Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Fecal microbiota profiles across growth predict carcass fat traits and reveal host genetic signals in Iberian pigs

**Lamiae Azouggagh**
Universitat Politècnica de Valencia
**Cristina Casto-Rebollo**
Lohmann Breeders
**Pilar Hernández-Pérez**
Universitat Politècnica de Valencia
**Luis Varona**
Universidad de Zaragoza
**Joaquim Casellas**
Universitat Autònoma de Barcelona
**Sara Negro**
Inga Food S.A
**Noelia Ibáñez-Escriche**
noeibes@dca.upv.es

Universitat Politècnica de Valencia

**Research Article**

**Keywords:** Fatty acids, Gut microbiome, Pigs, Meat quality, Genetic correlation

**Posted Date:** September 11th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9578385/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Fecal microbiota profiles across growth predict carcass fat traits and reveal host genetic signals in Iberian pigs

Lamiae Azouggagh[1], Cristina Casto-Rebollo[2], Pilar Hernández-Pérez[1], Luis Varona[3], Joaquim Casellas[4], Sara Negro[5] and Noelia Ibáñez-Escriche[1*]

[1] Institute for Animal Science and Technology, Universitat Politècnica de Valencia, 46022 Valencia, Spain

[2] Lohmann Breeders, 27472 Cuxhaven, Germany

[3] Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Zaragoza, Spain,

[4] Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain,

[5] Inga Food S.A, 06200, Almendralejo, Spain.

* Corresponding author

Email addresses:

LA: lazolam@upv.edu.es

CCR: cricasre@posgrado.upv.es

PHP: phernan@dca.upv.es

LV: lvarona@unizar.es

JC: joaquim.casellas@uab.cat

SN: sara.negro@nutreco.com

NIE: noeibes@dca.upv.es

## Abstract

### Background

Fat composition in pigs is a key determinant of meat quality and nutritional value and is influenced by both genetic and environmental factors. However, reliable methods to assess fat and fatty acid traits in live animals remain limited, restricting early selection in breeding programs. The gut microbiota has emerged as a promising source of predictive information due to its role in host metabolism and fat deposition. In this context, this study aimed to evaluate the potential and genetic utility of fecal microbiota as a pre-slaughter predictor of fat content and fatty acid composition in Iberian pigs across different developmental stages.

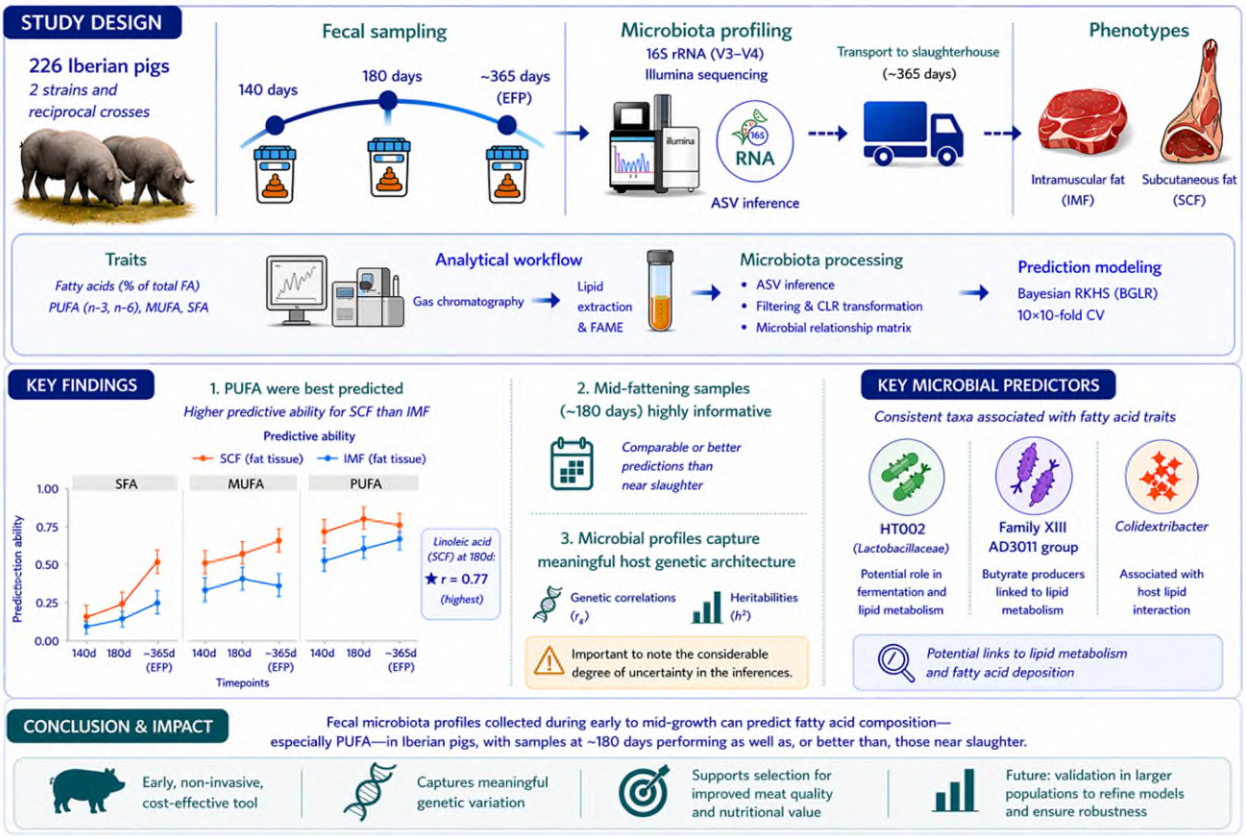
### Results

A total of 226 Iberian pigs, including two strains and their reciprocal crosses, were analyzed. Fecal samples were collected at 140 days, 180 days, and approximately 365 days of age. Microbial communities were characterized through sequencing of the V3–V4 region of the 16S ribosomal RNA gene. Microbial features were processed and used to build relationship matrices, which were then incorporated into Bayesian models to predict fat-related traits. Predictive ability varied across traits, with polyunsaturated fatty acids showing the highest accuracy, particularly linoleic acid in subcutaneous fat (0.77). In general, predictions were more accurate for subcutaneous fat than for intramuscular fat. Samples collected at mid-fattening (around 180 days) often provided predictive performance comparable to or better than those obtained at the end of the fattening period. Genetic correlations between predicted and observed traits suggested that microbial profiles captured biologically relevant host variation. Several microbial taxa, including HT002, Colidextribacter, and members of the Family XIII AD3011 group, were identified as important contributors to prediction.

## Conclusions

Fecal microbiota profiles show strong potential as a pre-slaughter predictor of fat and fatty acid traits in pigs, including at early life stages. These results support their genetic utility and highlight their potential application in breeding programs to improve meat quality and nutritional value. However, further validation in larger and independent populations is required to confirm the robustness and general applicability of these predictive approaches.

## Graphical Abstract





## Keywords

*Fatty acids, Gut microbiome, Pigs, Meat quality, Genetic correlation*


## Background

Iberian pig products, particularly dry-cured ham, are internationally renowned for their distinctive sensory attributes, including tenderness, flavor, and aroma [1]. These qualities are closely linked to the high intramuscular fat (IMF) content and favorable fatty acid (FA)

composition [2]. Fat content and FA profiles are central not only to meat quality but also to nutritional value. The relative proportions of saturated (SFA), monounsaturated (MUFA), and polyunsaturated fatty acids (PUFA) influence both product characteristics and consumer health. While greater PUFA concentrations enhance nutritional value, their susceptibility to oxidative degradation can negatively impact meat stability and shelf life [3]. Importantly, not all PUFA exert the same effects: n-6 are essential for growth and immune function but can promote pro-inflammatory pathways when present in excess [4], whereas n-3 PUFA provide cardioprotective and anti-inflammatory benefits [5]. Consequently, the n-6:n-3 ratio is a recognized indicator of nutritional quality, and the predominance of n-6 PUFA in western diets underscores the need for a more balanced profile. However, the enrichment of n-3 PUFA also increases susceptibility to lipid oxidation [6], which is particularly relevant for dry-cured products where fat stability is critical. Conversely, SFA contribute positively to fat texture and oxidative stability, yet excessive intake has been associated with adverse health effects [7]. Moreover, reducing SFA in favor of PUFA may produce softer fat, creating processing and storage challenges in Iberian pig production [8]. On the other hand, MUFA, particularly oleic acid, are generally considered beneficial for cardiovascular health [9], while also enhancing oxidative resistance, thus representing a desirable balance between sensory, technological, and nutritional attributes.

Fatty acid profiles result from a complex interplay of diet, genetics, and other environmental factors, which have been shown to shape fat deposition and carcass composition in pigs. Numerous studies have highlighted the influence of diet on carcass and muscle FA profiles [10]. Genetic contributions are also well documented [11–14], with moderate to high heritability estimates reported for intramuscular fat (IMF) and FA composition traits. These findings indicate considerable potential for selection, making the optimization of FA classes essential to reconcile breeding objectives with product quality and health-related traits. Notably,

as consumer preferences shift toward healthier meat products, breeding programs are increasingly prioritizing the improvement of the nutritional value of meat without compromising technological or sensory qualities [15].

More recently, the gut microbiota has emerged as an additional factor contributing to variation in pig traits, including feed efficiency [16], carcass composition and meat quality [17], and immune function [18]. Additionally, the microbiota interacts with host genetics to regulate fat deposition, with Tiezzi et al. (2021) providing evidence that the effects of certain genetic variants may be partially mediated by the gut microbiota.

Despite this growing body of research, the contribution of microbial features to the regulation or prediction of FA composition remains insufficiently characterized.

At the same time, advances in phenotyping technologies highlight the need for reliable, early-life indicators of fat quality. Near-infrared spectroscopy (NIRS) and hyperspectral imaging show promise in estimating FA profiles [20,21], but these approaches largely depend on postmortem samples or superficial tissue analysis, limiting their use for selection in live animals. Ultrasound, though non-invasive, provides inconsistent IMF content estimates in Iberian pigs under routine conditions (unpublished observations). These limitations underscore the importance of alternative predictive data sources. In this context, fecal samples collected through rectal swabs reflect gut microbiota composition and may offer a real-time window into host biology. Utilizing these microbiota profiles could therefore provide breeders with a practical tool to anticipate fat quality traits earlier in the production cycle.

The objective of this study was to evaluate the potential and genetic utility of gut microbiota composition, derived from fecal samples collected at different stages of the fattening period, as a pre-slaughter predictor of intramuscular fat content and fatty acid composition in both intramuscular and subcutaneous fat of Iberian pigs. In addition, we assessed the genetic relationships between microbiota-based predictions and phenotypes measured using gas

chromatography, in order to determine the feasibility of incorporating microbiota-derived information into breeding programs aimed at improving meat quality and nutritional value.

## Material and methods

### *Animals and phenotypes*

Data were collected from 226 purebred Iberian pigs of two strains (Entrepelado and Retinto) and their reciprocal crosses, all fattened at the Iberian Testing Center (CTI) of Inga Food S.A. (Almendralejo-Extremadura, Spain). Castrated males entered the CTI facility at 95 ± 13 days of age, with an average weight of 33 ± 10 kg, and were fed a commercial diet ad libitum using automatic feeders.

The diet was formulated as a complete feed for growing pigs and was primarily based on cereals and plant protein sources (including maize, wheat, barley, soybean meal, and rapeseed meal), supplemented with minerals, vitamins, and technological additives. According to the manufacturer's specifications, the diet contained 14% crude protein, 3% crude fat, 4% crude fiber, and 5.1% crude ash.

Animals were introduced into the facility in consecutive batches between May 2021 and June 2022, and were housed in pens, with a maximum capacity of 13 animals per pen. Slaughter took place at a commercial abattoir at an average age of 365 ± 35 days, after a 12-hour fasting period, with animals weighing approximately 160 ± 10 kg at the time of slaughter.

After slaughter, a 200 g sample of the Longissimus thoracis at the fourth rib location was collected from the carcass to analyze intramuscular fat (IMF) content and its fatty acid (FA) profile. Fatty acid methyl esters (FAMEs) were prepared by transesterification in the presence of sulfuric acid (5% sulfuric acid in methanol), following the method described by Sandler & Karo (1992). Separation and identification of fatty acid methyl esters (FAMEs) were performed by gas chromatography using a gas chromatograph (Hewlett-Packard 4890 Series II) equipped

with a split/splitless injector and a flame ionization detector. Fatty acids were identified by comparing the retention times of the methyl esters with those of standards analyzed under the same conditions [23]. Subcutaneous fat (SCF) analysis was conducted on a 50 g sample of adipose tissue collected from the sacral region, processed using the same analytical workflow. A total of 15 fatty acids were identified and quantified in both IMF and SCF, and their concentrations were expressed as percentages of the total identified fatty acids. The fatty acids analysed included myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1), margaric acid (C17:0), margaroleic acid (C17:1), stearic acid (C18:0), oleic acid (C18:1n-9), linoleic acid (C18:2n-6), α-linolenic acid (C18:3n-3), arachidic acid (C20:0), eicosenoic acid (C20:1n-9), eicosadienoic acid (C20:2n-9), dihomo-γ-linolenic acid (C20:3n-6), arachidonic acid (C20:4n-6), and dihomo-α-linolenic acid (C20:3n-3). In addition to individual fatty acids, the overall composition was grouped into major classes, including saturated fatty acids (SFA), monounsaturated fatty acids (MUFA), and polyunsaturated fatty acids (PUFA). Furthermore, several fat quality indices were derived, including nutritional ratios (PUFA:SFA and n-6:n-3 PUFA) and desaturation ratios (MUFA:SFA and the oleic:stearic acid ratio, C18:1n-9/C18:0).

***Microbial abundance data***

A total of 554 rectal swabs were collected from the animals across three sampling rounds at three different timepoints: 171 samples at 140 days (143 ± 6 days), 171 samples at 180 days (183 ± 6 days), and 212 samples at the end of the fattening period (EFP, ~365 ± 35 days). Samples were obtained at the CTI facilities and stored at -80°C to preserve integrity for subsequent analyses.

Bacterial DNA extraction and sequencing of V3–V4 region of the 16S rRNA gene were performed following standardized protocols as described in a previously published work [24]. The primers used for amplification were: forward (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNCCTACGGGNGGCWGCAG

-3') and reverse (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNGACTACHVGGGTATCTAATCC-3'). Raw reads were quality-checked using FastQC [25] and MultiQC [26], and then processed through the nf-core/ampliseq pipeline [27], employing DADA2 software [28] for Amplicon Sequence Variant (ASV) inference.

Reads were truncated at 280 bp (forward) and 220 bp (reverse) to remove low-quality bases, with a minimum read length of 100 bp and minimum ASV length of 400 bp. The sample inference mode was set to "pooled", and taxonomic classification was performed against the SILVA v.138 database, with non-bacterial ASVs (including mitochondria, chloroplasts, archaea, and eukarya) excluded from downstream analysis. DADA2 inferred a total of 15,579 ASVs across the full dataset. After subsetting samples by timepoint and excluding non-bacterial ASVs (mitochondria, chloroplasts, archaea, and eukarya), 9,296, 10,006, and 11,474 bacterial ASVs were observed for the 140d, 180d, and EFP timepoints, respectively. These initial ASV tables served as the input for downstream filtering in R. To reduce noise from spurious or low-prevalence features, ASVs with a relative abundance below 0.001% were excluded. This prevalence filtering resulted in a final dataset of 2,762 ASVs for 140d, 2,863 ASVs for 180d, and 3,159 ASVs for EFP.

Animals were included if they had phenotypic records for at least one of the traits under study and microbiota data available for at least one of the three sampling timepoints (140d, 180d, or EFP). Phenotypes were recorded at a single post-slaughter timepoint. Table 1 summarizes the number of animals with both phenotypic records at slaughter and microbial data available at each timepoint.

**Table 1.** Number of animals with phenotypic and microbial data across timepoints

| Phenotype category | 140d | 180d | EFP |
|---|---|---|---|
| IMF content (%) | 138 | 137 | 186 |
| Fatty acid profile (%) and fat quality indices in IMF | 164 | 163 | 215 |
| Fatty acid profile (%) and fat quality indices in SCF | 165 | 164 | 225 |

EFP End of the fattening period, IMF Intramuscular fat, SCF Subcutaneous fat.

***Predictive modeling using Bayesian regression***

A Bayesian Reproducing kernel Hilbert space (RKHS) linear regression model [29] was used to develop predictive models using microbiota data at each time point. Models were implemented using the BGLR package in R [30], and were defined for each phenotype-timepoint combination as follows:

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{W}\boldsymbol{m} + \boldsymbol{\varepsilon} \quad (1)$$

Where $\boldsymbol{y}$ is the vector of records for the trait under study without adjustment, except for IMF, which was previously adjusted at 300 days to account for the variability at slaughter age. The adjustment was performed only for IMF because it is significantly affected by age, whereas fatty acid composition shows little to no variation with age. IMF was adjusted to 300 days rather than the average age because this later corresponds to the typical slaughter age for Iberian pigs, which is the stage of practical interest for meat quality evaluation. $\mu$ is the overall mean, $\mathbf{1}$ is a vector of ones of length *n*, $\boldsymbol{W}$ is the incidence matrix linking microbiota effects to phenotypic records and equals the identity matrix; $\boldsymbol{m}$ is the vector of random microbiota effects (137-225 levels, depending on the dataset) and $\boldsymbol{\varepsilon}$ is the vector of residuals with $n \times 1$ dimension, where n is the number of records.

The phenotypic data (**y**) were assumed to be conditionally normally distributed as :

$$\boldsymbol{y}|\mu, \boldsymbol{m}, \sigma_e^2 \sim \boldsymbol{N}(\mu + \boldsymbol{W}\boldsymbol{m}, \boldsymbol{I}\sigma_e^2),$$

where a flat prior distribution was assumed for $\boldsymbol{\mu}$. Normal distributions were assumed for $\boldsymbol{m} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{M}\sigma_m^2)$ and $\boldsymbol{\varepsilon} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{I}\sigma_e^2)$, where $\sigma_m^2$ is the associated microbial variance and $\sigma_e^2$ is associated to the residual variance. Both $\sigma_m^2$ and $\sigma_e^2$ were assigned a scaled-inverse chi-squared prior. $\boldsymbol{M}$ is the microbial kernel matrix and $\mathbf{I}$ an identity matrix. $\boldsymbol{M}$ was computed separately for each phenotype-timepoint combination from the Centered log-ratio (CLR)-transformed ASV abundance matrix $\boldsymbol{X}$, following the method proposed by Difford et al. (2018) :

$$\boldsymbol{M} = \frac{\boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}}{n} \quad (2)$$

Where $\boldsymbol{X}$ is the matrix of CLR-transformed ASV abundances, with dimension $p \times n$, $p$ is the number of animals and $n$ the number of ASVs.

The BGLR models were run for 50,000 iterations, with a 5,000-iteration burn-in period and a thinning interval of 5. A single MCMC chain was used for each model, and convergence was visually assessed through inspection of trace plots.

***Inference, model training and validation***

Posterior means of the intercept ($\hat{\mu}$) and microbiota effects ($\widehat{\boldsymbol{m}}$) were used to compute the predicted phenotypes:

$$\widehat{\boldsymbol{y}} = \mathbf{1}\hat{\mu} + \boldsymbol{W}\widehat{\boldsymbol{m}} \quad (3)$$

Predictive ability of the models was assessed using a repeated 10-fold cross-validation approach. The dataset was randomly partitioned into ten mutually exclusive groups of approximately equal size. In each iteration, nine groups were used to train the Bayesian RKHS model, and the remaining group was used for validation. This procedure was repeated ten times so that each group served once as the validation set (10-fold cross-validation). To reduce the influence of random partitioning and obtain more stable estimates of predictive ability, the entire 10-fold cross-validation procedure was repeated ten times using different random seeds. Predictive ability was determined by calculating the Pearson correlation between observed ($\boldsymbol{y}$) and predicted phenotypes ($\widehat{\boldsymbol{y}}$), averaged across the ten cross-validation rounds. Additionally,

the mean absolute error (MAE) of predictions was computed to assess the magnitude of prediction error.

***Genetic analysis of observed and predicted phenotypes***

A Bayesian bivariate animal model was used to estimate genetic correlations ($r_g$) between observed phenotypes ($\boldsymbol{y}$) and the predicted values using the microbiota information ($\hat{\boldsymbol{y}}$), as follows:

$$\begin{bmatrix} \boldsymbol{y} \\ \hat{\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_a} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_a} \end{bmatrix} \begin{bmatrix} \boldsymbol{a_1} \\ \boldsymbol{a_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_l} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_l} \end{bmatrix} \begin{bmatrix} \boldsymbol{l_1} \\ \boldsymbol{l_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e_1} \\ \boldsymbol{e_2} \end{bmatrix} \quad (4)$$

Where $\boldsymbol{y}$ is the vector of observed phenotypes and $\hat{\boldsymbol{y}}$ the vector of predicted phenotypes using microbiota, $\boldsymbol{X}$, $\boldsymbol{Z_a}$ and $\boldsymbol{Z_l}$ are the corresponding incidence matrices associated with fixed effects ($\boldsymbol{b_1}$,$\boldsymbol{b_2}$) of contemporary group (12-21 levels, according to animals in each phenotype-timepoint dataset), random additive genetic effects ($\boldsymbol{a_1}$,$\boldsymbol{a_2}$) with 137-225 levels, and random litter effects ($\boldsymbol{l_1}$,$\boldsymbol{l_2}$) with 92-149 levels respectively. ($\boldsymbol{e_1}$,$\boldsymbol{e_2}$) are the vectors of random errors.

Data were assumed to be conditionally distributed as:

$$\begin{bmatrix} \boldsymbol{y} \\ \hat{\boldsymbol{y}} \end{bmatrix} | \, \boldsymbol{b_1}, \boldsymbol{b_2}, \boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{l_1}, \boldsymbol{l_2}, \boldsymbol{R} \sim \boldsymbol{N}\left( \boldsymbol{X} \begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{bmatrix} + \boldsymbol{Z_a} \begin{bmatrix} \boldsymbol{a_1} \\ \boldsymbol{a_2} \end{bmatrix} + \boldsymbol{Z_l} \begin{bmatrix} \boldsymbol{l_1} \\ \boldsymbol{l_2} \end{bmatrix}, \boldsymbol{R} \right)$$

where $\boldsymbol{R}$ is the residual (co)variance matrix between the two traits.

The (co)variances were assumed to be :

$$V = Var\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{l} \\ \boldsymbol{e} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} \otimes \boldsymbol{G_0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I_l} \otimes \boldsymbol{L_0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I_e} \otimes \boldsymbol{R_0} \end{pmatrix}$$

Where $\boldsymbol{A}$ is the relationship matrix of the same order as the number of animals in the pedigree; $\boldsymbol{I_l}$ is an identity matrix of the same order as the number of levels of litter effects; and $\boldsymbol{I_e}$ is an identity matrix of the same order as the number of records. $\boldsymbol{G_0}$, $\boldsymbol{L_0}$,and $\boldsymbol{R_0}$ are the 2 × 2 additive genetic, litter, and residual (co)variance matrices between the two traits.

The model was solved by Gibbs sampling using GIBBSF90 software of the BLUPF90 family programs [32]. Flat priors were considered for systematic effects and co(variances) in all the models. The Gibbs samplers were run for a total of 2,000,000 iterations, with the first 1,000,000 discarded as burn-in. Sampling was performed with a thinning interval of 50 samples, yielding 20,000 samples for post-Gibbs analysis, which was conducted using POSTGIBBSF90 program. Convergence of parameter estimates was assessed using Geweke's diagnostic [33] and visual inspection of MCMC chain distributions. For each marginal posterior distribution of the genetic correlation ($r_g$), we obtained the posterior mean, the 95% highest posterior density $HPD_{95\%}$ interval, the probability of the estimation being higher than 0.00 ($P_0$) and the probability of relevance which is the probability of the estimation being higher than 0.30 ($P_r$).

***Feature Importance analysis***

To identify the most influential microbial taxa in predicting fatty acid traits, a permutation-based feature importance analysis was conducted using the Random Forest algorithm [34]. The analysis was performed independently for each trait and tissue combination.

Briefly, the predictor matrix consisted of CLR-transformed ASV counts, while the response variable was the phenotypic value of the target trait (%IMF, fatty acid or ratio). For each trait, a Random Forest Regressor with 250 trees was trained using 10-fold cross-validation. Within each fold, the model was fitted on the training set, and permutation importance was computed on the test set by randomly shuffling each feature and measuring the resulting decrease in model performance ($R^2$). This process was repeated 5 times per feature to improve stability, and the mean importance score was recorded. Feature importance was averaged across all 10 folds to generate a final, robust estimate for each ASV. The overall importance of each taxon was assessed by ranking features based on their mean permutation importance (decrease in $R^2$). Results were sorted in descending order to identify the most relevant predictors. Top-

277 performing ASVs and their taxonomic assignments were reported. All analyses were
278 implemented in Python (v.3.11.5) using scikit-learn module [35].

279 **Results**

280 ***Characterization of fatty acid composition***

281 Fatty acids with 16 and 18 carbon atoms were the predominant components in both
282 intramuscular fat (IMF) and subcutaneous fat (SCF). Oleic acid was the most abundant,
283 averaging 51.54% in IMF and 49.62% in SCF, followed by palmitic acid (24.95% in IMF, and
284 22.33% in SCF). The third most prevalent fatty acid differed between depots: stearic acid
285 ranked third in IMF (10.39%), whereas linoleic acid was more prominent in SCF (10.30%).
286 These compositional differences resulted in SCF showing slightly greater desaturation indices
287 (oleic:stearic and MUFA:SFA ratios) compared with IMF.
288 At the group level, monounsaturated fatty acids (MUFA) were most abundant, representing
289 57.23% in IMF and 54.02% in SCF, followed by saturated fatty acids (SFA: 37.11% in IMF,
290 and 33.92% in SCF). Polyunsaturated fatty acids (PUFA) were the least represented, though
291 more prevalent in SCF (12.07%) than in IMF (5.66%). The intramuscular fat percentage
292 (%IMF) averaged 10.53%, ranging from 1.88–23.24% across individuals. Basic statistics of all
293 traits are presented in Additional file 1 (Tables 1 and 2).

294 ***Microbial predictive ability of fat quality traits***

295 Predictive abilities were highly dependent on trait, timepoint, and tissue, with values ranging
296 from -0.11 to 0.77 and standard deviations of 0.17–0.27 across validation folds. Corresponding
297 predictive abilities and MAE values are provided in Additional file 2 (Tables 1–4).
298 At the group level (SFA, MUFA, PUFA including n-3 and n-6), predictive ability increased
299 with the degree of unsaturation, as shown in Fig. 1. For example, in IMF at 180d, SFA showed

low predictive ability (0.11 ± 0.27), MUFA moderate predictive ability (0.39 ± 0.24), and PUFA the greatest predictive ability (0.62 ± 0.19). Across tissues, SCF generally yielded greater predictive abilities for FA group totals than IMF, with the exception of omega-3 PUFA, which were better predicted in IMF (predictive abilities ranging from 0.40-0.48 in SCF vs. 0.55-0.63 in IMF). Predictive ability also improved with age, being higher at 180d compared with 140d. However, at the EFP timepoint, predictive abilities did not consistently improve and in some cases declined (e.g., MUFA in IMF).

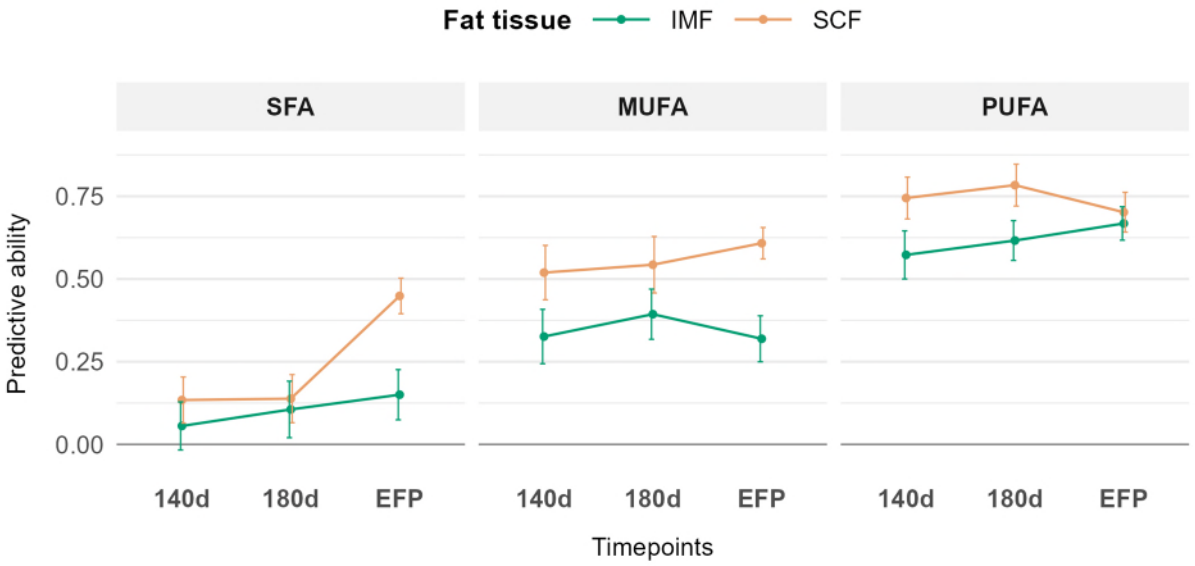



**Fig. 1 Predictive ability of fatty acid groups**.

Predictive ability of saturated (SFA), monounsaturated (MUFA), and polyunsaturated (PUFA) fatty acids using microbiota data (amplicon sequence variants, ASVs) collected at 140 days (140d), 180 days (180d), and the end of the fattening period (EFP) in intramuscular fat (IMF) and subcutaneous fat (SCF) of Iberian pigs. Points represent the mean predictive ability across 10 cross-validation replicates, and error bars indicate the standard error (SE).

At the individual FA level, the same pattern was observed (Additional file 1, tables 1-2). PUFA such as linoleic acid and α-linolenic acid were the best predicted fatty acids in both fat tissues. For example, linoleic acid in IMF reached 0.60 ± 0.20 at 140d, 0.65 ± 0.18 at 180d, and 0.69 ±

317 0.16 at EFP. Across tissues, SCF yielded better predictions for most individual fatty acids,
318 particularly at the final timepoint. On average, predictive ability in SCF was 0.40 ± 0.21,
319 compared with 0.28 ± 0.21 in IMF. The largest difference in prediction between tissues, was
320 observed for eicosenoic acid at 180d, with a predictive ability of 0.53 ± 0.23 in SCF compared
321 with –0.11 ± 0.23 in IMF.
322 Nutritional and desaturation ratios also showed variable predictive abilities depending on tissue
323 and timepoint (Additional file 2, tables 3-4). Among PUFA-based ratios, PUFA:SFA was
324 predicted with moderate to moderately high predictive ability (0.49–0.59 in IMF vs. 0.48–0.56
325 in SCF). In contrast, the n-6:n-3 ratio was predicted much more accurately in SCF (0.67–0.77)
326 than in IMF (0.27–0.45). Desaturation indices (MUFA:SFA and oleic:stearic) generally yielded
327 lower values overall (0.12–0.27 in IMF vs. 0.09–0.51 in SCF). In SCF, predictive abilities of
328 all nutritional and desaturation ratios improved at the EFP timepoint compared with earlier
329 ages; for instance, MUFA:SFA increased from 0.30 ± 0.23 at 180d to 0.51 ± 0.15 at EFP
330 (~70%). In IMF, PUFA-based ratios increased at EFP (by 8-50% with respect to 180d), whereas
331 desaturation indices decreased (by 14-48% with respect to 180d) (Additional file 2, tables 3-4).
332 The prediction of intramuscular fat percentage (%IMF) was consistently low across all
333 timepoints, with the best predictive ability reaching only 0.17 ± 0.25 at EFP (Additional file 2,
334 table 1). In several folds, predictive abilities approached zero or became negative, underscoring
335 the difficulty of predicting this trait with the current microbiome-based models.

336 ***Genetic parameters of microbiota-predicted traits***

337 Heritability estimates ($h^2$) of microbiota-predicted phenotypes were generally comparable to
338 those obtained from gold standard measurements across most FA traits and tissues (Additional
339 file 3, tables 1-2). Estimates were not always more precise at the later timepoint (Fig. 2), despite
340 the increased amount of data used for estimation. A slight tissue-specific patterns in heritability

emerged. Microbiota-predicted phenotypes related to PUFA (including individual fatty acids, group totals and ratios) showed in average greater posterior mean heritability in IMF than in SCF (average $h^2$ of 0.33 in IMF vs. 0.21 in SCF for PUFA; Fig. 2). In contrast, MUFA and SFA related traits exhibited lower $h^2$ in IMF than in SCF (average $h^2$ of 0.29 vs. 0.37 for MUFA and 0.24 vs. 0.41 for SFA; Fig. 2). Similarly, microbiota-predicted %IMF showed moderate $h^2$, ranging from 0.34 to 0.38 across timepoints (Additional file 3). It is important to note that all estimates were characterized by wide HPD95% intervals, reflecting a considerable degree of uncertainty.

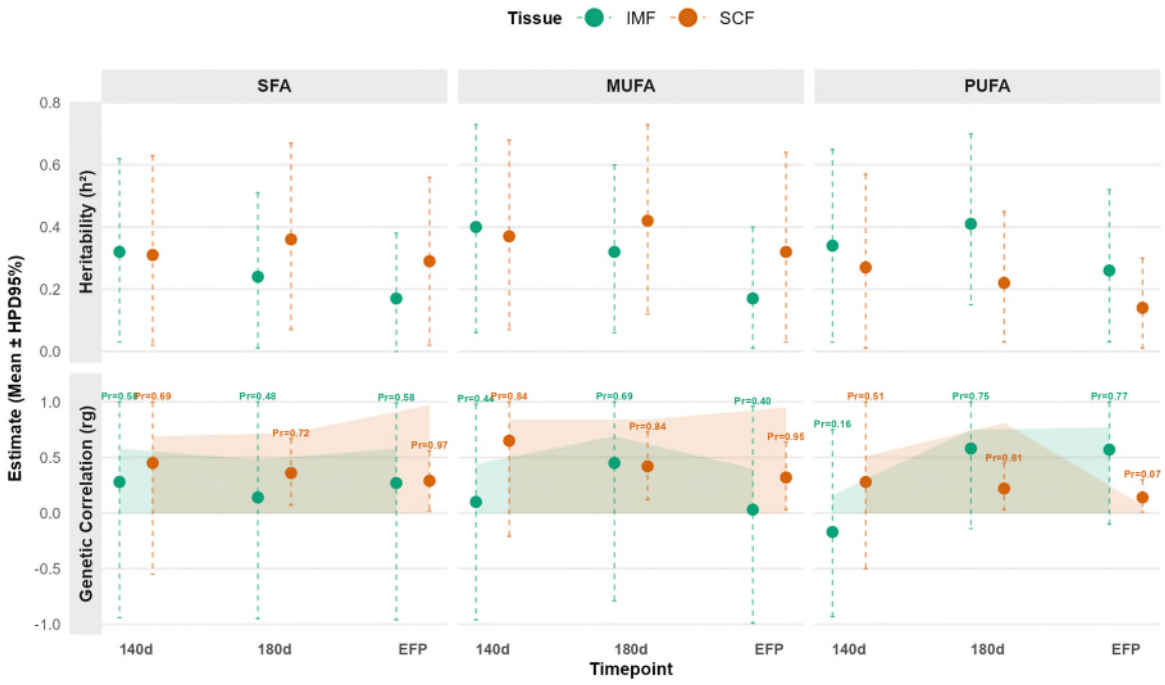



**Fig. 2 Posterior distributions of heritability and genetic correlations**.

Posterior distributions of heritability (h²; top panels) and genetic correlations ($r_g$; bottom panels) for microbiota-predicted fatty acid groups (SFA, MUFA, PUFA) in intramuscular fat (IMF; green) and subcutaneous fat (SCF; orange) of Iberian pigs. Estimates were obtained using microbiota data collected at 140 days (140d), 180 days (180d), and the end of the fattening period (EFP). Points represent posterior means, and vertical bars indicate 95% highest posterior density intervals (HPD95%). Shaded areas represent the posterior probability that $r_g$ exceeds

0.30 ($P_r$, probability of relevance), with numeric labels indicating the corresponding probability values.

Genetic correlation estimates ($r_g$) were more precise for SCF than for IMF, as reflected by narrower HPD95% intervals (Fig. 2 and Table 2). Consistent with this higher precision, the probabilities of relevance ($P_r$) were generally more robust in SCF, whereas in IMF, they displayed considerable heterogeneity. While PUFA in IMF showed increasing relevance over time, rising from a weak genetic signal at 140d ($P_r$ = 0.16) to a moderate-to-strong signal at 180d and EFP ($P_r$ ≈ 0.75–0.77), SFA and MUFA traits consistently exhibited a modest signal ($P_r$ generally ranging from 0.40 to 0.60) across all stages. In contrast, SCF exhibited consistently robust probabilities of relevance for SFA and MUFA traits. Notably, total MUFA maintained a stable, high probability ($P_r$ = 0.84–0.95) across all developmental stages, and the MUFA:SFA ratio reached near-certainty at EFP ($P_r$ = 0.97). Although PUFA traits in SCF also showed temporal variability ($P_r$ ranging from 0.51 to 0.81), the initial probabilities of relevance at 140d were notably higher than in IMF.

Table 2 presents microbial predictive ability, $r_g$ estimates, HPD95% intervals, $P_0$ and $P_r$ for %IMF, the most accurately predicted fatty acids within each major group, and the group totals in both adipose tissues. Values are reported for the timepoints representing the best compromise between early prediction ability (140d or 180d) and favorable genetic alignment, with 180d generally providing the most balanced estimates. It is important to note, however, that all genetic correlation estimates were characterized by wide HPD95% intervals, indicating a substantial degree of uncertainty in the inferences.

**Table 2.** Microbial predictive ability and genetic parameters for fat traits

| Tissue | Trait | Timepoint[1] | Predictive ability[2] | $r_g$ | HPD95% | | $P_0$ | $P_r$ |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower limit | Upper limit | | |
| **IMF** | IMF (%) | 140d | 0.12 (±0.26) | 0.61 | -0.30 | 1.00 | 0.89 | 0.79 |
| | Linoleic | 180d | 0.65 (±0.15) | 0.61 | -0.13 | 1.00 | 0.91 | 0.78 |
| | α-Linolenic | 180d | 0.65 (±0.19) | 0.49 | -0.20 | 1.00 | 0.86 | 0.67 |
| | Eicosadienoic | 180d | 0.61 (±0.19) | 0.61 | -0.04 | 1.00 | 0.93 | 0.79 |
| | Margaric | 180d | 0.43 (±0.17) | 0.72 | 0.13 | 1.00 | 0.97 | 0.90 |
| | Oleic | 180d | 0.37 (±0.24) | 0.53 | -0.58 | 1.00 | 0.83 | 0.74 |
| | Total PUFA | 180d | 0.62 (±0.19) | 0.58 | -0.10 | 1.00 | 0.90 | 0.75 |
| | Total MUFA | 180d | 0.39 (±0.20) | 0.45 | -0.79 | 1.00 | 0.78 | 0.69 |
| | Total SFA | 180d | 0.11 (±0.25) | 0.14 | -0.94 | 1.00 | 0.58 | 0.48 |
| **SCF** | Linoleic | 180d | 0.77 (±0.20) | 0.67 | -0.05 | 1.00 | 0.94 | 0.85 |
| | α-Linolenic | 180d | 0.64 (±0.27) | 0.78 | 0.25 | 1.00 | 0.98 | 0.93 |
| | Margaric | 180d | 0.41 (±0.20) | 0.86 | 0.51 | 1.00 | 0.99 | 0.98 |
| | Oleic | 180d | 0.53 (±0.27) | 0.69 | -0.11 | 1.00 | 0.93 | 0.86 |
| | Eicosenoic | 140d | 0.53 (±0.23) | 0.77 | 0.21 | 1.00 | 0.88 | 0.93 |
| | Total PUFA | 180d | 0.75 (±0.20) | 0.61 | -0.15 | 1.00 | 0.92 | 0.81 |
| | Total MUFA | 180d | 0.51 (±0.27) | 0.67 | -0.20 | 1.00 | 0.92 | 0.84 |
| | Total SFA | 180d | 0.13 (±0.23) | 0.51 | -0.42 | 1.00 | 0.85 | 0.72 |

IMF Intramuscular fat, SCF Subcutaneous fat, $r_g$ Genetic correlation, HPD95% 95% Highest posterior density intervals, $P_0$ Probability that $r_g > 0$, $P_r$ Probability that $r_g > 0.30$.

Predictive ability was calculated as the Pearson correlation between observed and predicted phenotypes, averaged across 10 cross-validation replicates (mean ± SD). Results correspond to intramuscular fat content (IMF%), key fatty acids, and fatty acid groups (SFA, MUFA, PUFA) at the timepoint representing the optimal compromise between early predictive ability and favorable genetic correlation.

Beyond FA traits, %IMF showed a contrasting profile. Despite its consistently low predictive ability across timepoints, microbiota-predicted %IMF displayed moderate-to-high genetic correlations with direct measurements. Correlations were already moderate at 140d and 180d (0.61, $P_r$ = 0.79) and increased further at EFP (0.77, $P_r$ = 0.91), where the HPD95% interval excluded zero. Complete results for all traits and timepoints are provided in Additional file 3 (Tables 3-4).

***Feature importance analysis of microbial predictors***

After collapsing the most important ASVs at the genus level, recurrent contributors were identified across traits and tissues; notably, ASVs of the HT002 genus consistently ranked among the most influential predictors for PUFA traits in both IMF and SCF at 180d. In SCF, Cellulosilyticum and Colidextribacter were also among the most influential taxa for PUFA predictions. For MUFA, the most important predictors in IMF were ASVs from the Family XIII AD3011 group and Ruminococcus, whereas in SCF, ASVs from Lachnospira and an unidentified genus within Ruminococcaceae were most prominent. For SFA, Ruminococcus again appeared as a key predictor in IMF, while in SCF the Christensenellaceae R-7 group was identified as the most influential genus. Extended lists of most important ASVs and their associated genera for each trait are provided in Additional file 4.

## Discussion

The main objective of this study was to evaluate the potential and the genetic utility of fecal microbiota as a pre-slaughter predictor of fat and fatty acid (FA) traits in Iberian pigs. The use of microbial biomarkers for complex traits in livestock has gained traction in recent years, particularly in the context of animal breeding and live-animal sampling strategies that avoid slaughter. Here, we used fecal microbiota sampled at three timepoints to predict FAs traits from

intramuscular and subcutaneous fat, measured through gas chromatography and explored how this potential evolves over time. Our results provide evidence that fecal microbiota may serve to predict key FA traits with moderate predictive ability, particularly in the case of polyunsaturated fatty acids (PUFA). Given that these predictions capture meaningful genetic signals, this highlights their potential relevance for selection purposes. The analysis of feature importance further revealed biologically relevant taxa consistently associated with lipid metabolism, reinforcing the functional plausibility of these predictions in breeding programs.

***Fecal microbiota as an early predictor of fat quality traits***

Based on earlier studies, we expected that microbiota composition measured closer to slaughter would give better prediction results. This has been observed in pigs, where predictive ability for growth and carcass traits improves when microbiome data are collected at mid-test or off-test stages, rather than earlier in the production cycle [36].

Consistent with this trend, most traits in our study showed improving predictive abilities toward the end of the fattening period (EFP). However, we observed that for certain traits, earlier timepoints yielded predictive abilities comparable to, or even better than, those obtained at EFP, e.g. n-3 PUFA. This suggests that microbial profiles established during early to mid-growth phases may already capture signals relevant to later fatty acid deposition. These results align with emerging evidence indicating that the gut microbiota plays a role in host lipid metabolism beyond immediate nutritional responses [37], possibly through long-term host-microbiota interactions and shared genetic control mechanisms.

The value of microbiota-based predictions in genetic breeding programs goes beyond phenotypic predictive ability. Their usefulness depends on whether these predictions capture heritable variation and are genetically correlated with the true trait of interest. A model may perform well statistically, but if it does not capture the underlying host genetic architecture, its

usefulness in selection is limited. The results of this manuscript showed that for several FA traits (especially PUFA), microbiota-predicted values share a moderate genetic basis with observed phenotypes, exhibiting heritabilities in a similar range. Although the genetic correlations were estimated with considerable uncertainty, the consistency of these patterns indicates that microbial profiles may act as proxies for host genetic effects. This positions fecal microbiota as a candidate indicator for the host's genetic regulation of fat metabolism, rather than solely a tool for phenotypic prediction.

Across both IMF and SCF, group-level traits followed a similar predictability pattern: PUFA were the best predicted, already evident at 180d, with predictive abilities reaching 0.60–0.77 for individual FAs like linoleic acid (C18:2n-6), followed by MUFA and SFA. From a nutritional standpoint, accurate predictive ability of PUFA and related indices, such as the PUFA:SFA ratio, a marker of fat healthiness, is highly valuable. The n-6:n-3 ratio, while strongly predictable in SCF, showed lower predictive ability in IMF, possibly due to error propagation inherent in ratio traits and/or tissue-specific metabolic differences. Among the individual PUFA, linoleic acid stands out not only for its high predictive ability but also for its relevance to meat quality. High levels are associated with reduced firmness in cured products [3], yet it has also been proposed as a biomarker for increased intramuscular fat content without compromising lean growth [38]. This dual interest makes it a compelling target for early-life selection.

Importantly, genetic correlations for PUFA traits were moderate to moderately high, and, despite wide $HPD_{95\%}$ intervals due to limited sample size, remained positive with high posterior probability. These results suggest that microbial profiles encode part of the host’s genetic influence on PUFA metabolism. Although PUFA are largely influenced by dietary intake, moderate to high heritabilities have been reported [39], consistent with our findings. In particular, the genetic determinism observed for n-3 PUFA is thought to arise from variation in

462 genes regulating intestinal absorption from the diet or organ-specific distribution [40]. Our
463 results support this hypothesis and suggest that gut microbiota may mediate genetic effects on
464 fatty acid processing and absorption.
465 In contrast to PUFA, monounsaturated (MUFA) and saturated fatty acids (SFA) are
466 predominantly synthesized *de novo* and are less influenced by dietary lipids [41]. As such, one
467 might expect these traits to exhibit stronger host genetic control compared to diet-sensitive
468 PUFA. However, our results reveal a more nuanced picture: PUFA traits showed greater
469 heritability than MUFA and SFA in IMF with estimates reaching up to 0.42 while the reverse
470 was observed in SCF, where MUFA and SFA exhibited moderate-to-moderately high
471 heritabilities ($h^2$ = 0.35–0.42) and PUFA showed lower values ($h^2$ = 0.22–0.29). This depot-
472 specific pattern aligns with previous findings indicating distinct genetic regulation of fat
473 deposition between adipose tissues [39]. Although predictive ability was generally lower for
474 MUFA than for PUFA in both tissues, likely due to weaker microbial mediation of endogenous
475 lipid synthesis, this pattern did not translate into a proportional decrease in the genetic
476 correlations. In fact, MUFA correlations were greater than those for PUFA in SCF, where they
477 were also more precisely estimated. This suggests that although the microbiota may explain
478 less phenotypic variance in these traits, the predictions that are generated still capture
479 meaningful host genetic signals. Oleic acid (C18:1n-9), the major contributor to total MUFA
480 content, showed prediction performance and genetic parameters consistent with the group-level
481 results though less predictable in IMF compared to SCF (at 180d, predictive ability was 0.37
482 with $r_g$=0.54 for IMF and 0.53 with $r_g$=0.69 for SCF). Notably, oleic acid in SCF has been
483 reported to exhibit a strong genetic correlation with oleic acid in IMF [39]. Therefore,
484 microbiota-based prediction of SCF oleic acid could serve as a predictor for IMF oleic acid
485 content in breeding programs.

Total SFA had inconclusive results both in predictive ability and genetic correlation, but among SFA, margaric acid (C17:0) emerged as one of the best predicted FA in both IMF and SCF with favorable genetic parameter estimates (at 180d, predictive ability was 0.43 with $r_g$=0.73 for IMF and 0.41 with $r_g$=0.86 for SCF). However, its very low abundance in Iberian pig fat limits its practical relevance as a selection target.

We also assessed the oleic:stearic acid (C18:1n-9/C18:0) and the MUFA:SFA ratios, which have been used as an indicator of stearoyl-CoA desaturase activity, the enzyme responsible for converting saturated to monounsaturated FAs [42], and linked to metabolic conditions like obesity [43]. However, in our study, predictive abilities and genetic correlations of these traits were inconclusive particularly in the IMF, limiting their utility as microbiota-informed traits for early selection.

Similarly, the %IMF which is a critical trait for meat quality, as a minimum threshold is essential to ensure meat tenderness and flavor, also showed low predictive ability. However surprising genetic correlations of 0.61-0.77 were found. This mirrors findings for MUFA and indicates that microbiota can rank animals genetically, even when precise phenotypic prediction is limited.

These results suggest two strategies for predicting %IMF: a direct approach via microbes linked to %IMF deposition, or an indirect approach through biomarkers like linoleic acid, as discussed earlier in relation to its potential as an IMF predictor without affecting lean growth. The latter may enhance early prediction, although its reliability requires further validation in Iberian pigs. Overall, predictive ability tended to be higher for SCF than for IMF, reflecting a stronger association between fecal microbiota and SCF composition. Some exceptions were observed, such as n-3 PUFA, which were better predicted in IMF. Together with prior heritability patterns, these results highlight depot-specific regulation, consistent with evidence that part of the genetic control of SCF operates independently of IMF [14,39,44].

511 ***Microbial signatures for fatty acid traits***

512 At the prediction level, microbiota composition emerged as a potential predictor of fatty acid
513 profiles, with several recurrent taxa identified as key contributors to fat trait prediction. In IMF,
514 linoleic acid stood out as a key trait. Our findings align with Sebastià et al. (2024) who reported
515 correlations between linoleic acid levels in the longissimus dorsi muscle and fecal genera such
516 as *Clostridium sensu stricto 1* and *Terrisporobacter*. In our model, ASVs from these genera
517 ranked among the top 10 most important predictors of linoleic acid at 180d. However, ASVs
518 assigned to the genus *HT002* (family *Lactobacillaceae*) exhibited the greatest variable
519 importance at this timepoint. Known for its role in carbohydrate fermentation and lactate
520 production [45], *HT002* contributes to short-chain fatty acid (SCFA) formation and may
521 influence host metabolic status. Notably, *HT002* ASVs also showed the greatest importance
522 scores for multiple PUFA traits, including n-3 and n-6 PUFA in IMF. In subcutaneous fat
523 (SCF), this genus remained among the top five predictors for PUFA-related traits, though it was
524 surpassed by ASVs from *Colidextribacter* and *Cellulosilyticum*. *Colidextribacter* was
525 previously identified as a key discriminator between purebred and crossbred Iberian pigs in our
526 population [24], suggesting a potential link to breed-specific microbial profiles. Meanwhile,
527 *Cellulosilyticum* has been associated with improved feed efficiency and growth performance in
528 pigs [46,47], indicating a broader physiological role that may indirectly modulate lipid
529 metabolism.
530 The predictive performance for MUFA was moderate, yet the identification of influential taxa
531 may provide valuable insights into the microbial drivers of monounsaturated lipid metabolism,
532 and several of these associations are supported by findings from other studies in pigs. Microbial
533 signatures for MUFA also revealed distinct, depot-specific predictors. In IMF, ASVs assigned
534 to the *Family XIII AD3011 group* genus and the *Ruminococcus* genus were ranked as the most
535 important features. The relevance of the *Family XIII AD3011 group* is particularly notable, as

it has been previously associated with increased fat deposition and disease resistance in pigs [48], suggesting a potential direct link to host lipid metabolism. In contrast, the predictive landscape for MUFA in SCF was dominated by ASVs from the genus *Lachnospira*, a known producer of short-chain fatty acids (SCFAs) [49], and an unidentified genus within the Ruminococcaceae family. This aligns with previous findings in Iberian pigs, where Ruminococcaceae OTUs were positively correlated with oleic acid content in backfat [50].

Together, these findings demonstrate that specific members of the gut microbiota may contribute meaningfully to fat composition and highlight the potential to leverage microbial profiles as predictive tools. The use of fecal microbiota therefore presents a promising approach for predicting FA composition. While technologies such as near-infrared spectroscopy (NIRS) and hyperspectral imaging have shown good predictive performance for FA profiles in fat and muscle [20,21,51], their application is largely limited to postmortem or superficial tissues, constraining their utility for in vivo, early-life selection. Our findings suggest that longitudinally sampled gut microbiota captures genetically meaningful signals for both IMF and SCF, and the favorable predictive abilities for key traits by mid-fattening period highlight its potential to complement existing phenotyping strategies in breeding programs.

## Conclusions, limitations and future perspectives

This study provides evidence that gut microbiota composition, assessed via fecal samples collected during early to mid-growth, holds promise as an early predictor of fatty acid profiles in Iberian pigs. Notably, PUFA exhibited the most consistent predictive signals, with microbial profiles at 180 days of age performing comparably to—or in some cases better than—those collected closer to slaughter. While prediction performance was higher for subcutaneous fat than intramuscular fat, this tissue-specific pattern offers valuable insight into where microbiota-based approaches may be most effective. Key microbial taxa, including *HT002*

(Lactobacillaceae), *Family XIII AD3011 group*, and *Colidextribacter*, emerged as consistent predictors, suggesting biologically relevant links to lipid metabolism worthy of further exploration.

The estimated genetic correlations and heritabilities indicate that microbial profiles capture meaningful aspects of the host's genetic architecture, particularly for PUFA and MUFA traits. That said, the modest sample size and resulting wide HPD95% intervals reflect the preliminary nature of these genetic estimates and highlight the value of validation in larger, independent cohorts. Similarly, while the observed predictive patterns are encouraging, refining model performance, especially for intramuscular fat traits, remains an important next step. Overall, these results support using microbiota as an additional tool to help select animals with better meat quality at an early stage, with future work expected to improve its reliability and use in practice.

**List of abbreviations**

**ASV**: Amplicon sequence variant

**CLR**: Centered log-ratio transformation

**CTI**: Iberian Testing Center

**EFP**: End of the fattening period

**FA**: Fatty acid(s)

**FAMEs**: Fatty acid methyl esters

**HPD95%**: 95% highest posterior density interval

**IMF**: Intramuscular fat

**MAE**: Mean absolute error

**MCMC**: Markov chain Monte Carlo

**MUFA**: Monounsaturated fatty acids

**n-3 PUFA**: Omega-3 polyunsaturated fatty acids

**n-6 PUFA**: Omega-6 polyunsaturated fatty acids

**n-6:n-3 ratio**: Ratio of omega-6 to omega-3 fatty acids

**NIRS**: Near-infrared spectroscopy

$\boldsymbol{P_0}$ **:** Probability of the estimation being higher than 0.00

$\boldsymbol{P_r}$ : Probability of the estimation being higher than 0.30 (probability of relevance)

**PUFA**: Polyunsaturated fatty acids

**RKHS**: Reproducing kernel Hilbert space

$\boldsymbol{r_g}$**:** Genetic correlation

**SCF**: Subcutaneous fat

**SE**: Standard error

**SFA**: Saturated fatty acids

## Ethics approval

All management procedures under standard farm management were approved by the Committee of Ethics and Animal Welfare of the Polytechnical University of Valencia, according to Council Directives 98/58/EC and 2010/63/EU.

## Consent for publication

Not applicable.

## Availability of data and material

The phenotypic data (fatty acid composition, intramuscular fat content, and metadata) supporting the findings of this study are available in the Zenodo repository under the DOI: 10.5281/zenodo.20957877. The corresponding 16S rRNA sequencing dataset is available in a companion Zenodo repository under the DOI: 10.5281/zenodo.14013900. All data are publicly accessible without restriction.

## Competing interests

609 The authors declare that they have no competing interests

610 **Funding**

611 This study was funded by the Spanish Ministry of Science and Innovation in the Project
612 PID2020-114705RB-I00.

613 **Authors' contributions**

614 The study was conceived and designed by NIE and SNR. Data collection was carried out by
615 SNR and NIE. Bioinformatic analyses of the 16S and WGS sequencing data were performed
616 by LA and CCR. Methodological aspects were discussed and supervised by NIE, CCR, PHP,
617 LV, and JC. LA conducted the statistical analyses and predictive modelling, interpreted the
618 results, and drafted the manuscript. The results were discussed by LA, CCR, PHP, LV, and
619 NIE. All authors reviewed the manuscript and approved the final version.

620 **Acknowledgements**

621 LA gratefully acknowledges the receipt of the GRISOLIA scholarship [CIGRIS/2021/098]
622 awarded by the Generalitat Valenciana. The authors would like to thank the company INGA
623 FOOD S.A. for providing the animals and the facilities used for data collection.

624 **References**

625 1. Ventanas S, Ventanas J, Ruiz J. Sensory characteristics of Iberian dry-cured loins: Influence
626 of crossbreeding and rearing system. Meat Sci. 2007;75:211–9.
627 https://doi.org/10.1016/j.meatsci.2006.07.003
628 2. Webb EC, O'Neill HA. The animal fat paradox and meat quality. Meat Sci. 2008;80:28–36.
629 https://doi.org/10.1016/j.meatsci.2008.05.029

630 3. Wood JD, Richardson RI, Nute GR, Fisher AV, Campo MM, Kasapidou E, et al. Effects of
631 fatty acids on meat quality: a review. Meat Sci. 2004;66:21–32. https://doi.org/10.1016/S0309-
632 1740(03)00022-6
633 4. Radzikowska U, Rinaldi AO, Çelebi Sözener Z, Karaguzel D, Wojcik M, Cypryk K, et al.
634 The Influence of Dietary Fatty Acids on Immune Responses. Nutrients. 2019;11:2990.
635 https://doi.org/10.3390/nu11122990
636 5. Simopoulos AP. The importance of the ratio of omega-6/omega-3 essential fatty acids.
637 Biomedicine & Pharmacotherapy. 2002;56:365–79. https://doi.org/10.1016/S0753-
638 3322(02)00253-6
639 6. Kamal-Eldin A, Yanishlieva N V. N-3 fatty acids for human nutrition: stability
640 considerations. European Journal of Lipid Science and Technology. 2002;104:825–36.
641 https://doi.org/10.1002/1438-9312(200212)104:12<825::AID-EJLT825>3.0.CO;2-N
642 7. Zock PL. Health problems associated with saturated and trans fatty acids intake. Improving
643 the Fat Content of Foods. Elsevier; 2006. p. 3–24. https://doi.org/10.1533/9781845691073.1.3
644 8. Maw SJ, Fowler VR, Hamilton M, Petchey AM. Physical characteristics of pig fat and their
645 relation to fatty acid composition. Meat Sci. 2003;63:185–90. https://doi.org/10.1016/S0309-
646 1740(02)00069-4
647 9. Terés S, Barceló-Coblijn G, Benet M, Álvarez R, Bressani R, Halver JE, et al. Oleic acid
648 content is responsible for the reduction in blood pressure induced by olive oil. Proceedings of
649 the National Academy of Sciences. 2008;105:13811–6.
650 https://doi.org/10.1073/pnas.0807500105
651 10. Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, et al. Fat deposition,
652 fatty acid composition and meat quality: A review. Meat Sci. 2008;78:343–58.
653 https://doi.org/10.1016/j.meatsci.2007.07.019

11. Fernández A, de Pedro E, Núñez N, Silió L, García-Casco J, Rodríguez C. Genetic parameters for meat and fat quality and carcass composition traits in Iberian pigs. Meat Sci. 2003;64:405–10. https://doi.org/10.1016/S0309-1740(02)00207-3

12. Sellier P, Maignel L, Bidanel JP. Genetic parameters for tissue and fatty acid composition of backfat, perirenal fat and longissimus muscle in Large White and Landrace pigs. Animal. 2010;4:497–504. https://doi.org/10.1017/S1751731109991261

13. Suzuki K, Ishida M, Kadowaki H, Shibata T, Uchida H, Nishida A. Genetic correlations among fatty acid compositions in different sites of fat tissues, meat production, and meat quality traits in Duroc pigs. J Anim Sci. 2006;84:2026–34. https://doi.org/10.2527/jas.2005-660

14. Davoli R, Catillo G, Serra A, Zappaterra M, Zambonelli P, Zilio DM, et al. Genetic parameters of backfat fatty acids and carcass traits in Large White pigs. Animal. 2019;13:924–32. https://doi.org/10.1017/S1751731118002082

15. Mortimer S, Przybylski W. Breeding strategies for improving meat quality. Meat Quality Genetic and Environmental Factors. In: CRC Press Taylor & Francis Group, editor. Meat quality :genetic and environmental factors. Boca Raton; 2016. p. 257–86.

16. Nuñez P, Casto-Rebollo C, Martinez-Boggio G, Negro Rama S, Casellas J, Varona L, et al. The gut microbiota as a predictor of feed efficiency and feeding behavior in Iberian pigs. J Anim Sci. 2025;103. https://doi.org/10.1093/jas/skaf268

17. Khanal P, Maltecca C, Schwab C, Fix J, Bergamaschi M, Tiezzi F. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. Genetics Selection Evolution. 2020;52:41. https://doi.org/10.1186/s12711-020-00561-7

18. Calle-García J, Ramayo-Caldas Y, Zingaretti LM, Quintanilla R, Ballester M, Pérez-Enciso M. On the holobiont 'predictome' of immunocompetence in pigs. Genetics Selection Evolution. 2023;55:29. https://doi.org/10.1186/s12711-023-00803-4

678 19. Tiezzi F, Fix J, Schwab C, Shull C, Maltecca C. Gut microbiome mediates host genomic
679 effects on phenotypes: a case study with fat deposition in pigs. Comput Struct Biotechnol J.
680 2021;19:530–44. https://doi.org/10.1016/j.csbj.2020.12.038
681 20. Kucha CT, Liu L, Ngadi M, Gariépy C. Assessment of Intramuscular Fat Quality in Pork
682 Using Hyperspectral Imaging. Food Engineering Reviews. 2021;13:274–89.
683 https://doi.org/10.1007/s12393-020-09246-9
684 21. Parrini S, Sirtori F, Čandek-Potokar M, Charneca R, Crovetti A, Kušec ID, et al. Prediction
685 of fatty acid composition in intact and minced fat of European autochthonous pigs breeds by
686 near infrared spectroscopy. Sci Rep. 2023;13:7874. https://doi.org/10.1038/s41598-023-34996-
687 x
688 22. Sandler SR, Karo W. Sourcebook of advanced organic laboratory preparations. San Diego,
689 CA, USA: Academic Press; 1992.
690 23. Eder K. Gas chromatographic analysis of fatty acid methyl esters. J Chromatogr B Biomed
691 Sci Appl. 1995;671:113–31. https://doi.org/10.1016/0378-4347(95)00142-6
692 24. Azouggagh L, Ibáñez-Escriche N, Martínez-Álvaro M, Varona L, Casellas J, Negro S, et
693 al. Characterization of microbiota signatures in Iberian pig strains using machine learning
694 algorithms. Anim Microbiome. 2025;7:13. https://doi.org/10.1186/s42523-025-00378-z
695 25. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data.
696 http://www.bioinformatics.babraham.ac.uk/projects/fastqc. 2010.
697 26. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for
698 multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8.
699 https://doi.org/10.1093/bioinformatics/btw354
700 27. Straub D, Blackwell N, Langarica-Fuentes A, Peltzer A, Nahnsen S, Kleindienst S.
701 Interpretations of Environmental Microbial Community Studies Are Biased by the Selected 16S

rRNA (Gene) Amplicon Sequencing Pipeline. Front Microbiol. 2020;11. https://doi.org/10.3389/fmicb.2020.550420

28. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3. https://doi.org/10.1038/nmeth.3869

29. Wahba G. Spline Models for Observational Data. Society for Industrial and Applied Mathematics; 1990. https://doi.org/10.1137/1.9781611970128

30. Pérez P, de los Campos G. Genome-Wide Regression and Prediction with the BGLR Statistical Package. Genetics. 2014;198:483–95. https://doi.org/10.1534/genetics.114.164442

31. Difford GF, Plichta DR, Løvendahl P, Lassen J, Noel SJ, Højberg O, et al. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. PLoS Genet. 2018;14:e1007580. https://doi.org/10.1371/journal.pgen.1007580

32. Lourenco D, Tsuruta S, Aguilar I, Masuda Y, Bermann M, Legarra A, et al. 366. Recent updates in the BLUPF90 software suite. Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP). The Netherlands: Wageningen Academic Publishers; 2022. p. 1530–3. https://doi.org/10.3920/978-90-8686-940-4_366

33. Geweke J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. 1991. https://doi.org/10.21034/sr.148

34. Breiman L. Random Forests. Mach Learn. 2001;45:5–32. https://doi.org/10.1023/A:1010933404324

35. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. 2012;

36. Khanal P, Maltecca C, Schwab C, Fix J, Bergamaschi M, Tiezzi F. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. Genetics Selection Evolution. 2020;52:41. https://doi.org/10.1186/s12711-020-00561-7

727 37. Sebastià C, Folch JM, Ballester M, Estellé J, Passols M, Muñoz M, et al. Interrelation
728 between gut microbiota, SCFA, and fatty acid composition in pigs. mSystems. 2024;9.
729 https://doi.org/10.1128/msystems.01049-23
730 38. Gol S, González-Prendes R, Bosch L, Tor M, Reixach J, Pena RN, et al. Linoleic acid
731 metabolic pathway allows for an efficient increase of intramuscular fat content in pigs. J Anim
732 Sci Biotechnol. 2019;10:33. https://doi.org/10.1186/s40104-019-0343-8
733 39. Ibáñez-Escriche N, Magallón E, Gonzalez E, Tejeda JF, Noguera JL. Genetic parameters
734 and crossbreeding effects of fat deposition and fatty acid profiles in Iberian pig lines1. J Anim
735 Sci. 2016;94:28–37. https://doi.org/10.2527/jas.2015-9433
736 40. Casellas J, Noguera JL, Reixach J, Díaz I, Amills M, Quintanilla R. Bayes factor analyses
737 of heritability for serum and muscle lipid traits in Duroc pigs1. J Anim Sci. 2010;88:2246–54.
738 https://doi.org/10.2527/jas.2009-2205
739 41. Wood JD, Enser M. Factors influencing fatty acids in meat and the role of antioxidants in
740 improving meat quality. British Journal of Nutrition. 1997;78:S49–60.
741 https://doi.org/10.1079/BJN19970134
742 42. Estany J, Ros-Freixedes R, Tor M, Pena RN. A Functional Variant in the Stearoyl-CoA
743 Desaturase Gene Promoter Enhances Fatty Acid Desaturation in Pork. PLoS One.
744 2014;9:e86177. https://doi.org/10.1371/journal.pone.0086177
745 43. Ntambi J. Regulation of stearoyl-CoA desaturases and role in metabolism. Prog Lipid Res.
746 2004;43:91–104. https://doi.org/10.1016/S0163-7827(03)00039-0
747 44. Tyra M, Ropka-Molik K, Terman A, Piórkowska K, Oczkowicz M, Bereta A. Association
748 between subcutaneous and intramuscular fat content in porcine ham and loin depending on age,
749 breed and FABP3 and LEPR genes transcript abundance. Mol Biol Rep. 2013;40:2301–8.
750 https://doi.org/10.1007/s11033-012-2311-7

751 45. Forwood DL, Holman DB, Meale SJ, Chaves A V. Ensiled Mixed Vegetables Enriched
752 Carbohydrate Metabolism in Heterofermentative Lactic Acid Bacteria. Fermentation.
753 2022;8:699. https://doi.org/10.3390/fermentation8120699
754 46. Bergamaschi M, Tiezzi F, Howard J, Huang YJ, Gray KA, Schillebeeckx C, et al. Gut
755 microbiome composition differences among breeds impact feed efficiency in swine.
756 Microbiome. 2020;8:110. https://doi.org/10.1186/s40168-020-00888-9
757 47. Hou J, Ji X, Chu X, Wang B, Sun K, Wei H, et al. Mulberry Leaf Dietary Supplementation
758 Can Improve the Lipo-Nutritional Quality of Pork and Regulate Gut Microbiota in Pigs: A
759 Comprehensive Multi-Omics Analysis. Animals. 2024;14:1233.
760 https://doi.org/10.3390/ani14081233
761 48. Shang P, Wei M, Duan M, Yan F, Chamba Y. Healthy Gut Microbiome Composition
762 Enhances Disease Resistance and Fat Deposition in Tibetan Pigs. Front Microbiol. 2022;13.
763 https://doi.org/10.3389/fmicb.2022.965292
764 49. Biddle A, Stewart L, Blanchard J, Leschine S. Untangling the Genetic Basis of Fibrolytic
765 Specialization by Lachnospiraceae and Ruminococcaceae in Diverse Gut Communities.
766 Diversity (Basel). 2013;5:627–40. https://doi.org/10.3390/d5030627
767 50. López-García A, Benítez R, Núñez Y, Gómez-Izquierdo E, de Mercado E, García-Casco
768 JM, et al. Influence of genetic background and dietary oleic acid on gut microbiota composition
769 in Duroc and Iberian pigs. PLoS One. 2021;16:e0251804.
770 https://doi.org/10.1371/journal.pone.0251804
771 51. Zamora-Rojas E, Garrido-Varo A, De Pedro-Sanz E, Guerrero-Ginel JE, Pérez-Marín D.
772 Prediction of fatty acids content in pig adipose tissue by near infrared spectroscopy: At-line
773 versus in-situ analysis. Meat Sci. 2013;95:503–11.
774 https://doi.org/10.1016/j.meatsci.2013.05.020
775

**Additional files**

**File name : Additional File 1**

File format: .xlsx

Title: Descriptive statistics for intramuscular and subcutaneous fatty acid composition.

Description: Table 1: Descriptive statistics for intramuscular fatty acid composition (%): mean, SD, minimum (Min), and maximum (Max) values. Table 2: Descriptive statistics for subcutaneous fatty acid composition (%): mean, SD, minimum (Min), and maximum (Max) values.

**File name : Additional File 2**

File format: .xlsx

Title: Predictive ability of fatty acid composition and fat quality traits in intramuscular and subcutaneous fat across timepoints.

Description: Table 1 : Predictive performance of fatty acid composition in intramuscular fat across three timepoints (140d, 180d, and end of fattening period, EFP), showing predictive ability (SD) and mean absolute error (MAE, SD). Table 2 : Predictive performance of fatty acid composition in subcutaneous fat across three timepoints (140d, 180d, and end of fattening period, EFP), showing predictive ability (SD) and mean absolute error (MAE, SD). Table 3 : Predictive performance of desaturation and nutritional indices of intramuscular fat (IMF) across different timepoints (140d, 180d and EFP). Table 4 : Predictive performance of desaturation and nutritional indices of subcutaneous fat (SCF) across different timepoints (140d, 180d and EFP).

**File name : Additional File 3**

File format: .xlsx

Title: Genetic parameters for fatty acid composition and fat quality traits in intramuscular and subcutaneous fat.

Description: Table 1 : Heritabilities (h2) and 95% highest posterior density intervals (lower and higher limits) for fat quality traits, including fatty acids, group totals, and ratios, in intramuscular fat (IMF). Table 2 : Heritabilities (h2) and 95% highest posterior density intervals (lower and higher limits) for fat quality traits, including fatty acids, group totals, and ratios, in subcutaneous fat (SCF). Table 3 : Genetic correlations (rg), 95% highest posterior density intervals (lower and higher limits), P0 and Pr for fat quality traits, including fatty acids, group totals, and ratios, in intramuscular fat (IMF). Table 4 : Genetic correlations (rg), 95% highest posterior density intervals (lower and higher limits), P0 and Pr for fat quality traits, including fatty acids, group totals, and ratios, in subcutaneous fat (SCF).

**File name : Additional File 4**

File format: .xlsx

Title: Key ASVs associated with fatty acid composition and fat quality traits in intramuscular and subcutaneous fat.

Description: Table 1. Most important ASVs associated with individual fatty acids, group totals, and ratios in intramuscular fat (IMF) at 180d, and with %IMF at 140d. Table 2. Most important ASVs associated with individual fatty acids, group totals, and ratios, in subcutaneous fat (SCF) at 180d.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Graphicalabtract.jpeg
- Graphicalabtract.jpeg
- Graphicalabtract.jpeg
- Additionalfile1.xlsx
- Additionalfile1.xlsx
- Additionalfile1.xlsx
- Additionalfile2.xlsx
- Additionalfile2.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx
- Additionalfile3.xlsx
- Additionalfile3.xlsx
- Additionalfile4.xlsx
- Additionalfile4.xlsx
- Additionalfile4.xlsx