV, Lee JA, Li Y, Zhao J. 2019. Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. Microbiome 7:109.

38. Xiao L, Estellé J, Kiilerich P, Ramayo-Caldas Y, Xia Z, Feng Q, Liang S, Pedersen AØ, Kjeldsen NJ, Liu C, Maguin E, Doré J, Pons N, Le Chatelier E, Prifti E, Li J, Jia H, Liu X, Xu X, Ehrlich SD, Madsen L, Kristiansen K, Rogel-Gaillard C, Wang J. 2016. A reference gene catalogue of the pig gut microbiome. Nat Microbiol 1:16161.

39. Crespo-Piazuelo D, Estellé J, Revilla M, Criado-Mesas L, Ramayo-Caldas Y, Óvilo C, Fernández AI, Ballester M, Folch JM. 2018. Characterization of bacterial microbiota compositions along the intestinal tract in pigs and their interactions and functions. Sci Rep 8:12727.

40. Zhang L, Wu W, Lee Y-K, Xie J, Zhang H. 2018. Spatial Heterogeneity and Co-occurrence of Mucosal and Luminal Microbiome across Swine Intestinal Tract. Front Microbiol 9:48.

41. Zhao W, Wang Y, Liu S, Huang J, Zhai Z, He C, Ding J, Wang J, Wang H, Fan W, Zhao J, Meng H. 2015. The dynamic distribution of porcine microbiota across different ages and gastrointestinal tract segments. PLOS ONE 10:e0117441.

42. Han GG, Lee J-Y, Jin G-D, Park J, Choi YH, Kang S-K, Chae BJ, Kim EB, Choi Y-J. 2018. Tracing of the fecal microbiota of commercial pigs at five growth stages from birth to shipment. Sci Rep 8:6012.

43. Shin D, Chang SY, Bogere P, Won K, Choi J-Y, Choi Y-J, Lee HK, Hur J, Park B-Y, Kim Y, Heo J. 2019. Beneficial roles of probiotics on the modulation of gut microbiota and immune response in pigs. PLOS ONE 14:e0220843.

44. Niu Q, Li P, Hao S, Zhang Y, Kim SW, Li H, Ma X, Gao S, He L, Wu W, Huang X, Hua J, Zhou B, Huang R. 2015. Dynamic distribution of the gut microbiota and the relationship with apparent crude fiber digestibility and growth stages in pigs. Sci Rep 5:9938.

45. Le Sciellour M, Labussière E, Zemb O, Renaudeau D. 2018. Effect of dietary fiber content on nutrient digestibility and fecal microbiota composition in growing-finishing pigs. PLOS ONE 13:e0206159.

46. Precup G, Vodnar D-C. 2019. Gut *Prevotella* as a possible biomarker of diet and its eubiotic versus dysbiotic roles: a comprehensive literature review. Br J Nutr 122:131–140.

47. Hua Y, Fan R, Zhao L, Tong C, Qian X, Zhang M, Xiao R, Ma W. 2020. Trans -fatty acids alter the gut microbiota in high-fat-diet-induced obese rats. Br J Nutr 124:1251–1263.

48. Solano-Aguilar G, Lakshman S, Shao J, Chen C, Beshah E, Dawson H, Vinyard B, Schroeder S, Jang S, Molokin A, Urban J. 2021. Fruit and Vegetable Supplemented Diet Modulates the Pig Transcriptome and Microbiome after a Two-Week Feeding Intervention. Nutrients 13:4350.

49. Vital M, Howe AC, Tiedje JM. 2014. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. mBio 5:e00889-14.

50. Franke T, Deppenmeier U. 2018. Physiology and central carbon metabolism of the gut bacterium *Prevotella copri*. Mol Microbiol 109:528–540.

51. Louis P, Flint HJ. 2017. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol 19:29–41.

52. Amat S, Lantz H, Munyaka PM, Willing BP. 2020. Prevotella in Pigs: The Positive and Negative Associations with Production and Health. Microorganisms 8:1584.

53. Oliphant K, Allen-Vercoe E. 2019. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome 7:91.

54. McCormack UM, Curião T, Metzler-Zebeli BU, Magowan E, Berry DP, Reyer H, Prieto ML, Buzoianu SG, Harrison M, Rebeiz N, Crispie F, Cotter PD, O'Sullivan O, Gardiner GE, Lawlor PG. 2019. Porcine feed efficiency-associated intestinal microbiota and physiological traits: Finding consistent cross-locational biomarkers for residual feed intake. mSystems 4:e00324-18.

55. Lindberg JE. 2014. Fiber effects in nutrition and gut health in pigs. J Anim Sci Biotechnol 5:15.

56. Ke S, Fang S, He M, Huang X, Yang H, Yang B, Chen C, Huang L. 2019. Age-based dynamic changes of phylogenetic composition and interaction networks of health pig gut microbiome feeding in a uniformed condition. BMC Vet Res 15:172.

57. Ramayo-Caldas Y, Mach N, Lepage P, Levenez F, Denis C, Lemonnier G, Leplat J-J, Billon Y, Berri M, Doré J, Rogel-Gaillard C, Estellé J. 2016. Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. ISME J 10:2973–2977.

58. Liu H, Ivarsson E, Dicksved J, Lundh T, Lindberg JE. 2012. Inclusion of Chicory (Cichorium intybus L.) in Pigs' Diets Affects the Intestinal Microenvironment and the Gut Microbiota. Appl Environ Microbiol 78:4102–4109.

59. Zhou L, Xiao X, Zhang Q, Zheng J, Li M, Yu M, Wang X, Deng M, Zhai X, Li R. 2018. Improved glucose and lipid metabolism in the early life of female offspring by maternal dietary genistein is associated with alterations in the gut microbiota. Front Endocrinol 9:516.

60. Kang Y, Li Y, Du Y, Guo L, Chen M, Huang X, Yang F, Hong J, Kong X. 2019. Konjaku flour reduces obesity in mice by modulating the composition of the gut microbiota. Int J Obes 43:1631–1643.

61. Wang B, Kong Q, Li X, Zhao J, Zhang H, Chen W, Wang G. 2020. A high-fat diet increases gut microbiota biodiversity and energy expenditure due to nutrient difference. Nutrients 12:3197.

62. Quan J, Cai G, Ye J, Yang M, Ding R, Wang X, Zheng E, Fu D, Li S, Zhou S, Liu D, Yang J, Wu Z. 2018. A global comparison of the microbiome compositions of three

gut locations in commercial pigs with extreme feed conversion ratios. Sci Rep 8:4536.

63. Ottman N, Reunanen J, Meijerink M, Pietilä TE, Kainulainen V, Klievink J, Huuskonen L, Aalvink S, Skurnik M, Boeren S, Satokari R, Mercenier A, Palva A, Smidt H, de Vos WM, Belzer C. 2017. Pili-like proteins of *Akkermansia muciniphila* modulate host immune responses and gut barrier function. PLOS ONE 12:e0173004.

64. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Doré J, MetaHIT Consortium (additional members), Weissenbach J, Ehrlich SD, Bork P. 2011. Enterotypes of the human gut microbiome. Nature 473:174–180.

65. Mokhtarzade M, Molanouri Shamsi M, Abolhasani M, Bakhshi B, Sahraian MA, Quinn LS, Negaresh R. 2021. Home-based exercise training influences gut bacterial levels in multiple sclerosis. Complement Ther Clin Pract 45:101463.

66. Derrien M, Belzer C, de Vos WM. 2017. *Akkermansia muciniphila* and its role in regulating host functions. Microb Pathog 106:171–181.

67. Xu Y, Wang N, Tan H-Y, Li S, Zhang C, Feng Y. 2020. Function of Akkermansia muciniphila in Obesity: Interactions With Lipid Metabolism, Immune Response and Gut Systems. Front Microbiol 11:219.

68. Jayachandran M, Chung SSM, Xu B. 2020. A critical review of the relationship between dietary components, the gut microbe *Akkermansia muciniphila* , and human health. Crit Rev Food Sci Nutr 60:2265–2276.

69. Mathur R, Kim G, Morales W, Sung J, Rooks E, Pokkunuri V, Weitsman S, Barlow GM, Chang C, Pimentel M. 2013. Intestinal *Methanobrevibacter smithii* but not total bacteria is related to diet-induced weight gain in rats. Obesity 21:748–754.

70. Catlett JL, Carr S, Cashman M, Smith MD, Walter M, Sakkaff Z, Kelley C, Pierobon M, Cohen MB, Buan NR. 2022. Metabolic Synergy between Human Symbionts *Bacteroides* and *Methanobrevibacter*. Microbiol Spectr 10:e01067-22.

71. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023.

KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res 51:D587–D592.

72. Walz A-C, Demel RA, De Kruijff B, Mutzel R. 2002. Aerobic sn-glycerol-3-phosphate dehydrogenase from Escherichia coli binds to the cytoplasmic membrane through an amphipathic α-helix. Biochem J 365:471–479.

73. Doi Y. 2019. Glycerol metabolism and its regulation in lactic acid bacteria. Appl Microbiol Biotechnol 103:5079–5093.

74. Viana MB, Freitas AV, Leitão RC, Pinto GAS, Santaella ST. 2012. Anaerobic digestion of crude glycerol: a review. Environ Technol Rev 1:81–92.

75. Matsuoka S. 2017. Biological functions of glucolipids in *Bacillus subtilis*. Genes Genet Syst 92:217–221.

76. Percy MG, Gründling A. 2014. Lipoteichoic Acid Synthesis and Function in Gram-Positive Bacteria. Annu Rev Microbiol 68:81–100.

77. Jize Z, Zhuoga D, Xiaoqing Z, Na T, Jiacuo G, Cuicheng L, Bandan P. 2022. Different feeding strategies can affect growth performance and rumen functions in Gangba sheep as revealed by integrated transcriptome and microbiome analyses. Front Microbiol 13:908326.

78. Mostafavi M, Wang L, Xie L, Takeoka KT, Richie DL, Casey F, Ruzin A, Sawyer WS, Rath CM, Wei J-R, Dean CR. 2018. Interplay of Klebsiella pneumoniae *fabZ* and *lpxC* Mutations Leads to LpxC Inhibitor-Dependent Growth Resulting from Loss of Membrane Homeostasis. mSphere 3:e00508-18.

79. Hoffmeister M, Piotrowski M, Nowitzki U, Martin W. 2005. Mitochondrial trans-2-Enoyl-CoA Reductase of Wax Ester Fermentation from Euglena gracilis Defines a New Family of Enzymes Involved in Lipid Synthesis. J Biol Chem 280:4329–4338.

80. Tucci S, Martin W. 2007. A novel prokaryotic *trans* -2-enoyl-CoA reductase from the spirochete *Treponema denticola*. FEBS Lett 581:1561–1566.

Paper III

Metagenomic and metatranscriptomic interplay between the pig gut microbiota and short-chain fatty acids

Sebastià, C., Folch, J.M., Estellé, J., Ballester, M., Castelló, A., Sánchez, A., and Crespo-Piazuelo, D.

(in preparation)

Metagenomic and metatranscriptomic interplay between the pig gut microbiota and short-chain fatty acids

Sebastià, C.; Folch, J.M.; Estellé, J.; Ballester, M., Castelló, A., Sánchez, A., and Crespo-Piazuelo, D.

Abstract

Bacterial communities are usually studied with 16S rRNA gene sequencing methodologies, which often provide taxonomic information only up to the genus level. In this study, we compared the abundance and activity between the metagenome and metatranscriptome of pigs with extreme differences in SCFAs content in their rectal content. At phylum level, Firmicutes and Bacteroidetes were the most active and abundant, but Spirochaetes showed higher activity than expected relative to its abundance. At genus level, *Clostridium* was the most active, despite being third in abundance after *Bacteroides* and *Prevotella*. Differential metatranscriptomic analysis reported the potential role of *Clostridium* and *Lactobacillus* species, as acetate producers, in modulating SCFA content. Additionally, *Akkermansia glycaniphila* was also identified as an acetate producer, while *Prevotella copri* was more active in samples with high butyric acid content, supporting its role in butyrate production via the butyryl-CoA pathway. Overall, our findings highlight the potential limitations of relying solely on metagenome datasets rather than metatranscriptome data for functional and taxonomic analyses.

Introduction

The gut microbiota is composed of trillions of microorganisms that play an essential role in the health and disease of the host (Thursby & Juge, 2017). While the disruption of the homeostasis by pathogens causes a dysbiosis and has different negative consequences in the host such as enteric or systemic diseases (Carding et al., 2015), the symbiotic microbiota has multiple and beneficial functions in immunity and metabolism (Koh & Bäckhed, 2020; Rooks & Garrett, 2016). The gut microbiome changes along the intestine, depending on the diet and the age of the individual, but also on the genetic background and immunity status (Crespo-Piazuelo et al., 2018; Kurilshikov et al., 2017; Rinninella et al., 2019). Individuals with an imbalance in the gut microbiota, also referred to as dysbiosis, can manifest symptoms such as diarrhoea. In pigs, these diarrhoeas are typically found at weaning, one of the most challenging moments in pig production, where the piglets worsen their performance, resulting in important economic losses (Gresse et al., 2017; Lallès et al., 2007).

Gut microorganisms are able to break down dietary fibres and produce metabolites such as short-chain fatty acids (SCFAs) (Ramakrishna, 2013), from which the most abundant ones in the rectal content are acetate, propionate and butyrate. They can act as an energy source for colonocytes or pass through the enterocytes to blood circulation and participate in different processes such as lipid metabolism (Morrison & Preston, 2016).

Gut bacterial communities are usually analysed through 16S rRNA gene sequencing techniques, which examine the hyper-variable regions of the bacterial and archaeal 16S subunit of the rRNA gene, allowing the taxonomic classification of a high number of samples at genus or species level (Johnson et al., 2019). On the other hand, metagenomic analyses sequence the whole genome of all the different microorganisms contained within a sample, quantifying and giving a deeper taxonomic classification at species and strain level than 16S rRNA gene sequencing analyses (Jovel et al., 2016; Ranjan et al., 2016). Metagenomic analyses also have the advantage of inferring the functions of the species present within the sample (Ranjan et al., 2016), but their main drawback is that the sequenced microbiota may not be active. Thus, metatranscriptomic techniques target the mRNA of the samples as it is more accurate for representing active species and their expressed genes (Bashiardes et al., 2016). These studies become very important when trying to understand the microbial gut populations, as they interact closely with the host (Patil et al., 2020; Quince et al., 2017; Y. Zhang et al., 2021).

The aim of this work was to study the differences between the abundance and the activity of the gut microbiota in two groups of pigs differentiated by their SCFAs content.

Materials and methods

Animal material and DNA/RNA sequencing

A total of 26 healthy pigs belonging to a commercial F1 Iberian × Duroc population were used in this study. The rearing conditions and the sampling are defined in (Sebastià et al., 2024). In brief, the animals were kept in intensive conditions and fed with a cereal-based commercial diet. Samples from the rectal content were collected at the slaughterhouse, snap-frozen in liquid nitrogen and preserved at -80 ºC until further use. 20 of the pigs were used for metagenomic analyses and were selected from the original population based on SCFA content of the rectal content samples, which was evaluated with gas chromatography. For each sample, 0.2 g of rectal content were used to extract the microbial DNA with the PowerFecal kit (MoBio Laboratories, Carlsbad, CA, USA) following the manufacturer's recommendations. DNA purity and concentration were measured through a ND-1000 spectrophotometer. Samples were sequenced in two runs of an Illumina Hi-Seq 3000/4000 instrument (Illumina, San Diego, CA, USA) in the Sequencing Unit of the CNAG (*Centre Nacional d'Anàlisi Genòmica*) using 2 × 150 bp paired-end reads. An average of 25.64 million of read-pairs per sample [20.00 - 31.07 million] was obtained, with a total size of 154.88 Gb (having an average size of 7.74 Gb per sample).

Metatranscriptome analyses were performed on 24 pigs based on SCFAs content (18 animals were selected for both metagenomic and metatranscriptomic analyses). For each sample, microbial RNA was extracted using the Macherey-Nagel NucleoSpin RNA stool kit (Macherey-Nagel GmbH & Co. KG, Germany). The extracted RNA was sequenced in an Illumina NextSeq High Output kit (Illumina, San Diego, CA, USA) in the Sequencing Service of the FISABIO (*Fundació per al Foment de la Investigació Sanitària i Biomèdica de la Comunitat Valenciana*) using 2 × 150 bp paired-end reads. There were obtained an average number of 16.38 million reads per sample [10.92 - 19.19 million], and with a total size of 95.77 Gb (with an average of 3.99 Gb per sample). FISABIO Sequencing Service performed a first quality control of the samples, establishing a minimum read length of 50 bp and removing the samples with a Phred score lower than 20. Therefore, the remaining reads on average were 15.52 million per sample [9.82 - 18.55 million], and with a total size of 88.77 Gb (with an average of 3.7 Gb per sample).

Bioinformatics analyses

Taxonomic and functional classification

For metagenomic samples, bioinformatics analyses were performed using the software metaWRAP v1.3.2 (Uritskiy et al., 2018). Quality control was carried out on the raw fastq files generated in the sequencing with the metaWRAP module read_qc and its default parameters, which uses the programmes Trim-galore v0.6.6, bmtagger v3.101 and FASTQC v0.11.9. This step removed the Illumina adaptors as well as the host reads. Quality control discarded an average 20.89% of the sequences, which mainly belonged to the pig genome. An average of 20.28 million reads per sample was conserved.

For each sample, the two Illumina runs were later co-assembled, binned, and processed with the different metaWRAP modules from the standard pipeline and their default values. A complete list of all the modules included in the software, as well as their function and the programmes they use, can be found in the Supplementary Table S1a. The Blobology module assigned a putative taxonomy for each contig, and they were filtered to keep only the Bacteria and Archaea domains.

Prior to the metatranscriptomic analysis, the module read_qc from metaWRAP was applied to filter out the RNA sequences that belonged to the host. This step detected an average 19.02% of pig sequences, leaving an average of 12.57 million reads per sample for the subsequent analyses. Then, SAMSA v2.2.0 (Westreich et al., 2018) was used to obtain a taxonomic and functional classification. A second quality control was carried out with SAMSA2 integrated program Trimmomatic v0.36, using the default values (a minimum Phred score of 33 and a minimum sequence length of 70 bp), which filtered out the low-quality reads and the Illumina adaptors of the metatranscriptomic samples. Forward and reverse files of each sample were merged with the program PEAR v0.9.8, and the merged files were analysed with SortMeRNA v2.1, which removed the rRNA sequences. The ribodepleted files were annotated with DIAMOND v0.8.3 based on RefSeq database, which was used to annotate the microbial species and was later aggregated with the Python script DIAMOND_analysis_counter.py and analysed with different SAMSA2 R scripts. For a complete description of SAMSA2 integrated programs and functions, see Supplementary Table S1b.

For metatranscriptomics functional analysis, filtered samples were processed with kallisto v0.48.0 software (Bray et al., 2016), using the reference catalogue of the pig gut microbiome (Xiao et al., 2016). First, the reference genome was indexed using a kmer length of 21, and

after that the samples were quantified with the indexed reference genome. The reference catalogue included 7.7 million genes, which were merged into 7,012 KOs (KEGG Orthologies).

Differential expression analysis

Metatranscriptomic data was explored with a differential expression analysis. Samples were divided in two groups (high vs. low) based on their abundance of each of the three main SCFA (i.e., acetic, propionic, and n-butyric acids), resulting in six samples per group. The raw microbial taxonomic and functional classifications were filtered to remove low-abundant taxa and KOs (0.005% minimum abundance threshold). Filtered data was analysed with the DESeq2 v1.36 package (Love et al., 2014). Taxa or KOs with an adjusted *p*-value ≤ 0.05 and |log2 fold change (FC)| ≥ 1.5 were considered significant. MA and Volcano plots were created with the EnhancedVolcano v1.13.2 R package (Blighe et al., 2019).

Table 1. Average content (%) of high and low groups for the three studied SCFAs.

	High group	**Low group**
Acetic acid	64.61	52.42
Propionic acid	22.10	16.67
Butyric acid	14.24	7.48

Results

Taxonomic classification and comparison

In the 20 metagenomic samples, Blobology found 59 bacterial and archaeal phyla, which were divided into 1,466 bacterial and archaeal genera and 6,476 species. Firmicutes, Bacteroidetes and Proteobacteria were the three most abundant phyla, with 44.74%, 24.82% and 11.48% of the relative abundance, respectively (Figures 1a and 2a). Out of all the bacterial and archaeal sequences, 4.67% remained unclassified at phylum level. At genus level, *Prevotella*, *Bacteroides* and *Clostridium* were the three most abundant taxa, with 6.54%, 6.43% and 4.90% of the total abundance, respectively (Figures 1b, 2b and 3).

In the metatranscriptomic analysis, 38 bacterial and archaeal phyla were found in the 24 samples, which were divided into 1,895 bacterial and archaeal genera and 8,282 species. The most abundant phyla were Firmicutes and Bacteroidetes, with 49.65% and 27.33% of the total relative abundance (Figures 1c and 2b). Other abundant phyla were Spirochaetes (8.35%) and Proteobacteria (7.70%). Only 0.18% of the sequences remained not annotated at phylum level. The most abundant genus was *Clostridium*, representing 18.15% of the total abundance (Figures 1d, 2b, and 3). Other abundant genera were *Treponema* (7.15%), *Prevotella* (5.37%), *Bacteroides* (5.30%) and *Lactobacillus* (5.21%).

When comparing the metagenome and the metatranscriptome datasets at phylum level, Spirochaetes showed higher activity than expected relative to its abundance (Figure 2a). Conversely, Actinobacteria and Proteobacteria were less active than expected. At genus level, *Clostridium* was the most active, despite being third in abundance (Figures 2b and 3). The most abundant genera were *Bacteroides* and *Prevotella*, although *Treponema* was the second most active genus. Altogether, with the inclusion of *Lactobacillus*, the top 5 genera were also the 5 most active. Notably, some genera such as *Dorea*, *Dysgonomonas* and *Butyriciccocus* were more active than abundant, whereas *Acutalibacter* showed lower activity than expected relative to its abundance.

Although fewer species were initially found in the metagenome dataset (n = 6,476) compared to the metatranscriptome dataset (n = 8,282), after filtering out those genera with a relative abundance lower than 0.005% in both datasets, 4,809 species remained in the metagenome dataset and 2,571 in the metatranscriptome dataset. All the species found in the metatranscriptome dataset were already found in the metagenome analysis (Figure 4).

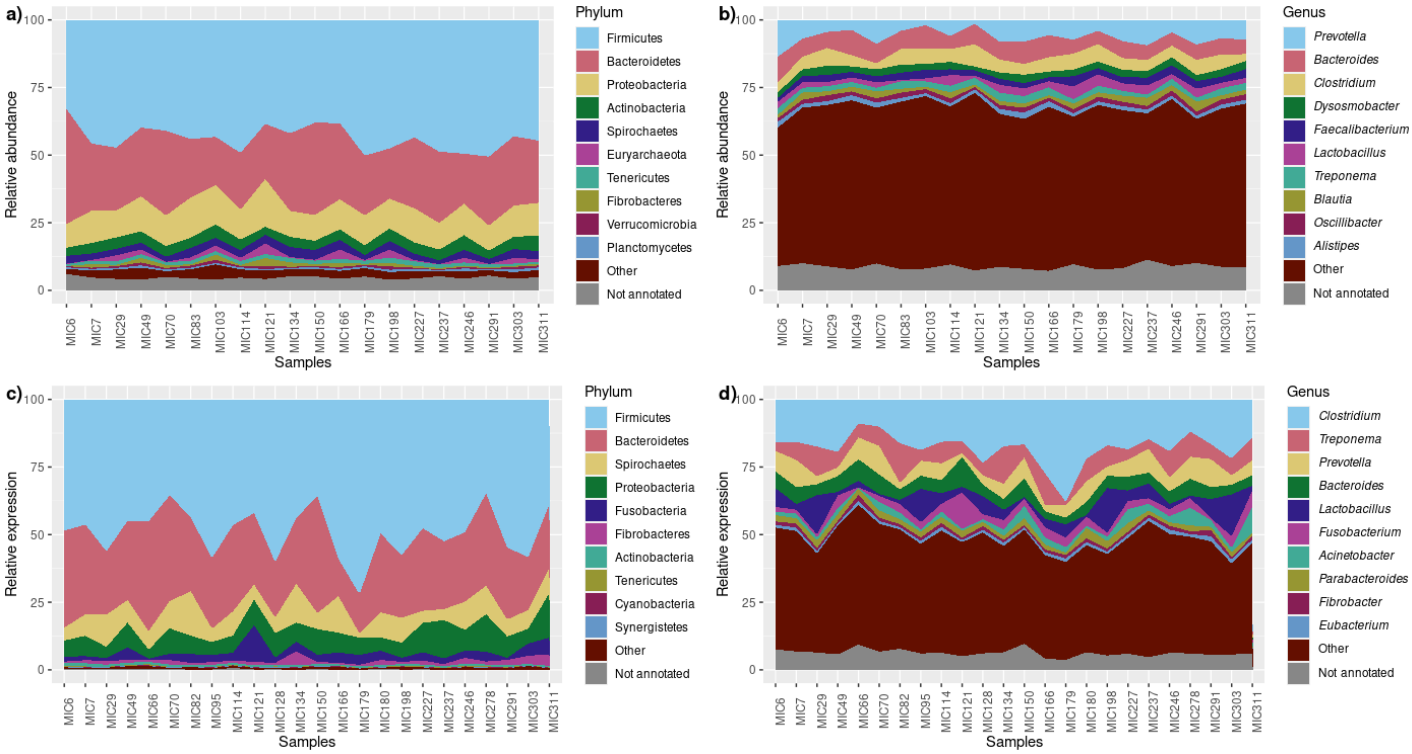


Figure 1. Stacked area plots representing the relative abundance of top 10 classified taxa of the metagenomic and metatranscriptomic analyses. a) Metagenomic results at phylum level. The rest of the 49 annotated phyla were grouped as “Other”, which accounted for 2.74% of the total abundance. b) Metagenomic results at genus level. The rest of the 1,456 annotated genera were grouped as “Other”, which accounted for 54.76% of the total abundance. c) Metatranscriptomic results at phylum level. The rest of the 28 annotated phyla were grouped as “Other”, which accounted for 0.76% of the total expression. d) Metatranscriptomic results at genus level. The rest of the 1,884 annotated taxa were grouped as “Other”, which accounted for 37.07% of the total expression.

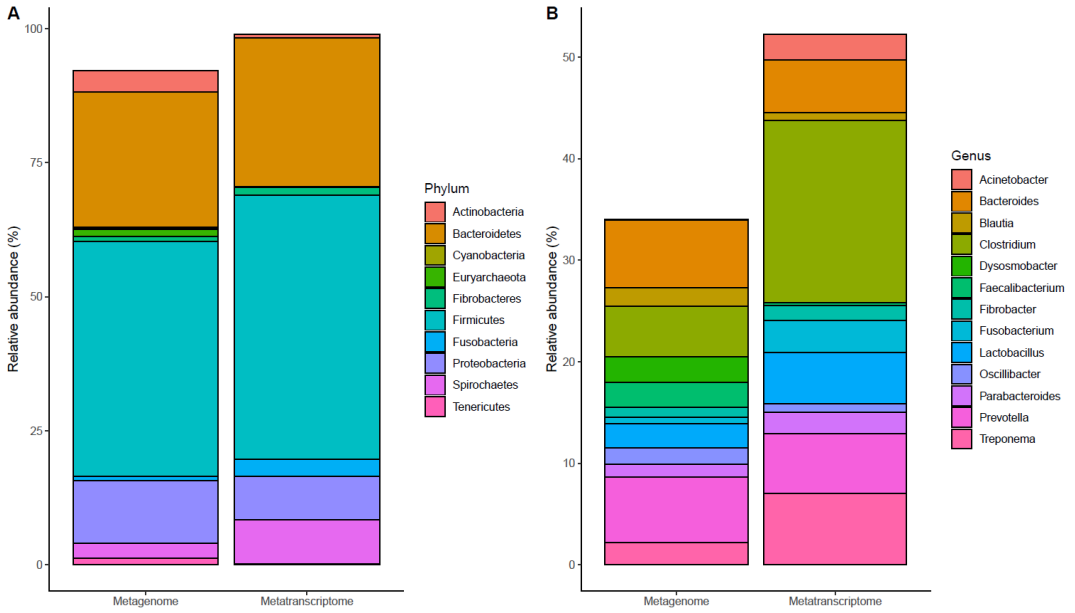


Figure 2. Stacked bar plots representing the mean of the relative abundance of (A) phyla and (B) genera found in the metagenome (n = 20) and metatranscriptome (n = 24) of pigs.

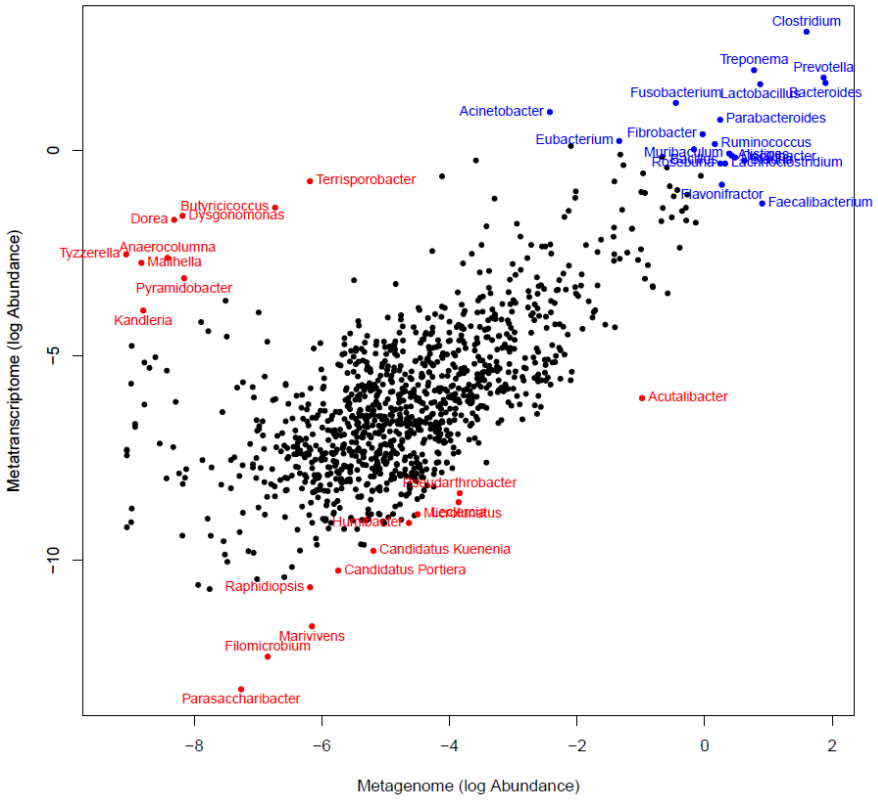


Figure 3. Scatter plot showing the abundance of genera in metagenome and metatranscriptome results. Blue dots represent the 20 most abundant genera, while red dots highlight the 20 genera with the greatest differences between their metagenomic and metatranscriptomic abundances.

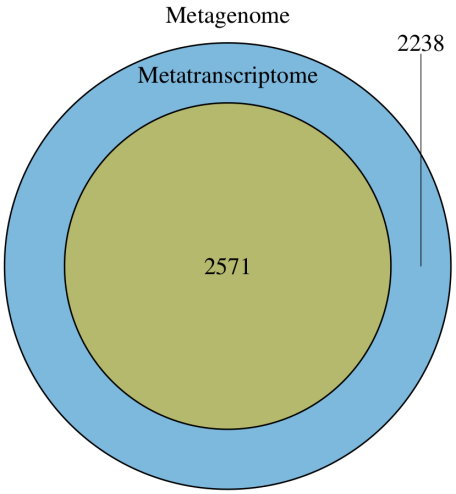


Figure 4. Venn diagram representing the shared species between the metagenome and metatranscriptome datasets. Species with a relative abundance lower than 0.005% were filtered in both datasets.

Metatranscriptomic analysis

In addition to the taxonomy classification extracted from the metatranscriptomic dataset, KOs were also identified and quantified to further understand the metabolic functions and biological processes within the metatranscriptome. Of the 7,012 KOs in the pig gut reference catalogue (Xiao et al., 2016), 5,544 were found at least once in the metatranscriptomic samples. The most abundant KO was K06421, which codifies for the small acid-soluble spore protein D (minor alpha/beta-type SASP), with almost half of the total estimated counts (47.51%). Other abundant KOs codified for an anti-repressor protein (K07741; 8.66%); peroxiredoxin 2/4 (K03386; 2.59%), GTPase (K07588; 1.63%), and the elongation factor Tu (K02358; 1.07%).

The metatranscriptomic results were analysed to assess the differences in activity of the species and in expression of KOs based on the content of the three main SCFA in the samples. In general, samples with high acetic acid content had low propionic and butyric acid content, and vice versa. This opposite behaviour is illustrated in Figure 5, where the arrows indicate this relationship. The results of the metratranscriptomic analysis can be found in Supplementary Table S2.

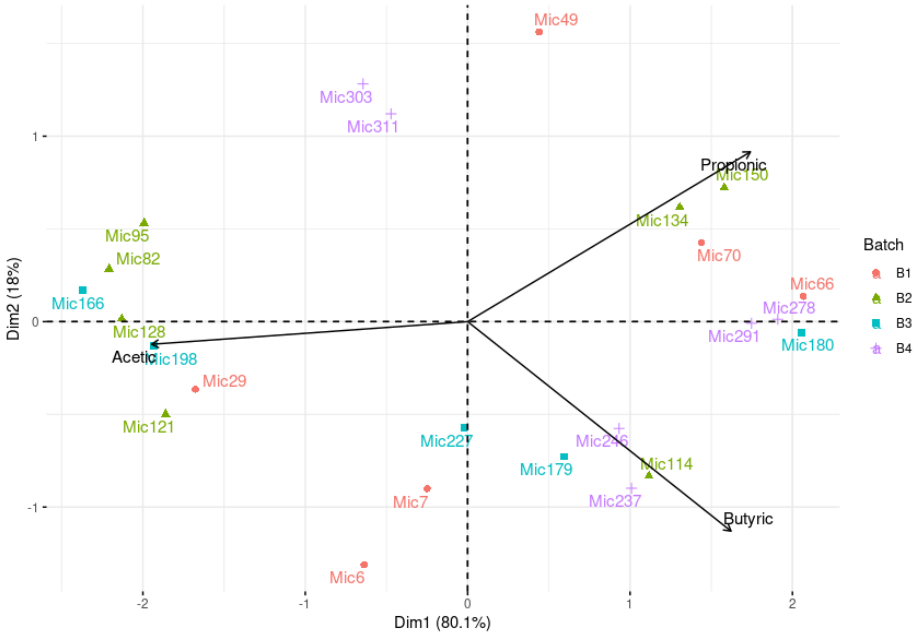


Figure 5. PCA biplot with the sample distribution based on the three main SCFAs content. Samples are coloured depending on their batch, which was taken into account to choose the high and low animals for each SCFA.

Acetic acid

There were 24 species that showed differential activity between samples with high and low acetic acid content: 19 more active in samples with high acetic acid content and five were more active in samples with low acetic acid content (Figure 6a). Out of these 19 species more active in high-acetic samples, one was *Akkermansia glycaniphila* and another seven were from the genus *Lactobacillus*, including *L. reuteri*. The five species more active in low-acetic samples had a relatively low abundance. The most significantly active species between groups was *Clostridium hydrogeniformans*.

Regarding the functional results, 11 KOs were differentially expressed between groups, being nine of them over-expressed in samples with high acetic acid content (Figure 6b). The two most significantly over-expressed KOs were K00356 and K06346.

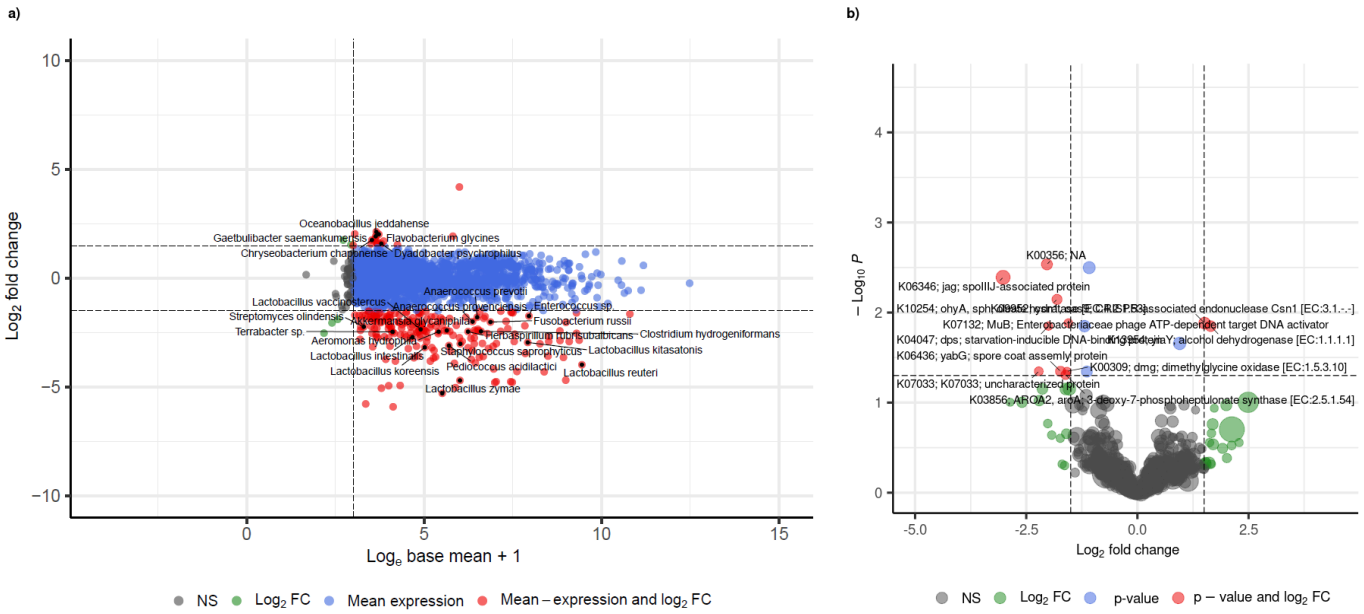


Figure 6. Differential abundance analysis between samples with high and low acetic acid content. a) MA plot showing the differences in bacterial activity. Species with log2FC ≤ -1.5 were more active in high-acetic gut samples. Significantly active species (adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-acetic gut samples.

Propionic acid

The study of the differentially active species between samples with differing propionic acid content followed a mirrored-like distribution as the one observed for the acetic acid. A total of 33 species exhibited significant differential activity (Figure 7a). Among these, 28 species showed higher activity in samples with low propionic acid content. These included *Akkermansia glycaniphila*, six species of *Clostridium*, *Escherichia coli*, four species of *Lactobacillus* and *Shigella sp.* The most significantly active species was *Klebsiella pneumoniae*, followed by *C. hydrogeniformans*. The five genera that were more active in high-propionic samples were also the same ones that showed higher activity in low-acetic samples.

In the functional analysis, 14 KOs were differentially expressed between groups: five of them over-expressed in high propionic acid content samples and nine in low-propionic samples (Figure 7b). The two most significantly over-expressed KOs were K04047 and K10254, which had a higher expression in low-propionic samples and were also found over-expressed in high-acetic samples.

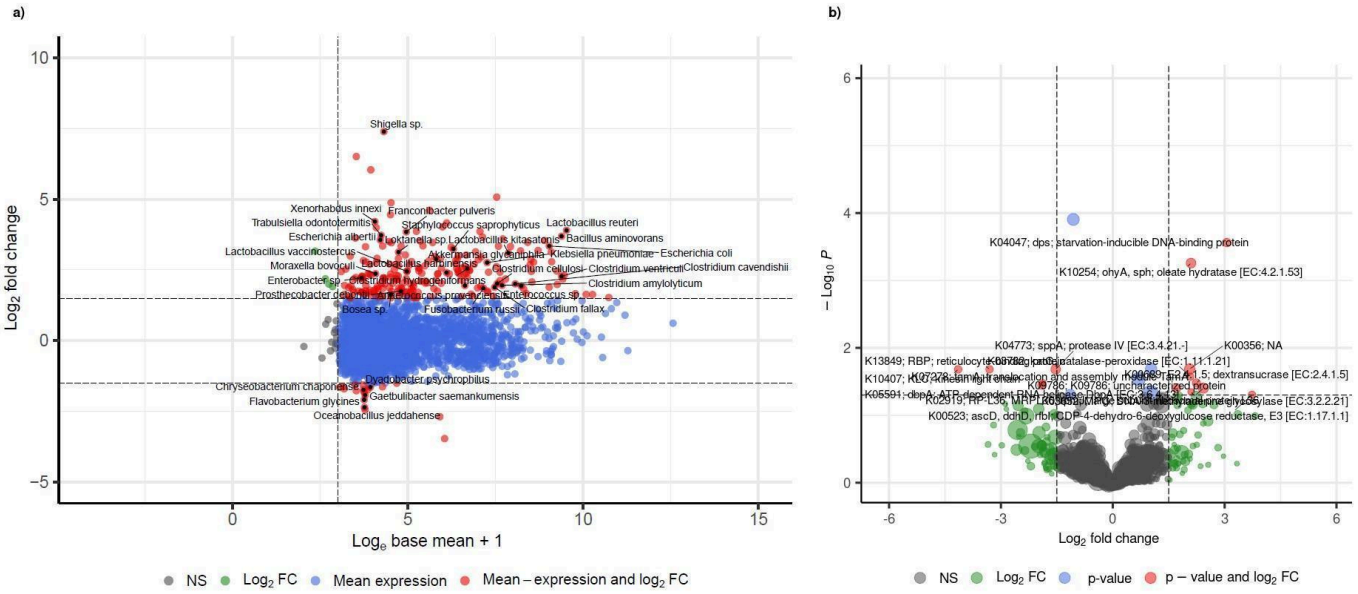


Figure 7. Differential abundance analysis between samples with high and low propionic acid content. a) MA plot showing the differences in bacterial activity. Species with log2FC ≤ -1.5 were more active in high-propionic gut samples. Significantly active species (adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-propionic gut samples.

Butyric acid

Similar to the distribution observed in the propionic acid analysis, 25 species exhibited differential activity between samples with high and low butyric acid content (Figure 8a). Samples with high butyric acid content had four significant active species, including *Prevotella copri* and *Prevotella stercorea*. The 21 significantly active species in low-butyric samples included *Akkermansia glycaniphila*, *Escherichia coli*, and *Shigella dysenteriae*, which was the most significant one.

The functional results showed six KOs over-expressed in samples with low butyric acid content, while no under-expressed KOs were obtained (Figure 8b). The most significantly over-expressed KO was K06346.

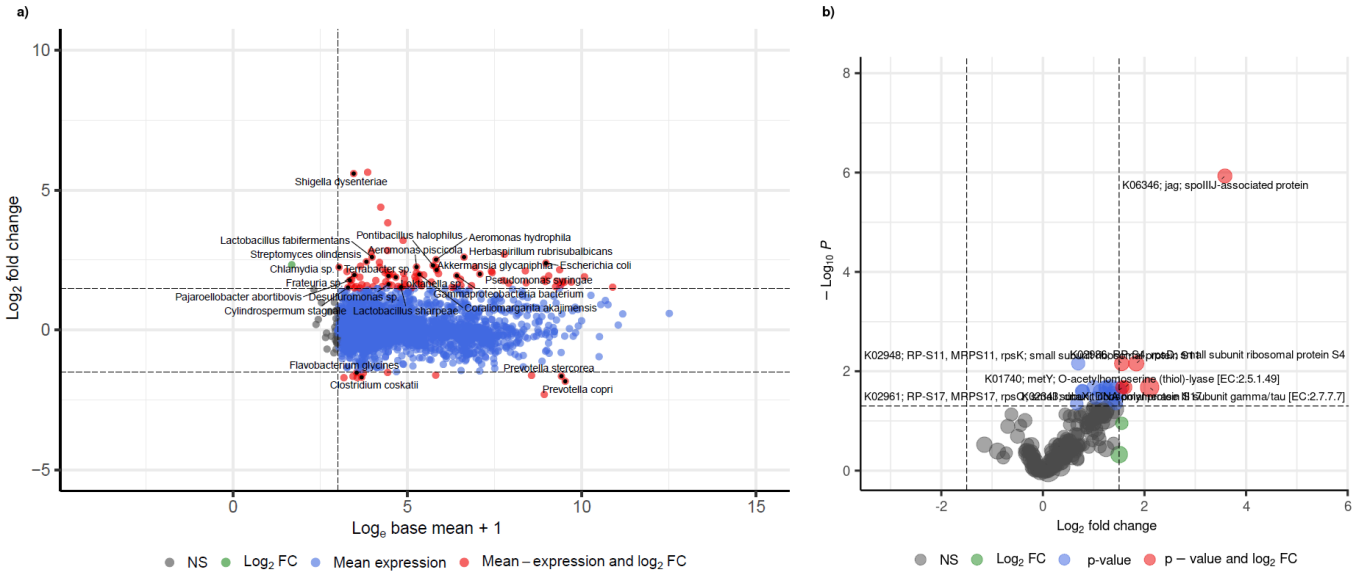


Figure 8. Differential abundance analysis between samples with high and low butyric acid content. a) MA plot showing the differences in bacterial activity. Species with log2FC ≤ -1.5 were more active in high-butyric gut samples. Significantly active species (adjusted *p*-value ≤ 0.05) are marked in black. b) Volcano plot representing the differentially expressed KOs. KOs with log2FC ≤ -1.5 were more expressed in high-butyric gut samples.

Discussion

In this study, we compared the metagenome and the metatranscriptome of pigs with extreme differences of SCFAs content in their rectal content. The most abundant and active phylum was Firmicutes, which represented almost half of the microbial populations, followed by Bacteroidetes. These two phyla are usually the most abundant in the pig hindgut; Firmicutes is the dominant taxon in the small bowel along with Proteobacteria, while Bacteroidetes abundance is lower in the small intestine, increasing its presence in colon (Crespo-Piazuelo et al., 2018; L. Zhang et al., 2018; Zhao et al., 2015). The ratio Firmicutes/Bacteroidetes was similar in both metagenome and metatranscriptome datasets. However, the Spirochaetes phylum, mainly represented by *Treponema*, showed higher activity than expected relative to its abundance. In addition, other phyla were less active than expected relative to their abundance, mainly Actinobacteria and Proteobacteria. In this context, further studies are required to unveil these discrepancies as bacterial processes, such as genetic regulation, microbial and host interaction, sporulation or environmental effects could modulate the activity of certain bacteria. Furthermore, some of the sequenced bacteria in the metagenome dataset could be residual from the upper parts of the gut, and thus, not be active. Supporting this hypothesis, all the species found in the metatranscriptome dataset were also found in the metagenome results.

At genus level, the most active genus was *Clostridium*, representing almost a fifth of the total active genera, despite being third in terms of abundance. *Clostridium* is considered a core genus of the pig gut microbiota, and its abundance is higher in grown pigs than in piglets (Wang et al., 2022). This clade is divided into 19 clusters, but the clostridia found in the healthy gut of humans and animals, and specifically in the colon, belong mainly to clusters IV and XIVa (Guo et al., 2020). These clusters produce a wide variety of metabolites, including different SCFA among them, such as acetate (Zhu et al., 2022). In agreement with *Clostridium* being the most active genus, the most abundant KO was K06421 (sspD), a gene typically found in *Bacillus* and *Clostridium* (B. Setlow et al., 1992).

In the differential metatranscriptomic analysis, where bacterial activity can be compared through the gene expression profiles of the bacterial communities, the most active species found for the samples with high acetic acid content was *Clostridium hydrogeniformans*. Despite *C. hydrogeniformans* being characterised by its production of hydrogen, it can act as an acetate producer, and to a lesser extent, butyrate (Bowman et al., 2010). *Lactobacillus reuteri* was also more active in samples with high acetic acid content. *L. reuteri*, like other obligate heterofermentative lactobacilli, produces lactic acid together with ethanol and acetic acid (Hammes & Hertel, 2009; Oh et al., 2019; Slavica et al., 2015). Furthermore, *L. reuteri*

has been used as a probiotic to prevent the development of acetic acid-induced colitis in rats (Fabia et al., 1993). Following with the propionic results, we have already described a mirrored distribution in the MA plot of acetic acid, consistent with the negative correlation between these two SCFA and the use of the relative abundances for the analysis. Here, *Escherichia coli* and *Shigella sp.* were more active in low-propionic samples. *E. coli* is one of the most well-known bacteria in pig production, and it has been reported that its excretion of acetate can inhibit bacterial growth (Pinhal et al., 2019). *Shigella sp.*, another member of the Enterobacteriaceae family, behaves similarly to *E. coli* (Zha et al., 2024).

Species already described as acetic acid producers, such as *E. coli* and *Shigella dysenteriae*, were more active in low-butyric samples. However, *Prevotella copri* and *Prevotella stercorea* were few of the species more active in animals with high content in butyric acid. *Prevotella* was the third most expressed genus (after *Clostridium* and *Treponema*) and the most abundant one, which agrees with the fact that this taxon is frequently found as one of the most abundant in the distal colon of pigs (Crespo-Piazuelo et al., 2018; Holman et al., 2017; Quan et al., 2020). *P. copri* is one of the most common species of this genus in the pig gut, followed by *P. stercorea* and *P. ruminicola* (Amat et al., 2020). This species produces succinate (Franke & Deppenmeier, 2018), which may generate propionate or, alternatively, be transformed into butyrate through the butyryl-CoA pathway (Oliphant & Allen-Vercoe, 2019). In our previous work (Sebastià et al., 2024), we found the levels of the *Prevotella* genus positively correlated with those of the butyric acid in the rectal content.

Through the three differential metatranscriptomic analyses, *Akkermansia glycaniphila* was more active in pigs with high content of acetic acid and with low content of propionic and butyric acids. The *Akkermansia* genus is one of the few genera within the Verrucomicrobia phylum, and was found to be positively correlated with the levels of acetic acid measured in a larger population (n = 285), which included the same animal material used for this study (Sebastià et al., 2024). Accordingly, the members of the *Akkermansia* genus have been reported as acetate producers (Abbasi et al., 2023; Dao et al., 2016; Derrien et al., 2004). Firstly described in the digestive tract of pythons (Ouwerkerk et al., 2017), *A. glycaniphila* was later found in the human gut (Glover et al., 2022; Lv et al., 2022). The fact that *A. glycaniphila* was more active in pigs with a high abundance of acetic acid suggests that this species can be an acetate producer as *Akkermansia muciniphila*. To the best of our knowledge, this is the first study reporting *A. glycaniphila* as a species of the pig gut microbiota.

The functional analysis showed common proteins found in many bacterial species as the most abundant ones. The small acid-soluble spore protein D (minor α/β-type SASP) was the most expressed KO by far, representing almost half of the KO abundance. α/β-type SASP proteins are synthesised when the spore is formed and are among the most abundant ones in the core of the structure. By binding to the spore DNA, they protect it from damage caused by chemicals or heat (Koopman et al., 2022; P. Setlow, 2006, 2007). Spore-forming bacteria belong to the Firmicutes phylum (Hutchison et al., 2016) and are part of the commensal bacteria of the gut, in which can be found in vegetative or spore form depending on the environment and the microorganism traits (Browne et al., 2016).

Interestingly, the second most expressed KO was an anti-repressor protein (antB) found in a wide variety of bacteriophages, which allows the over-expression of the phages (Azulay et al., 2022). The gut virome is still largely unknown, but in the recent years there have been published several human gut virome databases (Li et al., 2022). Even though our analysis did not study viral data, this finding hints an important activity of bacteriophages in the microbial cells of the pig gut.

The next most expressed proteins were the bacterial peroxiredoxin AhpC, which is involved in cell protection (Parsonage et al., 2008); the GTPase argK, involved in amino acid transport (Dobson et al., 2002); and the elongation factor Tu (as tuf gene), which performs various functions both in the cytosol and on the cell surface of prokaryotic and eukaryotic cells (Harvey et al., 2019).

Regarding the differential expression analysis, the KOs found most expressed when there was low content in propionic or butyric acids were also found in a high acetic content, as this last SCFA abundance follows a negative correlation with the other two. The most expressed KO in animals with high acetic content was a hypothetical protein encoded by the KEGG ortholog K00356, which was also found over-expressed in low propionic content. In a similar way, starvation-inducible DNA-binding protein (K04047, dps) and oleate hydratase (K10254, ohyA, sph) were the most differentially expressed proteins in samples with low propionic content, and were also found over-expressed in high acetic content. In turn, spoIIIJ-associated protein (K06346, jag) was over-expressed in low butyric content, while it was found over-expressed in samples with high acetic content.

Both jag and dps proteins are expressed in harsh environment conditions, in response to starvation that results in sporulation in some species. The ohyA protein is an unsaturated fatty acid hydratase that produces hydroxy-fatty acids that suppress the inflammatory response and therefore they create a better environment for the bacteria (Miyamoto et al., 2015; Radka et al., 2021). These harsh environment adaptations are found in samples with

high content in acetic acid and low in propionic and butyric acids, which agrees with the finding of acetate as substrate for sporulating bacteria (Sarrafzadeh et al., 2014).

Conclusions

This work aimed to compare bacterial abundances and activity between the metagenomic and metatranscriptomic datasets obtained for the pig gut microbiota and their potential relationship with SCFA production. As expected, the two main phyla in both datasets were Firmicutes and Bacteroidetes. However, Spirochaetes showed higher activity than expected relative to its abundance, while Actinobacteria and Proteobacteria were less active than expected. At genus level, *Clostridium* was the most active, but *Bacteroides* and *Prevotella* were the most abundant. The differential metatranscriptomic analysis revealed that *Clostridium* and *Lactobacillus* species, and *C. hydrogeniformans* and *L. reuteri* in particular, were more active in samples with high acetic acid and low propionic acid content. Additionally, *Akkermansia glycaniphila* was identified as an acetate producer, being less active in high-propionic and high-butyric samples, while *Prevotella copri* was more active in samples with high butyric acid content, supporting its role in butyrate production via the butyryl-CoA pathway. Functional analysis highlighted significant expression of spore-related proteins and bacteriophage activity, such as antB. Overall, our findings underscore the potential bias of relying exclusively on metagenome datasets rather than metatranscriptome data for functional and taxonomic analyses. Additionally, we also reported some species such as *A. glycaniphila* and *P. copri* that could be used as probiotics to modulate SCFA gut profiles.

References

Abbasi, A., Bazzaz, S., Da Cruz, A. G., Khorshidian, N., Saadat, Y. R., Sabahi, S., Ozma, M. A., Lahouty, M., Aslani, R., & Mortazavian, A. M. (2023). A Critical Review on Akkermansia muciniphila: Functional Mechanisms, Technological Challenges, and Safety Issues. *Probiotics and Antimicrobial Proteins*. https://doi.org/10.1007/s12602-023-10118-x

Amat, S., Lantz, H., Munyaka, P. M., & Willing, B. P. (2020). Prevotella in Pigs: The Positive and Negative Associations with Production and Health. *Microorganisms*, 8(10), 1584. https://doi.org/10.3390/microorganisms8101584

Azulay, G., Pasechnek, A., Stadnyuk, O., Ran-Sapir, S., Fleisacher, A. M., Borovok, I., Sigal, N., & Herskovits, A. A. (2022). A dual-function phage regulator controls the response of cohabiting phage elements via regulation of the bacterial SOS response. *Cell Reports*, 39(3), 110723. https://doi.org/10.1016/j.celrep.2022.110723

Bashiardes, S., Zilberman-Schapira, G., & Elinav, E. (2016). Use of Metatranscriptomics in Microbiome Research. *Bioinformatics and Biology Insights*, 10, BBI.S34610. https://doi.org/10.4137/BBI.S34610

Blighe, K., Rana, S., & Lewis, M. (2019). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R Package Version, 1(0), 10.18129.

Bowman, K. S., Dupré, R. E., Rainey, F. A., & Moe, W. M. (2010). Clostridium hydrogeniformans sp. Nov. And Clostridium cavendishii sp. Nov., hydrogen-producing bacteria from chlorinated solvent-contaminated groundwater. *International Journal of Systematic and Evolutionary Microbiology*, 60(2), 358–363. https://doi.org/10.1099/ijs.0.013169-0

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, 34(5), 525–527. https://doi.org/10.1038/nbt.3519

Browne, H. P., Forster, S. C., Anonye, B. O., Kumar, N., Neville, B. A., Stares, M. D., Goulding, D., & Lawley, T. D. (2016). Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature*, 533(7604), 543–546. https://doi.org/10.1038/nature17645

Carding, S., Verbeke, K., Vipond, D. T., Corfe, B. M., & Owen, L. J. (2015). Dysbiosis of the gut microbiota in disease. *Microbial Ecology in Health & Disease*, 26(0). https://doi.org/10.3402/mehd.v26.26191

Crespo-Piazuelo, D., Estellé, J., Revilla, M., Criado-Mesas, L., Ramayo-Caldas, Y., Óvilo, C., Fernández, A. I., Ballester, M., & Folch, J. M. (2018). Characterization of bacterial microbiota compositions along the intestinal tract in pigs and their interactions and functions. *Scientific Reports*, 8(1), 12727. https://doi.org/10.1038/s41598-018-30932-6

Dao, M. C., Everard, A., Aron-Wisnewsky, J., Sokolovska, N., Prifti, E., Verger, E. O., Kayser, B. D., Levenez, F., Chilloux, J., Hoyles, L., MICRO-Obes Consortium, Dumas, M.-E., Rizkalla, S. W., Doré, J., Cani, P. D., & Clément, K. (2016). Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: Relationship with gut microbiome richness and ecology. *Gut*, 65(3), 426–436. https://doi.org/10.1136/gutjnl-2014-308778

Derrien, M., Vaughan, E. E., Plugge, C. M., & De Vos, W. M. (2004). Akkermansia muciniphila gen. Nov., sp. Nov., a human intestinal mucin-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology*, 54(5), 1469–1476. https://doi.org/10.1099/ijs.0.02873-0

Dobson, C. M., Wai, T., Leclerc, D., Wilson, A., Wu, X., Doré, C., Hudson, T., Rosenblatt, D. S., & Gravel, R. A. (2002). Identification of the gene responsible for the cblA complementation group of vitamin B 12 -responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proceedings of the National Academy of Sciences*, 99(24), 15554–15559. https://doi.org/10.1073/pnas.242614799

Fabia, R., Ar'rajab, A., Johansson, M. L., Willén, R., Andersson, R., Molin, G., & Bengmark, S. (1993). The Effect of Exogenous Administration of Lactobacillus reuteri R2LC and Oat Fiber on Acetic Acid-Induced Colitis in the Rat. *Scandinavian Journal of Gastroenterology*, 28(2), 155–162. https://doi.org/10.3109/00365529309096063

Franke, T., & Deppenmeier, U. (2018). Physiology and central carbon metabolism of the gut bacterium Prevotella copri. *Molecular Microbiology*, 109(4), 528–540. https://doi.org/10.1111/mmi.14058

Glover, J. S., Ticer, T. D., & Engevik, M. A. (2022). Characterizing the mucin-degrading capacity of the human gut microbiota. *Scientific Reports*, 12(1), 8456. https://doi.org/10.1038/s41598-022-11819-z

Gresse, R., Chaucheyras-Durand, F., Fleury, M. A., Van de Wiele, T., Forano, E., & Blanquet-Diot, S. (2017). Gut Microbiota Dysbiosis in Postweaning Piglets: Understanding the Keys to Health. *Trends in Microbiology*, 25(10), 851–873. https://doi.org/10.1016/j.tim.2017.05.004

Guo, P., Zhang, K., Ma, X., & He, P. (2020). Clostridium species as probiotics: Potentials and challenges. *Journal of Animal Science and Biotechnology*, 11(1), 24. https://doi.org/10.1186/s40104-019-0402-1

Hammes, W. P., & Hertel, C. (2009). Genus I. Lactobacillus Beijerinck 1901, 212AL. *Bergey's Manual of Systematic Bacteriology*, 3, 465–490.

Harvey, K. L., Jarocki, V. M., Charles, I. G., & Djordjevic, S. P. (2019). The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Frontiers in Microbiology*, 10, 2351. https://doi.org/10.3389/fmicb.2019.02351

Holman, D. B., Brunelle, B. W., Trachsel, J., & Allen, H. K. (2017). Meta-analysis To Define a Core Microbiota in the Swine Gut. *mSystems*, 2(3), e00004-17. https://doi.org/10.1128/mSystems.00004-17

Hutchison, E. A., Miller, D. A., & Angert, E. R. (2016). Sporulation in Bacteria: Beyond the Standard Model. In A. Driks & P. Eichenberger (Eds.), The Bacterial Spore (pp. 87–102). *ASM Press*. https://doi.org/10.1128/9781555819323.ch4

Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, 10(1), 5029. https://doi.org/10.1038/s41467-019-13036-1

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A. L., Madsen, K. L., & Wong, G. K.-S. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Frontiers in Microbiology*, 7. https://doi.org/10.3389/fmicb.2016.00459

Koh, A., & Bäckhed, F. (2020). From Association to Causality: The Role of the Gut Microbiota and Its Functional Products on Host Metabolism. *Molecular Cell*, 78(4), 584–596. https://doi.org/10.1016/j.molcel.2020.03.005

Koopman, N., Remijas, L., Seppen, J., Setlow, P., & Brul, S. (2022). Mechanisms and Applications of Bacterial Sporulation and Germination in the Intestine. *International Journal of Molecular Sciences*, 23(6), 3405. https://doi.org/10.3390/ijms23063405

Kurilshikov, A., Wijmenga, C., Fu, J., & Zhernakova, A. (2017). Host Genetics and Gut Microbiome: Challenges and Perspectives. *Trends in Immunology*, 38(9), 633–647. https://doi.org/10.1016/j.it.2017.06.003

Lallès, J.-P., Bosi, P., Smidt, H., & Stokes, C. R. (2007). Nutritional management of gut health in pigs around weaning. *Proceedings of the Nutrition Society*, 66(2), 260–268. https://doi.org/10.1017/S0029665107005484

Li, J., Yang, F., Xiao, M., & Li, A. (2022). Advances and challenges in cataloging the human gut virome. *Cell Host & Microbe*, 30(7), 908–916. https://doi.org/10.1016/j.chom.2022.06.003

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lv, Q.-B., Li, S., Zhang, Y., Guo, R., Wang, Y.-C., Peng, Y., & Zhang, X.-X. (2022). A thousand metagenome-assembled genomes of Akkermansia reveal phylogroups and geographical and functional variations in the human gut. *Frontiers in Cellular and Infection Microbiology*, 12, 957439. https://doi.org/10.3389/fcimb.2022.957439

Miyamoto, J., Mizukure, T., Park, S.-B., Kishino, S., Kimura, I., Hirano, K., Bergamo, P., Rossi, M., Suzuki, T., Arita, M., Ogawa, J., & Tanabe, S. (2015). A Gut Microbial Metabolite of Linoleic Acid, 10-Hydroxy-cis-12-octadecenoic Acid, Ameliorates Intestinal Epithelial Barrier Impairment Partially via GPR40-MEK-ERK Pathway. *Journal of Biological Chemistry*, 290(5), 2902–2918. https://doi.org/10.1074/jbc.M114.610733

Morrison, D. J., & Preston, T. (2016). Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes*, 7(3), 189–200. https://doi.org/10.1080/19490976.2015.1134082

Oh, J.-H., Alexander, L. M., Pan, M., Schueler, K. L., Keller, M. P., Attie, A. D., Walter, J., & Van Pijkeren, J.-P. (2019). Dietary Fructose and Microbiota-Derived Short-Chain Fatty Acids Promote Bacteriophage Production in the Gut Symbiont Lactobacillus

reuteri. *Cell Host & Microbe*, 25(2), 273-284.e6. https://doi.org/10.1016/j.chom.2018.11.016

Oliphant, K., & Allen-Vercoe, E. (2019). Macronutrient metabolism by the human gut microbiome: Major fermentation by-products and their impact on host health. *Microbiome*, 7(1), 91. https://doi.org/10.1186/s40168-019-0704-8

Ouwerkerk, J. P., Koehorst, J. J., Schaap, P. J., Ritari, J., Paulin, L., Belzer, C., & de Vos, W. M. (2017). Complete Genome Sequence of Akkermansia glycaniphila Strain Pyt T , a Mucin-Degrading Specialist of the Reticulated Python Gut. *Genome Announcements*, 5(1), e01098-16. https://doi.org/10.1128/genomeA.01098-16

Parsonage, D., Karplus, P. A., & Poole, L. B. (2008). Substrate specificity and redox potential of AhpC, a bacterial peroxiredoxin. *Proceedings of the National Academy of Sciences*, 105(24), 8209–8214. https://doi.org/10.1073/pnas.0708308105

Patil, Y., Gooneratne, R., & Ju, X.-H. (2020). Interactions between host and gut microbiota in domestic pigs: A review. *Gut Microbes*, 11(3), 310–334. https://doi.org/10.1080/19490976.2019.1690363

Pinhal, S., Ropers, D., Geiselmann, J., & De Jong, H. (2019). Acetate Metabolism and the Inhibition of Bacterial Growth by Acetate. *Journal of Bacteriology*, 201(13). https://doi.org/10.1128/JB.00147-19

Quan, J., Wu, Z., Ye, Y., Peng, L., Wu, J., Ruan, D., Qiu, Y., Ding, R., Wang, X., Zheng, E., Cai, G., Huang, W., & Yang, J. (2020). Metagenomic Characterization of Intestinal Regions in Pigs With Contrasting Feed Efficiency. *Frontiers in Microbiology*, 11, 32. https://doi.org/10.3389/fmicb.2020.00032

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, 35(9), 833–844. https://doi.org/10.1038/nbt.3935

Radka, C. D., Batte, J. L., Frank, M. W., Rosch, J. W., & Rock, C. O. (2021). Oleate Hydratase (OhyA) Is a Virulence Determinant in Staphylococcus aureus. *Microbiology Spectrum*, 9(3), e01546-21. https://doi.org/10.1128/Spectrum.01546-21

Ramakrishna, B. S. (2013). Role of the gut microbiota in human nutrition and metabolism: Role of the gut microbiota. *Journal of Gastroenterology and Hepatology*, 28, 9–17. https://doi.org/10.1111/jgh.12294

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and Biophysical Research Communications*, 469(4), 967–977. https://doi.org/10.1016/j.bbrc.2015.12.083

Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G., Gasbarrini, A., & Mele, M. (2019). What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms*, 7(1), 14. https://doi.org/10.3390/microorganisms7010014

Rooks, M. G., & Garrett, W. S. (2016). Gut microbiota, metabolites and host immunity. *Nature Reviews Immunology*, 16(6), 341–352. https://doi.org/10.1038/nri.2016.42

Sarrafzadeh, M. H., Schorr-Galindo, S., La, H.-J., & Oh, H.-M. (2014). Aeration effects on metabolic events during sporulation of Bacillus thuringiensis. *Journal of Microbiology*, 52(7), 597–603. https://doi.org/10.1007/s12275-014-3547-9

Sebastià, C., Folch, J. M., Ballester, M., Estellé, J., Passols, M., Muñoz, M., García-Casco, J. M., Fernández, A. I., Castelló, A., Sánchez, A., & Crespo-Piazuelo, D. (2024). Interrelation between gut microbiota, SCFA, and fatty acid composition in pigs. *mSystems*, 9(1), e01049-23. https://doi.org/10.1128/msystems.01049-23

Setlow, B., Sun, D., & Setlow, P. (1992). Interaction between DNA and ox/,3-Type Small, Acid-Soluble Spore Proteins: A New Class of DNA-Binding Protein. *J. BACTERIOL.*, 174.

Setlow, P. (2006). Spores of Bacillus subtilis: Their resistance to and killing by radiation, heat and chemicals. *Journal of Applied Microbiology*, 101(3), 514–525. https://doi.org/10.1111/j.1365-2672.2005.02736.x

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. *Trends in Microbiology,* 15(4), 172–180. https://doi.org/10.1016/j.tim.2007.02.004

Slavica, A., Trontel, A., Jelovac, N., Kosovec, Ž., Šantek, B., & Novak, S. (2015). Production of lactate and acetate by Lactobacillus coryniformis subsp. Torquens DSM 20004 T in comparison with Lactobacillus amylovorus DSM 20531 T. *Journal of Biotechnology*, 202, 50–59. https://doi.org/10.1016/j.jbiotec.2015.01.014

Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal*, 474(11), 1823–1836. https://doi.org/10.1042/BCJ20160510

Uritskiy, G. V., DiRuggiero, J., & Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*, 6(1), 158. https://doi.org/10.1186/s40168-018-0541-1

Wang, C., Wei, S., Chen, N., Xiang, Y., Wang, Y., & Jin, M. (2022). Characteristics of gut microbiota in pigs with different breeds, growth periods and genders. *Microbial Biotechnology*, 15(3), 793–804. https://doi.org/10.1111/1751-7915.13755

Westreich, S. T., Treiber, M. L., Mills, D. A., Korf, I., & Lemay, D. G. (2018). SAMSA2: A standalone metatranscriptome analysis pipeline. *BMC Bioinformatics*, 19(1), 175. https://doi.org/10.1186/s12859-018-2189-z

Xiao, L., Estellé, J., Kiilerich, P., Ramayo-Caldas, Y., Xia, Z., Feng, Q., Liang, S., Pedersen, A. Ø., Kjeldsen, N. J., Liu, C., Maguin, E., Doré, J., Pons, N., Le Chatelier, E., Prifti, E., Li, J., Jia, H., Liu, X., Xu, X., … Wang, J. (2016). A reference gene catalogue of the pig gut microbiome. *Nature Microbiology*, 1(12), 16161. https://doi.org/10.1038/nmicrobiol.2016.161

Zha, A., Li, W., Wang, J., Bai, P., Qi, M., Liao, P., Tan, B., & Yin, Y. (2024). Trimethylamine oxide supplementation differentially regulates fat deposition in liver, longissimus dorsi muscle and adipose tissue of growing-finishing pigs. *Animal Nutrition*, 17, 25–35. https://doi.org/10.1016/j.aninu.2023.12.006

Zhang, L., Wu, W., Lee, Y.-K., Xie, J., & Zhang, H. (2018). Spatial Heterogeneity and Co-occurrence of Mucosal and Luminal Microbiome across Swine Intestinal Tract. *Frontiers in Microbiology*, 9, 48. https://doi.org/10.3389/fmicb.2018.00048

Zhang, Y., Thompson, K. N., Branck, T., Yan Yan, Nguyen, L. H., Franzosa, E. A., & Huttenhower, C. (2021). Metatranscriptomics for the Human Microbiome and Microbial Community Functional Profiling. *Annual Review of Biomedical Data Science*, 4(1), 279–311. https://doi.org/10.1146/annurev-biodatasci-031121-103035

Zhao, W., Wang, Y., Liu, S., Huang, J., Zhai, Z., He, C., Ding, J., Wang, J., Wang, H., Fan, W., Zhao, J., & Meng, H. (2015). The dynamic distribution of porcine microbiota across different ages and gastrointestinal tract segments. *PLOS ONE*, 10(2), e0117441. https://doi.org/10.1371/journal.pone.0117441

Zhu, Z., Zhu, L., & Jiang, L. (2022). Dynamic regulation of gut Clostridium-derived short-chain fatty acids. *Trends in Biotechnology,* 40(3), 266–270. https://doi.org/10.1016/j.tibtech.2021.10.005

Paper IV

Gut microbial genera as predictors of liver gene expression in pigs

Sebastià, C., Crespo-Piazuelo, D., Ramayo-Caldas, Y., Ballester, M., Liu, J., Llobet-Cabau, F., Castelló, A., and Folch, J.M.

(in preparation)

Gut microbial genera as predictors of liver gene expression in pigs

Sebastià, C., Crespo-Piazuelo, D., Ramayo-Caldas, Y., Liu, J., Llobet-Cabau, F., Castelló, A., Ballester, M., Sánchez, A., Folch, J.M.

Abstract

Meat quality in pigs is highly influenced by the fatty acid content and profile, where the liver has a key role. In turn, the gut microbiota has an impact in the fatty acid composition and maintains a close interrelation with immunity, given the negative effects of dysbiosis and disease on growth and carcass performance. In this study, we aimed to identify potential microbial predictors of gene expression in the liver at the time of slaughter. We found several bacterial genera such as *Caccousia*, *Eubacterium*, *Floccifex*, *Mediterraneibacter*, *SFLA01*, and *Streptococcus*, which may influence the expression of key genes related with liver metabolism and functions of genes such as *ABCD3*, *CXCL8*, *CYP8B1*, *DEFB1*, *ELOVL2*, *FADS2*, *HMGCS2*, *MZB1* and *SCD*. The validation of these relationships holds potential to modulate these genes through the microbiota.

Introduction

The pig is one of the most important livestock species, as it provides one of the most widely consumed meats worldwide (FAOSTAT, 2023). Meat quality is highly influenced by the fatty acid (FA) amount and profile in subcutaneous fat, and especially in intramuscular fat (Pietruszka et al., 2015; Wood et al., 2008). Differences in saturated FAs (SFAs), monounsaturated FAs (MUFAs) and polyunsaturated FAs (PUFAs) can affect the organoleptic properties of meat, and have a significant impact on the health of the consumer. For example, meats with higher content in SFA are related with cardiovascular diseases, while MUFAs and PUFAs are known for increasing palatability and act as cardioprotectors (Hammad et al., 2016). In addition, PUFAs of the omega-6 type are pro-inflammatory, while omega-3 PUFAs have anti-inflammatory properties (Marion-Letellier et al., 2015).

Liver is an important organ for energy and fatty acid metabolism. While the primary site for de novo FA biosynthesis in pigs is the adipose tissue (Corominas et al., 2013; O'Hea & Leveille, 1969), the liver plays a key role in the synthesis of long-chain PUFA from the essential omega-6 and omega-3 PUFAs (linoleic and α-linolenic acids, respectively) (Smink et al., 2012). By analysing the liver transcriptome, gene activity can be assessed, as well as its potential influence in lipid metabolism (Freeman et al., 2012; Summers et al., 2020). High-throughput sequencing methods, such as RNA-Seq, have been instrumental in mapping and quantifying the transcriptome for this purpose (Z. Wang et al., 2009).

On the other hand, the gut microbiota is considered another "organ" of the host due to its complex and important relationship with it. The microbiota composition is largely influenced by environmental factors, especially by diet (Frese et al., 2015; Heinritz et al., 2016), but is also influenced by host genetics (Crespo-Piazuelo et al., 2019). In turn, the gut microbiota can modulate gene expression and metabolic pathways in the host, influencing growth efficiency and meat production (Maltecca et al., 2020; Ramayo-Caldas et al., 2016). From birth, gut microbiota suffers drastic changes, particularly during critical periods of animal growth such as weaning, when animals are more susceptible to bacterial dysbiosis and diarrhoea (Guevarra et al., 2018). After weaning, *Prevotella* spp. are one of the predominant genera in the gut (Chen et al., 2017), reducing its presence as the animal grows, yet remaining as part of the microbial core (Luo et al., 2022). The analysis of the 16S rRNA gene via high throughput sequencing allows for the identification of prokaryotic species through the analysis of their hyper-variable regions that differ at species level (Kim et al., 2011). In research, rectal or faecal microbiota is among the most extensively studied tissues due to its ease of collection and its ability to provide an informative overview of the gut population.

The main goal of this study was to analyse the microbiota of the pig gut at an early stage of life and integrate it with the liver transcriptome at the moment of slaughter in order to find plausible bacterial predictors of the active genes in the liver related with immunity, energy or lipid metabolism.

Materials and methods

Animal material

A total of 287 Duroc pigs from a commercial population (144 females and 143 males) were used in this study. The animals were raised under intensive conditions and fed a cereal-based diet, with *ad libitum* access to food and water. Faeces samples were taken when the piglets were 60 ± 8 days old, and animals were slaughtered at 181 – 228 days of age in a commercial abattoir, where samples from liver were collected with RNAlater (ThermoFisher, Spain) and snap-frozen at -80 ºC immediately after slaughter until further use.

Microbial classification

Faecal DNA was extracted with the DNeasy PowerSoil Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions, as detailed in (Ramayo-Caldas et al., 2020). The V3-V4 rRNA gene sequences were processed into amplicon sequence variants (ASVs) at 99% of identity with QIIME2 software version 2021.11.0 (Bolyen et al., 2019). ASVs with an abundance lower than 0.005% of the total counts were removed. The Greengenes2 database (McDonald et al., 2023) was used to classify the ASVs at genus level. We retained only those genera present in more than 60% of the samples, resulting in a total of 147 genera.

Liver transcriptome analysis

Liver samples were homogenised with the Precellys Tissue Homogenizer and bead tubes (Lysing matrix D) and the RNA was extracted using a Maxwell RSC instrument with its Maxwell RSC simplyRNA Tissue Kit. Then, RNA quantification was carried out in a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). In addition, RNA purity and integrity were checked in a Bioanalyzer-2100 (Agilent Technologies, Santa Clara, CA, USA), ensuring that only RNA with a RIN higher than 7 was used. The libraries were prepared with the TruSeq Stranded mRNA Sample Preparation kit (Illumina, San Diego, CA, USA), which were constructed with an insert size of 285 bp. The samples were paired-end sequenced (2 × 150 bp) in an Illumina NovaSeq6000 platform at the Centro Nacional de Análisis Genómico (CNAG, Barcelona, Spain), with a sequencing depth of >45 million of paired-end reads.

Raw reads were quality checked with FastQC (Andrews, 2010). The sequences were mapped against the pig reference genome *Sscrofa11.1* and the Ensembl Genes 107 annotation database with STAR v2.5.3a (Dobin et al., 2013), and the counts were quantified with RSEM v1.3.0 (Li & Dewey, 2011). From the 35,670 obtained genes, those expressed in less than the 80% of the samples were filtered out, resulting in 14,594 genes being retained. A logCPM transformation was performed on the data, which was then corrected for sex (2 levels: male or female) and slaughter batch (6 levels) variables.

Data integration

Pearson correlation analysis were performed between the expression of the selected genes and the relative abundance of the bacterial and archaeal genera. To refine these results and avoid spurious relationships, mixOmics package v6.26.0 (Lê Cao et al., 2016; Rohart et al., 2017) from R (R Core Team, 2022) was used to integrate both datasets. Due to the large amount of data, we selected the 5% most variable genes and the 147 filtered genera for practical reasons, and then conducted a Regularised Canonical Correlation Analysis (rCCA) to explore the results. The network of the results was filtered for a weight of the connections > |0.15| and exported to Cytoscape (Shannon et al., 2003) for visualisation.

We used the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v2023q4 (Huang et al., 2009; Sherman et al., 2022) to infer the functions of the genes and filter the network of the rCCA based on the genes that belonged to the lipid metabolism, energy metabolism, and immunity. The information regarding these genes can be found in Supplementary Table S1.

Results

The taxonomic classification showed 18 bacterial and archaeal phyla, but the four most abundant accounted for more than 97% of the total phyla. Firmicutes and Bacteroidota were by far the most abundant phyla, with 49.83% and 42.72% of the total relative abundance, while Spirochaetota and Proteobacteria had 3.47% and 1.17% of the relative abundance. The remaining phyla had a relative abundance lower than 1%. Despite being Firmicutes the most abundant phylum, the most abundant genus was the Bacteroidota *Prevotella*, with 27.88% of the total relative abundance. *Lactobacillus*, an unassigned Lachnospiraceae and *Cryptobacteroides* followed in the highest relative abundances, with 6.08%, 6.02% and 5.41%. Table 1 shows the relative abundance of the genera with a relative abundance > 1%.

Table 1. Relative abundance (%) of the main genera found and the phylum to which they belong. Genera with a lower relative abundance than 1% were grouped as “Others”.

Genus	Phylum	Relative abundance
Prevotella	Bacteroidota	27.88
Lactobacillus	Firmicutes	6.08
Unassigned Lachnospiraceae	Firmicutes	6.02
Cryptobacteroides	Bacteroidota	5.41
Faecalibacterium	Firmicutes	2.90
Gemmiger	Firmicutes	2.52
Treponema	Spirochaetota	2.48
Paraprevotella	Bacteroidota	2.17
Clostridium	Firmicutes	2.16
Limosilactobacillus	Firmicutes	2.07
Sodaliphilus	Bacteroidota	2.04
Unassigned Butyricicoccaceae	Firmicutes	1.66
Blautia	Firmicutes	1.55
Anaerovibrio	Firmicutes	1.53
Terrisporobacter	Firmicutes	1.41
Alloprevotella	Bacteroidota	1.25
Others	-	30.87

Pearson correlation analysis found the strongest positive correlation between the abundance of *Cryptobacteroides* and the liver expression of a novel annotated gene, ENSSSCG00000018063 (ρ = 0.39), a mitochondrial ribosomal RNA gene. The strongest negative correlation was found between *SFLA01* and ENSSSCG00000048233 (ρ = −0.37), a lncRNA located in *Sus scrofa* chromosome (SSC)1.

The abundance of the bacterial genera *SFLA01* and *UBA6398* showed the highest number of significant correlations with the liver expression of different genes, with 39 and 19 correlations, respectively. Fifteen other genera demonstrated at least one significant correlation with an expressed gene (Fig. 1).

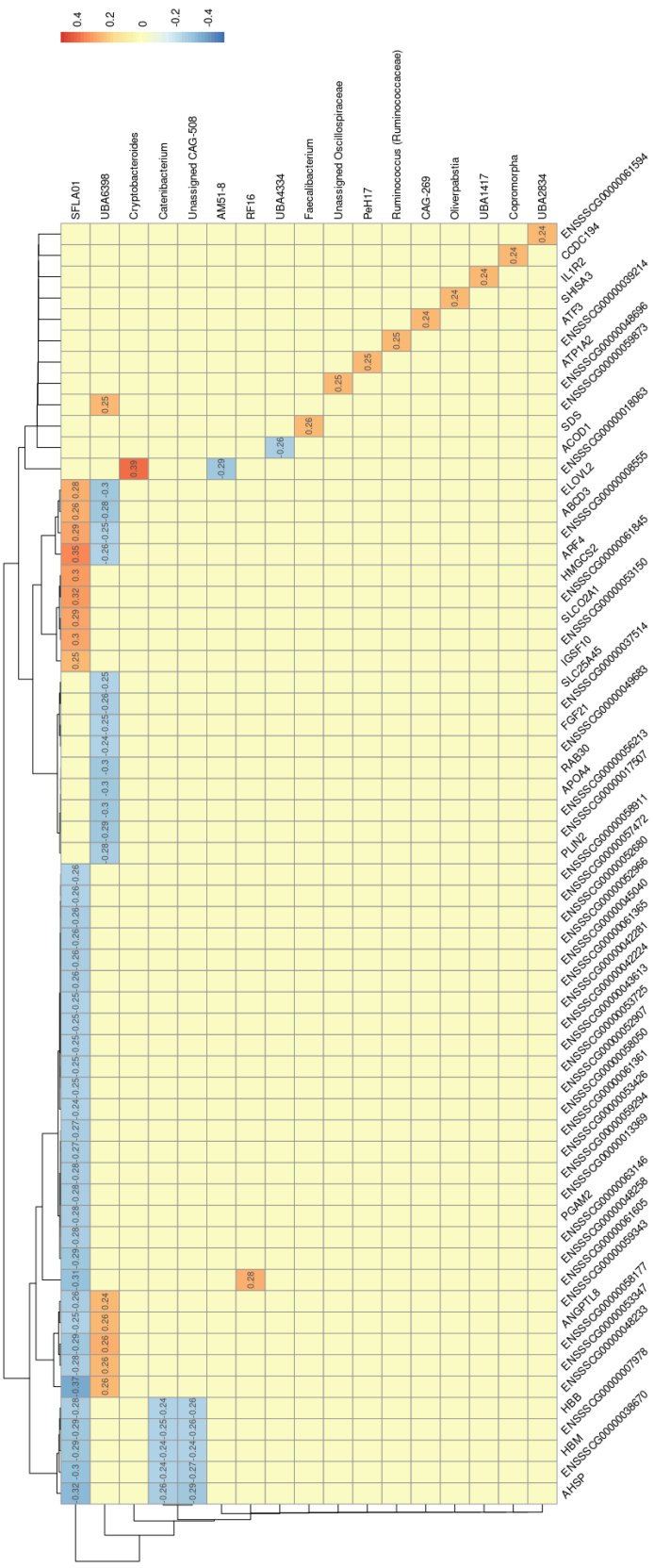


Fig. 1. Pearson correlation results. Significance threshold was set at an adjusted p-value < 0.05. For visualisation purposes, we only included the gene expressions and genera abundances that had at least one significant correlation.

In rCCA analysis, the expression of the mitochondrial rRNA gene ENSSSCG00000018063 also had the highest number of connections (27 genera). These results can be found in Supplementary Table S2. When filtering for relevant genes of liver-related pathways, keeping genes related with immunity, energy or lipid metabolism, the expression of *HMGCS2* had the highest number of connections (13 genera). Other relevant genes from liver-related pathways were *DEFB1* (10 genera), *ELOVL2* (10 genera), *SCD* (10 genera), *MZB1* (8 genera), *FADS2* (7 genera), *ABCD3* (6 genera), and *CYP8B1* (6 genera). These results can be found in Table 2.

Table 2. List of rCCA relevant genes and the genera with which they are connected.

Gene	Related genera
ABCD3	*Agathobaculum*, *Copromorpha*, *Fimousia*, *SFDB01*, *SFLA01*, Unassigned CAG-508
AGXT	*Eubacterium*, *Mediterraneibacter*, *SFLA01*
ANGPTL8	*Eubacterium*, *Mediterraneibacter*, *SFLA01*
APOA4	*SFLA01*
COL12A1	*SFLA01*
CXCL8	*Floccifex*
CYP7A1	*SFLA01*
CYP8B1	*SFDB01*, *SFLA01*, Unassigned CAG-508
DEFB1	*Caccousia*, *Caproiciproducens*, *Ellagibacter*, *Eubacterium*, *Floccifex*, *Mediterraneibacter*, *Ornithospirochaeta*, *Sutterella*, *UBA1259*, Unassigned Peptostreptococcaceae
DIO2	*Caccousia*, *Floccifex*, *Ornithospirochaeta*, *Streptococcus*
DISP3	*SFLA01*
ELOVL2	*Agathobaculum*, *CAG-349*, *Catenibacterium*, *Collinsella*, *Copromorpha*, *Fimousia*, *SFDB01*, *SFLA01*, Unassigned CAG-508, Unassigned UBA932
ENSSSCG00000031037	*Floccifex*
ENSSSCG00000036224	*Caccousia*, *Eubacterium*, *Floccifex*, *Mediterraneibacter*
ENSSSCG00000038719	*CAG-533*, *Floccifex*, *Streptococcus*, Unassigned Gastranaerophilaceae
ENSSSCG00000041880	*Floccifex*
ENSSSCG00000058639	*SFLA01*
ENSSSCG00000060584	*Caccousia*, *Floccifex*, *Streptococcus*

Gene	Related genera
ENSSSCG00000061605	*SFLA01*
FADS2	*CAG-349*, *CAG-533*, *Floccifex*, *Onthomonas*, *Streptococcus*, *UBA1259*, Unassigned Gastranaerophilaceae
HMGCS2	*Agathobaculum*, *Collinsella*, *Eubacterium*, *Helicobacter*, *Limivivens*, *Mediterraneibacter*, *Paludicola*, *Phascolarctobacterium*, *SFDB01*, *SFLA01*, *UBA636*, Unassigned CAG-508, Unassigned Erysipelotrichaceae
KIF1A	*Eubacterium*
LEPR	*SFLA01*
LIF	*SFLA01*
MFSD2A	*SFLA01*
MZB1	*Caccousia*, *Caproiciproducens*, *Floccifex*, *Ornithospirochaeta*, *Streptococcus*, *UBA1259*, Unassigned Gastranaerophilaceae, Unassigned Peptostreptococcaceae
PGAM2	*SFLA01*
PLIN2	*SFLA01*
PTGS2	*Eubacterium*, *SFLA01*
SCD	*Caccousia*, *Caproiciproducens*, *Floccifex*, *Hallerella*, *Ornithospirochaeta*, *Streptococcus*, *Sutterella*, *UBA1259*, Unassigned Gastranaerophilaceae, Unassigned Peptostreptococcaceae
THBS1	*Eubacterium*, *SFLA01*
TSKU	*Eubacterium*, *Mediterraneibacter*, *SFLA01*

The rCCA analysis also identified *SFLA01* as the most relevant bacterial genus. Its abundance showed the strongest correlation with the first component ($\rho = -0.66$), as seen in Fig. 2, and also had the highest number of associations with gene expression (20 genes), depicted in the network from Fig. 3. *CAG-349* and *Streptococcus* showed the strongest positive correlations with the first component ($\rho = 0.45$ and $\rho = 0.41$ respectively), while *Floccifex* had the strongest negative correlation with the second component ($\rho = -0.6$) and an unassigned Gastranaerophilaceae had the strongest positive correlation with the second component ($\rho = 0.47$). Besides *SFLA01*, other genera with a high number of connections in the rCCA network were *Floccifex* (11 genes), *Eubacterium* (10 genes), *Mediterraneibacter* (7 genes), *Caccousia* (6 genes), and *Streptococcus* (6 genes).

Figure 3 depicted *SFLA01* showing the strongest relationships, with the expression of the genes *ELOVL2*, *ABCD3* and *HMGCS2* (with a weight of 0.30, 0.29 and 0.28 respectively). The network also pointed out that the expression of *MZB1* and *DEFB1* had a high number of common relationships and opposite to the expression of *SCD*, with the genera *Floccifex*, *Caccousia*, *Ornithospirochaeta*, *UBA1259*, *Caproiciproducens*, *Sutterella* and an unassigned Peptostreptococcaceae. The expression of *DIO2* also showed a similar pattern of relationships than the expression of *SCD*, and the expressions of *THBS1* and *PTGS2*, and *ANGPTL8* and *TSKU* had the same relationships with genera.

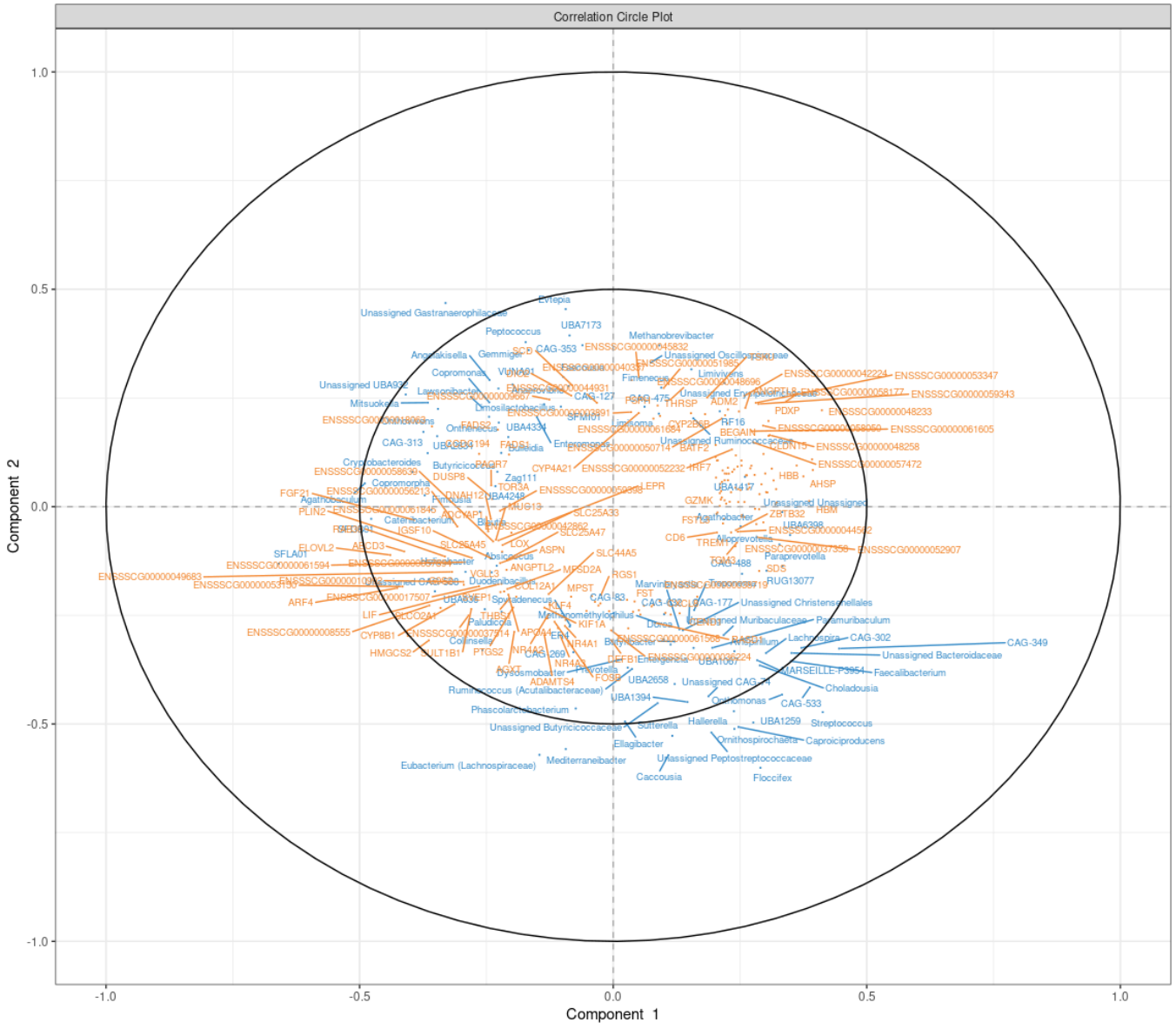


Fig. 2. Circular plot of the rCCA results. For visualisation purposes, we established a 0.2 cut-off to keep the more relevant genes (n = 203, represented in orange) and genera (n = 105, represented in blue).

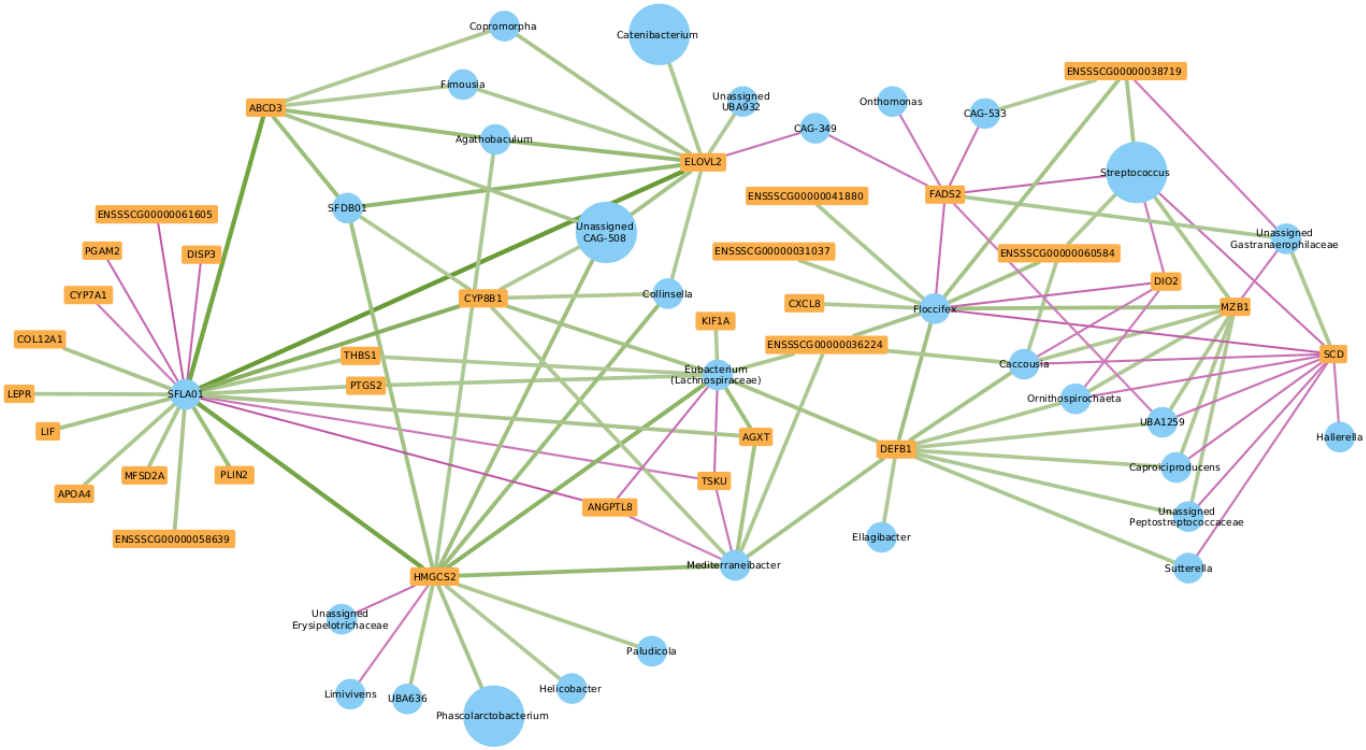


Fig. 3. Correlation network of the rCCA results. We filtered the expressed genes by lipid metabolism, energy metabolism and immunity based on the functional annotation analysis with DAVID and established a 0.15 cut-off to keep the more relevant genes (n = 32, represented as orange squared nodes) and genera (n = 31, represented as blue round nodes with a different size depending on their relative abundance). Green edges indicate positive relationships, while negative relationships are marked by magenta edges.

Discussion

Gut microbiota is known to have an impact on the host (Valdes et al., 2018). However, this impact is complex and influenced by other factors. While a basic correlation analysis may suffice for closely associated data, the two datasets used in this study required methods capable of excluding the spurious relationships to reveal the more relevant ones. Deepening the understanding of the microbial species that affect the gene expression in the host metabolism and overall health can help to promote or avoid them through the diet, thereby influencing the animal traits.

After the correlation analysis, the gene with the highest number of connections was *HMGCS2*. This gene has an important role in lipid metabolic processes, particularly in mitochondrial ketogenesis, as this synthase is a key rate-limiting enzyme in the process, which mainly takes place in hepatocytes (Hwang et al., 2022). *HMGCS2* converts acetyl-CoA to HMG-CoA, which is then used to produce ketone bodies (Cotter et al., 2013). Under fed conditions, ketone body levels were similar between microbiota-transplanted and germ-free mice, but during fasting, microbiota-transplanted mice increased hepatic ketogenesis production through an upregulation of the PPARA-HMGCS2 axis (Cotter et al., 2013; Crawford et al., 2009). In our results, *HMGCS2* expression was positively connected to nine genera, including SCFA-producing bacteria such as *Agathobaculum*, *Collinsella*, *Phascolarctobacterium*. The abundance of *Agathobaculum* was reduced in patients with non-alcoholic fatty liver disease and in patients with liver cirrhosis compared to those with chronic hepatitis (Ahn et al., 2019; Midori et al., 2024). *Collinsella* is considered a potential beneficial genus, improving host metabolism and colonic mucosal immunity, although its contribution is not fully understood. In humans, *Collinsella* is reduced with low dietary fibre intake and it is correlated to insulin levels, but an overgrowth may disturb gut integrity by increasing gut permeability (Gomez-Arango et al., 2018; Teng et al., 2023). Our group already reported the association between *Phascolarctobacterium* levels and the pig genome (Crespo-Piazuelo et al., 2019), but diet also modulates the abundance of this genus. Inulin intake promotes *Phascolarctobacterium* growth reducing hepatic steatosis by restoring the integrity of the intestinal barrier (Z. Yang et al., 2023)

The remaining genes with a substantial number of connections, *ABCD3*, *ELOVL2*, *FADS2*, and *SCD*, are known to be involved in fatty acid synthesis. In the correlation network, *ABCD3* and *ELOVL2* were both positively correlated with six genera, including the aforementioned Agathobaculum and *SFLA01*, a member of the Oscillospirales order. *ABCD3* gene is a peroxisomal transporter, it transports long-chain FAs into the peroxisomes for β-oxidation and it is also thought to transport bile acids and branched chain FAs (Imanaka et

al., 1999; Morita & Imanaka, 2012). *ELOVL2* catalyses the elongation of very long omega-6 and omega-3 PUFAs, which are related with inflammation. Oxylipins produced from omega-6 PUFA promote inflammation, whereas oxylipins derived from omega-3 inhibit it (Innes & Calder, 2018; Liput et al., 2021). Both of the analyses conducted in this study showed the *SFLA01* genus as the more relevant one. It had the highest number of connections in both the Pearson and the regularised correlation analyses, including associations with *ABCD3*, *ELOVL2*, and *HMGCS2* expression. Although *SFLA01* is a relatively new genus, some bacteria from the same order, such as *Oscillospira*, are closely linked to obesity and chronic inflammatory and metabolic diseases (J. Yang et al., 2021). Thus, apart from the ones already described, *SFLA01* was correlated with genes related to lipid metabolism, such as *APOA4*, *CYP7A1*, *MFSD2A*, and *PLIN2*. Notably, *APOA4* and *PLIN2* also showed negative Pearson correlations with the *UBA6398* genus, a member of the Bacteroidaceae family.

It is worth noting that the two desaturases found, *FADS2* and *SCD*, had both negative correlations with *Floccifex*, *Streptococcus* and *UBA1259* (Borkfalkiaceae family). *FADS2* is a Δ6 desaturase that transforms the essential fatty acids from the diet, linoleic (C18:2n-6) and α-linolenic acids (C18:3n-3), into γ-linolenic (C18:3n-6) and stearidonic (C18:4n-3) acids respectively. These fatty acids are transformed in long-chain PUFA arachidonic acid (20:4n-6) and dihomo-γ-linolenic acid (20:3n-6), which are precursor molecules capable of influencing the immune system and inflammatory responses (Liput et al., 2021). However, the role of *SCD* is more indirect; *SCD* is a Δ9 desaturase able to convert SFA to MUFA such as palmitic acid (C16:0) to palmitoleic acid (C16:1n-7) or stearic acid (C18:0) to oleic acid (C18:1n-9) (Bond et al., 2016). In TLR5-deficient mice, elevated levels of oleic acid in the liver were attributed to an upregulation of hepatic *SCD*, partially driven by an increased production of SCFAs by their gut microbiota (Singh et al., 2015).

The rCCA analysis also reported some genes worth discussing further, *CXCL8*, *CYP8B1*, *DEFB1*, *MZB1*, and *ZDHHC19*. The protein encoded by *CXCL8*, also referred as interleukin-8 (IL-8), participates in immune response. In a study conducted on CXCL8-knockout porcine cell lines, they exhibited significantly increased resistance to *Pasteurella multocida* toxin (Yuan et al., 2024). *CYP8B1*, another highly correlated gene, is involved in the bile acid synthesis. It encodes for the enzyme sterol 12α-hydrolase, which is key for the cholic acid synthesis. In mice, the lack of this enzyme leads to a resistance to diabetes and to hepatic steatosis, but also produces a change in the gut microbiota (Patankar et al., 2018). However, adult pigs lack cholic acid and produce hyocholic acid instead, it is only found in pig foetus and it was considered that *CYP8B1* gene was mainly expressed before birth (Lundell & Wikvall, 2008). However, in recent studies the expression of *CYP8B1* in weaned piglets has been found increased in diseased livers (Maj et al., 2023;

H. Wang et al., 2021). In addition, *Lactobacillus gasseri* has been found to activate the hepatic bile acid biosynthesis in germ-free mice, stimulating the production of proteins encoded by genes such as *CYP8B1* (Hu et al., 2023). Nevertheless, this bacterium was not found in our data, so we hypothesise that other lactobacilli could activate *CYP8B1*. On the opposite side of *CYP8B1* in the rCCA, there was a lncRNA gene named ENSSSCG00000051985. This gene is located adjacent to the genomic region of the *ZDHHC19* gene, and the whole genomic region (5 Mb) has been found associated with *E. coli* F4ab/ac susceptibility in pigs (Jacobsen et al., 2010).

DEFB1 and *MZB1* were the other highly neighbouring genes, they are related with mucosal immunity. *DEFB1* encodes for an antimicrobial peptide (Polesello et al., 2015), while *MZB1* plays a role in IgM and IgA secretion (Suzuki et al., 2019; H. Wei & Wang, 2021). As a β-defensin, *DEFB1* participates in innate immunity producing an extracellular bacteria-entrapping net that protects against bacterial transmigration and invasion (Raschig et al., 2017). In pigs, *DEFB1* confers protection against *Bordetella pertussis*, although experimental intrapulmonary infection is typically required since this pathogen is exclusively a human pathogen (Elahi et al., 2006). In MZB1-deficient mice, inefficient IgA secretion is associated with increased levels of Proteobacteria and Verrucomicrobia phyla, and a reduction in Bacteroidetes (Suzuki et al., 2019). In addition, IgA also maintains gut homeostasis and promotes symbiosis between *Bacteroides*, Firmicutes and Clostridiales, as *Bacteroides* serve as anchors in the mucin layer, facilitating interactions with other bacteria (Nakajima et al., 2018).

Overall, there are many different genes expressed in the liver at the moment of slaughter that are correlated to microbial genera when the pig is younger. However, there are various factors involved in modulating both liver gene expression and the composition of the microbiota. For instance, the genetic differences of the animals influence both liver gene expression (C. Wei et al., 2023) and the microbiota (Ramayo-Caldas et al., 2023). Variations in the microbiota, in turn, impact the animal's energy metabolism and immunity, affecting liver gene expression at the time of slaughter (Peng et al., 2021; H. Wang et al., 2020; Xie et al., 2022). Additionally, the variability of these factors influences liver gene expression both in young animals and at the age of slaughter, which in turn affects the microbiota (Ayuso et al., 2016). For these reasons it is necessary to validate the relationships we found in this work, which will allow us to modulate the desired liver gene expression through the shaping of the microbiota with the diet or to partially predict the behaviour of the microbiota based on the genetic variations of the animal.

Conclusions

Several gut microbial genera had significant associations with the expression levels of genes in the liver. *SFLA01* and *Floccifex* were the most connected genera with diverse genes involved in lipid metabolism or immunity, while *DEFB1*, *HMGCS2*, *ELOVL2* and *SCD* were the most connected genes, showing interactions with different bacterial genera. Further in-depth analyses are required to examine these relationships and validate the different genera as potential modulators of gene expression in the liver.

References

Ahn, S. B., Jun, D. W., Kang, B.-K., Lim, J. H., Lim, S., & Chung, M.-J. (2019). Randomized, Double-blind, Placebo-controlled Study of a Multispecies Probiotic Mixture in Nonalcoholic Fatty Liver Disease. *Scientific Reports*, 9(1), 5688. https://doi.org/10.1038/s41598-019-42059-3

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Ayuso, M., Fernández, A., Núñez, Y., Benítez, R., Isabel, B., Fernández, A. I., Rey, A. I., González-Bulnes, A., Medrano, J. F., Cánovas, Á., López-Bote, C. J., & Óvilo, C. (2016). Developmental Stage, Muscle and Genetic Type Modify Muscle Transcriptome in Pigs: Effects on Gene Expression and Regulatory Factors Involved in Growth and Metabolism. *PLOS ONE*, 11(12), e0167858. https://doi.org/10.1371/journal.pone.0167858

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bond, L. M., Miyazaki, M., O'Neill, L. M., Ding, F., & Ntambi, J. M. (2016). Fatty Acid Desaturation and Elongation in Mammals. *In Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 185–208). Elsevier. https://doi.org/10.1016/B978-0-444-63438-2.00006-7

Chen, L., Xu, Y., Chen, X., Fang, C., Zhao, L., & Chen, F. (2017). The Maturing Development of Gut Microbiota in Commercial Piglets during the Weaning Transition. *Frontiers in Microbiology*, 8, 1688. https://doi.org/10.3389/fmicb.2017.01688

Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Estellé, J., Castelló, A., Alves, E., Pena, R. N., Ballester, M., & 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. https://doi.org/10.1186/1471-2164-14-843

Cotter, D. G., Schugar, R. C., & Crawford, P. A. (2013). Ketone body metabolism and cardiovascular disease. *American Journal of Physiology-Heart and Circulatory Physiology*, 304(8), H1060–H1076. https://doi.org/10.1152/ajpheart.00646.2012

Crawford, P. A., Crowley, J. R., Sambandam, N., Muegge, B. D., Costello, E. K., Hamady, M., Knight, R., & Gordon, J. I. (2009). Regulation of myocardial ketone body metabolism by the gut microbiota during nutrient deprivation. *Proceedings of the National Academy of Sciences*, 106(27), 11276–11281. https://doi.org/10.1073/pnas.0902366106

Crespo-Piazuelo, D., Migura-Garcia, L., Estellé, J., Criado-Mesas, L., Revilla, M., Castelló, A., Muñoz, M., García-Casco, J. M., Fernández, A. I., Ballester, M., & Folch, J. M. (2019). Association between the pig genome and its gut microbiota composition. *Scientific Reports*, 9(1), 8791. https://doi.org/10.1038/s41598-019-45066-6

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Elahi, S., Buchanan, R. M., Attah-Poku, S., Townsend, H. G. G., Babiuk, L. A., & Gerdts, V. (2006). The Host Defense Peptide Beta-Defensin 1 Confers Protection against Bordetella pertussis in Newborn Piglets. *Infection and Immunity*, 74(4), 2338–2352. https://doi.org/10.1128/IAI.74.4.2338-2352.2006

FAOSTAT. (2023). FAOSTAT. FAOSTAT. https://www.fao.org/faostat/en/#home

Freeman, T. C., Ivens, A., Baillie, J. K., Beraldi, D., Barnett, M. W., Dorward, D., Downing, A., Fairbairn, L., Kapetanovic, R., Raza, S., Tomoiu, A., Alberio, R., Wu, C., Su, A. I., Summers, K. M., Tuggle, C. K., Archibald, A. L., & Hume, D. A. (2012). A gene expression atlas of the domestic pig. *BMC Biology*, 10(1), 90. https://doi.org/10.1186/1741-7007-10-90

Frese, S. A., Parker, K., Calvert, C. C., & Mills, D. A. (2015). Diet shapes the gut microbiome of pigs during nursing and weaning. *Microbiome*, 3(1), 28. https://doi.org/10.1186/s40168-015-0091-8

Gomez-Arango, L. F., Barrett, H. L., Wilkinson, S. A., Callaway, L. K., McIntyre, H. D., Morrison, M., & Dekker Nitert, M. (2018). Low dietary fiber intake increases Collinsella abundance in the gut microbiota of overweight and obese pregnant

women. *Gut Microbes*, 9(3), 189–201. https://doi.org/10.1080/19490976.2017.1406584

Guevarra, R. B., Hong, S. H., Cho, J. H., Kim, B.-R., Shin, J., Lee, J. H., Kang, B. N., Kim, Y. H., Wattanaphansak, S., Isaacson, R. E., Song, M., & Kim, H. B. (2018). The dynamics of the piglet gut microbiome during the weaning transition in association with health and nutrition. *Journal of Animal Science and Biotechnology*, 9(1), 54. https://doi.org/10.1186/s40104-018-0269-6

Hammad, S., Pu, S., & Jones, P. J. (2016). Current Evidence Supporting the Link Between Dietary Fatty Acids and Cardiovascular Disease. *Lipids*, 51(5), 507–517. https://doi.org/10.1007/s11745-015-4113-x

Heinritz, S., Weiss, E., Eklund, M., Aumiller, T., Heyer, C., Messner, S., Rings, A., Louis, S., Bischoff, S., & Mosenthin, R. (2016). Impact of a High-Fat or High-Fiber Diet on Intestinal Microbiota and Metabolic Markers in a Pig Model. *Nutrients*, 8(5), 317. https://doi.org/10.3390/nu8050317

Hu, J., Hou, Q., Zheng, W., Yang, T., & Yan, X. (2023). Lactobacillus gasseri LA39 promotes hepatic primary bile acid biosynthesis and intestinal secondary bile acid biotransformation. *Journal of Zhejiang University-SCIENCE B*, 24(8), 734–748. https://doi.org/10.1631/jzus.B2200439

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Hwang, C. Y., Choe, W., Yoon, K.-S., Ha, J., Kim, S. S., Yeo, E.-J., & Kang, I. (2022). Molecular Mechanisms for Ketone Body Metabolism, Signaling Functions, and Therapeutic Potential in Cancer. *Nutrients*, 14(22), 4932. https://doi.org/10.3390/nu14224932

Imanaka, T., Aihara, K., Takano, T., Yamashita, A., Sato, R., Suzuki, Y., Yokota, S., & Osumi, T. (1999). Characterization of the 70-kDa Peroxisomal Membrane Protein, an ATP Binding Cassette Transporter. *Journal of Biological Chemistry*, 274(17), 11968–11976. https://doi.org/10.1074/jbc.274.17.11968

Innes, J. K., & Calder, P. C. (2018). Omega-6 fatty acids and inflammation. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 132, 41–48. https://doi.org/10.1016/j.plefa.2018.03.004

Jacobsen, M., Kracht, S. S., Esteso, G., Cirera, S., Edfors, I., Archibald, A. L., Bendixen, C., Andersson, L., Fredholm, M., & Jørgensen, C. B. (2010). Refined candidate region specified by haplotype sharing for Escherichia coli F4ab/F4ac susceptibility alleles in pigs. *Animal Genetics*, 41(1), 21–25. https://doi.org/10.1111/j.1365-2052.2009.01952.x

Kim, M., Morrison, M., & Yu, Z. (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *Journal of Microbiological Methods*, 84(1), 81–87. https://doi.org/10.1016/j.mimet.2010.10.020

Lê Cao, K.-A., Rohart, F., González, I., Déjean, S., Gautier, B., Bartolo, F., Monget, P., Coquery, J., Yao, F., & Liquet, B. (2016). mixOmics: Omics Data Integration Project (6.1.1) [Computer software]. http://mixomics.org/

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12, 1–16. https://doi.org/10.1186/1471-2105-12-323

Liput, K. P., Lepczyński, A., Ogłuszka, M., Nawrocka, A., Poławska, E., Grzesiak, A., Ślaska, B., Pareek, C. S., Czarnik, U., & Pierzchała, M. (2021). Effects of Dietary n–3 and n–6 Polyunsaturated Fatty Acids in Inflammation and Cancerogenesis. *International Journal of Molecular Sciences*, 22(13), 6965. https://doi.org/10.3390/ijms22136965

Lundell, K., & Wikvall, K. (2008). Species-Specific and Age-Dependent Bile Acid Composition: Aspects on CYP8B and CYP4A Subfamilies in Bile Acid Biosynthesis. *Current Drug Metabolism*, 9(4), 323–331. https://doi.org/10.2174/138920008784220574

Luo, Y., Ren, W., Smidt, H., Wright, A.-D. G., Yu, B., Schyns, G., McCormack, U. M., Cowieson, A. J., Yu, J., He, J., Yan, H., Wu, J., Mackie, R. I., & Chen, D. (2022). Dynamic Distribution of Gut Microbiota in Pigs at Different Growth Stages: Composition and Contribution. *Microbiology Spectrum,* 10(3), e00688-21. https://doi.org/10.1128/spectrum.00688-21

Maj, M. A., Burrin, D. G., & Manjarín, R. (2023). Decreased FXR Agonism in the Bile Acid Pool Is Associated with Impaired FXR Signaling in a Pig Model of Pediatric NAFLD. *Biomedicines*, 11(12), 3303. https://doi.org/10.3390/biomedicines11123303

Maltecca, C., Bergamaschi, M., & Tiezzi, F. (2020). The interaction between microbiome and pig efficiency: A review. *Journal of Animal Breeding and Genetics*, 137(1), 4–13. https://doi.org/10.1111/jbg.12443

Marion-Letellier, R., Savoye, G., & Ghosh, S. (2015). Polyunsaturated fatty acids and inflammation. *IUBMB Life*, 67(9), 659–667. https://doi.org/10.1002/iub.1428

McDonald, D., Jiang, Y., Balaban, M., Cantrell, K., Zhu, Q., Gonzalez, A., Morton, J. T., Nicolaou, G., Parks, D. H., Karst, S. M., Albertsen, M., Hugenholtz, P., DeSantis, T., Song, S. J., Bartko, A., Havulinna, A. S., Jousilahti, P., Cheng, S., Inouye, M., … Knight, R. (2023). Greengenes2 unifies microbial data in a single reference tree. *Nature Biotechnology*. https://doi.org/10.1038/s41587-023-01845-1

Midori, Y., Nosaka, T., Hiramatsu, K., Akazawa, Y., Tanaka, T., Takahashi, K., Naito, T., Matsuda, H., Ohtani, M., & Nakamoto, Y. (2024). Isolation of mucosa-associated microbiota dysbiosis in the ascending colon in hepatitis C virus post-sustained virologic response cirrhotic patients. *Frontiers in Cellular and Infection Microbiology*, 14, 1371429. https://doi.org/10.3389/fcimb.2024.1371429

Morita, M., & Imanaka, T. (2012). Peroxisomal ABC transporters: Structure, function and role in disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1822(9), 1387–1396. https://doi.org/10.1016/j.bbadis.2012.02.009

Nakajima, A., Vogelzang, A., Maruya, M., Miyajima, M., Murata, M., Son, A., Kuwahara, T., Tsuruyama, T., Yamada, S., Matsuura, M., Nakase, H., Peterson, D. A., Fagarasan, S., & Suzuki, K. (2018). IgA regulates the composition and metabolic function of gut microbiota by promoting symbiosis between bacteria. *Journal of Experimental Medicine*, 215(8), 2019–2034. https://doi.org/10.1084/jem.20180427

O'Hea, E. K., & Leveille, G. A. (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. https://doi.org/10.1093/jn/99.3.338

Patankar, J. V., Wong, C. K., Morampudi, V., Gibson, W. T., Vallance, B., Ioannou, G. N., & Hayden, M. R. (2018). Genetic ablation of Cyp8b1 preserves host metabolic function by repressing steatohepatitis and altering gut microbiota composition. *American Journal of Physiology-Endocrinology and Metabolism*, 314(5), E418–E432. https://doi.org/10.1152/ajpendo.00172.2017

Peng, J., Tang, Y., & Huang, Y. (2021). Gut health: The results of microbial and mucosal immune interactions in pigs. *Animal Nutrition*, 7(2), 282–294. https://doi.org/10.1016/j.aninu.2021.01.001

Pietruszka, A., Jacyno, E., Kawęcka, M., & Biel, W. (2015). The Relation Between Intramuscular Fat Level in the Longissimus Muscle and the Quality of Pig Carcasses and Meat. *Annals of Animal Science*, 15(4), 1031–1041. https://doi.org/10.1515/aoas-2015-0046

Polesello, V., Zupin, L., Di Lenarda, R., Biasotto, M., Ottaviani, G., Gobbo, M., Cecco, L., Alberi, G., Pozzato, G., Crovella, S., & Segat, L. (2015). Impact of DEFB1 gene regulatory polymorphisms on hBD-1 salivary concentration. *Archives of Oral Biology*, 60(7), 1054–1058. https://doi.org/10.1016/j.archoralbio.2015.03.009

R Core Team. (2022). R: A language and environment for statistical computing (4.1.3) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Ramayo-Caldas, Y., Crespo-Piazuelo, D., Morata, J., González-Rodríguez, O., Sebastià, C., Castello, A., Dalmau, A., Ramos-Onsins, S., Alexiou, K. G., Folch, J. M., Quintanilla, R., & Ballester, M. (2023). Copy Number Variation on ABCC2-DNMBP Loci Affects the Diversity and Composition of the Fecal Microbiota in Pigs. *Microbiology Spectrum*, 11(4), e05271-22. https://doi.org/10.1128/spectrum.05271-22

Ramayo-Caldas, Y., Mach, N., Lepage, P., Levenez, F., Denis, C., Lemonnier, G., Leplat, J.-J., Billon, Y., Berri, M., Doré, J., Rogel-Gaillard, C., & Estellé, J. (2016). Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. *The ISME Journal*, 10(12), 2973–2977. https://doi.org/10.1038/ismej.2016.77

Ramayo-Caldas, Y., Prenafeta-Boldú, F., Zingaretti, L. M., Gonzalez-Rodriguez, O., Dalmau, A., Quintanilla, R., & Ballester, M. (2020). Gut eukaryotic communities in pigs: Diversity, composition and host genetics contribution. *Animal Microbiome*, 2(1), 18. https://doi.org/10.1186/s42523-020-00038-4

Raschig, J., Mailänder-Sánchez, D., Berscheid, A., Berger, J., Strömstedt, A. A., Courth, L. F., Malek, N. P., Brötz-Oesterhelt, H., & Wehkamp, J. (2017). Ubiquitously expressed Human Beta Defensin 1 (hBD1) forms bacteria-entrapping nets in a redox dependent mode of action. *PLOS Pathogens*, 13(3), e1006261. https://doi.org/10.1371/journal.ppat.1006261

Rohart, F., Gautier, B., Singh, A., & Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Computational Biology*, 13(11), e1005752. https://doi.org/10.1371/journal.pcbi.1005752

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, 50(W1), W216–W221. https://doi.org/10.1093/nar/gkac194

Singh, V., Chassaing, B., Zhang, L., San Yeoh, B., Xiao, X., Kumar, M., Baker, M. T., Cai, J., Walker, R., Borkowski, K., Harvatine, K. J., Singh, N., Shearer, G. C., Ntambi, J. M., Joe, B., Patterson, A. D., Gewirtz, A. T., & Vijay-Kumar, M. (2015). Microbiota-Dependent Hepatic Lipogenesis Mediated by Stearoyl CoA Desaturase 1 (SCD1) Promotes Metabolic Syndrome in TLR5-Deficient Mice. *Cell Metabolism*, 22(6), 983–996. https://doi.org/10.1016/j.cmet.2015.09.028

Smink, W., Gerrits, W. J. J., Gloaguen, M., Ruiter, A., & van Baal, J. (2012). Linoleic and α-linolenic acid as precursor and inhibitor for the synthesis of long-chain polyunsaturated fatty acids in liver and brain of growing pigs. *Animal*, 6(2), 262–270. https://doi.org/10.1017/S1751731111001479

Summers, K. M., Bush, S. J., Wu, C., Su, A. I., Muriuki, C., Clark, E. L., Finlayson, H. A., Eory, L., Waddell, L. A., Talbot, R., Archibald, A. L., & Hume, D. A. (2020). Functional Annotation of the Transcriptome of the Pig, Sus scrofa, Based Upon Network Analysis of an RNAseq Transcriptional Atlas. *Frontiers in Genetics*, 10, 1355. https://doi.org/10.3389/fgene.2019.01355

Suzuki, K., Vogelzang, A., & Fagarasan, S. (2019). MZB1 folding and unfolding the role of IgA. *Proceedings of the National Academy of Sciences*, 116(27), 13163–13165. https://doi.org/10.1073/pnas.1908012116

Teng, T., Sun, G., Ding, H., Song, X., Bai, G., Shi, B., & Shang, T. (2023). Characteristics of glucose and lipid metabolism and the interaction between gut microbiota and colonic mucosal immunity in pigs during cold exposure. *Journal of Animal Science and Biotechnology*, 14(1), 84. https://doi.org/10.1186/s40104-023-00886-5

Valdes, A. M., Walter, J., Segal, E., & Spector, T. D. (2018). Role of the gut microbiota in nutrition and health. *BMJ*, k2179. https://doi.org/10.1136/bmj.k2179

Wang, H., Han, Q., Chen, Y., Hu, G., & Xing, H. (2021). Novel insights into cytochrome P450 enzyme and solute carrier families in cadmium-induced liver injury of pigs. *Ecotoxicology and Environmental Safety*, 211, 111910. https://doi.org/10.1016/j.ecoenv.2021.111910

Wang, H., Xu, R., Zhang, H., Su, Y., & Zhu, W. (2020). Swine gut microbiota and its interaction with host nutrient metabolism. *Animal Nutrition*, 6(4), 410–420. https://doi.org/10.1016/j.aninu.2020.10.002

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1), 57–63. https://doi.org/10.1038/nrg2484

Wei, C., Zeng, H., Zhong, Z., Cai, X., Teng, J., Liu, Y., Zhao, Y., Wu, X., Li, J., & Zhang, Z. (2023). Integration of non-additive genome-wide association study with a multi-tissue transcriptome analysis of growth and carcass traits in Duroc pigs. *Animal*, 17(6), 100817. https://doi.org/10.1016/j.animal.2023.100817

Wei, H., & Wang, J.-Y. (2021). Role of Polymeric Immunoglobulin Receptor in IgA and IgM Transcytosis. *International Journal of Molecular Sciences*, 22(5), 2284. https://doi.org/10.3390/ijms22052284

Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Sheard, P. R., Richardson, R. I., Hughes, S. I., & Whittington, F. M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Science,* 78(4), 343–358. https://doi.org/10.1016/j.meatsci.2007.07.019

Xie, C., Zhu, X., Xu, B., Niu, Y., Zhang, X., Ma, L., & Yan, X. (2022). Integrated analysis of multi-tissues lipidome and gut microbiome reveals microbiota-induced shifts on lipid metabolism in pigs. *Animal Nutrition*, 10, 280–293. https://doi.org/10.1016/j.aninu.2022.04.011

Yang, J., Li, Y., Wen, Z., Liu, W., Meng, L., & Huang, H. (2021). Oscillospira—A candidate for the next-generation probiotics. *Gut Microbes*, 13(1), 1987783. https://doi.org/10.1080/19490976.2021.1987783

Yang, Z., Su, H., Lv, Y., Tao, H., Jiang, Y., Ni, Z., Peng, L., & Chen, X. (2023). Inulin intervention attenuates hepatic steatosis in rats via modulating gut microbiota and

maintaining intestinal barrier function. *Food Research International*, 163, 112309. https://doi.org/10.1016/j.foodres.2022.112309

Yuan, J., Zhao, Q., Li, J., Wen, Y., Wu, R., Zhao, S., Lang, Y.-F., Yan, Q.-G., Huang, X., Du, S., & Cao, S.-J. (2024). CXCL8 Knockout: A Key to Resisting Pasteurella multocida Toxin-Induced Cytotoxicity. *International Journal of Molecular Sciences*, 25(10), 5330. https://doi.org/10.3390/ijms25105330

4. General discussion

The pig is closely linked to humans as one of the most important livestock species and, more recently, as a key animal in research. Its physiology and gut microbiome, similar to those of humans, make the pig a valuable model species (Heinritz et al., 2013). Through domestication, pigs have undergone significant changes and adaptations to diverse environments. These adaptations have been the consequence of selective breeding aimed at optimising traits in accordance with consumer preferences across different times and regions, leaving distinct signatures of selection in the pig genome (Bovo et al., 2020).

As a livestock species, one of the main foci of selective breeding has been on enhancing pork quality. Meat quality is a complex trait influenced by various factors, among which fatty acid composition plays a crucial role. The balance of saturated and unsaturated fatty acids significantly affects the flavour, tenderness, and overall nutritional value of the meat (J. D. Wood et al., 2004). In this regard, diet is the main factor influencing the fatty acid composition (J. D. Wood et al., 1999), however, the genetic background of the animal is also significant (Ryu et al., 2008). While lean commercial breeds have less fat content and a lower percentage of SFA and MUFA, rustic or local breeds like the Iberian pig have more fat deposition and, in the case of the Iberian breed, more percentage of SFA and MUFA, markedly in oleic acid (Barea et al., 2013). In addition, gene expression and lipid metabolism pathways vary depending on the tissue type, resulting in a distinct fatty acid profile for each tissue (J. D. Wood et al., 2008).

Another factor that modulates meat quality, albeit to a lesser extent, is the gut microbiota. The gut microbiota modifies the fatty acid composition in various tissues by incorporating different metabolites into the host bloodstream (Bergamaschi et al., 2020; Lei et al., 2021; G. Zhao et al., 2022). Additionally, the gut microbiota influences the immune system, as some species produce metabolites such as SCFAs that can improve the intestinal barrier, and the presence or proliferation of other species can cause dysbiosis, inflammation, and different diseases (Frank et al., 2007; R. Hu et al., 2023). Reciprocally, the diet of the host strongly impacts the composition of the gut microbiota, thereby indirectly affecting meat quality (Zhu et al., 2022). Hence, elucidating the microbial species associated with lipid metabolism could enhance meat quality through dietary modulations, promoting the beneficial microorganisms and reducing the prevalence of the harmful ones (Niu et al., 2023). Conversely, host genetics also influences the microbial composition of the gut microbiota, leading to the potential identification of microbial or host biomarkers associated with specific traits (Chen et al., 2018).

In this thesis, we used an integrative approach to deepen the understanding of the interplay between fatty acid composition, gut microbiota, and gene expression in different tissues. In

the following sections, we undertake a comparative analysis of the outcomes from various studies with the aim of obtaining a comprehensive perspective.

4.1. Association between gene expression and fatty acid composition

As previously stated, host genetics, and consequently, gene expression, serve as pivotal factors in shaping the fatty acid profile of various tissues. Transcriptomics elucidates the gene expression patterns in a specific tissue and time frame, contrasting with genomics, which provides the genetic information of the organism. By knowing the active genes, we can infer relationships with the resulting metabolites, lipid or protein profiles, or immune responses, among others. In the first study conducted during this thesis, a transcriptome analysis was conducted to explore the associations between gene expression and the fatty acid profile in pig muscle, a trait of interest for pork meat quality.

We used an experimental population consisting of an Iberian × Duroc backcross (BC1_DU), with 25% Iberian and 75% Duroc, which had been previously used in studies of our research group. These studies aimed to compare gene expression in *longissimus dorsi* muscle with its fatty acid profile, as well as various fatty acid ratios and metabolic indices. In an initial approach conducted by Valdés-Hernández and colleagues, a variable number of genes (with a total of 1,022 identified genes) were observed correlating with different fatty acids, ratios, and indices. Furthermore, an analysis of enriched metabolic pathways was conducted (Valdés-Hernández et al., 2023). After this study, the mixOmics tool was employed to refine the correlation method and identify the most significant relationships, revealing 365 genes associated with 15 fatty acids (Valdés-Hernández et al., 2024). Additionally, eGWAS analyses of 45 candidate genes in muscle were performed by Passols and colleagues, using this backcross and two others, with a 25% of Iberian background and a 75% of Landrace or Pietrain breeds. Notably, the BC1_DU backcross showed a statistically significant higher expression of various genes related to fatty acid synthesis, including different elongases (*ELOVL1*, *ELOVL5*, and *ELOVL6*), the *ADIPOQ* gene, *PLIN1*, and *FASN*, among others (Passols et al., 2023).

The main goal of reanalysing this population using WGCNA and graphical Gaussian models was to reduce data dimensionality to obtain modules of co-expressed genes potentially associated with the fatty acid profile, while also conducting an enrichment analysis to identify relevant metabolic pathways within each module. As a result, we identified four modules with significant correlations with more than five fatty acids (colour named as “Magenta”, “Yellow”, “Green”, and “Brown”). Remarkably, the genes within these modules had been found in previous studies. The expression of *ADIPOQ* and *PLIN1* genes was found in a module that showed positive correlations with several SFAs and MUFAs, while displaying negative

correlations with PUFAs. In line with this, the analysis conducted by Valdés-Hernández et al. (2024) reported a positive correlation between *ADIPOQ* expression and palmitoleic acid content, whereas the expression of *PLIN1* was positively correlated with MUFAs (palmitoleic and oleic acids) and negatively correlated with PUFAs (eicosadienoic and α-linolenic acids). The *ADIPOQ* gene encodes for the adiponectin, and in pigs is mapped in the chromosome 13, near the location of QTLs for the palmitoleic content in *longissimus dorsi* muscle (Čepica et al., 2005; Guo et al., 2009). On the other hand, *PLIN1* is differentially expressed in animals with high IMF content (B. Li et al., 2018). The study by Passols et al. (2023) revealed that these genes exhibit higher expression levels in animals with Duroc background compared to backcrosses of leaner breeds. Furthermore, we observed that both genes were associated with the PPAR signalling pathway, which was found enriched in the Magenta module and was also found relevant in the study by Valdés-Hernández and colleagues (2024). PPAR signalling pathway is important in meat quality, as the different Peroxisome Proliferator-Activated Receptors (PPARs) are activated by fatty acids and modulate different genes involved in different processes including lipid metabolism (He et al., 2013).

The expression of the gene *ELOVL6* was also found relevant in the Magenta module. Similar to the previously mentioned genes, its expression was higher in animals with Duroc background compared to backcrosses of leaner breeds (Passols et al., 2023). Previous studies of our research group with Duroc, Landrace and Pietrain backcrosses have linked *ELOVL6* expression to SFAs and MUFAs (Corominas et al., 2015; Corominas, Ramayo-Caldas, Puig-Oliveras, Pérez-Montarelo, et al., 2013; Crespo-Piazuelo et al., 2020). This enzyme is involved in the elongation of saturated or monounsaturated long-chain fatty acids (C12-C16) (Matsuzaka & Shimano, 2009) and is important in backfat fatty acid composition (Zappaterra et al., 2018).

The expression of the genes *PPARA* and *PPARGC1A* was found significant in the Green module, which was positively correlated with PUFAs. These genes are involved in lipolytic pathways, as they participate in the mitochondrial β-oxidation of fatty acids (Vega et al., 2000). In a previous study of our research group, the expression of these two genes was higher in the Duroc backcross than in the backcrosses of leaner breeds (Criado-Mesas et al., 2020). While *PPARG* belongs to the same family as *PPARA* and they are involved in the PPAR signalling pathway (He et al., 2013), their function is different. The expression of *PPARG* was found relevant in another module, positively correlated with SFAs and MUFAs, which agrees with the fact that this gene is involved in adipogenesis (Lehrke & Lazar, 2005).

The expression of *SCD* was found in the same module as *PPARG*, *ELOVL6*, *ADIPOQ* and *PLIN1*, with positive correlations with SFAs and MUFAs. *SCD* is a desaturase that converts

SFAs to MUFAs, and is important for the production of oleic acid. In previous studies of our research group, it was found related to the expression of *PPARG*, as they both are lipogenic genes (Criado-Mesas et al., 2020).

4.2. Association between microbiota and liver gene expression

Fatty acid composition differs markedly between liver, muscle and adipose tissues, which reflects their different physiological functions. While the adipose tissue is the primary site for *de novo* fatty acid biosynthesis in pigs, the liver plays a key role in other lipid metabolism pathways, such as the production of PUFAs, cholesterol, and bile acids (O'Hea & Leveille, 1969). Muscle tissue, on the other hand, serves primarily as a reservoir for energy storage and fatty acid oxidation (Gardan et al., 2006).

Table 4.1. Mean relative abundance (percentage) ± SD of the main fatty acids of the three studied tissues in the three different populations. Backfat and one set of the muscle samples belong to the Duroc × Iberian population (Du × Ib, 288 animals), while another set of the muscle samples belong to the backcross BC1_DU (129 animals), and the liver samples belong to Duroc pigs (310 animals).

Abbrev.	**Name**	**Backfat**	***Longissimus dorsi***		**Liver**
		(Du × Ib)	**(Du × Ib)**	**(BC1_DU)**	**(Duroc)**
C14:0	Myristic acid	1.57±0.142	1.53±0.150	1.27±0.228	0.61±0.262
C16:0	Palmitic acid	24.95±0.884	27.64±1.955	23.91±1.650	22.19±2.078
C16:1n-7	Palmitoleic acid	2.85±0.481	4.63±0.639	2.79±0.526	1.10±0.396
C18:0	Stearic acid	10.71±0.844	12.43±1.274	14.38±1.695	29.96±5.248
C18:1n-9	Oleic acid	49.10±1.410	47.82±3.257	35.93±5.823	21.31±3.993
C18:2n-6	Linoleic acid	7.65±0.700	3.81±0.781	12.13±5.823	11.20±2.394
C18:3n-3	α-linolenic acid	0.43±0.043	0.14±0.037	0.40±0.131	0.25±0.115
C20:0	Arachidic acid	0.20±0.028	0.18±0.107	0.23±0.078	0.24±0.110
C20:1n-9	Gondoic acid	1.28±0.159	0.67±0.232	0.73±0.164	0.26±0.054
C20:2n-6	Eicosadienoic acid	0.46±0.046	0.10±0.095	0.43±0.123	0.39±0.085
C20:3n-6	Dihomo-γ-linolenic acid	0.11±0.040	0.07±0.113	0.45±0.288	0.36±0.162
C20:4n-6	Arachidonic acid	0.14±0.025	0.54±0.289	2.58±1.938	5.82±2.263
SFA	Saturated fatty acids	37.43±1.391	41.78±3.044	39.79±3.094	53.01±4.981
MUFA	Monounsaturated fatty acids	53.22±1.268	53.12±3.177	39.46±6.087	22.67±4.346
PUFA	Polyunsaturated fatty acids	8.79±0.758	4.67±1.042	16.00±8.114	18.01±4.469

As observed in our data, MUFA content was notably high in backfat and muscle, primarily represented by oleic acid. Conversely, liver samples exhibited higher contents of SFA (mainly stearic acid) and PUFA (mainly arachidonic and linoleic acids). Adipose tissue, both subcutaneous and intramuscular, had higher levels of the shorter SFA (C14:0 and C16:0), which agrees with the fact that adipose tissue is the more efficient place for *de novo* fatty acid biosynthesis in pigs.

The expression of two desaturases, *FADS2* and *SCD*, was negatively linked to the abundance of genera that promote inflammation, such as *Streptococcus* or *Sutterella*. In the Duroc × Iberian population, GWAS analysis reported five genomic regions associated with the relative abundance of *Streptococcus* (Crespo-Piazuelo et al., 2019). Inside these regions there are genes encoding for immunity-related traits or genes that might favour the presence of *Streptococcus*. Some of the identified genes are affected by specific genera, such as the *RFTN1* gene and metabolites produced by *Lactobacillus* (Zhou et al., 2023), and other genes affecting the gut microbiota composition, like *BIN1*, which improves gut health (Thomas et al., 2022). *Streptococcus* increases its abundance in the pig gut after weaning, with the species *S. suis* emerging as the most prevalent and potentially harmful (Su et al., 2008). *Sutterella*, in turn, has been associated with a lower backfat thickness in pigs (X. Li et al., 2022) and with a mild inflammatory capacity in the human gut (Hiippala et al., 2016). These facts could lead to a decrease in the synthesis of SFA and MUFA, thus explaining the negative correlation between the abundance of these genera and the expression of *FADS2* and *SCD*.

In the previous section we discussed the expression of *ELOVL6*, but in liver we found a relevant expression of *ELOVL2*. The expression of this gene was positively associated with the abundance of *Collinsella*, which was also found to have positive relationships with the expression of two immunity-related genes (*CYP8B1* and *HMGCS2*). This genus has been associated with lipid metabolism and it has also been considered a pathobiont and proposed as a contributor to the fatty liver disease in pigs (J. Wang et al., 2023). The relationship with the expression of *ELOVL2* could be due to the effect of this elongase in the production of proinflammatory and anti-inflammatory cytokines as it is a key regulator of the elongation of PUFAs starting from the essential linoleic and α-linolenic acids (Smink et al., 2012).

Although we detected various genera involved in inflammatory responses or SFCA production, many of them have not yet been characterised in the porcine intestine. Here it is worth mentioning *SFLA01*, as its relative abundance was positively correlated with genes related to lipid metabolism or immunity (table 4.2).

Table 4.2. List of genes whose expression was significantly correlated with the abundance of *SFLA01*.

Gene	Correl.	Function	Reference
ABCD3	+	Peroxisomal transporter, involved in lipid β-oxidation	(Ranea-Robles et al., 2021)
ELOVL2	+	PUFA elongator	(Leonard et al., 2002)
CYP8B1	+	Bile acid synthesis	(W. Fang et al., 2019)
THBS1	+	Immunity modulator	(Tietze et al., 2024)
PTGS2	+	Prostaglandine synthesis	(Herschman, 1996)
HMGCS2	+	Ketonic bodies synthesis	(Newman & Verdin, 2014)
PLIN2	+	Lipid droplets metabolism	(Tsai et al., 2017)
ENSSSCG00000058639	+		
COL12A1	+		
LEPR	+	Energy balance, growth and fatty acid composition	(Pérez-Montarelo et al., 2013)
LIF	+	Immunity regulator	(Gao et al., 2009)
APOA4	+	Lipid transport, lipid oxidation inhibitor, potential role in immunity	(Qin et al., 1998)
MFSD2A	+	Fatty acid transport	(Nguyen et al., 2014)
TSKU	-		
ANGPTL8	-		
DISP3	-		
ENSSSCG00000061605	-		
PGAM2	-	Glycolysis, muscle growth	(Qiu et al., 2008)
CYP7A1	-	Bile acid synthesis	(W. Fang et al., 2019)

The genus *SFLA01* belongs to the order Oscillospirales, and according to the Greengenes2 database, it has been predominantly found in faecal samples from animals (public samples from redbiom). However, there is currently no literature reporting it, and to the best of our

knowledge, this is the first study that characterises this genus and shows its putative role in different pathways mainly related to lipid metabolism or immunity.

4.3. Gut microbial composition comparison

The gut microbiota is a complex and dynamic system that significantly influences the host. It releases byproducts that play crucial roles in the metabolism of the host, and any imbalance in this ecosystem can have negative repercussions for the host. The microbiota composition is influenced by dietary factors and environmental conditions, and it also evolves throughout the life of the host, particularly during the early growth stages (X. Wang et al., 2019). Furthermore, the genome of the host also impacts the microbiota composition, leading to differences between breeds (C. Wang et al., 2022).

In this thesis, we studied two different commercial pig populations, analysing the microbiota composition of rectal content samples through V3-V4 region sequencing of the 16S rRNA gene. One of the studied populations comprised 60-day-old Duroc pigs, while the other consisted of Duroc × Iberian crossbred pigs at slaughter age (209 days old on average, ranging 181 - 228). As the age difference implies, the two populations received different diets, even though both formulations were cereal-based, the 60 day-old piglets were fed a fattening diet, whereas the 6-7 month-old pigs received a finishing diet. The fattening diet had a higher content in protein for muscle deposition, while the finishing diet was focused on weight gaining and had a higher content in complex carbohydrates (Le Dinh et al., 2022). The different type of diet implies modulations in gut microbiota, as a fattening diet will favour the presence of proteolytic bacteria and the finishing diet promotes bacteria that ferment complex carbohydrates (Luo et al., 2022).

In addition, two different databases were used for the taxonomic classification: SILVA138 for the Duroc × Iberian pigs, and the newer Greengenes2 database for the Duroc piglets. Thus, these methodological differences contributed to the observed variability when comparing the two studies, as seen in Figures 4.1 and 4.2.

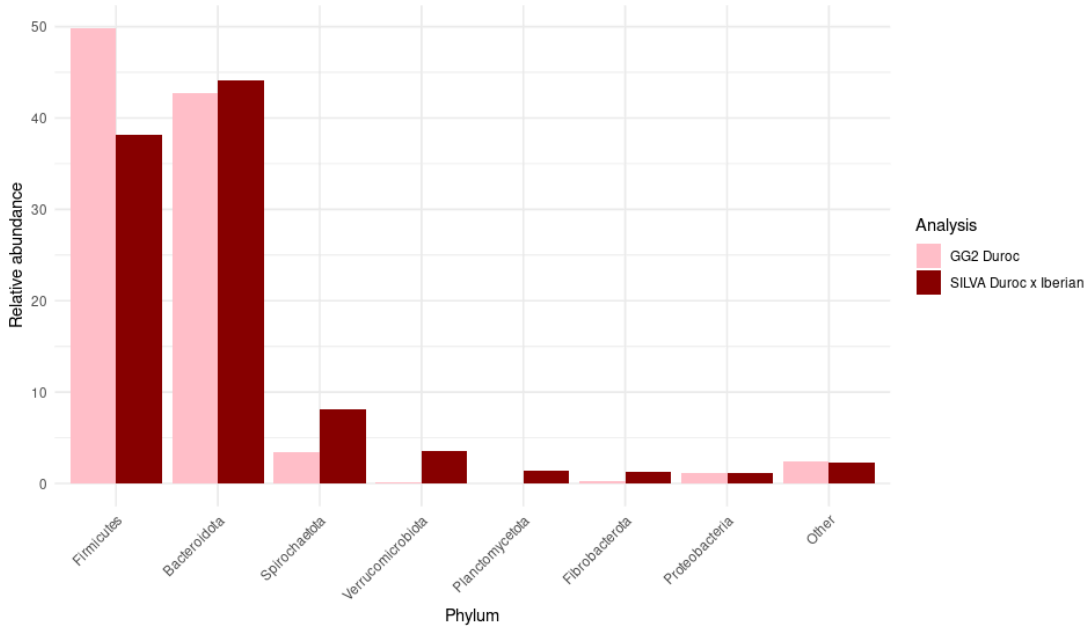


Figure 4.1. Relative abundance (%) barplot of the main bacterial and archaeal phyla in the two studied populations. Phyla with a relative abundance lower than 1% have been grouped as "Other". Light pink colour represents the samples taken from Duroc piglets and analysed with Greengenes2 (GG2), while the dark red colour represents the Duroc × Iberian crossbred pigs analysed with SILVA138.

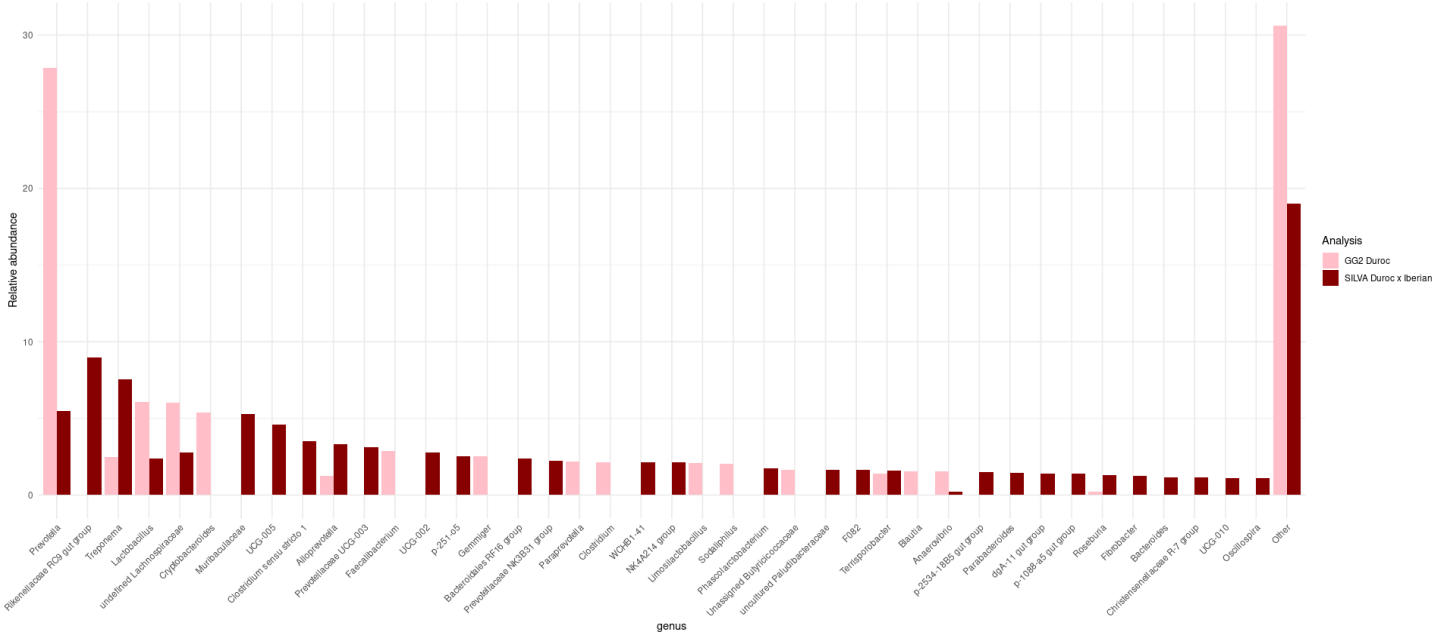


Figure 4.2. Relative abundance (%) barplot of the main bacterial and archaeal genera in the two studied populations. Genera with a relative abundance lower than 1% have been grouped as "Other". Light pink colour represents the samples taken from Duroc piglets and analysed with Greengenes2 (GG2), while the dark red colour represents the Duroc × Iberian crossbred pigs analysed with SILVA138.

In two-months-old piglets, the most abundant bacterial phylum was Firmicutes, representing nearly half of the relative abundance of all phyla found, while the most abundant genus was *Prevotella*, comprising nearly 28% of the relative abundance. The most abundant genera of the Firmicutes phylum were *Lactobacillus* and an unassigned genus of the Lachnospiraceae family, each representing more than a 6% of the total relative abundance. Conversely, in pigs at slaughter age, *Prevotella* was the third most abundant genus with a relative abundance of 5.5%, while the most abundant was another Bacteroidota, *Rikenellaceae RC9 gut group*, followed in abundance by the Spirochaetes member *Treponema*.

It is noteworthy the high relative abundance of *Prevotella* in the samples from 60 days-old piglets. This genus is a part of the core microbiota in almost all the growth stages (Holman et al., 2017) and is one of the most abundant genera, markedly in fattening pigs (Luo et al., 2022). *Lactobacillus* genus was also found to be more abundant in younger pigs, in agreement with previous studies, in which this genus was found at higher levels in two-months-old piglets than in pigs at slaughter age (H. B. Kim et al., 2011).

The gut content of the colon and rectum of the pig is predominantly inhabited by bacteria belonging to the phyla Firmicutes and Bacteroidota, with lesser proportions of Spirochaetes and Proteobacteria, in contrast to the small intestine, where the most abundant phyla are Firmicutes and Proteobacteria. In the rectum, Firmicutes may have a higher prevalence than Bacteroidota, but along the colon, the latter phylum is found in a higher abundance (Crespo-Piazuelo et al., 2018, 2019).

Thus, due to the differences in age and diet, it is difficult to compare the two populations, although they are indicative of the dynamism of the microbiota throughout the life of the animals. Weaning implies an important change in the diet, and the gut microbiota follows this change. Before weaning, lactic acid bacteria are highly prevalent, while after weaning the genera abundances start to switch to a different profile that will be more stable in adult life with bacteria capable of degrading polysaccharides of the diet and obtaining metabolites profitable for the host (Luo et al., 2022; X. Wang et al., 2019). Furthermore, weaning is a critical moment in the life of the piglet, where it is more likely that pathogens and opportunistic bacteria may cause dysbiosis harmful to the host, potentially leading to diarrhoea and a loss of growth efficiency (Bauer et al., 2006).

4.4. Comparison between 16S, metagenome and metatranscriptome

The analysis of a microbial population can be addressed through different techniques. While microbiological culture can be useful to study certain species and their metabolites, it has the main disadvantage of being unable to detect a huge number of species that cannot grow in lab conditions (such as strict anaerobes). Sequencing a fragment of the bacterial genome or characterising the whole metagenome of the microorganisms enables a comprehensive understanding of the population rather than solely focusing on specific microorganisms.

However, there are huge differences between 16S rRNA gene sequencing and whole-metagenome shotgun sequencing. As the 16S rRNA gene method only sequences a small fragment of the metagenome, it is more cost effective and faster than the whole metagenome sequencing. However, it is less accurate and only allows the analysis of prokaryotes (the microorganisms that possess a small subunit 16S of the ribosomal RNA). Bacteria and archaea are analysed with this method, but fungi, protists and viruses are not considered even though they are part of the microbiome. Sequencing the whole metagenome addresses this issue, with the advantage of being able to study the functions of the encoded proteins (Brumfield et al., 2020). In turn, metatranscriptome sequencing allows not only the taxonomic classification, but truly reports the active genes of the microbial population in the moment of sampling. Whole-metagenome and whole-metatranscriptome sequencing also have their drawbacks when compared to the 16S rRNA gene sequencing, as they are more demanding in terms of cost and computational resources (Guyomar & Lemaitre, 2022). However, the sequencing costs are being reduced each year as the techniques evolve.

When analysing the whole Duroc × Iberian commercial population, we obtained 759 bacterial and archaeal genera with metagenomic analysis and 422 genera in the metatranscriptomic analysis, while in the 16S rRNA gene sequencing results only 119 genera appeared and some of them were undefined. At phylum level, the number of taxa found was also different, with 17 phyla in the 16S rRNA gene sequence analysis, 34 in metagenome and 22 in metatranscriptome.

We found a different ranking of bacterial and archaeal phyla and genera between 16S rRNA gene analysis and the whole metagenomic study, and also a different one when compared to the metatranscriptomic analysis. Figure 4.3 shows the differences between the most abundant phyla of the three types of analysis. In the 16S rRNA gene sequence analysis, we use the whole pig population (285 animals), while in metagenomics and metatranscriptomics

analyses we only sampled a subset of the original population, selected based on the amount of SCFA in the rectal content (20 and 24 animals respectively). Running a comparative analysis with a subset of the original samples that only had the common samples of the three studies, we found a similar result. Therefore, we could discard the differences in abundance due to the number of animals.

We must take into account that 16S rRNA gene sequence analysis may have considerable biases, as only an amplified region of the bacterial genome is sequenced. This can lead to the non-identification of low-abundance species or an incorrect classification. Even the selection of the hypervariable region to be sequenced can introduce errors (Martínez-Porchas et al., 2016). To avoid the problems of a misclassification of low-abundant taxa, we filtered the three datasets for taxa that had a mean relative abundance higher than 0.005%.

In addition, another reason to explain these differences, besides the analysis technique, is that we used different databases to classify the taxa. Different databases use different ways to classify the sequences and have different taxonomic arrangements, therefore the results may vary even though the samples are the same.

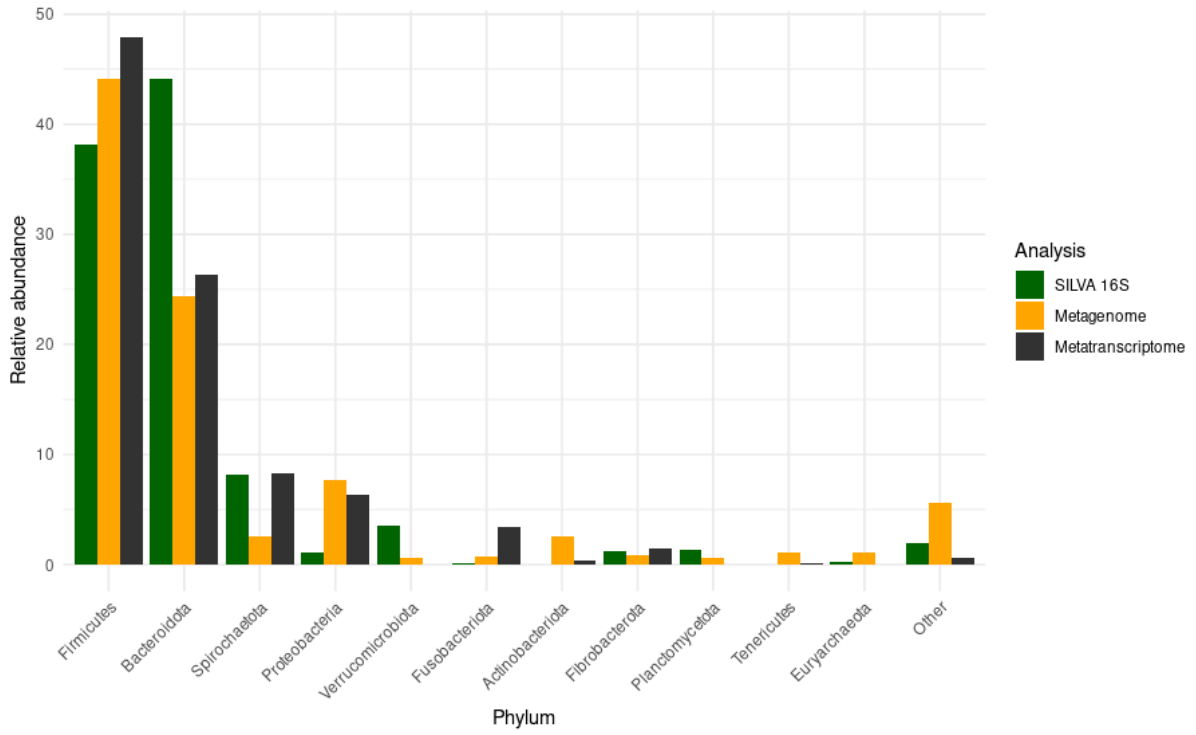


Figure 4.3. Relative abundance (%) barplot of the main bacterial and archaeal phyla based on the analysis in the crossbred Duroc × Iberian population. Phyla with a relative abundance/activity lower than 1% were grouped as “Other”.

Both results from 16S rRNA gene analysis and metagenome showed the genera considered part of the faecal core microbiota (Holman et al., 2017) among the most abundant taxa. However, *Rikenellaceae RC9 gut group* did not appear in the database used for the metagenomic analysis, probably due to the fact that this genus was first reported recently (Sun et al., 2015). The characterisation of this genus of the order Bacteroidales and the presence of non-annotated taxa in the 16S rRNA gene sequence analysis could partly explain the differences in the relative abundance of genera such as *Prevotella*.

The most abundant Firmicutes genus in the metagenomic study was *Clostridium*, surpassed by two genera from the phylum Bacteroidota (*Prevotella* and *Bacteroides*). However, in the metatranscriptomic analysis, the genus *Clostridium* was by far the most active, accounting for over 20% of the total activity, whereas its relative abundance in the metagenome analysis was much lower. In previous studies by our group, it was found to be more abundant in the small intestine (Crespo-Piazuelo et al., 2018). While some species of the *Clostridium* genus are pathogenic (e.g. *Clostridium botulinum* or *Clostridium perfringens*), there are species such as *Clostridium butyricum* that are commensals and have beneficial effects on gut health, such as the production of butyrate. This species is found at high levels in different mammals and acts as a probiotic (Peng et al., 2023). The metatranscriptomic differential analysis showed *Clostridium hydrogeniformans* as an overabundant species in samples with high content in acetic acid, which agrees with the fact that it can act as an acetate producer (Bowman et al., 2010).

This higher activity of the genus *Clostridium* also agrees with the most expressed gene in the differential analysis, *sspD*, which encodes for a sporulation protein commonly found in clostridia that gives the bacteria resistance to heat, UV radiation and chemical components (Paredes-Sabja et al., 2008).

4.5. Association between the gut microbiota and the fatty acid composition

Our study was one of the first studies assessing the relationship between gut microbiota and fatty acid composition (Sebastià et al., 2024). Although some moderate correlations between gut bacteria and SCFAs were found, the correlation between gut microbiota and the fatty acid composition of backfat and muscle were low. One of the main reasons we hypothesised for these lower correlations was the indirect relationship between gut microbiota and fatty acid deposition in backfat and muscle, as opposed to the direct relationship between gut microbiota and SCFAs. While our metagenomic analysis did not show relevant differential abundances, the metatranscriptomic study confirmed some correlations observed between microbial genera and fatty acid composition in the 16S rRNA gene sequence analysis. 16S rRNA gene sequence analysis showed *Prevotella* as the genus with the highest correlation with butyric acid, and *Prevotella copri* was found to more active in pigs with high levels of this SCFA in rectal content in the metatranscriptomic analysis. This species is considered a succinate producer (Franke & Deppenmeier, 2018), which could be converted to butyrate, and that could be the reason why it had a higher activity in the metatranscriptomic analysis.

Conversely, the *Akkermansia* genus described in the 16S rRNA gene sequence study showed a positive correlation with the abundance of acetic acid, and *Akkermansia glycaniphila* was found to have more activity in samples with higher levels of acetic acid in the metatranscriptomic analysis. It is worth to mention that rather than *Akkermansia muciniphila*, a common inhabitant of the colon (Dao et al., 2016), the species found more active was *Akkermansia glycaniphila*, an acetate producer. In addition, *A. glycaniphila* has been recently described as a gut inhabitant and is also a mucin degrader (Lv et al., 2022; Ouwerkerk et al., 2016).

In the 16S rRNA gene sequence analysis the *Methanobrevibacter* genus was positively correlated with the abundance of acetic acid and negatively correlated with butyric acid, similar to *Akkermansia*. In accordance, acetate supports the growth of the species *M. smithii*, as it simulates the use of formate to produce methane, and this genus is the most abundant methanogen in the colon of finishing pigs (Mi et al., 2019). However, the other archaeon identified in the 16S rRNA gene sequence analysis, a member of the Methanomethylophilaceae family, only reported a negative correlation with butyric acid levels.

Altogether, the correlations do not imply a direct causality. The microbiota promotes the presence of certain SCFAs, as in the lactic acid bacteria found more active in the

metatranscriptomic analysis, which had a higher activity in samples with high levels of acetic acid. However, the reverse situation also occurs, as the presence of certain SCFAs promotes the proliferation of different genera that use them as a substrate.

Furthermore, while we have observed that the microbiota may have a certain influence in host gene expression, the opposite situation also occurs, as different bacterial genera are influenced by the genetics and gene expression of the individual. In a parallel analysis of a subset of the Duroc × Iberian population for a related article, it was observed that the genera *Phascolarctobacterium*, *Roseburia*, and *Faecalibacterium* were more abundant when the sampled animals had a higher number of copies of a copy number variant (CNV) located on SSC14. This CNV partially contains the genes *ABCC2* and *DNMBP* and has been associated with fatty acid composition (Ramayo-Caldas et al., 2023). *Phascolarctobacterium*, in addition, was found relevant in a previous study of Crespo-Piazuelo and colleagues, in which they detected candidate genes in a region of SSC9 that could be related to the abundance of this genus (Crespo-Piazuelo et al., 2019). These three genera are present in the pig intestine and are associated with the production of SCFA, propionate in the case of *Phascolarctobacterium*, and butyrate in the case of *Roseburia* and *Faecalibacterium*. All three genera have been linked to a better intestinal health (W. Li et al., 2022; Machiels et al., 2014), and the species *Faecalibacterium prausnitzii* is known to have anti-inflammatory properties (Lopez-Siles et al., 2017).

4.6. Microbiota - gene expression - lipid metabolism

Throughout the different studies of the present thesis, we have observed an interrelation between the fatty acid composition of different tissues, gene expression, and gut microbiota, among other factors. Gene expression in backfat, liver and muscle significantly influences the lipid metabolism of the host and their fatty acid composition within tissue. Meanwhile the microbiota, shaped by the diet, produces various metabolites that may be incorporated into the lipid metabolism of the host, while the genetics of the hosts and this same metabolism can also shape the gut microbiota.

While intramuscular fat serves as a storage and production site for fatty acids in pigs, with a higher expression of genes involved in the synthesis of SFAs and MUFAs, such as *FASN* and long-chain fatty acid desaturases like *SCD*, *ELOVL6*, or *FADS2*, the liver is characterised by a prevalence of genes involved in the elongation of PUFAs, such as *ELOVL2*. Moreover, the liver plays a significant role in the regulation of inflammation and immunity, partly due to its production of PUFAs. The liver expression of the previous genes has been found modulated by changes in the gut microbiota in ducks (Gu et al., 2022) and in mice (Albouery et al., 2020), where a dysregulated microbiota caused a reduction in the production of PUFA.

Many commensal genera produce SCFAs as metabolites that may be influencing lipid metabolism and immunity. While *Prevotella* has been associated both positively and negatively with intestinal health, the most abundant species we have found (*P. copri*) is a succinate producer. Succinate is an intermediate metabolite in the pathway of SCFAs production by microbial fermentation and also regulates mucosal immunity in the intestine of pigs (X. Li et al., 2019; Louis & Flint, 2017). Other bacteria such as *Phascolarctobacterium* or *Clostridium* can utilise this substrate and produce SCFAs (Wei et al., 2023).

Butyrate is one of the metabolites derived from the utilisation of succinate, although it can also be produced from carbohydrate fermentation. This SCFA serves as the primary energy source for colonocytes and contributes to intestinal health by modulating the immune system, maintaining the integrity of the intestinal barrier, and exerting anti-inflammatory effects (Klampfer et al., 2003; Roediger, 1982).

Acetate and propionate are absorbed to a greater extent by the host. While in the liver, they can influence fatty acid oxidation through the activation of *PPARGC1A*, acetate serves as an initial molecule in de novo fatty acid synthesis. It is converted into acetyl-CoA and then into malonyl-CoA by the *ACACA* gene. From here, the *FASN* gene elongates two carbons in

each reaction until producing C16:0 (Corominas, Ramayo-Caldas, Puig-Oliveras, Pérez-Montarelo, et al., 2013).

The previously mentioned bacteria can be considered commensal microorganisms that produce beneficial metabolites for the host, but, conversely, potentially pathogenic bacteria can activate the animal's immunity and also, they may be modulated by it. For instance, we found a negative correlation between *Escherichia-Shigella* and propionate, and this SCFA limits the expansion of pathogens like *Salmonella* and *Shigella* by disrupting their intracellular pH balance (Jacobson et al., 2018).

4.7. Future perspectives

The study of the influence of the microbiota on the host and its interactions is an expanding field that has received more attention in recent years. While in human research the aim is to apply this knowledge to improve intestinal health and reduce the severity or incidence of diseases such as IBD or certain types of cancer, in production animals, the main goal is to improve growth efficiency and meat quality, as well increasing their resilience. Therefore, characterising the intestinal microbiota and its relationships with gene expression or fatty acid profiles allows for the identification of taxa involved in desirable traits, which can be supplemented through diet. Moreover, certain microbial species can be considered biomarkers for traits such as intestinal health, so determining their presence or abundance provides insight into the animal's immune status and can pre-empt more serious health issues that could impact its productive efficiency. This approach also helps avoid the use of antibiotics, which can promote the proliferation of resistant bacteria. The validation of these biomarkers will enable their application in the nutrition programs of production animals.

The WGCNA study was conducted in an Iberian and Duroc backcross. However, we have observed in previous studies that the expression of many lipid metabolism genes varies depending on the breed. Therefore, it could be interesting to apply this study to different backcrosses available or to analyse commercial populations since this analysis was carried out in experimental populations that do not fully reflect the commercial lines.

One of the main drawbacks of the comparative analysis between the intestinal microbiota and gene expression in the liver was that the two datasets were collected at different ages of the animals, increasing the difficulty of interpreting their relationship. Additionally, since the gut microbiota sample was taken at two months of age, its profile will vary considerably until adulthood, making it difficult to establish genuine relationships between its composition and gene expression. To address this issue, longitudinal analyses of intestinal microbial composition could be performed to assess changes over time and identify genera that may be involved in liver gene expression. Another additional approach could involve comparing the intestinal microbiota with the liver fatty acid profile to determine if any genus can predict the abundance of certain fatty acids.

To get a comprehensive view of the effect of the microbiota on the host through SCFA production, we believe further studies should include the gene expression of the enterocytes of small and large intestines and the SCFA profile of blood, in addition to the rest of the animal material mentioned in this thesis: fatty acid profiles and gene expression in liver, backfat and muscle, as well as SCFA levels in the gut. This would allow us to assess and

trace more direct relationships between host metabolism and gut microbiota, and understand how they influence each other.

As research progresses in studying the microbial composition across different environments, newer updates to taxonomic databases will deepen our knowledge of unknown or not annotated species. At the moment, the Greengenes2 database identified genera that are still not well characterised, such as *SFLA01*, a genus with great relevance in our study and found to be associated with the expression of various genes in liver. This improvement in taxonomic classification would improve our understanding of this species and another species, their similarity to others and their functions.

Nowadays, there are constant advances in sequencing techniques (TGS), the development of new pipelines, and the computing power for analysis and storage is increasing constantly. Therefore, the study of different omics is becoming more affordable, which will allow for more complex studies with a greater number of samples, deeper dataset characterization, and more precise relationships through systems biology.

5. Conclusions

1. The *WGCNA* method identified four significant modules of co-expressed genes in *longissimus dorsi* muscle of a backcrossed Iberian × Duroc pig population. Two of the relevant modules had positive correlations with SFAs and MUFAs, and negative correlations with PUFAs and the omega-6/omega-3 ratio. Conversely, the other two relevant modules exhibited inverse relationships with the FA composition and omega-6/omega-3 ratio.

2. The two modules positively correlated with SFAs and MUFAs had enriched pathways related to insulin and FA biosynthesis. On the contrary, the modules positively correlated with PUFAs and the omega-6/omega-3 ratio had enriched pathways related to adipocytokine signalling and FA β-oxidation in the mitochondria.

3. The analysis of the region V3-V4 of the 16S rRNA gene in the rectal microbiota content in a commercial Duroc × Iberian population detected the *Rikenellaceae RC9 gut group* as the most abundant genus. Furthermore, the levels of *Rikenellaceae RC9 gut group* were negatively correlated with the stearic acid percentage in backfat. This genus also exhibited a positive correlation with palmitic acid and a negative correlation with oleic acid in muscle, suggesting a role for *Rikenellaceae RC9 gut group* in the modulation of the FA composition.

4. *Prevotella* was the third most abundant genus and its abundance showed a negative correlation with the levels of acetic acid and a positive correlation with the levels of butyric acid in the gut. Conversely, *Akkermansia* showed the opposite correlations with the levels of rectal acetic and butyric acids, highlighting the potential role of both *Akkermansia* and *Prevotella* as biomarkers for the SCFA profile in the gut.

5. The metagenomic and metatranscriptomic analysis of a subset of the Duroc × Iberian population with divergent content in SCFAs identified *Akkermansia glycaniphila* as an acetate producer, while *Prevotella copri* was positively associated with the abundance of butyric acid.

6. The *Clostridium* genus was the most active genus in the metatranscriptomic analysis, and together with *Lactobacillus* species, were more active in samples with high acetic acid and low propionic acid content. In agreement with their reported activity, the most abundant KO was K06421 (*sspD*), a gene commonly found in *Bacillus* and *Clostridium* species. In the functional analysis, most of the differentially expressed genes encoded for proteins related with sporulation and bacteriophage activity, which can be explained by the harsh conditions of the environment.

7. Bacterial genera taken from the rectal content of 60 day-old Duroc piglets, may predict the expression of certain liver genes at slaughter. Thus, *Caccousia, Eubacterium, Floccifex*, *Mediterraneibacter*, *SFLA01,* and *Streptococcus* were correlated with the expression of key genes related with liver metabolism and functions such as *ABCD3*, *CXCL8*, *CYP8B1*, *DEFB1*, *ELOVL2*, *FADS2*, *HMGCS2*, *MZB1* and *SCD*.

6. References

Aaslyng, M. D., Oksama, M., Olsen, E. V., Bejerholm, C., Baltzer, M., Andersen, G., Bredie, W. L. P., Byrne, D. V., & Gabrielsen, G. (2007). The impact of sensory quality of pork on consumer preference. *Meat Science*, *76*(1), 61–73. https://doi.org/10.1016/j.meatsci.2006.10.014

Abegaz, F., Martines, A.-C. M. F., Vieira-Lara, M. A., Rios-Morales, M., Reijngoud, D.-J., Wit, E. C., & Bakker, B. M. (2021). Bistability in fatty-acid oxidation resulting from substrate inhibition. *PLOS Computational Biology*, *17*(8), e1009259. https://doi.org/10.1371/journal.pcbi.1009259

Afzaal, M., Saeed, F., Shah, Y. A., Hussain, M., Rabail, R., Socol, C. T., Hassoun, A., Pateiro, M., Lorenzo, J. M., Rusu, A. V., & Aadil, R. M. (2022). Human gut microbiota in health and disease: Unveiling the relationship. *Frontiers in Microbiology*, *13*, 999001. https://doi.org/10.3389/fmicb.2022.999001

Albouery, M., Buteau, B., Grégoire, S., Cherbuy, C., Pais De Barros, J.-P., Martine, L., Chain, F., Cabaret, S., Berdeaux, O., Bron, A. M., Acar, N., Langella, P., & Bringer, M.-A. (2020). Age-Related Changes in the Gut Microbiota Modify Brain Lipid Composition. *Frontiers in Cellular and Infection Microbiology*, *9*, 444. https://doi.org/10.3389/fcimb.2019.00444

Amills, M., Clop, A., Ramírez, O., & Pérez‐Enciso, M. (2010). Origin and Genetic Diversity of Pig Breeds. In Wiley, *Encyclopedia of Life Sciences* (1st ed.). Wiley. https://doi.org/10.1002/9780470015902.a0022884

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Ao, P. (2009). Global view of bionetwork dynamics: Adaptive landscape. *Journal of Genetics and Genomics*, *36*(2), 63–73. https://doi.org/10.1016/S1673-8527(08)60093-4

Athanasopoulou, K., Boti, M. A., Adamopoulos, P. G., Skourou, P. C., & Scorilas, A. (2021).

Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. *Life*, *12*(1), 30. https://doi.org/10.3390/life12010030

Ballester, M., Ramayo-Caldas, Y., González-Rodríguez, O., Pascual, M., Reixach, J., Díaz, M., Blanc, F., López-Serrano, S., Tibau, J., & Quintanilla, R. (2020). Genetic parameters and associated genomic regions for global immunocompetence and other health-related traits in pigs. *Scientific Reports*, *10*(1), 18462.

Barea, R., Isabel, B., Nieto, R., López-Bote, C., & Aguilera, J. F. (2013). Evolution of the fatty acid profile of subcutaneous back-fat adipose tissue in growing Iberian and Landrace × Large White pigs. *Animal*, *7*(4), 688–698. https://doi.org/10.1017/S175173111200184X

Bartholome, A., Albin, D., Baker, D., Holst, J., & Tappenden, K. (2004). Supplementation of total parenteral nutrition with butyrate acutely increases structural aspects of intestinal adaptation after an 80% jejunoileal resection in neonatal piglets. *Journal of Parenteral and Enteral Nutrition*, *28*(4), 210–222. https://doi.org/10.1177/0148607104028004210

Bauer, E., Williams, B. A., Smidt, H., Verstegen, M. W., & Mosenthin, R. (2006). Influence of the gastrointestinal microbiota on development of the immune system in young animals. *Current Issues in Intestinal Microbiology*, *7*(2), 35–52.

Beauvieux, M.-C., Tissier, P., Gin, H., Canioni, P., & Gallis, J.-L. (2001). Butyrate Impairs Energy Metabolism in Isolated Perfused Liver of Fed Rats. *The Journal of Nutrition*, *131*(7), 1986–1992. https://doi.org/10.1093/jn/131.7.1986

Benet, I., Guàrdia, M. D., Ibañez, C., Solà, J., Arnau, J., & Roura, E. (2015). Analysis of SPME or SBSE extracted volatile compounds from cooked cured pork ham differing in intramuscular fat profiles. *LWT - Food Science and Technology*, *60*(1), 393–399. https://doi.org/10.1016/j.lwt.2014.08.016

Bergamaschi, M., Maltecca, C., Schillebeeckx, C., McNulty, N. P., Schwab, C., Shull, C., Fix, J., & Tiezzi, F. (2020). Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Scientific Reports*, *10*(1),

10134. https://doi.org/10.1038/s41598-020-66791-3

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bovo, S., Ribani, A., Muñoz, M., Alves, E., Araujo, J. P., Bozzi, R., Čandek-Potokar, M., Charneca, R., Di Palma, F., Etherington, G., Fernandez, A. I., García, F., García-Casco, J., Karolyi, D., Gallo, M., Margeta, V., Martins, J. M., Mercat, M. J., Moscatelli, G., … Fontanesi, L. (2020). Whole-genome sequencing of European autochthonous and commercial pig breeds allows the detection of signatures of selection for adaptation of genetic resources to different breeding and production systems. *Genetics Selection Evolution*, *52*(1), 33. https://doi.org/10.1186/s12711-020-00553-7

Bowman, K. S., Dupré, R. E., Rainey, F. A., & Moe, W. M. (2010). Clostridium hydrogeniformans sp. Nov. And Clostridium cavendishii sp. Nov., hydrogen-producing bacteria from chlorinated solvent-contaminated groundwater. *International Journal of Systematic and Evolutionary Microbiology*, *60*(2), 358–363. https://doi.org/10.1099/ijs.0.013169-0

Brestenský, M., Nitrayová, S., Bomba, A., Patráš, P., Strojný, L., Szabadošová, V., Pramuková, B., & Bertková, I. (2017). The content of short chain fatty acids in the jejunal digesta, caecal digesta and faeces of growing pigs. *Livestock Science*, *205*, 106–110. https://doi.org/10.1016/j.livsci.2017.09.015

Broad Institute. (2019). *Picard Toolkit* [Computer software].

Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L., & Leddy, M. B. (2020). Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLoS One*, *15*(2), e0228899.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. https://doi.org/10.1038/nmeth.3176

Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: A de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, *8*(9), giz100. https://doi.org/10.1093/gigascience/giz100

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Camarinha-Silva, A., Maushammer, M., Wellmann, R., Vital, M., Preuss, S., & Bennewitz, J. (2017). Host Genome Influence on Gut Microbial Composition and Microbial Prediction of Complex Traits in Pigs. *Genetics*, *206*(3), 1637–1644. https://doi.org/10.1534/genetics.117.200782

Čepica, S., Masopust, M., Knoll, A., & Rohrer, G. A. (2005). Linkage and RH mapping of the porcine adiponectin gene on chromosome 13. *Animal Genetics*, *36*(3). https://doi.org/10.1111/j.1365-2052.2005.01293.x

Chen, C., Huang, X., Fang, S., Yang, H., He, M., Zhao, Y., & Huang, L. (2018). Contribution of Host Genetics to the Variation of Microbial Composition of Cecum Lumen and Feces in Pigs. *Frontiers in Microbiology*, *9*, 2626. https://doi.org/10.3389/fmicb.2018.02626

Chen, C., Zhou, Y., Fu, H., Xiong, X., Fang, S., Jiang, H., Wu, J., Yang, H., Gao, J., & Huang, L. (2021). Expanded catalog of microbial genes and metagenome-assembled genomes from the pig gut microbiome. *Nature Communications*, *12*(1), 1106. https://doi.org/10.1038/s41467-021-21295-0

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., McGarrell, D. M., Marsh, T., Garrity, G. M., & Tiedje, J. M. (2009). The

Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Research*, *37*(Database), D141–D145. https://doi.org/10.1093/nar/gkn879

Corominas, J., Marchesi, J. A., Puig-Oliveras, A., Revilla, M., Estellé, J., Alves, E., Folch, J. M., & 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), 20. https://doi.org/10.1186/s12711-015-0111-y

Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Estellé, J., Castelló, A., Alves, E., Pena, R. N., Ballester, M., & 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. https://doi.org/10.1186/1471-2164-14-843

Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Pérez-Montarelo, D., Noguera, J. L., Folch, J. M., & Ballester, M. (2013). Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. *PLoS ONE*, *8*(1), e53687. https://doi.org/10.1371/journal.pone.0053687

Crespo-Piazuelo, D., Criado-Mesas, L., Revilla, M., Castelló, A., Noguera, J. L., Fernández, A. I., Ballester, M., & 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. https://doi.org/10.1038/s41598-020-70894-2

Crespo-Piazuelo, D., Estellé, J., Revilla, M., Criado-Mesas, L., Ramayo-Caldas, Y., Óvilo, C., Fernández, A. I., Ballester, M., & Folch, J. M. (2018). Characterization of bacterial microbiota compositions along the intestinal tract in pigs and their interactions and functions. *Scientific Reports*, *8*(1), 12727. https://doi.org/10.1038/s41598-018-30932-6

Crespo-Piazuelo, D., Migura-Garcia, L., Estellé, J., Criado-Mesas, L., Revilla, M., Castelló, A., Muñoz, M., García-Casco, J. M., Fernández, A. I., Ballester, M., & Folch, J. M.

(2019). Association between the pig genome and its gut microbiota composition. *Scientific Reports*, *9*(1), 8791. https://doi.org/10.1038/s41598-019-45066-6

Crespo-Piazuelo, D., Ramayo-Caldas, Y., González-Rodríguez, O., Pascual, M., Quintanilla, R., & Ballester, M. (2021). A Co-Association Network Analysis Reveals Putative Regulators for Health-Related Traits in Pigs. *Frontiers in Immunology*, *12*, 784978. https://doi.org/10.3389/fimmu.2021.784978

Criado-Mesas, L., Ballester, M., Crespo-Piazuelo, D., Castelló, A., Benítez, R., Fernández, A. I., & 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. https://doi.org/10.1371/journal.pone.0220708

Criado-Mesas, L., Ballester, M., Crespo-Piazuelo, D., Castelló, A., Fernández, A. I., & 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), 9845. https://doi.org/10.1038/s41598-020-67015-4

Crisol-Martínez, E., Stanley, D., Geier, M. S., Hughes, R. J., & Moore, R. J. (2017). Understanding the mechanisms of zinc bacitracin and avilamycin on animal production: Linking gut microbiota and growth performance in chickens. *Applied Microbiology and Biotechnology*, *101*(11), 4547–4559. https://doi.org/10.1007/s00253-017-8193-9

Dao, M. C., Everard, A., Aron-Wisnewsky, J., Sokolovska, N., Prifti, E., Verger, E. O., Kayser, B. D., Levenez, F., Chilloux, J., Hoyles, L., MICRO-Obes Consortium, Dumas, M.-E., Rizkalla, S. W., Doré, J., Cani, P. D., & Clément, K. (2016). *Akkermansia muciniphila* and improved metabolic health during a dietary intervention in obesity: Relationship with gut microbiome richness and ecology. *Gut*, *65*(3), 426–436. https://doi.org/10.1136/gutjnl-2014-308778

Das, U. N. (2006). Essential Fatty Acids—A Review. *Current Pharmaceutical Biotechnology*, *7*(6), 467–482.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T.,

Dalevi, D., Hu, P., & Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. https://doi.org/10.1128/AEM.03006-05

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dong, W., Ricker, N., Holman, D. B., & Johnson, T. A. (2023). Meta-analysis reveals the predictable dynamic development of the gut microbiota in commercial pigs. *Microbiology Spectrum*, *11*(6), e01722-23. https://doi.org/10.1128/spectrum.01722-23

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the Human Intestinal Microbial Flora. *Science*, *308*(5728), 1635–1638. https://doi.org/10.1126/science.1110591

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., & Dalal, R. (2009). *Real-Time DNA Sequencing from Single Polymerase Molecules*. *323*.

Estany, J., Ros-Freixedes, R., Tor, M., & Pena, R. N. (2014). A Functional Variant in the Stearoyl-CoA Desaturase Gene Promoter Enhances Fatty Acid Desaturation in Pork. *PLOS ONE*, *9*(1).

Estellé, J., Mercadé, A., Noguera, J. L., Pérez-Enciso, M., Óvilo, C., Sánchez, A., & Folch, J. M. (2005). Effect of the porcine IGF2-intron3-G3072A substitution in an outbred Large White population and in an Iberian × Landrace cross1. *Journal of Animal Science*, *83*(12), 2723–2728. https://doi.org/10.2527/2005.83122723x

Estellé, J., Pérez-Enciso, M., Mercadé, A., Varona, L., Alves, E., Sánchez, A., & Folch, J. M. (2006). Characterization of the porcine FABP5 gene and its association with the FAT1 QTL in an Iberian by Landrace cross. *Animal Genetics*, *37*(6), 589–591. https://doi.org/10.1111/j.1365-2052.2006.01535.x

Fan, Y., Arbab, A. A. I., Zhang, H., Yang, Y., Nazar, M., Han, Z., & Yang, Z. (2021). Lactation associated genes revealed in Holstein dairy cows by weighted gene co-expression network analysis (WGCNA). *Animals*, *11*(2), 314.

Fang, C. L., Sun, H., Wu, J., Niu, H. H., & Feng, J. (2014). Effects of sodium butyrate on growth performance, haematological and immunological characteristics of weanling piglets. *Journal of Animal Physiology and Animal Nutrition*, *98*(4), 680–685. https://doi.org/10.1111/jpn.12122

Fang, W., Wen, X., Meng, Q., Wu, W., Everaert, N., Xie, J., & Zhang, H. (2019). Alteration in bile acids profile in Large White pigs during chronic heat exposure. *Journal of Thermal Biology*, *84*, 375–383.

FAOSTAT. (2024). *FAOSTAT*. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QCL

Fenske, G. J., Ghimire, S., Antony, L., Christopher-Hennings, J., & Scaria, J. (2020). Integration of culture-dependent and independent methods provides a more coherent picture of the pig gut microbiome. *FEMS Microbiology Ecology*, *96*(3), fiaa022. https://doi.org/10.1093/femsec/fiaa022

Font-i-Furnols, M., & Guerrero, L. (2014). Consumer preference, behavior and perception about meat and meat products: An overview. *Meat Science*, *98*(3), 361–371. https://doi.org/10.1016/j.meatsci.2014.06.025

Fouhse, J. M., Zijlstra, R. T., & Willing, B. P. (2016). The role of gut microbiota in the health and disease of pigs. *Animal Frontiers*, *6*(3), 30–36. https://doi.org/10.2527/af.2016-0031

Frank, D. N., St. Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N., & Pace, N. R. (2007). Molecular-phylogenetic characterization of microbial community imbalances

in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences*, *104*(34), 13780–13785. https://doi.org/10.1073/pnas.0706625104

Franke, T., & Deppenmeier, U. (2018). Physiology and central carbon metabolism of the gut bacterium *Prevotella copri*. *Molecular Microbiology*, *109*(4), 528–540. https://doi.org/10.1111/mmi.14058

Franzosa, E. A., McIver, L. J., Rahnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., Lipson, K. S., Knight, R., Caporaso, J. G., Segata, N., & Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nature Methods*, *15*(11), 962–968. https://doi.org/10.1038/s41592-018-0176-y

Gao, W., Thompson, L., Zhou, Q., Putheti, P., Fahmy, T. M., Strom, T. B., & Metcalfe, S. M. (2009). Treg versus Th17 lymphocyte lineages are cross-regulated by LIF versus IL-6. *Cell Cycle*, *8*(9), 1444–1450.

Gardan, D., Gondret, F., & Louveau, I. (2006). Lipid metabolism and secretory function of porcine intramuscular adipocytes compared with subcutaneous and perirenal adipocytes. *American Journal of Physiology-Endocrinology and Metabolism*, *291*(2), E372–E380. https://doi.org/10.1152/ajpendo.00482.2005

Gillingham, L. G., Harris-Janz, S., & Jones, P. J. (2011). Dietary monounsaturated fatty acids are protective against metabolic syndrome and cardiovascular disease risk factors. *Lipids*, *46*, 209–228.

Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., & Field, K. G. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, *345*(6270), 60–63. https://doi.org/10.1038/345060a0

Giuffra, E., Kijas, J. M. H., Amarger, V., Carlborg, Ö., Jeon, J.-T., & Andersson, L. (2000). The Origin of the Domestic Pig: Independent Domestication and Subsequent Introgression. *Genetics*, *154*(4), 1785–1791. https://doi.org/10.1093/genetics/154.4.1785

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., & Johnston, M. (1996). Life with 6000 genes. *Science*,

274(5287), 546–567.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Groenen, M. A., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., & Megens, H.-J. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, *491*(7424), 393–398.

Gu, T., Duan, M., Zhang, R., Zeng, T., Xu, W., Feng, W., Jiang, C., Tian, Y., Chen, L., & Lu, L. (2022). Probiotic Fermented Feed Alleviates Liver Fat Deposition in Shaoxing Ducks via Modulating Gut Microbiota. *Frontiers in Microbiology*, *13*, 928670. https://doi.org/10.3389/fmicb.2022.928670

Guillou, H., Zadravec, D., Martin, P. G. P., & 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. https://doi.org/10.1016/j.plipres.2009.12.002

Guo, T., Ren, J., Yang, K., Ma, J., Zhang, Z., & Huang, L. (2009). Quantitative trait loci for fatty acid composition in longissimus dorsi and abdominal fat: Results from a White Duroc× Erhualian intercross F2 population. *Animal Genetics*, *40*(2), 185–191.

Guyomar, C., & Lemaitre, C. (2022). Metagenomics and Metatranscriptomics. In *From Sequences to Graphs: Discrete Methods and Structures for Bioinformatics* (pp. 147–184). John Wiley & Sons.

Hammad, S., Pu, S., & Jones, P. J. (2016). Current Evidence Supporting the Link Between Dietary Fatty Acids and Cardiovascular Disease. *Lipids*, *51*(5), 507–517. https://doi.org/10.1007/s11745-015-4113-x

Han, G. G., Lee, J.-Y., Jin, G.-D., Park, J., Choi, Y. H., Chae, B. J., Kim, E. B., & Choi, Y.-J. (2017). Evaluating the association between body weight and the intestinal microbiota of weaned piglets via 16S rRNA sequencing. *Applied Microbiology and Biotechnology*, *101*(14), 5903–5911. https://doi.org/10.1007/s00253-017-8304-7

He, K., Wang, Q., Wang, Z., & Pan, Y. (2013). Association study between gene polymorphisms in PPAR signaling pathway and porcine meat quality traits. *Mammalian Genome*, *24*, 322–331.

Heinritz, S. N., Mosenthin, R., & Weiss, E. (2013). Use of pigs as a potential model for research into dietary modulation of the human gut microbiota. *Nutrition Research Reviews*, *26*(2), 191–209. https://doi.org/10.1017/S0954422413000152

Herschman, H. R. (1996). Prostaglandin synthase 2. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, *1299*(1), 125–140.

Hiippala, K., Kainulainen, V., Kalliomäki, M., Arkkila, P., & Satokari, R. (2016). Mucosal prevalence and interactions with the epithelium indicate commensalism of Sutterella spp. *Frontiers in Microbiology*, *7*, 1706.

Holman, D. B., Brunelle, B. W., Trachsel, J., & Allen, H. K. (2017). Meta-analysis To Define a Core Microbiota in the Swine Gut. *mSystems*, *2*(3), e00004-17. https://doi.org/10.1128/mSystems.00004-17

Holman, D. B., Gzyl, K. E., Mou, K. T., & Allen, H. K. (2021). Weaning Age and Its Effect on the Development of the Swine Gut Microbiome and Resistome. *mSystems*, *6*(6), e00682-21. https://doi.org/10.1128/mSystems.00682-21

Hrdlickova, R., Toloue, M., & Tian, B. (2017). RNA‐Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA*, *8*(1), e1364.

Hu, R., Li, S., Diao, H., Huang, C., Yan, J., Wei, X., Zhou, M., He, P., Wang, T., Fu, H., Zhong, C., Mao, C., Wang, Y., Kuang, S., & Tang, W. (2023). The interaction between dietary fiber and gut microbiota, and its effect on pig intestinal health. *Frontiers in Immunology*, *14*, 1095740. https://doi.org/10.3389/fimmu.2023.1095740

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies:

An overview. *Human Immunology*, *82*(11), 801–811. https://doi.org/10.1016/j.humimm.2021.02.012

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, *431*(7011), 931–945. https://doi.org/10.1038/nature03001

Jacobson, A., Lam, L., Rajendram, M., Tamburini, F., Honeycutt, J., Pham, T., Van Treuren, W., Pruss, K., Stabler, S. R., Lugo, K., Bouley, D. M., Vilches-Moure, J. G., Smith, M., Sonnenburg, J. L., Bhatt, A. S., Huang, K. C., & Monack, D. (2018). A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. *Cell Host & Microbe*, *24*(2), 296-307.e7. https://doi.org/10.1016/j.chom.2018.07.002

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., & Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nature Methods*, *12*(4), 351–356. https://doi.org/10.1038/nmeth.3290

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biology*, *17*(1), 239. https://doi.org/10.1186/s13059-016-1103-0

Jami, E., & Mizrahi, I. (2012). Composition and Similarity of Bovine Rumen Microbiota across Individual Animals. *PLoS ONE*, *7*(3), e33306. https://doi.org/10.1371/journal.pone.0033306

Ji, F., & Sadreyev, R. I. (2018). RNA-seq: Basic Bioinformatics Analysis. *Current Protocols in Molecular Biology*, *124*(1), e68. https://doi.org/10.1002/cpmb.68

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., Ong, W. K., Subhraveti, P., Caspi, R., Fulcher, C., Keseler, I. M., & Paley, S. M. (2021). Pathway Tools version 23.0 update: Software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, *22*(1), 109–126.

https://doi.org/10.1093/bib/bbz104

Khanal, P., Maltecca, C., Schwab, C., Fix, J., Bergamaschi, M., & Tiezzi, F. (2020). Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. *Genetics Selection Evolution*, *52*(1), 41. https://doi.org/10.1186/s12711-020-00561-7

Kien, C. L., Blauwiekel, R., Bunn, J. Y., Jetton, T. L., Frankel, W. L., & Holst, J. J. (2007). Cecal Infusion of Butyrate Increases Intestinal Cell Proliferation in Piglets. *The Journal of Nutrition*, *137*(4), 916–922. https://doi.org/10.1093/jn/137.4.916

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kim, H. B., Borewicz, K., White, B. A., Singer, R. S., Sreevatsan, S., Tu, Z. J., & Isaacson, R. E. (2011). Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. *Veterinary Microbiology*, *153*(1–2), 124–133. https://doi.org/10.1016/j.vetmic.2011.05.021

Klampfer, L., Huang, J., Sasazuki, T., Shirasawa, S., & Augenlicht, L. (2003). Inhibition of interferon γ signaling by the short chain fatty acid butyrate. *Molecular Cancer Research*, *1*(11), 855–862.

Kohl, P., Crampin, E. J., Quinn, T. A., & Noble, D. (2010). Systems Biology: An Approach. *Clinical Pharmacology & Therapeutics*, *88*(1), 25–33. https://doi.org/10.1038/clpt.2010.92

Kohl, P., & Noble, D. (2009). Systems biology and the virtual physiological human. *Molecular Systems Biology*, *5*(1), 292. https://doi.org/10.1038/msb.2009.51

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, *28*(24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611

Kumar, A., Tomer, V., & Singh, S. (2019). Lipids: Classification, functions and role in human health. *Think India Journal*, *22*(34), 370–384.

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

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Lê Cao, K.-A., Rohart, F., González, I., Déjean, S., Gautier, B., Bartolo, F., Monget, P., Coquery, J., Yao, F., & Liquet, B. (2016). *mixOmics: Omics Data Integration Project* (6.1.1) [Computer software]. http://mixomics.org/

Le Dinh, P., Van Der Peet-Schwering, C., Ogink, N., & Aarnink, A. (2022). Effect of Diet Composition on Excreta Composition and Ammonia Emissions from Growing-Finishing Pigs. *Animals*, *12*(3), 229. https://doi.org/10.3390/ani12030229

Lehrke, M., & Lazar, M. A. (2005). The many faces of PPARγ. *Cell*, *123*(6), 993–999.

Lei, L., Wang, Z., Li, J., Yang, H., Yin, Y., Tan, B., & Chen, J. (2021). Comparative Microbial Profiles of Colonic Digesta between Ningxiang Pig and Large White Pig. *Animals*, *11*(7), 1862. https://doi.org/10.3390/ani11071862

Leonard, A. E., Kelder, B., Bobik, E. G., Chuang, L.-T., Lewis, C. J., Kopchick, J. J., Mukerji, P., & Huang, Y.-S. (2002). Identification and expression of mammalian long-chain PUFA elongation enzymes. *Lipids*, *37*, 733–740.

Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2014). *IDBA-MTP: A Hybrid MetaTranscriptomic Assembler Based on Protein Information*. 160–172.

Leung, H. C. M., Yiu, S.-M., Parkinson, J., & Chin, F. Y. L. (2013). IDBA-MT: *De Novo* Assembler for Metatranscriptomic Data Generated from Next-Generation Sequencing Technology. *Journal of Computational Biology*, *20*(7), 540–550. https://doi.org/10.1089/cmb.2013.0042

Li, B., Weng, Q., Dong, C., Zhang, Z., Li, R., Liu, J., Jiang, A., Li, Q., Jia, C., & Wu, W. (2018). A key gene, PLIN1, can affect porcine intramuscular fat content based on

transcriptome analysis. *Genes*, *9*(4), 194.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, P.-E., Lo, C.-C., Anderson, J. J., Davenport, K. W., Bishop-Lilly, K. A., Xu, Y., Ahmed, S., Feng, S., Mokashi, V. P., & Chain, P. S. G. (2017). Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform. *Nucleic Acids Research*, *45*(1), 67–80. https://doi.org/10.1093/nar/gkw1027

Li, W., Li, L., Yang, F., Hu, Q., & Xiong, D. (2022). Correlation between gut bacteria Phascolarctobacterium and exogenous metabolite α-linolenic acid in T2DM: a case-control study. *Annals of Translational Medicine*, *10*(19).

Li, X., Li, M., Han, J., Liu, C., Han, X., Wang, K., Qiao, R., Li, X.-L., & Li, X.-J. (2022). Correlation between fat accumulation and fecal microbiota in crossbred pigs. *Journal of Microbiology*, *60*(11), 1077–1085.

Li, X., Mao, M., Zhang, Y., Yu, K., & Zhu, W. (2019). Succinate modulates intestinal barrier function and inflammation response in pigs. *Biomolecules*, *9*(9), 486.

Liu, J., Tan, Y., Cheng, H., Zhang, D., Feng, W., & Peng, C. (2022). Functions of Gut Microbiota Metabolites, Current Status and Future Perspectives. *Aging and Disease*, *13*(4), 1106. https://doi.org/10.14336/AD.2022.0104

López-Bote, C. J. (1998). Sustained utilization of the Iberian pig breed. *Meat Science*, *49*(1), S17–S27. https://doi.org/10.1016/S0309-1740(98)90036-5

Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J., & Martinez-Medina, M. (2017). *Faecalibacterium prausnitzii*: From microbiology to diagnostics and prognostics. *The*

ISME Journal, *11*(4), 841–852. https://doi.org/10.1038/ismej.2016.176

Louis, P., & Flint, H. J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiology Letters*, *294*(1), 1–8. https://doi.org/10.1111/j.1574-6968.2009.01514.x

Louis, P., & Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environmental Microbiology*, *19*(1), 29–41. https://doi.org/10.1111/1462-2920.13589

Louis, P., Hold, G. L., & Flint, H. J. (2014). The gut microbiota, bacterial metabolites and colorectal cancer. *Nature Reviews Microbiology*, *12*(10), 661–672. https://doi.org/10.1038/nrmicro3344

Love, M. I., Anders, S., Kim, V., & Huber, W. (2015). RNA-Seq workflow: Gene-level exploratory analysis and differential expression. *F1000Research*, *4*, 1070. https://doi.org/10.12688/f1000research.7035.1

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lunney, J. K., Van Goor, A., Walker, K. E., Hailstock, T., Franklin, J., & Dai, C. (2021). Importance of the pig as a human biomedical model. *Science Translational Medicine*, *13*(621), eabd5758. https://doi.org/10.1126/scitranslmed.abd5758

Luo, Y., Ren, W., Smidt, H., Wright, A.-D. G., Yu, B., Schyns, G., McCormack, U. M., Cowieson, A. J., Yu, J., He, J., Yan, H., Wu, J., Mackie, R. I., & Chen, D. (2022). Dynamic Distribution of Gut Microbiota in Pigs at Different Growth Stages: Composition and Contribution. *Microbiology Spectrum*, *10*(3), e00688-21. https://doi.org/10.1128/spectrum.00688-21

Lv, Q.-B., Li, S., Zhang, Y., Guo, R., Wang, Y.-C., Peng, Y., & Zhang, X.-X. (2022). A thousand metagenome-assembled genomes of Akkermansia reveal phylogroups and geographical and functional variations in the human gut. *Frontiers in Cellular and Infection Microbiology*, *12*, 957439. https://doi.org/10.3389/fcimb.2022.957439

Macfarlane, G. T., & Macfarlane, S. (2012). Bacteria, Colonic Fermentation, and Gastrointestinal Health. *Journal of AOAC INTERNATIONAL*, *95*(1), 50–60. https://doi.org/10.5740/jaoacint.SGE_Macfarlane

Macfarlane, S., & Macfarlane, G. T. (2003). Regulation of short-chain fatty acid production. *Proceedings of the Nutrition Society*, *62*(1), 67–72. https://doi.org/10.1079/PNS2002207

Mach, N., Berri, M., Estellé, J., Levenez, F., Lemonnier, G., Denis, C., Leplat, J., Chevaleyre, C., Billon, Y., Doré, J., Rogel‐Gaillard, C., & Lepage, P. (2015). Early‐life establishment of the swine gut microbiome and impact on host phenotypes. *Environmental Microbiology Reports*, *7*(3), 554–569. https://doi.org/10.1111/1758-2229.12285

Machiels, K., Joossens, M., Sabino, J., De Preter, V., Arijs, I., Eeckhaut, V., Ballet, V., Claes, K., Van Immerseel, F., & Verbeke, K. (2014). A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut*, *63*(8), 1275–1283.

Maltecca, C., Lu, D., Schillebeeckx, C., McNulty, N. P., Schwab, C., Shull, C., & Tiezzi, F. (2019). Predicting Growth and Carcass Traits in Swine Using Microbiome Data and Machine Learning Algorithms. *Scientific Reports*, *9*(1), 6574. https://doi.org/10.1038/s41598-019-43031-x

Mao, S., Zhang, M., Liu, J., & Zhu, W. (2015). Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: Membership and potential function. *Scientific Reports*, *5*(1), 16116. https://doi.org/10.1038/srep16116

Marchesi, J. R., & Ravel, J. (2015). The vocabulary of microbiome research: A proposal. *Microbiome*, *3*(1), 31, s40168-015-0094–0095. https://doi.org/10.1186/s40168-015-0094-5

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., … Rothberg, J. M.

(2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376–380. https://doi.org/10.1038/nature03959

Marion-Letellier, R., Savoye, G., & Ghosh, S. (2015). Polyunsaturated fatty acids and inflammation. *IUBMB Life*, *67*(9), 659–667. https://doi.org/10.1002/iub.1428

Maron, D. F., Smith, T. J., & Nachman, K. E. (2013). Restrictions on antimicrobial use in food animal production: An international regulatory and economic survey. *Globalization and Health*, *9*(1), 48. https://doi.org/10.1186/1744-8603-9-48

Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S. K., Bignell, A., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Charkhchi, M., Cockburn, A., Da Rin Fiorretto, L., … Flicek, P. (2023). Ensembl 2023. *Nucleic Acids Research*, *51*(D1), D933–D941. https://doi.org/10.1093/nar/gkac958

Martinez, X., Pozuelo, M., Pascal, V., Campos, D., Gut, I., Gut, M., Azpiroz, F., Guarner, F., & Manichanh, C. (2016). MetaTrans: An open-source pipeline for metatranscriptomics. *Scientific Reports*, *6*(1), 26447. https://doi.org/10.1038/srep26447

Martínez-Porchas, M., Villalpando-Canchola, E., & Vargas-Albores, F. (2016). Significant loss of sensitivity and specificity in the taxonomic classification occurs when short 16S rRNA gene sequences are used. *Heliyon*, *2*(9).

Martin-Gallausiaux, C., Marinelli, L., Blottière, H. M., Larraufie, P., & Lapaque, N. (2021). SCFA: Mechanisms and functional importance in the gut. *Proceedings of the Nutrition Society*, *80*(1), 37–49. https://doi.org/10.1017/S0029665120006916

Matsuzaka, T., & Shimano, H. (2009). Elovl6: A new player in fatty acid metabolism and insulin sensitivity. *Journal of Molecular Medicine*, *87*(4), 379–384. https://doi.org/10.1007/s00109-009-0449-0

Maw, S. J., Fowler, V. R., Hamilton, M., & Petchey, A. M. (2003). Physical characteristics of pig fat and their relation to fatty acid composition. *Meat Science*, *63*(2), 185–190. https://doi.org/10.1016/S0309-1740(02)00069-4

McCormack, U. M., Curião, T., Buzoianu, S. G., Prieto, M. L., Ryan, T., Varley, P., Crispie, F.,

Magowan, E., Metzler-Zebeli, B. U., Berry, D., O'Sullivan, O., Cotter, P. D., Gardiner, G. E., & Lawlor, P. G. (2017). Exploring a Possible Link between the Intestinal Microbiota and Feed Efficiency in Pigs. *Applied and Environmental Microbiology*, *83*(15), e00380-17. https://doi.org/10.1128/AEM.00380-17

McDonald, D., Jiang, Y., Balaban, M., Cantrell, K., Zhu, Q., Gonzalez, A., Morton, J. T., Nicolaou, G., Parks, D. H., Karst, S. M., Albertsen, M., Hugenholtz, P., DeSantis, T., Song, S. J., Bartko, A., Havulinna, A. S., Jousilahti, P., Cheng, S., Inouye, M., … Knight, R. (2023). Greengenes2 unifies microbial data in a single reference tree. *Nature Biotechnology*. https://doi.org/10.1038/s41587-023-01845-1

Mi, J., Peng, H., Wu, Y., Wang, Y., & Liao, X. (2019). Diversity and community of methanogens in the large intestine of finishing pigs. *BMC Microbiology*, *19*(1), 83. https://doi.org/10.1186/s12866-019-1459-x

Ministerio de Agricultura, Pesca y Alimentación (MAPA). (2022). *Indicadores económicos carne de cerdo 2022*. https://www.mapa.gob.es/es/ganaderia/estadisticas/indicadoreseconomicossectorporcino2022_tcm30-564427.pdf

Morgan, X. C., & Huttenhower, C. (2012). Chapter 12: Human Microbiome Analysis. *PLoS Computational Biology*, *8*(12), e1002808. https://doi.org/10.1371/journal.pcbi.1002808

Mouse Genome Sequencing Consortium, Waterston, R. H., Lindblad-Toh, K., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., … Lander, E. S. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature*, *420*(6915), 520–562.

Muñoz, M., Bozzi, R., García, F., Núñez, Y., Geraci, C., Crovetti, A., García-Casco, J., Alves, E., Škrlep, M., Charneca, R., Martins, J. M., Quintanilla, R., Tibau, J., Kušec, G., Djurkin-Kušec, I., Mercat, M. J., Riquet, J., Estellé, J., Zimmer, C., … Óvilo, C. (2018). Diversity across major and candidate genes in European local pig breeds.

PLOS ONE, *13*(11), e0207475. https://doi.org/10.1371/journal.pone.0207475

Mutz, K.-O., Heilkenbrinker, A., Lönne, M., Walter, J.-G., & Stahl, F. (2013). Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology*, *24*(1), 22–30.

Naudí, A., Jové, M., Ayala, V., Portero-Otín, M., Barja, G., & Pamplona, R. (2013). Membrane lipid unsaturation as physiological adaptation to animal longevity. *Frontiers in Physiology*, *4*. https://doi.org/10.3389/fphys.2013.00372

Nearing, J. T., Douglas, G. M., Comeau, A. M., & Langille, M. G. I. (2018). Denoising the Denoisers: An independent evaluation of microbiome sequence error-correction approaches. *PeerJ*, *6*, e5364. https://doi.org/10.7717/peerj.5364

Neves, A. L. A., Chen, Y., Lê Cao, K.-A., Mandal, S., Sharpton, T. J., McAllister, T., & Guan, L. L. (2020). Taxonomic and functional assessment using metatranscriptomics reveals the effect of Angus cattle on rumen microbial signatures. *Animal*, *14*(4), 731–744.

Newman, J. C., & Verdin, E. (2014). Ketone bodies as signaling metabolites. *Trends in Endocrinology & Metabolism*, *25*(1), 42–52.

Nguyen, L. N., Ma, D., Shui, G., Wong, P., Cazenave-Gassiot, A., Zhang, X., Wenk, M. R., Goh, E. L., & Silver, D. L. (2014). Mfsd2a is a transporter for the essential omega-3 fatty acid docosahexaenoic acid. *Nature*, *509*(7501), 503–506.

Nieto, R., García-Casco, J., Lara, L., Palma-Granados, P., Izquierdo, M., Hernandez, F., Dieguez, E., Luis Duarte, J., & Batorek-Lukač, N. (2019). Ibérico (Iberian) Pig. In M. Candek-Potokar & R. M. Nieto Linan (Eds.), *European Local Pig Breeds—Diversity and Performance. A study of project TREASURE*. IntechOpen. https://doi.org/10.5772/intechopen.83765

Niu, J., Liu, X., Xu, J., Li, F., Wang, J., Zhang, X., Yang, X., Wang, L., Ma, S., Li, D., Zhu, X., Wang, C., Shi, Y., & Cui, Y. (2023). Effects of Silage Diet on Meat Quality through Shaping Gut Microbiota in Finishing Pigs. *Microbiology Spectrum*, *11*(1), e02416-22. https://doi.org/10.1128/spectrum.02416-22

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*(5), 824–834. https://doi.org/10.1101/gr.213959.116

O'Hea, E. K., & Leveille, G. A. (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. https://doi.org/10.1093/jn/99.3.338

Ojeda, A., Rozas, J., Folch, J. M., & Perez-Enciso, M. (2006). Unexpected High Polymorphism at the FABP4 Gene Unveils a Complex History for Pig Populations. *Genetics*, *174*(4), 2119–2127. https://doi.org/10.1534/genetics.106.063057

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Ouwerkerk, J. P., Aalvink, S., Belzer, C., & De Vos, W. M. (2016). Akkermansia glycaniphila sp. Nov., an anaerobic mucin-degrading bacterium isolated from reticulated python faeces. *International Journal of Systematic and Evolutionary Microbiology*, *66*(11), 4614–4620. https://doi.org/10.1099/ijsem.0.001399

PacBio. (2020, March 25). *Sequencing 101: The Evolution of DNA Sequencing Tools*. PacBio. https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/

Parada Venegas, D., De La Fuente, M. K., Landskron, G., González, M. J., Quera, R., Dijkstra, G., Harmsen, H. J. M., Faber, K. N., & Hermoso, M. A. (2019). Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Frontiers in Immunology*, *10*, 277. https://doi.org/10.3389/fimmu.2019.00277

Paredes-Sabja, D., Raju, D., Torres, J. A., & Sarker, M. R. (2008). Role of small, acid-soluble

spore proteins in the resistance of Clostridium perfringens spores to chemicals. *International Journal of Food Microbiology*, *122*(3), 333–335.

Passols, M., Llobet-Cabau, F., Sebastià, C., Castelló, A., Valdés-Hernández, J., Criado-Mesas, L., Sánchez, A., & Folch, J. M. (2023). Identification of genomic regions, genetic variants and gene networks regulating candidate genes for lipid metabolism in pig muscle. *Animal*, *17*(12), 101033. https://doi.org/10.1016/j.animal.2023.101033

Patil, Y., Gooneratne, R., & Ju, X.-H. (2020). Interactions between host and gut microbiota in domestic pigs: A review. *Gut Microbes*, *11*(3), 310–334. https://doi.org/10.1080/19490976.2019.1690363

Peng, Y., Chen, R., Zhang, Z., Jin, R., Xie, T., Liu, X., Chai, J., Howe, S., Zhao, J., & Li, Y. (2023). Metagenomic and meta-transcriptomic analysis reveal the colonization and expression profile of probiotic strains in humans and animals. *Fermentation*, *9*(5), 417.

Pereira-Marques, J., Hout, A., Ferreira, R. M., Weber, M., Pinto-Ribeiro, I., Van Doorn, L.-J., Knetsch, C. W., & Figueiredo, C. (2019). Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Frontiers in Microbiology*, *10*, 1277.

Pérez-Montarelo, D., Fernández, A., Barragán, C., Noguera, J. L., Folch, J. M., Rodríguez, M. C., Óvilo, C., Silió, L., & Fernández, A. I. (2013). Transcriptional characterization of porcine leptin and leptin receptor genes. *PLoS One*, *8*(6), e66398.

Petersen, L. M., Martin, I. W., Moschetti, W. E., Kershaw, C. M., & Tsongalis, G. J. (2019). Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *Journal of Clinical Microbiology*, *58*(1), e01315-19. https://doi.org/10.1128/JCM.01315-19

Poklukar, K., Mestre, C., Škrlep, M., Čandek-Potokar, M., Ovilo, C., Fontanesi, L., Riquet, J., Bovo, S., Schiavo, G., Ribani, A., Muñoz, M., Gallo, M., Bozzi, R., Charneca, R., Quintanilla, R., Kušec, G., Mercat, M.-J., Zimmer, C., Razmaite, V., … Servin, B.

(2023). A meta-analysis of genetic and phenotypic diversity of European local pig breeds reveals genomic regions associated with breed differentiation for production traits. *Genetics Selection Evolution*, *55*(1), 88. https://doi.org/10.1186/s12711-023-00858-3

Puig-Oliveras, A., Ballester, M., Corominas, J., Revilla, M., Estellé, J., Fernández, A. I., Ramayo-Caldas, Y., & Folch, J. M. (2014). A Co-Association Network Analysis of the Genetic Determination of Pig Conformation, Growth and Fatness. *PLoS ONE*, *9*(12), e114862. https://doi.org/10.1371/journal.pone.0114862

Puig-Oliveras, A., Revilla, M., Castelló, A., Fernández, A. I., Folch, J. M., & 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. https://doi.org/10.1038/srep31803

Qin, X., Swertfeger, D. K., Zheng, S., Hui, D. Y., & Tso, P. (1998). Apolipoprotein AIV: a potent endogenous inhibitor of lipid oxidation. *American Journal of Physiology-Heart and Circulatory Physiology*, *274*(5), H1836–H1840.

Qiu, H., Zhao, S., Xu, X., Yerle, M., & Liu, B. (2008). Assignment and expression patterns of porcine muscle-specific isoform of phosphoglycerate mutase gene. *Journal of Genetics and Genomics*, *35*(5), 257–260.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Ramayo-Caldas, Y., Crespo-Piazuelo, D., Morata, J., González-Rodríguez, O., Sebastià, C., Castello, A., Dalmau, A., Ramos-Onsins, S., Alexiou, K. G., & Folch, J. M. (2023). Copy Number Variation on ABCC2-DNMBP Loci Affects the Diversity and Composition of the Fecal Microbiota in Pigs. *Microbiology Spectrum*, *11*(4), e05271-22.

Ramayo-Caldas, Y., Mach, N., Lepage, P., Levenez, F., Denis, C., Lemonnier, G., Leplat,

J.-J., Billon, Y., Berri, M., Doré, J., Rogel-Gaillard, C., & Estellé, J. (2016). Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. *The ISME Journal*, *10*(12), 2973–2977. https://doi.org/10.1038/ismej.2016.77

Ramayo-Caldas, Y., Mercadé, A., Castelló, A., Yang, B., Rodríguez, C., Alves, E., Díaz, I., Ibáñez-Escriche, N., Noguera, J. L., & Pérez-Enciso, M. (2012). Genome-wide association study for intramuscular fatty acid composition in an Iberian× Landrace cross. *Journal of Animal Science*, *90*(9), 2883–2893.

Ramos-Onsins, S. E., Burgos-Paz, W., Manunza, A., & Amills, M. (2014). Mining the pig genome to investigate the domestication process. *Heredity*, *113*(6), 471–484. https://doi.org/10.1038/hdy.2014.68

Ranea-Robles, P., Chen, H., Stauffer, B., Yu, C., Bhattacharya, D., Friedman, S. L., Puchowicz, M., & Houten, S. M. (2021). The peroxisomal transporter ABCD3 plays a major role in hepatic dicarboxylic fatty acid metabolism and lipid homeostasis. *Journal of Inherited Metabolic Disease*, *44*(6), 1419–1433.

Regueira-Iglesias, A., Balsa-Castro, C., Blanco-Pintos, T., & Tomás, I. (2023). Critical review of 16S rRNA gene sequencing workflow in microbiome studies: From primer selection to advanced data analysis. *Molecular Oral Microbiology*, *38*(5), 347–399. https://doi.org/10.1111/omi.12434

Reverter, A., Ballester, M., Alexandre, P. A., Mármol-Sánchez, E., Dalmau, A., Quintanilla, R., & Ramayo-Caldas, Y. (2021). A gene co-association network regulating gut microbial communities in a Duroc pig population. *Microbiome*, *9*(1), 52. https://doi.org/10.1186/s40168-020-00994-8

Reverter, A., & Chan, E. K. (2008). Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics*, *24*(21), 2491–2497.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray

studies. *Nucleic Acids Research*, *43*(7), e47–e47. https://doi.org/10.1093/nar/gkv007

Roediger, W. (1982). Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology*, *83*(2), 424–429.

Rohart, F., Gautier, B., Singh, A., & Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Computational Biology*, *13*(11), e1005752. https://doi.org/10.1371/journal.pcbi.1005752

Rotmistrovsky, K., & Agarwala, R. (2011). BMTagger: Best Match Tagger for removing human reads from metagenomics datasets. *NCBI/NLM, National Institutes of Health*.

Rubin, C.-J., Megens, H.-J., Barrio, A. M., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, Ö., Jern, P., Jørgensen, C. B., Archibald, A. L., Fredholm, M., Groenen, M. A. M., & Andersson, L. (2012). Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences*, *109*(48), 19529–19536. https://doi.org/10.1073/pnas.1217149109

Ryu, Y. C., Choi, Y. M., Lee, S. H., Shin, H. G., Choe, J. H., Kim, J. M., Hong, K. C., & Kim, B. C. (2008). Comparing the histochemical characteristics and meat quality traits of different pig breeds. *Meat Science*, *80*(2), 363–369. https://doi.org/10.1016/j.meatsci.2007.12.020

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463

Schettini, G. P., Peripolli, E., Alexandre, P. A., Dos Santos, W. B., Pereira, A. S. C., de Albuquerque, L. G., Baldi, F., & Curi, R. A. (2022). Transcriptome profile reveals genetic and metabolic mechanisms related to essential fatty acid content of intramuscular longissimus thoracis in nellore cattle. *Metabolites*, *12*(5), 471.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing

and Comparing Microbial Communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schmerold, I., Van Geijlswijk, I., & Gehring, R. (2023). European regulations on the use of antibiotics in veterinary medicine. *European Journal of Pharmaceutical Sciences*, *189*, 106473. https://doi.org/10.1016/j.ejps.2023.106473

Schmidt, H., & Jirstrand, M. (2006). Systems Biology Toolbox for MATLAB: A computational platform for research in systems biology. *Bioinformatics*, *22*(4), 514–515. https://doi.org/10.1093/bioinformatics/bti799

Schook, L. B., Beever, J. E., Rogers, J., Humphray, S., Archibald, A., Chardon, P., Milan, D., Rohrer, G., & Eversole, K. (2005). Swine Genome Sequencing Consortium (SGSC): A strategic roadmap for sequencing the pig genome. *Comparative and Functional Genomics*, *6*(4), 251–255.

Schulz, M. H., Zerbino, D. R., Vingron, M., & Birney, E. (2012). *Oases:* robust *de novo* RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*, *28*(8), 1086–1092. https://doi.org/10.1093/bioinformatics/bts094

Scollan, N. D., Price, E. M., Morgan, S. A., Huws, S. A., & Shingfield, K. J. (2017). Can we improve the nutritional quality of meat? *Proceedings of the Nutrition Society*, *76*(4), 603–618. https://doi.org/10.1017/S0029665117001112

Sebastià, C., Folch, J. M., Ballester, M., Estellé, J., Passols, M., Muñoz, M., García-Casco, J. M., Fernández, A. I., Castelló, A., Sánchez, A., & Crespo-Piazuelo, D. (2024). Interrelation between gut microbiota, SCFA, and fatty acid composition in pigs. *mSystems*, *9*(1), e01049-23. https://doi.org/10.1128/msystems.01049-23

Sharpe, A., & McKenzie, M. (2018). Mitochondrial Fatty Acid Oxidation Disorders Associated with Short-Chain Enoyl-CoA Hydratase (ECHS1) Deficiency. *Cells*, *7*(6), 46. https://doi.org/10.3390/cells7060046

Simon, J.-C., Marchesi, J. R., Mougel, C., & Selosse, M.-A. (2019). Host-microbiota interactions: From holobiont theory to analysis. *Microbiome*, *7*(1), 5. https://doi.org/10.1186/s40168-019-0619-4

Siva, N. (2008). 1000 Genomes project. *Nature Biotechnology*, *26*(3), 256–257.

Smink, W., Gerrits, W. J. J., Gloaguen, M., Ruiter, A., & van Baal, J. (2012). Linoleic and α-linolenic acid as precursor and inhibitor for the synthesis of long-chain polyunsaturated fatty acids in liver and brain of growing pigs. *Animal*, *6*(2), 262–270. https://doi.org/10.1017/S1751731111001479

Su, Y., Yao, W., Perez-Gutierrez, O. N., Smidt, H., & Zhu, W.-Y. (2008). Changes in abundance of Lactobacillus spp. And Streptococcus suis in the stomach, jejunum and ileum of piglets after weaning. *FEMS Microbiology Ecology*, *66*(3), 546–555.

Sun, Y., Zhou, L., Fang, L., Su, Y., & Zhu, W. (2015). Responses in colonic microbial community and gene expression of pigs to a long-term high resistant starch diet. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00877

Tamames, J., & Puente-Sánchez, F. (2019). SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline. *Frontiers in Microbiology*, *9*, 3349. https://doi.org/10.3389/fmicb.2018.03349

The C. elegans Sequencing Consortium. (1998). Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. *Science*, *282*(5396), 2012–2018. https://doi.org/10.1126/science.282.5396.2012

The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Garmiri, P., Da Costa Gonzales, L. J., Hatton-Ellis, E., Hussein, A., … Zhang, J. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), D523–D531. https://doi.org/10.1093/nar/gkac1052

Thomas, S., Mercogliano, G., & Prendergast, G. C. (2022). Bin1 targeted immunotherapy alters the status of the enteric neurons and the microbiome during ulcerative colitis treatment. *PLOS ONE*, *17*(11), e0276910. https://doi.org/10.1371/journal.pone.0276910

Tietze, L., Christ, M., Yu, J., Stock, P., Nickel, S., Schulze, A., Bartels, M., Tautenhahn,

H.-M., & Christ, B. (2024). Approaching Thrombospondin-1 as a Potential Target for Mesenchymal Stromal Cells to Support Liver Regeneration after Partial Hepatectomy in Mouse and Humans. *Cells*, *13*(6), 529.

Truong, D. T., Franzosa, E. A., Tickle, T. L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., & Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods*, *12*(10), 902–903. https://doi.org/10.1038/nmeth.3589

Tsai, T.-H., Chen, E., Li, L., Saha, P., Lee, H.-J., Huang, L.-S., Shelness, G. S., Chan, L., & Chang, B. H.-J. (2017). The constitutive lipid droplet protein PLIN2 regulates autophagy in liver. *Autophagy*, *13*(7), 1130–1144.

Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., Solovyev, V. V., Rubin, E. M., Rokhsar, D. S., & Banfield, J. F. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, *428*(6978), 37–43. https://doi.org/10.1038/nature02340

Ugarte, A., Vicedomini, R., Bernardes, J., & Carbone, A. (2018). A multi-source domain annotation pipeline for quantitative metagenomic and metatranscriptomic functional profiling. *Microbiome*, *6*(1), 149. https://doi.org/10.1186/s40168-018-0532-2

Uritskiy, G. V., DiRuggiero, J., & Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*, *6*(1), 158. https://doi.org/10.1186/s40168-018-0541-1

Valdés-Hernández, J., Folch, J. M., Crespo-Piazuelo, D., Passols, M., Sebastià, C., Criado-Mesas, L., Castelló, A., Sánchez, A., & Ramayo-Caldas, Y. (2024). Identification of candidate regulatory genes for intramuscular fatty acid composition in pigs by transcriptome analysis. *Genetics Selection Evolution*, *56*(1), 12. https://doi.org/10.1186/s12711-024-00882-x

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. (2023). Global analysis of the association between pig muscle fatty acid composition

and gene expression using RNA-Seq. *Scientific Reports*, *13*(1), 535. https://doi.org/10.1038/s41598-022-27016-x

Van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The Third Revolution in Sequencing Technology. *Trends in Genetics*, *34*(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.008

Varona, L. (2017). Mejora genética porcina: Más allá de la genómica. *Suis*, *137*, 12–17.

Vasquez, R., Oh, J. K., Song, J. H., & Kang, D.-K. (2022). Gut microbiome-produced metabolites in pigs: A review on their biological functions and the influence of probiotics. *Journal of Animal Science and Technology*, *64*(4), 671–695. https://doi.org/10.5187/jast.2022.e58

Vega, R. B., Huss, J. M., & Kelly, D. P. (2000). The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor α in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Molecular and Cellular Biology*, *20*(5), 1868–1876.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., … Smith, H. O. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science*, *304*(5667), 66–74. https://doi.org/10.1126/science.1093857

Wallace, R. J., Sasson, G., Garnsworthy, P. C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A. R., Strozzi, F., Biscarini, F., Snelling, T. J., Saunders, N., Potterton, S. L., Craigon, J., Minuti, A., Trevisi, E., Callegari, M. L., Cappelli, F. P., Cabezas-Garcia, E. H., … Mizrahi, I. (2019). A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Science Advances*, *5*(7), eaav8391. https://doi.org/10.1126/sciadv.aav8391

Wang, C., Shi, C., Zhang, Y., Song, D., Lu, Z., & Wang, Y. (2018). Microbiota in fermented feed and swine gut. *Applied Microbiology and Biotechnology*, *102*(7), 2941–2948. https://doi.org/10.1007/s00253-018-8829-4

Wang, C., Wei, S., Chen, N., Xiang, Y., Wang, Y., & Jin, M. (2022). Characteristics of gut microbiota in pigs with different breeds, growth periods and genders. *Microbial Biotechnology*, *15*(3), 793–804. https://doi.org/10.1111/1751-7915.13755

Wang, J., Fan, H., Han, Y., Zhao, J., & Zhou, Z. (2016). Characterization of the microbial communities along the gastrointestinal tract of sheep by 454 pyrosequencing analysis. *Asian-Australasian Journal of Animal Sciences*, *30*(1), 100–110. https://doi.org/10.5713/ajas.16.0166

Wang, J., Li, Y., Cao, C., Yang, R., He, M., Yan, J., Huang, P., Tan, B., & Fan, Z. (2023). The periparturient gut microbiota's modifications in shaziling sows concerning bile acids. *Metabolites*, *13*(1), 68.

Wang, X., Tsai, T., Deng, F., Wei, X., Chai, J., Knapp, J., Apple, J., Maxwell, C. V., Lee, J. A., Li, Y., & Zhao, J. (2019). Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. *Microbiome*, *7*(1), 109. https://doi.org/10.1186/s40168-019-0721-7

Ward, D. M., Weller, R., & Bateson, M. M. (1990). *165 rRNA sequences reveal numerous uncultured microorganisms in a natural community*.

Wei, Y., Ma, X., Zhao, J., Wang, X., & Gao, C. (2023). Succinate metabolism and its regulation of host-microbe interactions. *Gut Microbes*, *15*(1), 2190300. https://doi.org/10.1080/19490976.2023.2190300

Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, *173*(2), 697–703. https://doi.org/10.1128/jb.173.2.697-703.1991

Westreich, S. T., Treiber, M. L., Mills, D. A., Korf, I., & Lemay, D. G. (2018). SAMSA2: A standalone metatranscriptome analysis pipeline. *BMC Bioinformatics*, *19*(1), 175. https://doi.org/10.1186/s12859-018-2189-z

White, S. (2011). From Globalized Pig Breeds to Capitalist Pigs: A Study in Animal Cultures and Evolutionary History. *Environmental History*, *16*(1), 94–120. https://doi.org/10.1093/envhis/emq143

Woese, C. (1998). The universal ancestor. *Proceedings of the National Academy of Sciences*, *95*(12), 6854–6859. https://doi.org/10.1073/pnas.95.12.6854

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*, *74*(11), 5088–5090. https://doi.org/10.1073/pnas.74.11.5088

Woese, C. R., Fox, G. E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B. J., & Stahl, D. (1975). Conservation of primary structure in 16S ribosomal RNA. *Nature*, *254*, 83–86.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of t He National Academy of Sciences*, *87*(12), 4576–4579.

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1), 257. https://doi.org/10.1186/s13059-019-1891-0

Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Richardson, R. I., & Sheard, P. R. (1999). Manipulating meat quality and composition. *Proceedings of the Nutrition Society*, *58*(2), 363–370. https://doi.org/10.1017/S0029665199000488

Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Sheard, P. R., Richardson, R. I., Hughes, S. I., & Whittington, F. M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Science*, *78*(4), 343–358. https://doi.org/10.1016/j.meatsci.2007.07.019

Wood, J. D., Nute, G. R., Richardson, R. I., Whittington, F. M., Southwood, O., Plastow, G., Mansbridge, R., Da Costa, N., & Chang, K. C. (2004). Effects of breed, diet and muscle on fat deposition and eating quality in pigs. *Meat Science*, *67*(4), 651–667. https://doi.org/10.1016/j.meatsci.2004.01.007

Wright, G. D. (2007). The antibiotic resistome: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, *5*(3), 175–186.

Xiao, L., Estellé, J., Kiilerich, P., Ramayo-Caldas, Y., Xia, Z., Feng, Q., Liang, S., Pedersen, A. Ø., Kjeldsen, N. J., Liu, C., Maguin, E., Doré, J., Pons, N., Le Chatelier, E., Prifti,

E., Li, J., Jia, H., Liu, X., Xu, X., … Wang, J. (2016). A reference gene catalogue of the pig gut microbiome. *Nature Microbiology*, *1*(12), 16161. https://doi.org/10.1038/nmicrobiol.2016.161

Yamashita, H., Kaneyuki, T., & Tagawa, K. (2001). Production of acetate in the liver and its utilization in peripheral tissues. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1532*(1–2), 79–87. https://doi.org/10.1016/S1388-1981(01)00117-2

Yang, B., Wang, Y., & Qian, P.-Y. (2016). Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics*, *17*(1), 135. https://doi.org/10.1186/s12859-016-0992-y

Yang, H., Huang, X., Fang, S., He, M., Zhao, Y., Wu, Z., Yang, M., Zhang, Z., Chen, C., & Huang, L. (2017). Unraveling the fecal microbiota and metagenomic functional capacity associated with feed efficiency in pigs. *Frontiers in Microbiology*, *8*, 1555.

Yang, H., Huang, X., Fang, S., Xin, W., Huang, L., & Chen, C. (2016). Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct fatness. *Scientific Reports*, *6*(1), 27427. https://doi.org/10.1038/srep27427

Yang, H., Yang, M., Fang, S., Huang, X., He, M., Ke, S., Gao, J., Wu, J., Zhou, Y., Fu, H., Chen, C., & Huang, L. (2018). Evaluating the profound effect of gut microbiome on host appetite in pigs. *BMC Microbiology*, *18*(1), 215. https://doi.org/10.1186/s12866-018-1364-8

Ye, Y., & Tang, H. (2016). Utilizing de Bruijn graph of metagenome assembly for metatranscriptome analysis. *Bioinformatics*, *32*(7), 1001–1008. https://doi.org/10.1093/bioinformatics/btv510

Yu, K., Zhang, Y., Chen, H., & Zhu, W. (2019). Hepatic Metabolomic and Transcriptomic Responses Induced by Cecal Infusion of Sodium Propionate in a Fistula Pig Model. *Journal of Agricultural and Food Chemistry*, *67*(47), 13073–13081. https://doi.org/10.1021/acs.jafc.9b05070

Yu, T., Tian, X., Li, D., He, Y., Yang, P., Cheng, Y., Zhao, X., Sun, J., & Yang, G. (2023). Transcriptome, proteome and metabolome analysis provide insights on fat deposition and meat quality in pig. *Food Research International*, *166*, 112550.

Zappaterra, M., Ros-Freixedes, R., Estany, J., & Davoli, R. (2018). Association study highlights the influence of ELOVL fatty acid elongase 6 gene region on backfat fatty acid composition in Large White pig breed. *Animal*, *12*(12), 2443–2452. https://doi.org/10.1017/S1751731118000484

Zeng, X., Li, S., Liu, L., Cai, S., Ye, Q., Xue, B., Wang, X., Zhang, S., Chen, F., Cai, C., Wang, F., & Zeng, X. (2023). Role of functional fatty acids in modulation of reproductive potential in livestock. *Journal of Animal Science and Biotechnology*, *14*(1), 24. https://doi.org/10.1186/s40104-022-00818-9

Zeng, X., Sunkara, L. T., Jiang, W., Bible, M., Carter, S., Ma, X., Qiao, S., & Zhang, G. (2013). Induction of Porcine Host Defense Peptide Gene Expression by Short-Chain Fatty Acids and Their Analogs. *PLoS ONE*, *8*(8), e72922. https://doi.org/10.1371/journal.pone.0072922

Zhang, Y., Chen, H., Zhu, W., & Yu, K. (2019). Cecal Infusion of Sodium Propionate Promotes Intestinal Development and Jejunal Barrier Function in Growing Pigs. *Animals*, *9*(6), 284. https://doi.org/10.3390/ani9060284

Zhang, Y. K., Zhang, X. X., Li, F. D., Li, C., Li, G. Z., Zhang, D. Y., Song, Q. Z., Li, X. L., Zhao, Y., & Wang, W. M. (2021). Characterization of the rumen microbiota and its relationship with residual feed intake in sheep. *Animal*, *15*(3), 100161. https://doi.org/10.1016/j.animal.2020.100161

Zhao, G., Xiang, Y., Wang, X., Dai, B., Zhang, X., Ma, L., Yang, H., & Lyu, W. (2022). Exploring the Possible Link between the Gut Microbiome and Fat Deposition in Pigs. *Oxidative Medicine and Cellular Longevity*, *2022*, 1–13. https://doi.org/10.1155/2022/1098892

Zhao, W., Wang, Y., Liu, S., Huang, J., Zhai, Z., He, C., Ding, J., Wang, J., Wang, H., Fan, W., Zhao, J., & Meng, H. (2015). The dynamic distribution of porcine microbiota

across different ages and gastrointestinal tract segments. *PLOS ONE*, *10*(2), e0117441. https://doi.org/10.1371/journal.pone.0117441

Zhou, Q., Tao, X., Guo, F., Wu, Y., Deng, D., Lv, L., Dong, D., Shang, D., & Xiang, H. (2023). Tryptophan metabolite norharman secreted by cultivated Lactobacillus attenuates acute pancreatitis as an antagonist of histone deacetylases. *BMC Medicine*, *21*(1), 329. https://doi.org/10.1186/s12916-023-02997-2

Zhu, Z., Zhu, L., & Jiang, L. (2022). Dynamic regulation of gut Clostridium-derived short-chain fatty acids. *Trends in Biotechnology*, *40*(3), 266–270. https://doi.org/10.1016/j.tibtech.2021.10.005

7. Annexes

7.1. Supplementary material Paper I: “Gene co-expression network analysis for porcine intramuscular fatty acid composition”

Supplementary Table S1. Gene pathways and source.

Pathway	**Source**
ABC transporters in lipid homeostasis pathway	Reactome-R-HSA-1369062
Activating transcription factor binding	BioSystem-GO:0033613
Adipocytokine signaling pathway	KEGG-hsa04920
Adipogenesis pathway	WikiPathways-WP236
Alanine aspartate and glutamate metabolism pathway	ConsensusPathDB_path_hsa00250_KEGG
Alditol metabolic process	BioSystem-GO:0019400
Amine catabolic process	BioSystem-GO:0009310
Amino Acid metabolism pathway	Pathcards-genecardsDB
AMPK signaling pathway	KEGG-hsa04152
Angiopoietin Like Protein 8 Regulatory pathway	WikiPathways-WP3915
Arachidonic acid metabolism pathway	KEGG-hsa00590
Beta-Oxidation pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa_M00087
Beta-Oxidation pathway of pristanoyl-CoA pathway	BioSystems_NCBI-REACTOME-R-HSA-389887
Beta-Oxidation pathway, acyl-CoA synthesis pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa_M00086
Biosynthesis of the N-glycan precursor(dolichol lipid-linked oligosaccharide-LLO) and transfer to a nascent protein pathway	Reactome-R-HSA-446193
Biosynthesis of unsaturated Fatty acids pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa01040
Canonical Wnt signaling pathway	BioSystems-Interaction-Database-BSID-138032
Carbon metabolism pathway	Pathcards-genecardsDB

Pathway	Source
Carboxylic acid biosynthetic process	BioSystem-GO:0046394
Carboxylic acid catabolic process	BioSystem-GO:0046395
Cellular amino acid catabolic process	BioSystem-GO:0009063
Cellular lipid metabolic process	BioSystem-GO:0044255
Cellular response to insulin stimulus	BioSystem-GO:0032869
Cholesterol homeostasis	BioSystem-GO:0042632
Cholesterol metabolic process	Informatics-jax-org-vocab-GO_0008203-MGI
Citrate cycle-first carbon oxidation-oxaloacetate=>2-oxoglutarate pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa_M00010
Cutin-suberine and wax biosynthesis pathway	KEGG-map00073
Cytochrome P450-arranged by substrate type pathway	Pathcards-genecardsDB
Fatty Acid Biosynthesis pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa00061
Fatty acid degradation pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa00071
Fatty acid elongation pathway	BioSystems_NCBI-KEGG-PATHWAY_hsa00062
Fatty acid metabolic process	BioSystem-GO:0006631
Fatty acid-triacylglycerol and Ketone body metabolism pathway	Reactome-R-HSA-535734
Folate biosynthesis pathway	ConsensusPathDB_path_hsa00790_KEGG
Free fatty acid receptors pathway	Reactome-R-HSA-444209
Glucocorticoid biosynthetic process	BioSystem-GO:0006704
Glucose homeostasis	BioSystem-GO:0042593
Glycerophospholipid biosynthesis pathway	Pathcards-genecardsDB
Growth hormone receptor signaling	BioSystem-GO:0060396
Insulin receptor signaling	BioSystems_NCBI-GO:0008286
Lipid biosynthetic process	BioSystem-GO:0008610

Pathway	Source
Lipid digestion-mobilization and transport pathway	Reactome-R-HSA-73923
Lipid homeostasis	BioSystem-GO:0055088
Lipid metabolic process	Informatics-jax-org-vocab-GO_00066 29-MGI
Lipid oxidation	BioSystems_NCBI-GO:0034440
Lipoprotein oxidation	BioSystems_NCBI-GO:0042161
Long-chain fatty acid metabolic process	BioSystem-GO:0001676
Lysosphingolipid and LPA receptors pathway	Reactome-R-HSA-419408
Metabolism of lipids pathway	Reactome-R-HSA-556833
mTOR signaling pathway	BioSystems_NCBI-mtor_4-pathway
Negative regulation of lipoprotein lipid oxidation	BioSystems_NCBI-GO:0060588
Non-alcoholic fatty liver disease (NAFLD) pathway	KEGG-hsa04932
Organic acid biosynthetic process	BioSystem-GO:0016053
Organic acid catabolic process	BioSystem-GO:0016054
Organophosphate metabolic process	BioSystem-GO:0019637
Plasma lipoprotein oxidation	BioSystems_NCBI-GO:0034441
PPAR signaling pathway	Wikipathways-WP3942
Pyruvate metabolism pathway	KEGG-hsa00620
Pyruvate oxidation-pyruvate=>acetyl-CoA pathway	BioSystems_NCBI-KEGG-PATHWAY _hsa_M00307
Regulation of Cholesterol biosynthesis pathway by SREBP (SREBF)	Reactome-Ptahway-R-HSA-1655829
Regulation of fatty acid Beta-Oxidation	BioSystems_NCBI-GO:0031998
Regulation of fatty acid oxidation	BioSystems_NCBI-GO:0046320
Regulation of insulin receptor signaling	BioSystems_NCBI-GO:0046626
Regulation of lipid biosynthetic process	BioSystem-GO:0046890
Regulation of lipid metabolism by PPARalpha pathway	Reactome-R-HSA-400206
Regulation of lipolysis in adipocytes	Pathcards-genecardsDB

Pathway	Source
Regulation of lipoprotein lipid oxidation	BioSystems_NCBI-GO:0060587
Response to retinoic acid	BioSystem-GO:0032526
Reverse cholesterol transport	BioSystem-GO:0043691
Role of phospholipids in phagocytosis pathway	Reactome-R-HSA-2029485
Steroid hormone biosynthesis pathway	Pathcards-genecardsDB
Steroid metabolic process	DAVID-GOTERM_BP_FAT-GO:0008202
Sterol metabolic process	DAVID-GOTERM_BP_FAT-GO:0016125
The fatty acid cycling model pathway	Reactome-R-HSA-167826
Transport of Fatty acids pathway	Reactome-R-HSA-804914
Triacylglyceride Synthesis pathway	BioSystems_NCBI-WikiPathways-WP325

Supplementary Table S2. LogCPM expression of lipid metabolism genes.

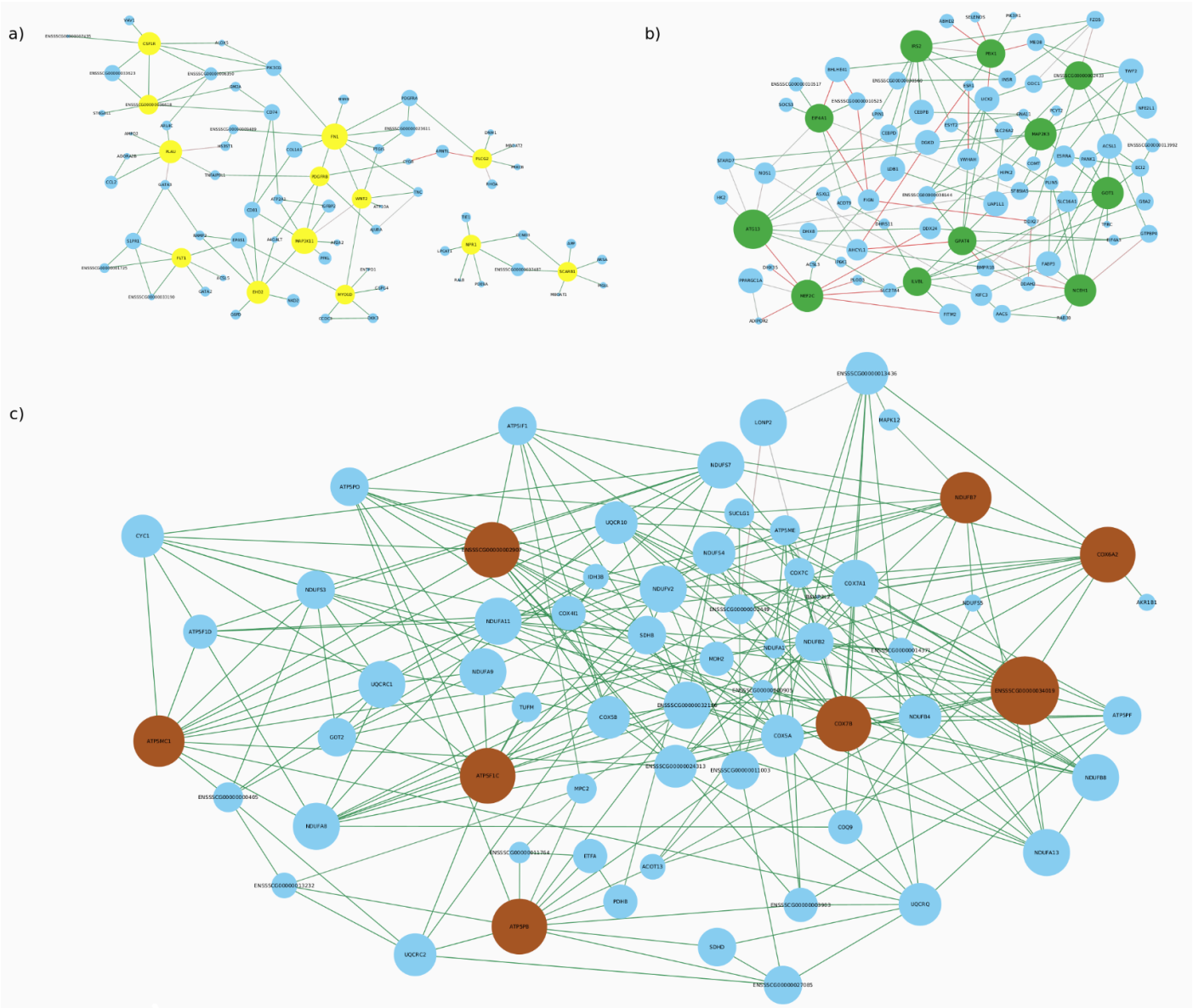
Available in GitHub: https://github.com/cristinasebastia294/Thesis-supplementary-stuff

Supplementary Table S3. Gene - module membership.

Available in GitHub: https://github.com/cristinasebastia294/Thesis-supplementary-stuff

Supplementary Table S4. Significantly enriched pathways

Available in GitHub: https://github.com/cristinasebastia294/Thesis-supplementary-stuff



Supplementary Figure S1. Co-expression networks of relevant modules. Edges in green show positive relationships and edges in red show negative relationships. a) Yellow module. b) Green module. c) Brown module.

7.2. Supplementary material Paper II: “Interrelation between gut microbiota, SCFA and fatty acid composition in pigs”

All the supplementary information regarding this article is already published in:

https://journals.asm.org/doi/full/10.1128/msystems.01049-23

7.3. Supplementary material Paper III: “Metagenomic and metatranscriptomic interplay between the pig gut microbiota and short-chain fatty acids

Supplementary Table S1. Main programs and functions of the bioinformatic tools used for the metagenomic and metatranscriptomic analyses.

a) Programs included in metaWRAP for metagenomic analysis.

Module	**Programs included**	**Function**	**Input**	**Output**
read_qc	Trim-galore v0.6.6, bmtagger v3.101, FASTQC v0.11.9	Filters out low quality reads and host sequences	Raw FASTQ files	Cleaned FASTQ files, QC reports
assembly	MegaHit v1.2.9, metaSPAdes v3.14.1, QUAST v4.1	Assembles the forward and reverse reads sets	Cleaned FASTQ files (forward and reverse)	A single FASTA file per sample, report
binning	CONCOCT v1.1.0, MaxBin v2.2.5, MetaBAT v2.12.1, bwa v0.7.17, Samtools v1.9, CheckM 1.0.18	Bins the assembly	FASTA file	Folders for each binner with the sequences divided in the different bins
bin_refinement	Own algorithm, CheckM v1.0.18	Chooses the best binner and uses an hybrid algorithm to improve the results	Folders with all the bins	Folders with the refined bins, statistics file
quant_bins	Salmon v0.13.1, Seaborn v0.9.0	Quantifies the abundance of each bin and the distribution of the genomes in the samples	Folder with the refined bins	File tab with the abundances, abundance heatmap
reassemble_bins	bwa v0.7.17, SPAdes v3.14.1, CheckM v1.0.18	Improves the binning results	Folders with the refined bins	Folders with the improved results, statistics file and graphs
classify_bins	MegaBLAST v2.6.0, taxator-tk v1.3.3e	Determines the taxonomy of each bin	Folder with the reassembled bins	Tab file with the classification of each bin
annotate_bins	PROKKA v1.13	Functionally annotates each contig of each bin	Folder with the reassembled bins	Different files with the annotations
kraken2 (optative)	KRAKEN2 v2.1.1, KronaTools v2.7.1	Estimates the taxonomy with the non-binned samples	FASTQ or FASTA files	Kraken2, krak2, and krona files, report

Module	Programs included	Function	Input	Output
blobology (optative)	MegaBLAST v2.6.0, bowtie2 v2.3.5.1	Estimates the taxonomy with the non-binned samples and creates plots of the abundance and GC content, annotating them with phylogenetic or bin information	FASTQ and FASTA files, folders with all the bins	Blob plots, taxonomy information files

b) Programs included in SAMSA2 for metatranscriptomic analysis.

Program/Scripts	Function	Input	Output
Trimmomatic v0.36	Filters out low quality sequences	Raw FASTQ files	Cleaned FASTQ files
PEAR v0.9.8	Merges forward and reverse files	Forward and reverse FASTQ files	Assembled FASTQ files
SortMeRNA v2.1	Removes the ribosomal sequences	Assembled FASTQ files	Ribodepleted FASTQ files
DIAMOND v0.8.3	Aligns sequences (similarly to BLAST, but faster)	Ribodepleted FASTQ files	Annotated files, against RefSeq and SEED Subsystems databases
Python scripts: *DIAMOND_analysis_counter.py*, *DIAMOND_subsystems_analysis_counter.py* and *DIAMOND_specific_organism_counter.py*	Aggregate DIAMOND results for each database result	Annotated files	Aggregated files, different tsv tables for RefSeq organism, RefSeq function and SEED Subsystems databases
R scripts: *diversity_stats.R*, *make_DESeq_PCA.R*, *make_DESeq_heatmap.R*, *make_combined_graphs.R*, *run_DESeq_stats.R* and *volcano_plot.R*	Perform some statistical analyses and plots of the different results	Aggregated files	Different tables and graphs

Supplementary Table S2. Results of the differential abundance and expression analysis for the content in acetic, propionic and butyric acids.

Available in GitHub: https://github.com/cristinasebastia294/Thesis-supplementary-stuff

7.4. Supplementary material Paper IV.

Supplementary Table S1. Genes involved in energy metabolism, lipid metabolism or immunity based on GO keywords retrieved with DAVID database.

ID	Gene Name	Short Name
ENSSSCG00000036808	3-hydroxy-3-methylglutaryl-CoA synthase 2	HMGCS2
ENSSSCG00000012347	5'-aminolevulinate synthase 2	ALAS2
ENSSSCG00000062489		
ENSSSCG00000041313		
ENSSSCG00000052903		
ENSSSCG00000053150		
ENSSSCG00000053251		
ENSSSCG00000006359	ADAM metallopeptidase with thrombospondin type 1 motif 4	ADAMTS4
ENSSSCG00000031791	ADP ribosylation factor 4	ARF4
ENSSSCG00000028620	ATP binding cassette subfamily D member 3	ABCD3
ENSSSCG00000056267	ATP synthase subunit a	
ENSSSCG00000061605	C-X-C motif chemokine	
ENSSSCG00000008953	C-X-C motif chemokine ligand 8	CXCL8
ENSSSCG00000041467	Coiled-coil domain containing 194	CCDC194
ENSSSCG00000001045	ELOVL fatty acid elongase 2	ELOVL2
ENSSSCG00000031616	FosB proto-oncogene, AP-1 transcription factor subunit	FOSB
ENSSSCG00000015617	G0/G1 switch 2	G0S2
ENSSSCG00000014725	Globin family profile domain-containing protein	
ENSSSCG00000017507	Grh/CP2 DB domain-containing protein	
ENSSSCG00000031037	Ig-like domain-containing protein	
ENSSSCG00000041880	Ig-like domain-containing protein	
ENSSSCG00000038719	Ig-like domain-containing protein	
ENSSSCG00000036224	Ig-like domain-containing protein	
ENSSSCG00000039102	Ig-like domain-containing protein	
ENSSSCG00000060584	Ig-like domain-containing protein	

ID	Gene Name	Short Name
ENSSSCG00000061568	Ig-like domain-containing protein	
ENSSSCG00000063511	Ig-like domain-containing protein	
ENSSSCG00000040961	LIF interleukin 6 family cytokine	LIF
ENSSSCG00000030655	MAM domain containing 2	MAMDC2
ENSSSCG00000003669	MFSD2 lysolipid transporter A, lysophospholipid	MFSD2A
ENSSSCG00000038811	MOB kinase activator 3B	MOB3B
ENSSSCG00000056213	Murine leukemia virus integrase C-terminal domain-containing protein	
ENSSSCG00000033735	PPARGC1 and ESRR induced regulator, muscle 1	PERM1
ENSSSCG00000017032	PTTG1 regulator of sister chromatid separation, securin	PTTG1
ENSSSCG00000014900	RAB30, member RAS oncogene family	RAB30
ENSSSCG00000043115	SLAM family member 9-like	
ENSSSCG00000013369	Serum amyloid A protein	
ENSSSCG00000052907	Serum amyloid A protein	
ENSSSCG00000045040	Transmembrane protein 258	
ENSSSCG00000009469	aconitate decarboxylase 1	ACOD1
ENSSSCG00000035186	adenylate cyclase activating polypeptide 1	ADCYAP1
ENSSSCG00000040337	adenylate kinase 4	AK4
ENSSSCG00000002814	adhesion G protein-coupled receptor G3	ADGRRG3
ENSSSCG00000059591	Adrenomedullin 2	ADM2
ENSSSCG00000016402	alanine--glyoxylate aminotransferase	AGXT
ENSSSCG00000037878	alpha hemoglobin stabilizing protein	AHSP
ENSSSCG00000005608	angiopoietin like 2	ANGPTL2
ENSSSCG00000036965	angiopoietin like 8	ANGPTL8
ENSSSCG00000040689	apolipoprotein A4	APOA4
ENSSSCG00000010992	Aquaporin 7	AQP7
ENSSSCG00000008072	asporin	ASPN
ENSSSCG00000023178	basic leucine zipper ATF-like transcription factor 2	BATF2
ENSSSCG00000032082	bone morphogenetic protein 8a	BMP8A

ID	Gene Name	Short Name
ENSSSCG00000037514	bone morphogenetic protein 8b	BMP8B
ENSSSCG00000036589	brain enriched guanylate kinase associated	BEGAIN
ENSSSCG00000012845	cell cycle exit and neuronal differentiation 1	CEND1
ENSSSCG00000008555	cell growth regulator with EF-hand domain 1	CGREF1
ENSSSCG00000004484	collagen type XII alpha 1 chain	COL12A1
ENSSSCG00000003006	cytochrome P450 family 2 subfamily B member 6A	CYP2B6A
ENSSSCG00000061845	cytochrome P450 family 2 subfamily C member 293	CYP2C293
ENSSSCG00000006238	cytochrome P450 family 7 subfamily A member 1	CYP7A1
ENSSSCG00000054436	cytochrome P450 family 8 subfamily B member 1	CYP8B1
ENSSSCG00000029990	defensin beta 1	DEFB1
ENSSSCG00000027555	defensin, beta 1	DEFB1
ENSSSCG00000003417	dispatched RND transporter family member 3	DISP3
ENSSSCG00000004675	dual oxidase 1	DUOX1
ENSSSCG00000004676	dual oxidase maturation factor 2	DUOXA2
ENSSSCG00000032911	dual specificity phosphatase 8	DUSP8
ENSSSCG00000020990	dynein axonemal heavy chain 12	DNAH12
ENSSSCG00000015212	fasciculation and elongation protein zeta 1	FEZ1
ENSSSCG00000013072	fatty acid desaturase 2	FADS2
ENSSSCG00000003140	fibroblast growth factor 21	FGF21
ENSSSCG00000016892	follistatin	FST
ENSSSCG00000009667	gulonolactone (L-) oxidase	GULO
ENSSSCG00000007978	hemoglobin subunit alpha	
ENSSSCG00000007977	hemoglobin subunit mu	HBM
ENSSSCG00000058639	hydroxymethylglutaryl-CoA synthase, mitochondrial-like	
ENSSSCG00000055724	immunoglobulin heavy variable 3-23(LOC100125542)	
ENSSSCG00000011717	immunoglobulin superfamily member 10	IGSF10
ENSSSCG00000028331	interleukin 1 receptor type 2	IL1R2
ENSSSCG00000040638	iodothyronine deiodinase 2	DIO2
ENSSSCG00000016401	kinesin family member 1A	KIF1A

ID	Gene Name	Short Name
ENSSSCG00000025188	leptin receptor	LEPR
ENSSSCG00000014232	lysyl oxidase	LOX
ENSSSCG00000004454	malic enzyme 1	ME1
ENSSSCG00000029239	marginal zone B and B1 cell specific protein	MZB1
ENSSSCG00000011862	mucin 13, cell surface associated	MUC13
ENSSSCG00000015871	nuclear receptor subfamily 4 group A member 2	NR4A2
ENSSSCG00000005385	nuclear receptor subfamily 4 group A member 3	NR4A3
ENSSSCG00000031321	nuclear receptor subfamily 4, group A, member 1	NR4A1
ENSSSCG00000035863	Perilipin 2	PLIN2
ENSSSCG00000016720	phosphoglycerate mutase 2	PGAM2
ENSSSCG00000003551	progestin and adipoQ receptor family member 7	PAQR7
ENSSSCG00000015579	prostaglandin-endoperoxide synthase 2	PTGS2
ENSSSCG00000034632	pyridoxal phosphatase	PDXP
ENSSSCG00000034656	reticulon 4 receptor	RTN4R
ENSSSCG00000036865	serine dehydratase	SDS
ENSSSCG00000014422	serine peptidase inhibitor Kazal type 5	SPINK5
ENSSSCG00000031595	serpin A3-8	
ENSSSCG00000038006	solute carrier family 25 member 33	SLC25A33
ENSSSCG00000012993	solute carrier family 25 member 45	SLC25A45
ENSSSCG00000002515	solute carrier family 25 member 47	SLC25A47
ENSSSCG00000011641	solute carrier organic anion transporter family member 2A1	SLCO2A1
ENSSSCG00000010554	stearoyl-CoA desaturase	SCD
ENSSSCG00000032374	sulfotransferase family cytosolic 1B member 1	
ENSSSCG00000005455	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	SVEP1
ENSSSCG00000004789	Thrombospondin 1	THBS1
ENSSSCG00000015525	torsin family 3 member A	TOR3A
ENSSSCG00000026043	Transglutaminase 3	TGM3
ENSSSCG00000014871	tsukushi, small leucine rich proteoglycan	TSKU

ID	Gene Name	Short Name
ENSSSCG00000032398	vestigial like family member 3	VGLL3
ENSSSCG00000002908	zinc finger and BTB domain containing 32	ZBTB32

Supplementary Table S2. rCCA connections between genes and genera filtered for an absolute weight > 0.15. A positive weight indicates a positive relationship and a negative weight indicates a negative relationship.

Gene	Genus	Weight
ENSSSCG00000018063	*UBA1259*	-0.1905
ENSSSCG00000018063	*CAG-302*	-0.1926
ENSSSCG00000018063	*Lachnospira*	-0.1699
ENSSSCG00000018063	*SFDB01*	0.1653
ENSSSCG00000018063	*Hallerella*	-0.1720
ENSSSCG00000018063	*Onthomonas*	-0.1988
ENSSSCG00000018063	Unassigned UBA932	0.1938
ENSSSCG00000018063	*Caproiciproducens*	-0.1819
ENSSSCG00000018063	*Cryptobacteroides*	0.1534
ENSSSCG00000018063	Unassigned Bacteroidaceae	-0.1879
ENSSSCG00000018063	*Mitsuokella*	0.1748
ENSSSCG00000018063	*Choladousia*	-0.1708
ENSSSCG00000018063	Unassigned Gastranaerophilaceae	0.2046
ENSSSCG00000018063	*CAG-533*	-0.2153
ENSSSCG00000018063	*SFLA01*	0.2113
ENSSSCG00000018063	Unassigned Peptostreptococcaceae	-0.1652
ENSSSCG00000018063	*Agathobaculum*	0.1848
ENSSSCG00000018063	*MARSEILLE-P3954*	-0.1667
ENSSSCG00000018063	*Faecalibacterium*	-0.1909
ENSSSCG00000018063	*Floccifex*	-0.2147
ENSSSCG00000018063	*Ornithospirochaeta*	-0.1798
ENSSSCG00000018063	*UBA2834*	0.1539
ENSSSCG00000018063	*Copromorpha*	0.1562
ENSSSCG00000018063	*CAG-313*	0.1655
ENSSSCG00000018063	*CAG-349*	-0.2195
ENSSSCG00000018063	*Streptococcus*	-0.2349

Gene	Genus	Weight
ENSSSCG00000018063	*Onthovivens*	0.1649
DEFB1	*UBA1259*	0.1515
DEFB1	*Caproiciproducens*	0.1548
DEFB1	*Sutterella*	0.1615
DEFB1	*Mediterraneibacter*	0.1717
DEFB1	Unassigned Peptostreptococcaceae	0.1589
DEFB1	*Eubacterium* (Lachnospiraceae)	0.1760
DEFB1	*Floccifex*	0.1835
DEFB1	*Ornithospirochaeta*	0.1562
DEFB1	*Caccousia*	0.1753
DEFB1	*Ellagibacter*	0.1518
ENSSSCG00000055724	*Floccifex*	0.1528
ENSSSCG00000061568	*Sutterella*	0.1511
ENSSSCG00000061568	*Mediterraneibacter*	0.1533
ENSSSCG00000061568	Unassigned Peptostreptococcaceae	0.1513
ENSSSCG00000061568	*Eubacterium* (Lachnospiraceae)	0.1555
ENSSSCG00000061568	*Floccifex*	0.1768
ENSSSCG00000061568	*Ornithospirochaeta*	0.1503
ENSSSCG00000061568	*Caccousia*	0.1634
ENSSSCG00000062467	*Floccifex*	0.1648
ENSSSCG00000060872	*Floccifex*	0.1526
ENSSSCG00000039102	*Mediterraneibacter*	0.1540
ENSSSCG00000039102	*Eubacterium* (Lachnospiraceae)	0.1624
ENSSSCG00000063511	*Mediterraneibacter*	0.1554
ENSSSCG00000063511	*Eubacterium* (Lachnospiraceae)	0.1656
ENSSSCG00000056525	*Mediterraneibacter*	0.1659
ENSSSCG00000056525	*Eubacterium* (Lachnospiraceae)	0.1753
ENSSSCG00000060584	*Floccifex*	0.1767

Gene	Genus	Weight
ENSSSCG00000060584	*Caccousia*	0.1509
ENSSSCG00000060584	*Streptococcus*	0.1581
ENSSSCG00000061361	*SFLA01*	-0.1832
ENSSSCG00000059271	*SFLA01*	-0.1821
ENSSSCG00000042281	*SFLA01*	-0.1639
ENSSSCG00000057472	*SFLA01*	-0.2015
ENSSSCG00000048258	*SFDB01*	-0.1519
ENSSSCG00000048258	*SFLA01*	-0.2251
ENSSSCG00000051549	*SFLA01*	-0.1677
ENSSSCG00000043613	*SFLA01*	-0.2032
PGAM2	*SFLA01*	-0.1720
ENSSSCG00000053426	*SFLA01*	-0.1995
ENSSSCG00000055843	*SFDB01*	-0.1528
ENSSSCG00000055843	*SFLA01*	-0.2117
ENSSSCG00000055843	*Agathobaculum*	-0.1564
ENSSSCG00000057894	*SFDB01*	0.1567
ENSSSCG00000057894	*SFLA01*	0.2286
ENSSSCG00000017507	*SFLA01*	0.1984
ENSSSCG00000056082	*SFLA01*	-0.1529
RASD1	*UBA1259*	0.1554
RASD1	*Caproiciproducens*	0.1558
RASD1	*Sutterella*	0.1504
RASD1	Unassigned Gastranaerophilaceae	-0.1523
RASD1	Unassigned Peptostreptococcaceae	0.1548
RASD1	*Floccifex*	0.1844
RASD1	*Ornithospirochaeta*	0.1563
RASD1	*Caccousia*	0.1617
RASD1	*Streptococcus*	0.1603

Gene	Genus	Weight
ENSSSCG00000055937	*SFLA01*	-0.1647
ENSSSCG00000060346	*SFLA01*	-0.1808
ENSSSCG00000052680	*SFLA01*	-0.1934
ENSSSCG00000044562	*SFLA01*	-0.1546
ENSSSCG00000010992	*SFLA01*	0.2176
MOB3B	*SFDB01*	-0.1517
MOB3B	*SFLA01*	-0.2176
ENSSSCG00000047427	*SFLA01*	-0.1640
ENSSSCG00000058351	*SFLA01*	-0.1602
ACOD1	*SFLA01*	-0.1675
ENSSSCG00000048696	*Eubacterium* (Lachnospiraceae)	-0.1571
ENSSSCG00000017032	*SFLA01*	-0.1749
FST	*Floccifex*	0.1598
ADM2	*SFLA01*	-0.1922
ADM2	*Eubacterium* (Lachnospiraceae)	-0.1611
NR4A1	*Mediterraneibacter*	0.1619
NR4A1	*Eubacterium* (Lachnospiraceae)	0.1705
PDXP	*SFDB01*	-0.1516
PDXP	*SFLA01*	-0.2277
PDXP	*Eubacterium* (Lachnospiraceae)	-0.1570
ENSSSCG00000053597	*SFLA01*	0.1560
ENSSSCG00000056213	*SFLA01*	0.2086
ENSSSCG00000058911	*SFLA01*	-0.1738
ELOVL2	*SFDB01*	0.2128
ELOVL2	Unassigned UBA932	0.1511
ELOVL2	*Fimousia*	0.1604
ELOVL2	*SFLA01*	0.3043
ELOVL2	*Agathobaculum*	0.2093

Gene	Genus	Weight
ELOVL2	Unassigned CAG-508	0.1761
ELOVL2	*Collinsella*	0.1500
ELOVL2	*Catenibacterium*	0.1584
ELOVL2	*Copromorpha*	0.1651
ELOVL2	*CAG-349*	-0.1595
BEGAIN	*SFLA01*	-0.1733
ENSSSCG00000031595	*Floccifex*	0.1567
SLC25A47	*SFLA01*	0.1538
DIO2	*Floccifex*	-0.1775
DIO2	*Ornithospirochaeta*	-0.1504
DIO2	*Caccousia*	-0.1551
DIO2	*Streptococcus*	-0.1550
ALAS2	*SFDB01*	-0.1544
ALAS2	*SFLA01*	-0.2171
ALAS2	*Agathobaculum*	-0.1552
ENSSSCG00000049683	*SFDB01*	0.1558
ENSSSCG00000049683	*SFLA01*	0.2292
ADAMTS4	*Mediterraneibacter*	0.1774
ADAMTS4	*Eubacterium* (Lachnospiraceae)	0.1884
ADAMTS4	*Caccousia*	0.1542
HMGCS2	*SFDB01*	0.1821
HMGCS2	*Helicobacter*	0.1520
HMGCS2	*Limivivens*	-0.1533
HMGCS2	Unassigned Erysipelotrichaceae	-0.1673
HMGCS2	*Mediterraneibacter*	0.2056
HMGCS2	*SFLA01*	0.2800
HMGCS2	*Eubacterium* (Lachnospiraceae)	0.2286
HMGCS2	*Agathobaculum*	0.1616

Gene	Genus	Weight
HMGCS2	*Paludicola*	0.1545
HMGCS2	Unassigned CAG-508	0.1877
HMGCS2	*Phascolarctobacterium*	0.1697
HMGCS2	*Collinsella*	0.1953
HMGCS2	*UBA636*	0.1647
ENSSSCG00000058639	*SFLA01*	0.1666
ENSSSCG00000061756	*SFLA01*	-0.1832
ENSSSCG00000043115	*SFLA01*	-0.1649
CYP7A1	*SFLA01*	-0.1545
ENSSSCG00000053251	*SFLA01*	-0.1705
ABCD3	*SFDB01*	0.1995
ABCD3	*Fimousia*	0.1504
ABCD3	*SFLA01*	0.2851
ABCD3	*Agathobaculum*	0.1963
ABCD3	Unassigned CAG-508	0.1649
ABCD3	*Copromorpha*	0.1549
ASPN	*SFLA01*	0.1638
HBM	*SFDB01*	-0.1847
HBM	Unassigned UBA932	-0.1548
HBM	*SFLA01*	-0.2564
HBM	*Agathobaculum*	-0.1886
HBM	*Copromorpha*	-0.1520
HBM	*CAG-349*	0.1680
HBM	*Streptococcus*	0.1537
ENSSSCG00000007978	*SFDB01*	-0.1809
ENSSSCG00000007978	*SFLA01*	-0.2563
ENSSSCG00000007978	*Agathobaculum*	-0.1799
CLDN15	*SFDB01*	-0.1508

Gene	Genus	Weight
CLDN15	*SFLA01*	-0.2234
ENSSSCG00000061594	*SFLA01*	0.2099
ENSSSCG00000008555	*SFDB01*	0.1794
ENSSSCG00000008555	*Mediterraneibacter*	0.1607
ENSSSCG00000008555	*SFLA01*	0.2688
ENSSSCG00000008555	*Eubacterium* (Lachnospiraceae)	0.1825
ENSSSCG00000008555	*Agathobaculum*	0.1654
ENSSSCG00000008555	Unassigned CAG-508	0.1717
ENSSSCG00000008555	*Collinsella*	0.1686
ENSSSCG00000036096	*SFLA01*	-0.1940
AHSP	*SFDB01*	-0.1902
AHSP	*SFLA01*	-0.2705
AHSP	*Agathobaculum*	-0.1883
AHSP	Unassigned CAG-508	-0.1548
IL1R2	*SFLA01*	-0.1551
ENSSSCG00000036224	*Mediterraneibacter*	0.1590
ENSSSCG00000036224	*Eubacterium* (Lachnospiraceae)	0.1630
ENSSSCG00000036224	*Floccifex*	0.1699
ENSSSCG00000036224	*Caccousia*	0.1623
ENSSSCG00000062489	*Floccifex*	0.1579
ENSSSCG00000059927	*SFLA01*	-0.1531
SULT1B1	*Mediterraneibacter*	0.1575
SULT1B1	*SFLA01*	0.2153
SULT1B1	*Eubacterium* (Lachnospiraceae)	0.1752
ENSSSCG00000061605	*SFLA01*	-0.2028
CXCL8	*Floccifex*	0.1566
ENSSSCG00000054701	*SFLA01*	-0.2076
RAB30	*SFDB01*	0.1615

Gene	Genus	Weight
RAB30	*SFLA01*	0.2338
RAB30	*Agathobaculum*	0.1564
ENSSSCG00000052232	*SFLA01*	-0.1715
PTGS2	*SFLA01*	0.1605
PTGS2	*Eubacterium* (Lachnospiraceae)	0.1612
APOA4	*SFLA01*	0.1627
ENSSSCG00000038670	*SFDB01*	-0.1902
ENSSSCG00000038670	*SFLA01*	-0.2728
ENSSSCG00000038670	*Agathobaculum*	-0.1862
ENSSSCG00000038670	Unassigned CAG-508	-0.1591
HBB	*SFDB01*	-0.1753
HBB	*SFLA01*	-0.2509
HBB	*Agathobaculum*	-0.1721
TSKU	*Mediterraneibacter*	-0.1574
TSKU	*SFLA01*	-0.1552
TSKU	*Eubacterium* (Lachnospiraceae)	-0.1704
ENSSSCG00000053725	*SFLA01*	-0.2113
TOR3A	*SFLA01*	0.1550
G0S2	*SFLA01*	0.2178
AGXT	*Mediterraneibacter*	0.1789
AGXT	*SFLA01*	0.1663
AGXT	*Eubacterium* (Lachnospiraceae)	0.1928
KIF1A	*Eubacterium* (Lachnospiraceae)	0.1524
ENSSSCG00000058050	*SFLA01*	-0.2193
NR4A2	*Mediterraneibacter*	0.1501
NR4A2	*Eubacterium* (Lachnospiraceae)	0.1601
ENSSSCG00000055100	*SFLA01*	-0.1802
ENSSSCG00000063146	*SFLA01*	-0.1609

Gene	Genus	Weight
SCD	*UBA1259*	-0.1613
SCD	*Hallerella*	-0.1510
SCD	*Caproiciproducens*	-0.1618
SCD	*Sutterella*	-0.1569
SCD	Unassigned Gastranaerophilaceae	0.1578
SCD	Unassigned Peptostreptococcaceae	-0.1610
SCD	*Floccifex*	-0.1916
SCD	*Ornithospirochaeta*	-0.1624
SCD	*Caccousia*	-0.1688
SCD	*Streptococcus*	-0.1657
ENSSSCG00000038719	Unassigned Gastranaerophilaceae	-0.1518
ENSSSCG00000038719	*CAG-533*	0.1504
ENSSSCG00000038719	*Floccifex*	0.1726
ENSSSCG00000038719	*Streptococcus*	0.1665
ENSSSCG00000031037	*Floccifex*	0.1544
ENSSSCG00000061845	*SFDB01*	0.1736
ENSSSCG00000061845	*SFLA01*	0.2428
ENSSSCG00000061845	*Agathobaculum*	0.1756
ENSSSCG00000061845	*CAG-349*	-0.1511
LIF	*SFLA01*	0.1725
ENSSSCG00000041880	*Floccifex*	0.1587
ENSSSCG00000044931	*Floccifex*	-0.1529
RTN4R	*SFDB01*	-0.1526
RTN4R	*SFLA01*	-0.2164
RTN4R	*Agathobaculum*	-0.1516
SDS	*UBA1259*	0.1530
SDS	*CAG-302*	0.1549
SDS	*Onthomonas*	0.1598

Gene	Genus	Weight
SDS	Unassigned UBA932	-0.1559
SDS	Unassigned Bacteroidaceae	0.1511
SDS	Unassigned Gastranaerophilaceae	-0.1644
SDS	*CAG-533*	0.1731
SDS	*SFLA01*	-0.1704
SDS	*Faecalibacterium*	0.1535
SDS	*Floccifex*	0.1725
SDS	*CAG-349*	0.1766
SDS	*Streptococcus*	0.1889
ENSSSCG00000009667	*UBA1259*	-0.1572
ENSSSCG00000009667	*Caproiciproducens*	-0.1560
ENSSSCG00000009667	Unassigned Gastranaerophilaceae	0.1570
ENSSSCG00000009667	*CAG-533*	-0.1513
ENSSSCG00000009667	Unassigned Peptostreptococcaceae	-0.1524
ENSSSCG00000009667	*Floccifex*	-0.1846
ENSSSCG00000009667	*Ornithospirochaeta*	-0.1561
ENSSSCG00000009667	*Caccousia*	-0.1546
ENSSSCG00000009667	*Streptococcus*	-0.1685
BATF2	*SFLA01*	-0.1518
SLC25A45	*SFLA01*	0.1788
FADS2	*UBA1259*	-0.1509
FADS2	*Onthomonas*	-0.1524
FADS2	Unassigned Gastranaerophilaceae	0.1582
FADS2	*CAG-533*	-0.1619
FADS2	*Floccifex*	-0.1725
FADS2	*CAG-349*	-0.1592
FADS2	*Streptococcus*	-0.1778
CEND1	*UBA1259*	0.1581

Gene	Genus	Weight
CEND1	*Caproiciproducens*	0.1568
CEND1	Unassigned Gastranaerophilaceae	-0.1583
CEND1	*CAG-533*	0.1529
CEND1	Unassigned Peptostreptococcaceae	0.1529
CEND1	*Floccifex*	0.1855
CEND1	*Ornithospirochaeta*	0.1568
CEND1	*Caccousia*	0.1546
CEND1	*Streptococcus*	0.1702
DUSP8	*SFLA01*	0.1568
ANGPTL8	*Mediterraneibacter*	-0.1511
ANGPTL8	*SFLA01*	-0.2016
ANGPTL8	*Eubacterium* (Lachnospiraceae)	-0.1677
ENSSSCG00000045040	*SFLA01*	-0.1920
CCDC194	*CAG-533*	-0.1530
CCDC194	*SFLA01*	0.1583
CCDC194	*CAG-349*	-0.1574
CCDC194	*Streptococcus*	-0.1667
SPINK5	*SFLA01*	-0.1890
ENSSSCG00000056267	*SFLA01*	-0.1730
ENSSSCG00000056267	*Eubacterium* (Lachnospiraceae)	-0.1521
MZB1	*UBA1259*	0.1547
MZB1	*Caproiciproducens*	0.1539
MZB1	Unassigned Gastranaerophilaceae	-0.1538
MZB1	Unassigned Peptostreptococcaceae	0.1510
MZB1	*Floccifex*	0.1821
MZB1	*Ornithospirochaeta*	0.1541
MZB1	*Caccousia*	0.1544
MZB1	*Streptococcus*	0.1643

Gene	Genus	Weight
ENSSSCG00000052907	*SFLA01*	-0.1783
ENSSSCG00000052907	*Streptococcus*	0.1509
ENSSSCG00000013369	*SFLA01*	-0.1905
LOX	*SFLA01*	0.1579
ENSSSCG00000052290	*SFLA01*	-0.1575
FGF21	*SFDB01*	0.1882
FGF21	Unassigned UBA932	0.1528
FGF21	*SFLA01*	0.2629
FGF21	*Agathobaculum*	0.1907
FGF21	*Copromorpha*	0.1530
FGF21	*CAG-349*	-0.1650
ENSSSCG00000055509	*SFLA01*	-0.1792
ENSSSCG00000050714	*SFLA01*	-0.1677
DISP3	*SFLA01*	-0.1806
ENSSSCG00000056084	*SFLA01*	-0.1838
PAQR7	*SFLA01*	0.1521
FOSB	*Mediterraneibacter*	0.1585
FOSB	*Eubacterium* (Lachnospiraceae)	0.1680
ADCYAP1	*SFLA01*	0.1699
ZBTB32	*SFLA01*	-0.1525
ADGRG3	*SFDB01*	-0.1542
ADGRG3	*SFLA01*	-0.2140
ADGRG3	*Agathobaculum*	-0.1575
ENSSSCG00000040337	*Mediterraneibacter*	-0.1656
ENSSSCG00000040337	*Eubacterium* (Lachnospiraceae)	-0.1714
ENSSSCG00000040337	*Floccifex*	-0.1642
ENSSSCG00000040337	*Caccousia*	-0.1624
LEPR	*SFLA01*	0.1592

Gene	Genus	Weight
ENSSSCG00000059695	*SFLA01*	-0.1528
ENSSSCG00000033735	*SFLA01*	-0.1784
ENSSSCG00000033735	*Eubacterium* (Lachnospiraceae)	-0.1605
SLC25A33	*SFLA01*	0.1553
MFSD2A	*SFLA01*	0.1622
ENSSSCG00000032082	*SFLA01*	0.1531
ENSSSCG00000037514	*Mediterraneibacter*	0.1664
ENSSSCG00000037514	*SFLA01*	0.2209
ENSSSCG00000037514	*Eubacterium* (Lachnospiraceae)	0.1845
ENSSSCG00000037514	*Collinsella*	0.1563
ENSSSCG00000045832	*Sutterella*	-0.1511
ENSSSCG00000045832	*Mediterraneibacter*	-0.1708
ENSSSCG00000045832	*Eubacterium* (Lachnospiraceae)	-0.1774
ENSSSCG00000045832	*Floccifex*	-0.1642
ENSSSCG00000045832	*Caccousia*	-0.1648
ENSSSCG00000059294	*SFLA01*	-0.2047
ENSSSCG00000041313	*SFLA01*	-0.1684
CYP8B1	*SFDB01*	0.1557
CYP8B1	*Mediterraneibacter*	0.1545
CYP8B1	*SFLA01*	0.2359
CYP8B1	*Eubacterium* (Lachnospiraceae)	0.1737
CYP8B1	Unassigned CAG-508	0.1538
CYP8B1	*Collinsella*	0.1549
ARF4	*SFDB01*	0.2085
ARF4	*Helicobacter*	0.1522
ARF4	*Fimousia*	0.1534
ARF4	*SFLA01*	0.3054
ARF4	*Eubacterium* (Lachnospiraceae)	0.1689

Gene	Genus	Weight
ARF4	*Agathobaculum*	0.1986
ARF4	*Unassigned CAG-508*	0.1862
ARF4	*Collinsella*	0.1717
ARF4	*Catenibacterium*	0.1541
ARF4	*UBA636*	0.1520
ARF4	*Copromorpha*	0.1537
ENSSSCG00000051985	*Mediterraneibacter*	-0.1616
ENSSSCG00000051985	*Eubacterium* (Lachnospiraceae)	-0.1704
DNAH12	*SFLA01*	0.1626
ENSSSCG00000042940	*SFLA01*	-0.1830
MUC13	*SFLA01*	0.1623
ENSSSCG00000059810	*SFLA01*	-0.1574
IGSF10	*SFLA01*	0.1734
SLCO2A1	*SFDB01*	0.1689
SLCO2A1	*Mediterraneibacter*	0.1617
SLCO2A1	*SFLA01*	0.2549
SLCO2A1	*Eubacterium* (Lachnospiraceae)	0.1825
SLCO2A1	*Agathobaculum*	0.1542
SLCO2A1	Unassigned CAG-508	0.1645
SLCO2A1	*Collinsella*	0.1647
VGLL3	*SFLA01*	0.1685
DUOXA2	*SFLA01*	0.1524
ANGPTL2	*SFLA01*	0.1576
THBS1	*SFLA01*	0.1635
THBS1	*Eubacterium* (Lachnospiraceae)	0.1515
ENSSSCG00000061365	*SFLA01*	-0.1653
ENSSSCG00000052966	*SFLA01*	-0.1899
MAMDC2	*SFLA01*	0.1514

Gene	Genus	Weight
ENSSSCG00000053150	*SFDB01*	0.2034
ENSSSCG00000053150	*SFLA01*	0.2980
ENSSSCG00000053150	*Eubacterium* (Lachnospiraceae)	0.1653
ENSSSCG00000053150	*Agathobaculum*	0.1936
ENSSSCG00000053150	Unassigned CAG-508	0.1818
ENSSSCG00000053150	*Collinsella*	0.1678
ENSSSCG00000053150	*Catenibacterium*	0.1503
SVEP1	*SFLA01*	0.1722
ENSSSCG00000059343	*Mediterraneibacter*	-0.1584
ENSSSCG00000059343	*SFLA01*	-0.2198
ENSSSCG00000059343	*Eubacterium* (Lachnospiraceae)	-0.1764
ENSSSCG00000059343	*Collinsella*	-0.1518
ENSSSCG00000048233	*SFDB01*	-0.2025
ENSSSCG00000048233	*Helicobacter*	-0.1531
ENSSSCG00000048233	Unassigned Erysipelotrichaceae	0.1533
ENSSSCG00000048233	*Mediterraneibacter*	-0.1621
ENSSSCG00000048233	*SFLA01*	-0.3003
ENSSSCG00000048233	*Eubacterium* (Lachnospiraceae)	-0.1864
ENSSSCG00000048233	*Agathobaculum*	-0.1896
ENSSSCG00000048233	Unassigned CAG-508	-0.1878
ENSSSCG00000048233	*Collinsella*	-0.1793
ENSSSCG00000048233	*UBA636*	-0.1565
ENSSSCG00000053347	*Mediterraneibacter*	-0.1594
ENSSSCG00000053347	*SFLA01*	-0.2167
ENSSSCG00000053347	*Eubacterium* (Lachnospiraceae)	-0.1772
ENSSSCG00000053347	*Collinsella*	-0.1513
ENSSSCG00000058177	*SFDB01*	-0.1698
ENSSSCG00000058177	*Mediterraneibacter*	-0.1671

Gene	Genus	Weight
ENSSSCG00000058177	*SFLA01*	-0.2570
ENSSSCG00000058177	*Eubacterium* (Lachnospiraceae)	-0.1881
ENSSSCG00000058177	*Agathobaculum*	-0.1543
ENSSSCG00000058177	Unassigned CAG-508	-0.1673
ENSSSCG00000058177	*Collinsella*	-0.1681
ENSSSCG00000042224	*Mediterraneibacter*	-0.1662
ENSSSCG00000042224	*SFLA01*	-0.2038
ENSSSCG00000042224	*Eubacterium* (Lachnospiraceae)	-0.1830
ENSSSCG00000042224	*Collinsella*	-0.1507
ME1	*Floccifex*	-0.1662
ME1	*Streptococcus*	-0.1607
COL12A1	*SFLA01*	0.1672
NR4A3	*Mediterraneibacter*	0.1747
NR4A3	*Eubacterium* (Lachnospiraceae)	0.1841
NR4A3	*Caccousia*	0.1577
PLIN2	*SFLA01*	0.1993
ENSSSCG00000052903	*SFLA01*	-0.1572

Acknowledgements

Now, if I may, I will switch to Spanish, as it is my mother tongue and how my heart speaks. But you have the English translation below.

Me gustaría empezar los agradecimientos por mis maestros jedi, los doctores JM y Dani, que me han dirigido la tesis y han tenido que manejar mi espíritu caótico y ansiedad inherente. Obviamente no estaría aquí sin ellos, y he tenido el honor de conocer a dos maravillosas personas que, además, surfean la ciencia como reyes. Gracias por confiar en mí. Llevaré conmigo todo lo que me habéis enseñado.

También tiene un lugar especial mi maestro jedi del extranjero, el doctor Jordi, que fue mi supervisor durante mi estancia en el INRAE de Jouy-en-Josas y seguramente me salvó la vida más de una vez. Gracias por eso y por acogerme. No pienso volver a preparar más tortillas de patatas, pero le estoy cogiendo el truco a las crêpes.

A mi grupo del CRAG, presente y pretérito, la secta de Animal Genomics (y asociados) cuyos miembros son más agentes del caos que yo: os quiero, aunque no lo parezca. Si bien os adoro a todos, hay unas cuantas personas que merecen una mención especial.

Tuve la suerte de compartir despacho y parte de la tesis con los doctores Magí, Jesús y Yron, que siempre me hicieron reír o salir a flote cuando la vida parecía demasiado. Aprendí a tomarme las cosas con tanta filosofía como Yron, a no cogerme lucha por las cosas que no importaban como me decía Jesús y a poner buen humor en mi día a día con Magí. Os echo de menos, pero veo dónde estáis ahora y estoy feliz por vosotros. Y el congreso de Oporto con Magí fue uno de los highlights de mi tesis, para qué engañarnos (había barra libre de vino).

Ferran, mi niño, hemos compartido todo este camino y te dejo enfrentarte al resto de tu tesis, que será fantástica. Te he visto pasar de ser un señor de veintiún años recién salido de la carrera a un hombre mucho más seguro de quién es y de lo que quiere, aprendiendo mil cosas por el camino y masterizando cada elemento en el que te movías. Cómo no voy a estar orgullosísima. Para estos próximos años: fuerza y honor.

Mi preciosa doctora Maria Gracia fue la luz de mis primeros años aquí en el CRAG y la echo de menos cada día. Gracias por adoptarme, a ti y a Lucas. Si alguien me pregunta alguna vez cuál es el gold standard de persona, por buena, por talentosa y por bella, sin duda diré que es mi MG.

La doctora Alice fue la última tesis de mi grupo en la que pude participar como organizadora del regalo. Y no lo habría querido de otra manera. Sufrí tu estrés contigo y fuiste un apoyo cuando más lo necesitaba, no sé si sabes cuánto. Grazie mille, dottoressa, sé muy feliz ahora que eres libre.

A más miembros de la facción italiana, las doctoras Elena y Antonia, gracias por ser tan preciosas personas y por acogernos en vuestra casa, Jesús no tendría tantos selfies con la Sagrada Familia de no haber sido así.

A mis reinas de la facción china, Liu y Lian, que han estado también durante casi toda mi tesis y me han enseñado a amar su cultura y los snacks picantes. Mucha fuerza para lo poquito que os queda.

Uri fue la persona más sabia de todos nosotros y salió a tiempo del doctorado, antes de que le hiciese daño. Fue un regalo encontrarte en el máster, sobrevivir al confinamiento estando bien lejos y aprender de lo valiente que eres.

Y Elisa, que llegó por un máster y también ha sabido huir a tiempo. Tu forma tan sencilla de estar ahí me ha ayudado a entender mejor el mundo. Te deseo lo mejor en tu nueva etapa.

A los doctores Yuli y Mélina, geniales personas que me enseñaron WGCNA: gracias por vuestra paciencia.

A nuestros familiares de genómica animal de Torre. Espero que vayáis muchas veces de birras, os dejo mi legado como la princesa Disney de los gorriones.

Si esta gente son quienes más me han marcado dentro de mi departamento, el resto del CRAG no se queda lejos. Mando un gran abrazo a 000 - Tot el CRAG, pero aquí también debo hacer menciones especiales.

Durante estos cuatro años he tenido los mejores compañeros de despacho, y eso incluye a mis maravillosas Lorena y Ari, que están sufriendo el final de tesis a la vez que yo. Lorena me ha salvado de forma evidente más de una vez (como con el Turnitin del demonio) y siempre ha estado ahí. Defender el mismo día será una forma perfecta de demostrar que el doctorado nos unió y saldremos adelante juntas aunque estemos en la otra punta del mundo. Y Ari, la reina del CRAG, nunca vi persona más buena y más preocupada por los demás, así como nunca podré agradecer suficientemente todo lo que me has enseñado siendo tal cual eres.

Y poco tiempo he compartido en ese despacho con Paloma, una de las personas más auténticas que he conocido. Nos llevamos unos cuantos memes de estos meses, sé que mantendrás alto el nivel de generación de contenido. Gracias por acogerme en tu casa y por no asustarte de mis conocimientos en ya-tú-sabes qué web.

A mi doctor Carlitos, que ha sido un gran apoyo los últimos años. Nuestras tardes de chisme y arreglar el mundo no deberían acabarse nunca. La carta de la luna siempre aparecerá para ti allá donde vaya.

Conocí a La Cris™ en una fiesta de tesis, gracias a que nos llamamos igual. Y desde entonces no he tenido una mejor amiga y tocaya. Gracias a ti y a Marc por enseñarnos Manresa.

María fue un descubrimiento que debo agradecer a mi alcoholismo social, así como su compañera de firma de contrato, Raquel. Estas ya doctoras maravillosas y creativas no paran de enseñarme que hay vida (y, por tanto, esperanza) más allá de la ciencia. Gracias por eso.

A las gemelas preciosas, Camila y Violeta, ojalá haber compartido más aire puro (del de verdad). Salí del cascarón en parte gracias a vosotras.

Nicole, la estrellita que más brilla, sé que serás capaz de las mejores cosas.

A los Cañitos™, a la facción IRTA, a todos los grupos que me han acogido una o más veces para comer o alcoholizarnos (Maroof et al., Jaime et al., Clara et al., Héctor et al…), nunca os podré agradecer lo suficiente que me hayáis incluido en esos momentos. He sido feliz con vosotros.

A Mª José, Alberto y todo el personal que siempre tienen una sonrisa o un momento para ayudar. Gracias por cada instante.

A Alexandra Elbakyan. If you know, you know.

A toda la gente que venga detrás, a los futuros miembros de los Folchies, los Marceliños, la Botigueta y un largo etcétera hasta incluir a todo el CRAG.

Es verdad que ese lugar llamado CRAG ha sido mi vida durante cuatro añazos, pero también ha habido gente que ha estado conmigo fuera de esas paredes coloridas, muchas veces a través de whatsapps y llamadas.

Si nos juntamos en el instituto porque no quedaba otra, con los años he visto que fue todo un acierto tener en mi vida cotidiana a Marta, a Anna, a Annamari y a Àngela. Llevamos más años separadas de los que estuvimos juntas, pero sé que, en realidad, siempre estaremos ahí para las demás. Mención especial a Marta por marujear día sí, día también por whatsapp. Me alegras los días.

La carrera me dio a Irene, cuyo nivel de sarcasmo es incluso mayor que el mío. Ánimos con el resto del doctorado, ya nos iremos después a ser críticas de cine de las mejores pelis mierder.

Y durante la carrera, gracias a un concurso literario, conocí a Arancha, escritora top con gran potencial para ser podcaster. A lo largo de los años hemos soportado muchísimos audios la una de la otra. Y los que nos quedan. Fuiste el mejor premio del concurso de La Caixa.

Hablando de escritoras top, el mundo de la literatura también me trajo a Laura, mi icono Escorpio, aquella que conoce seguramente mis mayores secretos y aun así decidió quedarse. Te quiero siempre en mi vida. Por más movidas como las de Bilbao. Ahora toca Helsinki. Y Chiang Mai.

Gracias a Judith, por escucharme cada jueves a las cinco desde hace casi dos años. Si no he perdido lo que me quedaba de salud mental es gracias a ella.

A la que fue mi casera y casi mi familia durante tres años, Neus, a sus niñas y a la Queen. Gracias por aceptarme.

A Elena (Elena Contel Maza, @helen.mushu.science), que ha diseñado la preciosa portada de esta tesis y que es incluso mejor persona que artista. No dejes nunca de crear.

A mi madre, el mayor ser de luz que vi jamás, la que me dio la vida, las alas para volar bien alto y las ganas de ser más y mejor persona. Si soy lo que soy ahora es gracias a ti, a tu aliento y a esa forma de apoyarme incluso cuando ninguna de las dos creíamos que sería posible. Si soy fuerte es por ti, si soy libre es gracias a ti. Te quiero.

A mi padre, el pilar de mi mundo, las raíces de mi tierra y la sabiduría encarnada en un señor con coleta. Aquel que me enseñó el auténtico significado de la palabra “amor” simplemente estando ahí para mi madre y para mí, y que me enseñó a vivir simplemente viviendo. Que me leía historias aunque no le gustase la literatura y que esperaba que yo las

escribiera porque era lo que a mí me gustaba. Si soy libre es por ti, si soy fuerte es gracias a ti. Te quiero, y te echaré de menos todos los días de mi vida. Esta tesis va por ti, Pare.

I would like to begin my acknowledgments by thanking my Jedi Masters, Drs. JM and Dani, who supervised my thesis and had to handle my chaotic spirit and inherent anxiety. I obviously wouldn't be here without them, and I've had the honour of getting to know two wonderful people who navigate the seas of science like kings. Thank you for believing in me. I will carry everything you've taught me with me.

A special place is reserved for my Jedi Master from abroad, Dr. Jordi, who was my supervisor during my internship at INRAE in Jouy-en-Josas and surely saved my life more than once. Thank you for that and for hosting me. I won't be making any more Spanish omelettes, but I'm getting the hang of crêpes.

To my group at CRAG, both present and past, the sect of Animal Genomics (and associates) whose members are even more agents of chaos than I am: I love you, even if it doesn't seem like it. While I adore you all, there are a few people who deserve a special mention.

I had the luck to share an office and part of my thesis with Drs. Magí, Jesús, and Yron, who always made me laugh or helped me stay afloat when life seemed too much. I learned to take things with as much philosophy as Yron, not to sweat the small stuff as Jesús advised, and to bring good humour into my daily life with Magí. I miss you, but seeing where you are now makes me happy. And the conference in Porto with Magí was one of the highlights of my thesis, let's be honest (there was an open bar of wine).

Ferran, my dear, we've shared this entire journey and now I leave you to face the rest of your thesis, which will be fantastic. I've seen you go from being a 21-year-old just out of university to a man much more confident in who he is and what he wants, learning a thousand things along the way and mastering every field you stepped into. How could I not be incredibly proud? For the coming years: strength and honour.

My dear Dr. Maria Gracia was the light of my first years here at CRAG and I miss her every day. Thank you for adopting me, you and Lucas. If anyone ever asks me who is the gold standard of a person, for their goodness, talent, and beauty, I would undoubtedly say it's my MG.

Dr. Alice was the last thesis in my group for which I could participate as the gift organiser. And I wouldn't have wanted it any other way. I suffered your stress with you and you were a support when I needed it most, I don't know if you realize how much. Grazie mille, dottoressa, be very happy now that you're free.

To more members of the Italian faction, Drs. Elena and Antonia, thank you for being such wonderful people and for welcoming us into your home, Jesús wouldn't have so many selfies with the Sagrada Familia otherwise.

To my queens from the Chinese faction, Liu and Lian, who have also been around for almost my entire thesis and have taught me to love their culture and spicy snacks. Much strength for the little bit you have left.

Uri was the wisest of all of us and left the PhD in time before it hurt him. It was a gift to meet you in the master's program, to survive the lockdown being far away and to learn from how brave you are.

And Elisa, who came for a master's program and also knew how to leave in time. Your simple way of being there has helped me understand the world better. I wish you all the best in your new stage.

To Drs. Yuli and Mélina, great people who taught me WGCNA: thanks for your patience.

To our animal genomics family from Torre. I hope you go for beers many times, I leave you my legacy as the Disney princess of the sparrows.

If these people are the ones who have marked me the most within my department, the rest of CRAG is not far behind. I send a big hug to 000 - Tot el CRAG, but here I must also make special mentions.

Over these four years, I've had the best office mates, and that includes my wonderful Lorena and Ari, who are suffering the end of the thesis alongside me. Lorena has clearly saved me more than once (like with that demon Turnitin) and has always been there. Defending on the same day will be a perfect way to show that the PhD brought us together and we will move forward together even if we're on opposite sides of the world. And Ari, the queen of CRAG, I've never seen a kinder person more concerned about others, and I can never thank you enough for everything you've taught me by being just the way you are.

And I've shared little time in that office with Paloma, one of the most authentic people I've met. We've got quite a few memes from these months, and I know you'll keep the level of content creation high. Thank you for welcoming me into your home and for not being scared by my knowledge of you-know-what website.

To my Dr. Carlitos, who has been a great support in recent years. Our afternoons of gossiping and fixing the world should never end. The card of the moon will always appear for you wherever I go.

I met La Cris™ at a thesis party, thanks to the fact that we share the same name. And since then I haven't had a better friend and namesake. Thanks to you and Marc for showing us Manresa.

María was a discovery I must thank my social alcoholism for, as well as her contract-signing partner, Raquel. These now wonderful and creative doctors keep teaching me that there is life (and therefore, hope) beyond science. Thank you for that.

To the beautiful twins, Camila and Violeta, I wish we had shared more pure air (the real kind). I came out of my shell partly thanks to you.

Nicole, the brightest star, I know you'll achieve the greatest things.

To the Cañitos™, the IRTA faction, to all the groups that have welcomed me once or more to eat or drink (Maroof et al., Jaime et al., Clara et al., Héctor et al…), I can never thank you enough for including me in those moments. I've been happy with you.

To Mª José, Alberto, and all the staff who always have a smile or a moment to help. Thank you for every moment.

To Alexandra Elbakyan. If you know, you know.

To all the people who come after, to the future members of the Folchies, the Marceliños, the Botigueta and a long etcetera to include all of CRAG.

It's true that this place called CRAG has been my life for four long years, but there have also been people who have been with me outside those colourful walls, often through WhatsApps and calls.

If we got together in high school because there was no other choice, over the years I've seen that it was a great decision to have Marta, Anna, Annamari and Àngela in my daily life. We've spent more years apart than we were together, but I know that, in reality, we will always be there for each other. A special mention to Marta for gossiping day in and day out on WhatsApp. You brighten my days.

University gave me Irene, whose level of sarcasm is even greater than mine. Good luck with the rest of your PhD, we'll go on to be movie critics of the best trash films afterwards.

And during university, thanks to a literary contest, I met Arancha, a top writer with great potential to be a podcaster. Over the years we've endured countless audios from each other. And the ones still to come. You were the best prize of the La Caixa contest.

Speaking of top writers, the world of literature also brought me Laura, my Scorpio icon, who surely knows my biggest secrets and still decided to stay. I always want you in my life. For more adventures like those in Bilbao. Next up is Helsinki. And Chiang Mai.

Thank you to Judith, for listening to me every Thursday at five for almost two years. If I haven't lost the rest of my mental health, it's thanks to her.

To the one who was my landlady and almost my family for three years, Neus, her little girls, and Queen. Thank you for accepting me.

To Elena (Elena Contel Maza, @helen.mushu.science), who designed the beautiful cover of this thesis and is an even better person than an artist. Never stop creating.

To my mother, the greatest being of light I've ever seen, who gave me life, the wings to fly high, and the desire to be more and a better person. If I am who I am now, it's thanks to you, your encouragement, and the way you support me even when neither of us thought it would be possible. If I am strong, it's for you; if I am free, it's thanks to you. I love you.

To my father, the pillar of my world, the roots of my land, and the wisdom embodied in a man with a ponytail. The one who taught me the true meaning of the word "love" simply by being there for my mother and me, and who taught me to live simply by living. Who read stories to me even though he didn't like literature and hoped I would write them because that's what I liked. If I am free, it's for you; if I am strong, it's thanks to you. I love you, and I will miss you every day of my life. This thesis is for you, Dad.