Animal 20 (2026) 101725

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Identification of biomarkers for feed efficiency and growth rate by exploring the plasma metabolome of divergent heavy pigs

F. Bertolini [a,1,*], S. Bovo [a,1], M. Bolner [a,1], G. Schiavo [a], A. Ribani [a], P. Zambonelli [a], S. Dall'Olio [a], M. Gallo [b], L. Fontanesi [a]

[a] Animal and Food Genomics Group, Division of Animal Sciences, Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy
[b] Associazione Nazionale Allevatori Suini, Roma, Italy

ARTICLE INFO

Article history:
Received 10 March 2025
Revised 11 November 2025
Accepted 13 November 2025
Available online 19 November 2025

Keywords:
Average daily gain
Feed conversion ratio
Machine learning
Metabolite
Sus scrofa

ABSTRACT

Feed represents the largest expense in pig farming and significantly affects the sustainability of the production system. Therefore, enhancing feed efficiency is a key strategy to mitigate these costs and environmental impacts. This is particularly relevant in the context of the heavy pig system in which animals are slaughtered at a heavier live weight than in many other production systems to follow the rules of Protected Designation of Origin (**PDO**) value chains. Since growth rate is correlated with feed efficiency, and under PDO rules, pigs cannot reach the slaughter weight earlier than a set age limit, the daily gain of the pigs needs to be controlled. In this study, we used untargeted metabolomics to identify plasma metabolites in Italian Large White heavy pigs that may differentiate between animals with divergent feed efficiency and growth rate, and that may constitute biomarkers for one or the other trait. From a starting cohort of 672 performance-tested pigs, two partially overlapping datasets of 200 pigs each, extreme and divergent for feed conversion ratio (**FCR**) and average daily gain (**ADG**), were selected. Approximately 700 metabolites were analysed in the plasma of these pigs. Metabolomic data were analysed with the Boruta machine learning algorithm. Discriminant metabolites were further evaluated through univariate and multivariate analyses. Boruta identified 10 and 7 metabolites that differentiate between FCR and ADG extreme pigs, respectively, with an additional metabolite shared by the two datasets. Most metabolites selected in the FCR dataset still show significant abilities to discriminate among high and low ADG pigs, even if they have not been selected in the Boruta analysis, showing medium to high values of Area Under the Curve, and highly significant Mann–Whitney test U P-values, while the opposite was not true. Among the metabolites detected, L-carnitine and O-adipoylcarnitine, both involved in fatty acid metabolism, were significantly higher in pigs with high FCR. Isoleucylhydroxyproline and prolylhydroxyproline, linked to collagen turnover, were higher in low FCR pigs, potentially reflecting more efficient protein metabolism. Other metabolites linked to gut microbiome activity significantly differentiate between high and low FCR and ADG pigs, suggesting a potential role of the microbiota in nutrient utilisation. The identified metabolomic profiles confirm that feed efficiency and growth rate are related yet distinct traits, whose independent consideration will enhance the accuracy of biomarker discovery and genetic selection in Italian heavy pigs.

Implications

Improving feed efficiency is crucial for the economic and environmental sustainability of pig breeding. Growth rate, which correlates with feed efficiency, should be balanced in heavy pig production systems. We used untargeted metabolomics in Italian Large White heavy pigs to identify biomarkers for two traits strictly related to these aspects: feed conversion ratio and average daily gain. This revealed metabolites and metabolic pathways involved in fatty acid oxidation, collagen turnover, bile acid metabolism, and gut microbiome activity that differentiate pigs by feed efficiency and growth rate, providing candidate biomarkers to support precise selection and nutritional strategies to improve production efficiency.

* Corresponding author.
E-mail address: francesca.bertolini3@unibo.it (F. Bertolini).
[1] Equal contribution

https://doi.org/10.1016/j.animal.2025.101725
1751-7311/

Introduction

Improving the sustainability of livestock systems is a major goal of the animal production sector. In pig production, feed may account for up to 72% of the production costs and significantly impacts greenhouse gas emissions, from crop production to feedstuff transportation and processing (Makkar, 2018; Noblet and van Milgen, 2004; Rocadembosch et al., 2016). One possible strategy to enhance the economic sustainability and reduce this specific impact in pig production is to improve the feed efficiency of the animals. Feed efficiency can be defined as the overall efficiency with which the pig utilises dietary nutrients for maintenance and growth (Patience et al., 2015). Feed efficiency is influenced by a combination of genetic, environmental, dietary, and physiological factors, which are also linked to age (Gardiner et al., 2020; Homma et al., 2021; Noblet and van Milgen, 2004). As pigs grow towards market weight, they require more energy to maintain basic biological processes and consequently become less efficient at converting feed into body weight gain (Patience et al., 2015). This is particularly important in the heavy pig production systems, where slaughter weight significantly exceeds that of standard commercial pigs, which, in contrast, are referred to as light carcass production systems (Bosi and Russo, 2004).

In pig farming, one of the most traditional ways to estimate feed efficiency is to calculate the Feed Conversion Ratio (**FCR**), which measures how effectively pigs convert feed into body mass by a mathematical relationship between feed intake and weight gain (Patience et al., 2015). This trait shows a moderate heritability (Cai et al., 2008; Gilbert et al., 2007; Homma et al., 2021; Kavlak and Uimari, 2019) and, in some cases, it is still the only trait included to target feed efficiency in breeding programmes. However, relying only on FCR may have some limitations, as it only captures the amount of feed required per unit of body weight gain, but it does not account for factors such as differences in body composition, metabolic efficiency, or energy expenditure for maintenance versus growth (Gilbert et al., 2007). Additionally, FCR is sensitive to variations in growth rate, making it challenging to determine whether changes are due to actual metabolic efficiency or simply accelerated growth (Patience et al., 2015). In heavy pig production systems that are linked to Protected Designation of Origin (**PDO**) dry–cured hams, pigs are required to reach a live weight of ~170 kg and a minimum age of 9 months, to have the maturity needed for a high-quality curing process (Disciplinare di produzione della Denominazione di Origine Protetta Prosciutto di Parma, 2023). In this system, growth rate is assessed through the average daily gain (**ADG**), which is calculated by dividing the change in body weight during the elapsed period (usually measured in pounds or kg/day) (Camp Montoro et al., 2020). Despite being negatively correlated, FCR and ADG exhibit distinct genetic backgrounds (Fu et al., 2020; Keel et al., 2020; Silva et al., 2019). This highlights the potential for developing selection strategies to improve FCR while controlling ADG and ensuring compliance with the PDO dry–cured ham regulations.

Exploring biomarkers to dissect FCR and ADG may be a valuable approach to provide specific indicators of the distinct biological processes that influence these traits. Biomarkers are measurable molecules that are associated with physiological states or responses and may provide insights into the characterisation of the molecular and biological pathways regulating each final trait (e.g. de Lima et al., 2020; Grubbs et al., 2016; Messad et al., 2021). Among those, metabolomics, an expanding field, has become a promising tool for a more precise characterisation of animal metabolism. By analysing a wide range of small molecules present in biofluids or tissues, metabolomics can identify biomarkers that help unravel the biological complexity of various phenomena across species, including production traits in livestock (Fontanesi, 2016; Goldansaz et al., 2017; Hao et al., 2021; Imaz et al., 2022; Li et al., 2022). Recent technological advancements in metabolomic profiling now offer several scalable approaches to characterise metabolites, including targeted and untargeted approaches (Gowda and Djukovic, 2014; Dudzik et al., 2018; Gertsman and Barshop, 2018).

In pigs, metabolomics has been used on several biofluids and tissues to characterise breed differentiation, to assess the metabolic impact of pig development, health, feeding regimes, responses to stress, and to explore the influence of genetic variability on metabolite profiles (e.g. Bovo et al., 2015, 2016, 2023, 2025b, 2025c; Dervishi et al., 2021, 2023; Luise et al., 2020; Metzler-Zebeli et al., 2023; Picone et al., 2018; Ruggeri et al., 2025). A limited number of metabolomics studies in plasma or faeces have already provided a first picture in crossbreed or Landrace and Duroc purebred, highlighting candidate biomarkers associated with feed efficiency and/or growth performance (Carmelo et al., 2020; Wu et al., 2021; Ye et al., 2021). However, the limited number of studies, along with variations in sample size, diets, breed, breeding systems and tissues analysed, means that more evidence is needed to achieve a reliable understanding of these traits. Moreover, there have been no studies specifically targeting Italian Large White pigs raised in heavy pig production systems for PDO ham production. Since these traits are genetically distinct yet physiologically interconnected, there is a need to clarify the degree of their interconnection at the metabolic level. In this study, we used untargeted metabolomics to identify plasma metabolites that could characterise feed efficiency and growth rate in Italian Large White heavy pigs. Our objectives were to compare metabolite abundances between pigs with high and low FCR, as well as between those with high and low ADG, and to describe metabolic pathways and candidate biomarkers that could enhance our understanding of the biological processes involved in feed efficiency and growth in pigs.

Material and methods

Animals and blood samples

All animals involved in this study were maintained in compliance with both Italian and European legislation governing pig production, and all procedures outlined here adhered to Italian and European Union regulations on animal care and slaughter. No experimental manipulation was performed by the researchers on live animals. The animals were slaughtered at a commercial abattoir, following standard protocols. The fasting of animals was not performed specifically for this study, but as part of routine preslaughter practices. All samples were collected postmortem from animals slaughtered in a commercial facility for meat production as a routine procedure, following standard procedures.

The overall structure of the experimental design is summarised in Fig. 1. A total of 672 Italian Large White pigs were included in the study, consisting of 458 females and 214 castrated males. The pigs were part of a sib-test programme, based on triplets of pigs from the same litter (two females and one castrated male), that underwent individual performance testing at the Central Station of the National Pig Breeders Association (ANAS). The test began at 100 days (approximately 30 kg) and ended at about 155 ± 5 kg live weight. During this period, pigs were housed individually in pens with a solid floor, one-third of which was slatted. Dry pelleted feed was provided via an automatic feeding system. The system distributes the predetermined individual feed ration twice a day into a hopper placed in each pen. The pig, by pushing a deflector with its snout, causes the feed to drop into the trough below. The pigs were fed and managed under identical conditions,

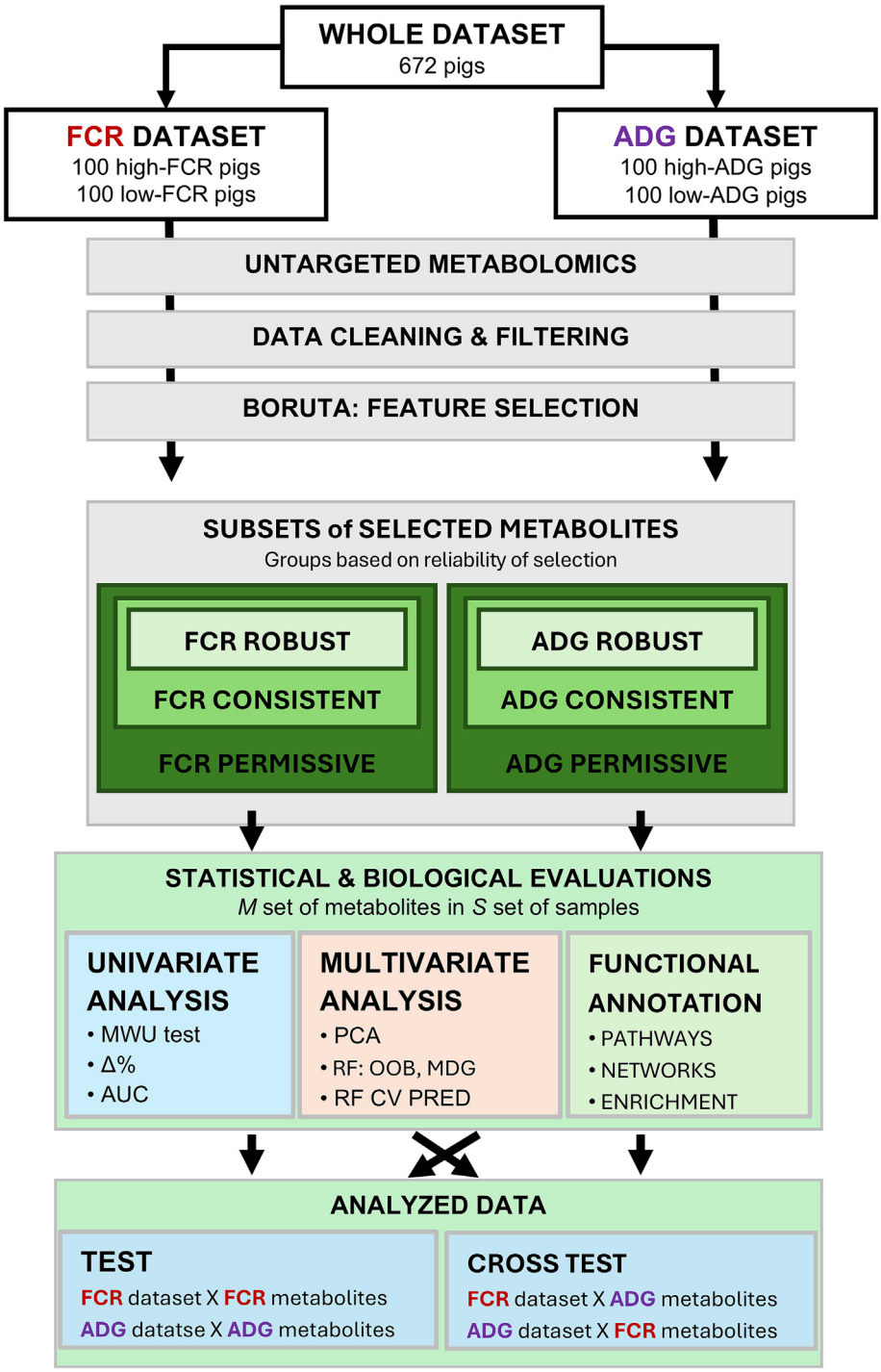

Fig. 1. Schematic representation of the overall experimental design and analytical pipeline for Italian Large White pigs. Abbreviations: FCR: Feed Conversion Ratio; ADG: Average Daily Gain; MWU: Mann-Whitney *U* test; Δ%: Relative percentage difference; AUC: Area under the curve; PCA: Principal Component Analyses; RF: Random Forest; OOB: Out-Of-Bag; MDG: Mean Decrease Gini; CV PRED: Cross Validation Prediction.

and the feeding level was quasi *ad libitum*. Diet is provided in Table S1. The feed distribution system records the individual daily amount distributed. The centre's staff collects and weighs any uneaten feed from each hopper/pen every 2 weeks. The difference between the distributed feed and the feed residue is assumed to be the individually consumed quantity. ADG was calculated as the regression coefficient of individual pig weights recorded every 2 weeks (9 weight measurements per individual during the trial). FCR was computed as the ratio between the average daily feed intake from 100 days to the end of the trial and the average daily gain over the same period. At the end of the test, when the animals reached 155 ± 5 kg live weight, they underwent a fasting period of ~12 h, were transported to a commercial abattoir, and slaughtered following electrical stunning. Slaughter occurred on 23 different days, with 17 to 37 animals collected each day. The procedures took place in the morning at around 0800 h. Blood samples were collected from the draining carotid artery into EDTA-containing tubes immediately after exsanguination. Each tube was inverted

8–10 times before being centrifuged within 2 h at 2 420 × g for 10 min at 4 °C to isolate the plasma, which was then stored in multiple aliquots (Bovo et al., 2016).

Feed conversion ratio and average daily gain estimates

For both FCR and ADG, random residuals (**RRs**) were calculated using a general linear model including fixed effects for sex, batch, date of slaughtering, age at slaughtering and inbreeding coefficient (Bertolini et al., 2018; Fontanesi et al., 2010). Spearman's correlation between the RRs of the two traits was calculated using the *cor* function in R v.4.2.3.3 (R Core Team, 2022). Two distinct datasets, FCR and ADG, each containing 200 samples with extreme and divergent values for the trait considered (100 with positive values and 100 with negative values, from here on referred to as high and low groups, respectively), were selected for subsequent analyses. The average RR values for the FCR dataset were 0.37 (± 0.12) for the high FCR group and −0.34 (± 0.09) for the low FCR group, while the average RRs for the high ADG group were 88.56 (± 22.09) and −104.33 (±3 4.14) for the low ADG group. The two extreme datasets were balanced in terms of male:female ratio and slaughter dates.

Metabolomic profiling of plasma samples

For the animals in the two extreme datasets, untargeted metabolomics profiling of plasma samples was conducted by Metabolon, Inc. (Durham, North Carolina, United States) using the HD4 metabolomics platform. Metabolite extraction was performed with a methanol-based procedure to eliminate proteins and isolate metabolites. The obtained extracts were analysed using ultra-performance liquid chromatography-tandem mass spectrometry with multiple techniques: reverse-phase chromatography with positive ion mode electrospray ionisation, reverse-phase chromatography with negative ion mode electrospray ionisation, and hydrophilic interaction liquid chromatography with negative ion mode electrospray ionisation. Samples were processed in four separate batches, each containing pooled plasma samples for quality control. Data extraction, peak identification, and quality control processing were carried out using Metabolon's proprietary hardware and software. For each metabolite, raw peak areas were normalised to the median raw peak area of the quality control samples. Metabolite annotations, including chemical names, database identifiers, and associated biological pathways, were provided by Metabolon.

Data quality control, imputation and cleaning

Data were processed following the methodology outlined in (Bovo et al., 2025a,b). In summary, data filtering was carried out as follows: (i) outlier metabolites were defined as those deviating by more than five times the interquartile range above or below the median of the metabolite values across all samples; (ii) these outlier values were removed from the dataset and recorded as missing data; (iii) metabolites with more than 25% of missing values were excluded from the dataset; and (iv) samples with over 30% missing values. Missing values were imputed within each dataset (and extreme group) by adapting the method described by Faquih et al., 2020, which employed the Multivariate Imputation by Chained Equations (MICE) approach (Azur et al., 2011; van Buuren, 2018). MICE uses a predictor matrix to determine which columns (i.e. metabolites) will be used to impute missing values in each column. For each metabolite with missing values, the ten metabolites with the highest Pearson correlation coefficient, along with sex and sampling date, were selected as predictors. Only metabolites of endogenous origin, as defined by Metabolon, were used as predictors. As recommended by Faquih et al. (2020), Predictive Mean Matching was used as the imputation algorithm, generating five datasets, as advised by the authors of MICE. To account for the random nature of the imputation process, the data were imputed five times, each with a different random state seed. This resulted in a total of 25 runs of imputation for each dataset (and extreme group). Extreme groups (high and low) were then merged within a random state seed and imputation cycle, and confounding factors were removed by regressing metabolite abundances on the sex and sampling date covariates, as previously described by Bovo et al. (2015, 2025a,b,c). The residuals were then used for further statistical analyses. All analyses were performed using R v.4.2.3 (R Core Team, 2022) and Python v3.11.7.

Identification of discriminant metabolites between extreme groups

For each dataset, the identification of metabolites discriminating between extreme groups of each trait (i.e. high FCR vs low FCR and high ADG vs low ADG) was carried out with the Boruta machine learning algorithm for feature selection (Kursa et al., 2010). Boruta is an all-relevant feature selection method implemented as a wrapper around the Random Forest classification algorithm and has been shown to be more robust and less impacted by the random component when compared to other algorithms for the analysis of untargeted metabolomics (Bovo et al., 2025b). The Boruta algorithm can be briefly summarised as the application of a series of Random Forest models. For each feature, i.e. metabolite, "shadow" features, i.e. random copies of the original data, are created; the original data are then compared with these shadow features, so as to evaluate whether the model performance is better when using the original data compared to the synthetic shadow features. After this comparison, features can be labelled as either "Confirmed", "Rejected" or "Tentative", depending on whether they significantly outperform the overshadow features. For this study, the analysis was implemented in Python v3.11.7, with the BORUTA_py and scikit_learn packages. Default parameters were used, with the exception ofmax_iter, which indicates the maximum number of iterations, for which the value was increased to 1 000 from the default value of 100, and alpha, which was reduced from 0.05 to a more stringent value of 0.01 for more robust statistics. This feature selection procedure was applied to each of the 25 datasets of residuals obtained from the missing value imputation and subsequent linear model. Each feature selection run was performed five times on each of the 25 datasets, with five unique random seeds, in order to take into account the random component of the Boruta algorithm. This resulted in a total of 125 runs of feature selection. We then applied a 10-fold cross-validation procedure (10CV) for evaluating the stability of each selected metabolite, and to confirm the reliability, quality and effectiveness of selection. To do so, we randomly split each of the 25 datasets into ten equal parts, making sure that each split dataset contained a balanced number of samples per group. This way, by rotating the dataset, removing one of the split datasets per iteration, we obtained a total of 250 CV datasets, each 9/10ths of the original size. The same Boruta analysis was applied to each of these datasets, again with five unique random seeds. Therefore, we tested the discriminative power of each metabolite an additional 1 250 times, again considering the random component nested in both missing value imputation and feature selection. For both datasets (FCR and ADG), discriminant metabolites were then categorised into three groups (subsets) based on their reliability of selection. (1) robust: the subset of metabolites that were labelled as "Confirmed" in 125 out of 125 Boruta runs (corresponding to 100% of runs) and also labelled as "Confirmed" in the 10CV procedure for at least 1 000 out of 1 250 runs. (2) consistent: the subset of metabolites that were classified as "Confirmed" in all of the feature selection runs (125 out of

125, i.e. 100%) without considering the 10CV procedure. (3) permissive: the subset of metabolites that were classified as "Confirmed" in at least 100 out of 125 feature selection runs (i.e. 80% of the runs) without considering the 10CV procedure.

Analyses of the discriminant metabolites

To evaluate the discriminatory power of the chosen metabolites, additional univariate and multivariate analyses were carried out. These analyses were conducted separately within each dataset for their specific discriminative metabolites and then applied and cross-tested to the other dataset. Specifically, metabolites found in the FCR dataset were examined in the ADG dataset and vice versa (Fig. 1).

The following univariate analyses were performed for each previously identified metabolite within its corresponding dataset, as well as when cross-applied to the alternate dataset:

1. Relative percentage difference in metabolite concentration [Δ%, (Bovo et al., 2016): the value was calculated and expressed as $\Delta\%_i = \frac{\bar{x}_i^H - \bar{x}_i^L}{\bar{x}_i^H} \times 100$, where $\bar{x}_i^H$ and $\bar{x}_i^L$ are the average metabolite abundance of the i^{th} metabolite in high and low extreme groups, respectively.
2. Mann–Whitney *U* test (**MWU**): The values were obtained with the function mannwhitneyu of scipy (Virtanen et al., 2020) in Python v3.11.7. For this analysis, thresholds of significance were set to $P < 0.05$ and $P < 0.005$ for significant and highly significant results, respectively (Benjamin et al., 2018).
3. Area Under the Curve (**AUC**): this metric summarises the Receiver Operating Characteristic (**ROC**) curve, which is typically used to assess the performance of a binary classification model with varying classification thresholds. Higher AUC values indicate better classification performance. ROC and AUC calculations were performed using the Python v3.11.7 roc_curve and auc functions from the scikit_learn package (Pedregosa et al., 2011)

Subsequently, multivariate analyses were performed on the various combinations of metabolites and datasets:

(1) Principal Component Analyses (**PCAs**): an approach used for linear dimensionality reduction. It reduces the number of dimensions in a dataset to a smaller set of principal components while preserving most of the original information. PCA is valuable for evaluating clustering and separation in a dataset, both qualitatively and quantitatively. To conduct PCA, we used the Python v3.11.7 scikit-learn's functions StandardScaler for scaling and centring the data, and PCA for computing the principal components (Pedregosa et al., 2011). Initially, PCA was applied to the entire metabolite dataset for both FCR and ADG samples, followed by various metabolite subsets. The MWU was also calculated using the first two Principal Components (**PCs**) as groups, following the same method described in the univariate analyses.

(2) Random Forest: Random: this analysis was conducted using the RandomForestClassifier function of the scikit_learn package (Pedregosa et al., 2011). Each set of metabolites was used as input to train a Random Forest Classifier. The trained model was then evaluated based on its Out-Of-Bag (**OOB**) score and the importance of each feature (i.e. metabolite) in classification was measured by the Mean Decrease in Impurity, also known as the Gini index. Random forest classifiers are trained using bootstrap aggregation, where each tree is fitted from a bootstrap sample of observations, and the remaining observations are used for prediction. The OOB error is calculated as the mean error of predictions during the training process, with the OOB score being computed as 1-(OOB error). A high OOB score indicates that during the training process, the model is correctly assigning most OOB samples to the extreme group, suggesting a good classification based on the metabolite set used.

Functional analyses

Annotations for super- and sub-pathways as defined by Metabolon were initially used to functionally evaluate the selected metabolites. Functional links were then studied through the construction of metabolite networks and pathway analyses. Networks were built within each dataset (FCR or ADG) based on selected metabolites (permissive subsets) and by computing Spearman's correlation coefficients (ρ). Similarly to our previous work (Bovo et al., 2025a,b,c), to include all animals without biasing the correlation coefficients, residuals of metabolites were computed, including also the extreme group information (i.e. high or low) as a fixed effect (in addition to sex and sampling day) in the regression model used to clean the data. Edges of the network were considered informative when presenting $\rho \geq 0.5$ (medium–high correlation).

Pathway analyses were conducted using MetaboAnalyst v6.0 (Pang et al., 2022), with the Human Metabolome Database identifiers of selected metabolites as input. Metabolite sets were sourced from RaMP-DB, a comprehensive database containing 3 694 metabolites and lipid pathways (Zhang et al., 2018). Pathways containing at least two metabolites from the input set and showing an FDR-corrected $P < 0.05$ were considered statistically significant.

Results

Overview of the two extreme and divergent datasets

The correlation between RRs of ADG and FCR was –0.69 ($P < 0.05$). This moderate negative correlation can be seen in the selection of samples that displayed extreme and divergent values for ADG and FCR. Specifically, 52% of the samples with contrasting traits (i.e. low FCR and high ADG, or vice versa) overlapped, while none of the samples with corresponding traits (i.e. both high and both low) overlapped. The untargeted metabolomics assay allowed the measurement of 722 metabolites (Table S2), of which 594 (83%) were of endogenous origin, while 60 (8%) were of xenobiotic origin; another 68 metabolites (9%) could not be identified and will be referred to as unnamed metabolites. The endogenous metabolites were grouped into eight super-pathways, most notably lipids (~half of the total) and amino acids (~one-third), with smaller proportions of nucleotides, peptides, carbohydrates, cofactors/vitamins, energy-related compounds, and partially characterised molecules. After excluding 56 metabolites with >25% missing values and removing xenobiotic compounds, 606 metabolites remained for analysis.

Discriminant metabolites for feed conversion ratio and average daily gain

The summary of the overall performance of the metabolites from the Boruta analyses for both traits is reported in Supplementary Table S3. As shown in Table 1, the analysis identified 11 discriminant metabolites for FCR that constitute the "permissive" subset (i.e., selected in at least 100 out of 125 feature selection runs, without considering the 10CV procedure). Based on their reliability of selection, nine metabolites fall into the "consistent" subset (i.e., selected in all 125 feature selection runs, without considering the 10CV procedure) and eight metabolites fall into

Table 1
List of discriminant metabolites obtained with the Boruta analysis using the Feed Conversion Ratio and the Average Daily Gain datasets for Italian Large White pigs. Statistics are reported for both test (features selected and tested within the dataset) and cross-test (features selected in one dataset and tested in the other dataset) analyses.

Metabolite			FCR dataset					ADG dataset				
Metabolite name	Super pathway	Sub pathway	Mean (±SD) high	Mean (±SD) low	Δ%	MWU test *P*	AUC	Mean (±SD) high	Mean (±SD) low	Δ%	MWU test *P*	AUC
FCR-derived												
Methionine sulfone[P,C,R]	Amino Acid	Methionine, Cysteine, SAM and Taurine Metabolism	1.04 (0.22)	1.22 (0.26)	−17.44	7.46×10^{-7}	0.70	1.18 (0.25)	1.11 (0.27)	6.04	0.33	0.54
Ceramide (d18:1/17:0, d17:1/18:0) [C,R]	Lipid	Ceramides	0.99 (0.3)	1.13 (0.34)	−14.98	4.78×10^{-4}	0.64	1.13 (0.32)	1.02 (0.35)	10.20	2.71×10^{-3}	0.62
Phenylacetylglycine[P,C,R]	Peptide	Acetylated Peptides	0.8 (0.27)	0.98 (0.27)	−21.71	3.72×10^{-6}	0.69	1.00 (0.28)	0.89 (0.29)	11.33	0.01	0.61
O-Adipoylcarnitine[P,C,R]	Lipid	Fatty Acid Metabolism (Acyl Carnitine, Dicarboxylate)	1.16 (0.53)	0.85 (0.31)	26.19	3.51×10^{-5}	0.67	0.83 (0.42)	1.03 (0.46)	−23.99	2.66×10^{-3}	0.62
X-16580[P,C,R]	Unnamed	Unnamed	0.99 (0.69)	0.62 (0.39)	37.85	1.78×10^{-4}	0.65	0.65 (0.48)	0.83 (0.53)	−27.28	0.10	0.57
P-cresol glucuronide[P,C,R]	Amino Acid	Tyrosine Metabolism	0.84 (0.35)	1.07 (0.31)	−26.66	1.71×10^{-6}	0.70	1.09 (0.32)	0.93 (0.38)	14.64	3.46×10^{-3}	0.62
X-23890[P,C,R]	Unnamed	Unnamed	1.03 (0.64)	1.36 (0.54)	−32.25	4.66×10^{-5}	0.67	1.27 (0.53)	1.13 (0.66)	10.82	0.06	0.58
Prolylhydroxyproline[P,C,R]	Amino Acid	Urea cycle; Arginine and Proline Metabolism	0.88 (0.24)	1.03 (0.23)	−16.19	1.21×10^{-5}	0.68	1.02 (0.23)	0.98 (0.26)	4.55	0.35	0.54
L-Carnitine[P,C,R]	Lipid	Carnitine Metabolism	0.9 (0.18)	0.79 (0.17)	12.81	8.88×10^{-7}	0.70	0.78 (0.15)	0.85 (0.17)	−8.88	0.03	0.59
Isoleucylhydroxyproline[P]	Peptide	Dipeptide Derivative	0.89 (0.29)	1.04 (0.30)	−16.81	2.05×10^{-4}	0.65	1.04 (0.27)	0.99 (0.32)	4.88	0.28	0.54
FCR- and ADG-derived												
X-17354[P]	Unnamed	Unnamed	1.07 (0.59)	1.65 (0.97)	−54.71	2.27×10^{-6}	0.69	1.68 (0.92)	1.20 (0.71)	28.61	2.48×10^{-4}	0.65
ADG-derived												
6beta-hydroxylithocholate [P,R]	Lipid	Secondary Bile Acid Metabolism	0.97 (0.17)	0.95 (0.18)	1.31	0.60	0.48	0.93 (0.19)	1.02 (0.21)	−9.93	1.64×10^{-5}	0.68
N6-methyllysine [P,R]	Amino Acid	Lysine Metabolism	0.70 (0.40)	0.69 (0.43)	1.21	0.46	0.53	0.67 (0.50)	0.86 (0.57)	−28.84	8.56×10^{-4}	0.64
X-12216[P,R]	Unnamed	Unnamed	0.87 (0.25)	0.88 (0.28)	−0.75	0.81	0.51	0.81 (0.26)	0.93 (0.29)	−15.79	5.58×10^{-4}	0.64
Trimethylamine N-oxide [P]	Lipid	Phospholipid Metabolism	0.75 (0.61)	0.86 (0.56)	−15.22	4.97×10^{-3}	0.62	0.84 (0.59)	0.71 (0.60)	15.48	1.37×10^{-4}	0.66
2-Hydroxydecanoate[P]	Lipid	Fatty Acid, Monohydroxy	0.95 (0.78)	1.08 (0.8)	−14.16	0.06	0.58	1.15 (0.72)	0.86 (0.60)	24.85	2.11×10^{-4}	0.65
1-Oleoyl-GPE (18:1) [P]	Lipid	Lysophospholipid	1.01 (0.51)	0.95 (0.54)	5.35	0.83	0.51	0.99 (0.60)	0.86 (0.41)	13.47	0.03	0.59
Vanillactate [P]	Amino Acid	Tyrosine Metabolism	0.69 (0.18)	0.64 (0.19)	8.19	0.07	0.57	0.62 (0.19)	0.71 (0.18)	−14.74	4.04×10^{-2}	0.62

Abbreviations: FCR: Feed Conversion Ratio; ADG: Average Daily Gain; Δ%: delta percentages; MWU test *P*: Mann-Whitney *U* test *P*; AUC: Area Under the Curve; SAM: S-adenosylmethionine.
[R] Metabolites belonging to the "robust" set.
[C] Metabolites belonging to the "consistent" set.
[P] Metabolites belonging to the "permissive" set.

the "robust" subset (i.e., selected in all feature selection runs and in at least 1 000 out of 1 250 10CV iterations). Excluding unnamed metabolites (n = 3) for which biological annotation is not available, the remaining eight metabolites belong to three known super-pathways, i.e. amino acids (n = 3), lipids (n = 3) and peptides (n = 2). For ADG, eight metabolites were identified: three fell in the "robust" subset, and five fell in the "permissive" or "consistent" subset only. These metabolites are respectively two amino acids, four lipids, and two unnamed. Among all identified metabolites, only one metabolite (unnamed: X-17354) overlapped between the ADG and FCR datasets.

The discriminative power of the 11 metabolites identified by Boruta in the FCR dataset was confirmed by the MWU test, as all of them presented highly significant *P*-values ($P < 0.005$). Most of these highly significant metabolites showed negative delta percentage values (Δ%), indicating lower concentrations in the low FCR group when compared to the high FCR group, with moderate to large values ranging from −14.98% (ceramide) to −54.71% in the unnamed metabolite selected in both FCR and ADG datasets. On the other hand, two lipids (L-carnitine and O-adipoylcarnitine) and one unnamed metabolite showed a moderate to large positive Δ% value (from 12.81 to 37.85%). The computed AUC values further confirmed the discriminative power of the metabolites, with medium to high values ranging from 0.64 to 0.70, with no distinction among the robust, consistent and permissive classes.

The eight metabolites obtained from the Boruta analysis on the ADG dataset confirmed their discriminative power in the same dataset, with all being highly significant (n = 6; $P < 0.005$) or significant (n = 2; $P < 0.05$) in the MWU test. The delta percentages (Δ%) for ADG metabolites reveal distinct patterns of change between high and low ADG groups. A total of four metabolites (two amino acids, one lipid and one unnamed) show a moderate to high increase from high to low ADG pigs, as indicated by the Δ% values (e.g. N6-methyllysine with Δ% = −28.84% and X-12216 with Δ% = −15.79%). Conversely, some metabolites exhibit the opposite trend, decreasing from high to low ADG groups (three lipid and one unnamed) (e.g. 1-oleoyl-GPE (18:1) with Δ% = 13.47 and 2-hydroxydecanoate with Δ% = 24.85%). The AUC values showed slightly lower levels of discriminatory power compared to the metabolites detected in the FCR dataset, ranging from 0.59 to 0.68.

In the cross-test analysis, after computing metabolite-level statistics in the FCR dataset for the ADG-derived metabolites, only trimethylamine N-oxide showed a highly significant *P–value*, with a moderate negative Δ% (−15.22%) and a moderate AUC (0.62). As for the FCR-derived metabolites tested against the ADG dataset, five of the metabolites identified in the FCR dataset exclusively (three lipids, one amino acid and one peptide) were also significant or highly significant, showing opposite delta trend.

Correlation network and functional analysis

Discriminant metabolites from the permissive sets were studied in the context of their relationships through the generation of correlation networks. Edges between nodes, indicating a functional link between metabolites, were drawn if correlation was moderate to high ($|\rho| > 0.5$) (Fig. 2). The network was obtained only for the FCR dataset, as all ADG-derived metabolites presented poor correlation coefficients ($|\rho| < 0.5$). In the FCR dataset, seven out of the 11 metabolites formed three clusters in total. The first cluster showed a moderate correlation ($\rho = 0.58$) between O-adipoylcarnitine and the X-16580 unnamed metabolite. The second cluster exhibited a very high correlation ($\rho = 0.94$) between prolylhydroxyproline and isoleucylhydroxyproline. A larger cluster of three metabolites consisted of a lipid (phenylacetylglycine) and an amino acid (p-cresol glucuronide) that were highly correlated ($\rho = 0.93$), both presenting moderate correlations ($\rho = 0.53$ and $\rho = 0.52$) with the unnamed metabolite X-17354. The clusters were composed of metabolites having the same Δ% trend (i.e. either all more abundant in the high FCR group compared to the low FCR group or vice versa) (Fig. 3A), while no clustering based on the metabolite subsets (i.e. robust, consistent and permissive) was observed (Fig. 3B).

To better evaluate the functional link among the selected metabolites, over-representation in biological pathways was carried out considering both FCR- or ADG-derived metabolite sets, either separately or jointly together. The MetaboAnalyst platform

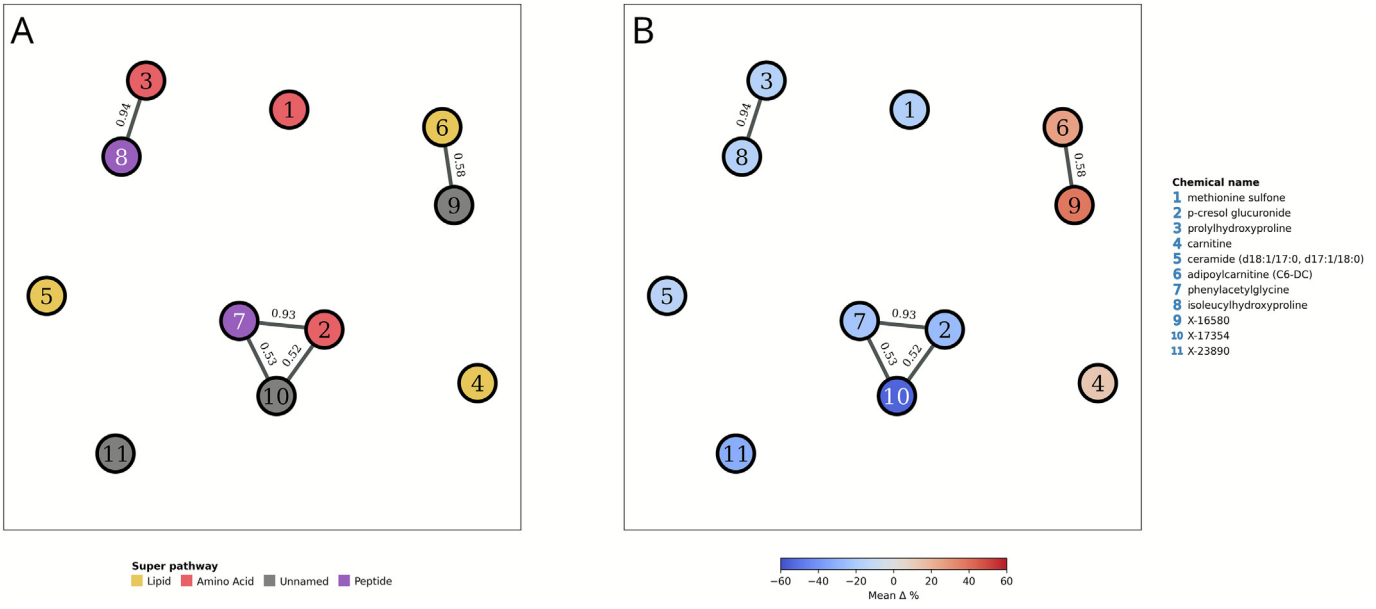

Fig. 2. Correlation network of FCR-selected metabolites for Italian Large White pigs; each node represents a metabolite and each edge a Spearman correlation with absolute value >=0.5. Metabolite names are shown in the numeric legend, and correlations are shown on each edge. (A) Metabolites coloured by the super pathway. (B) Metabolites coloured by the mean difference in abundance between high FCR and low FCR groups. Abbreviations: Mean Δ%: Mean difference in abundance between high and low FCR groups. FCR: Feed Conversion Ratio.

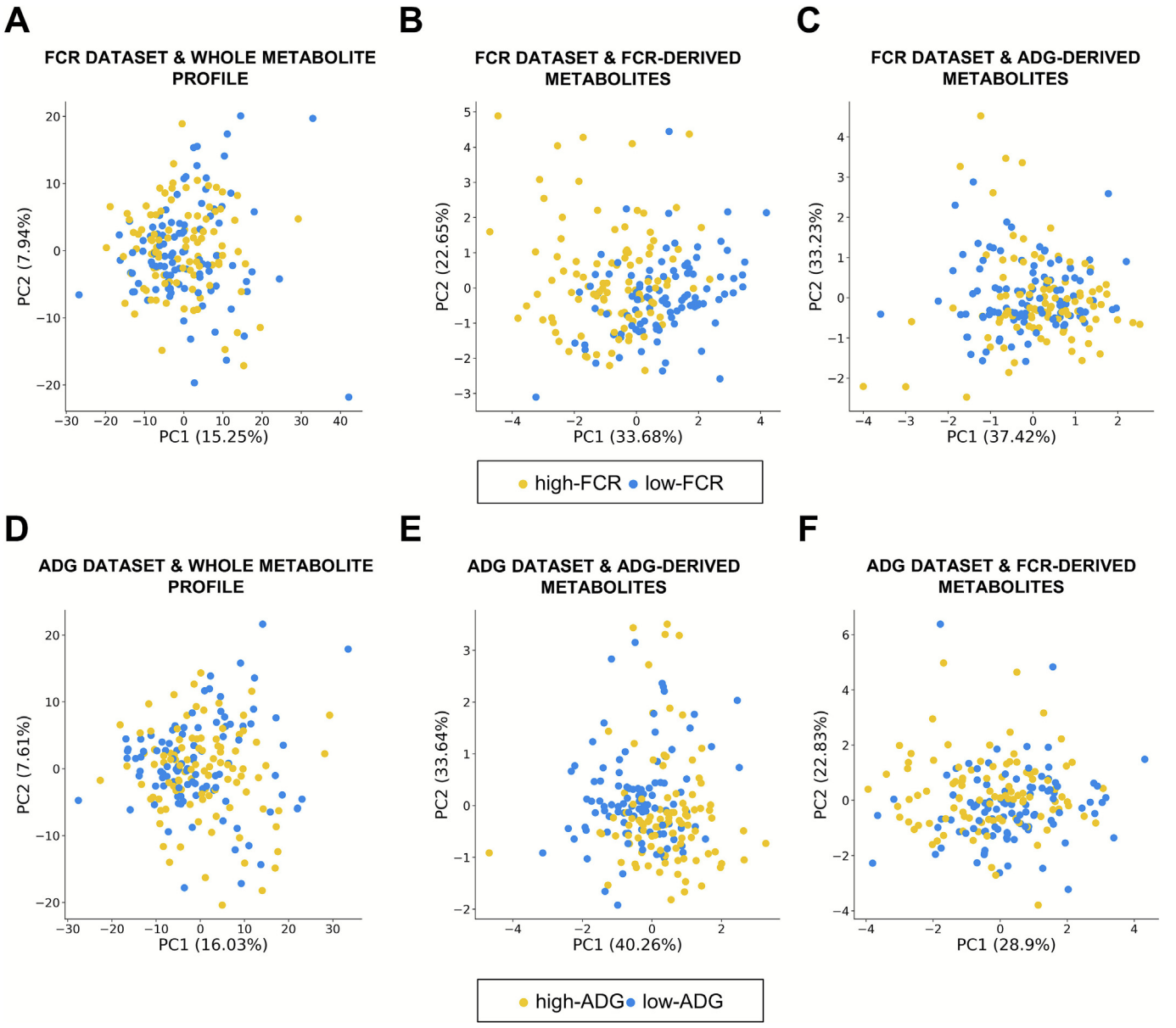

Fig. 3. Principal Component Analyses showing the Principal Components (PC) 1 and 2 of the samples under investigation in Italian Large White pigs, considering both the complete metabolite sets and the discriminant metabolites classified as permissive. For each subfigure, the explained variance of each component is reported in percentage alongside the axis label. In the FCR dataset, the PCA was conducted using (A) whole metabolite profile; (B) discriminant metabolites identified in the FCR analysis (FCR-derived metabolites); (C) discriminant metabolites identified in the ADG analysis (ADG-derived metabolites; cross-test analysis). In the ADG dataset, the PCA was conducted using (D) whole metabolite profile; (E) discriminant metabolites identified in the ADG analysis (ADG-derived metabolites); (F) discriminant metabolites identified in the FCR analysis (FCR-derived metabolites; cross-test analysis). Abbreviations: FCR: Feed Conversion Ratio; ADG: Average Daily Gain.

was able to map all the metabolites except for ceramide (d18:1/17:0, d17:1/18:0). No significant results were obtained when considering FDR-corrected *P-values.*

Principal component analysis of selected metabolites

The ability of all selected metabolites to collectively discriminate between high and low ADG and FCR groups is illustrated in Fig. 3. For FCR, although the overall metabolite profile (Fig. 3A) did not display group stratification, the metabolite set selected within the FCR dataset (Fig. 3B) exhibited a clustering pattern. Clustering was absent when using, in a cross-test analysis, the ADG-derived metabolite set (Fig. 3C). A similar but peculiar pattern was observed in the ADG dataset. In this case, the high and low groups did not show clustering when using all metabolites (Fig. 3D), but, interestingly, displayed clearer clustering with the metabolites selected from the FCR analysis (Fig. 3E), and showed a less pronounced level of clustering when using the metabolites selected for ADG (Fig. 3F).

MWU tests were then performed on the values of the PCs obtained from the PCA described above for both ADG and FCR datasets and analysed together with the Explained Variance (EV) of the PCs. The EV values and MWU test *P*-values for the PCs across different metabolite sets and traits are summarised in Table 2. For both FCR and ADG, as expected, the variance explained by PC1 and PC2 was higher for subsets of discriminant metabolites when compared to the whole metabolite profile. For FCR, the PCs in the "robust" subset of metabolites retrieved from the FCR analysis accounted for 33.68% (PC1) and 22.65% (PC2) of the explained variance, with significant *P* for PC1 ($P < 1.95 \times 10^{-12}$) and PC2 ($P = 0.02$). Similarly,

Table 2
Results of the Principal Component Analyses for Italian Large White Pigs. Explained variance and Mann-Whitney *U* test are provided considering the whole metabolite profile and each metabolite subset (based on reliability of selection) for both Feed Conversion Ratio and Average Daily Gain datasets.

	FCR dataset				ADG dataset			
	PC1		PC2		PC1		PC2	
Metabolite set	EV (%)	MWU test *P*	EV (%)	MWU test *P*	EV (%)	MWU test *P*	EV (%)	MWU test *P*
Whole metabolite profile	15.25	0.53	7.94	0.25	16.03	0.42	7.61	0.35
FCR robust	33.68	1.95×10^{-12}	22.65	0.02	28.90	5.91×10^{-03}	22.83	2.60×10^{-03}
FCR consistent	31.42	1.41×10^{-13}	20.43	0.05	26.44	1.08×10^{-03}	20.78	8.59×10^{-03}
FCR permissive	32.11	2.44×10^{-12}	18.71	2.23×10^{-03}	26.89	1.78×10^{-03}	19.63	4.82×10^{-04}
ADG robust	37.42	0.03	33.23	0.14	40.26	6.99×10^{-08}	33.64	3.49×10^{-04}
ADG permissive	23.65	0.66	17.12	3.70×10^{-05}	25.98	4.49×10^{-07}	16.41	5.10×10^{-07}

Abbreviations: PC: Principal Component; FCR: Feed Conversion Ratio; ADG: Average Daily Gain; EV: Explained Variance; MWU test *P*: Mann-Whitney *U* test *P*.

the "consistent" and "permissive" subsets showed substantial EV for PC1 (31.42 and 32.11%, respectively) and significant MWU test results ($P < 1.5 \times 10^{-13}$ for PC1 for both subsets). For ADG, PC1 explained 40.26% of the variance in the "robust" subset, with a highly significant MWU P-value ($P = 6.99 \times 10^{-8}$). Explained variance for PC2 was also notable for this subset (33.64%, $P = 3.50 \times 10^{-4}$). The "permissive" subset had slightly lower EV (25.98% for PC1 and 16.41% for PC2) but retained statistical significance ($P < 5.20 \times 10^{-7}$for both PCs). In the cross-test analysis, the PCA of metabolite subsets identified for ADG and tested against the FCR dataset, and vice versa, revealed significant patterns. For ADG-derived metabolites in the FCR dataset, the "robust" subset PCs explained 37.42% (PC1) and 33.23% (PC2) of the variance, with PC1 showing a significant difference ($P = 3.18 \times 10^{-2}$). Similarly, the "permissive" subset PCs explained 23.65% of the variance (PC1) with a non-significant *P-value*. Conversely, when FCR-derived subsets were tested against ADG, notable EV values were observed in the PCs of the "robust" subset (28.90% for PC1; $P = 5.91 \times 10^{-3}$) and "permissive" subset (26.89% for PC1; $P = 1.80 \times 10^{-3}$).

Random forest analyses

Fig. 4 summarises the OOB scores for the various models that were trained. Analysis of the mean decrease Gini scores for each model revealed minimal variability among metabolites, suggesting a balanced contribution of each metabolite to the OOB value calculation (Supplementary Table S4). In the FCR sample set, when considering FCR-derived metabolites (Fig. 4A), the OOB value ranged from 0.67 (whole metabolite set) to 0.72 (permissive metabolite set). As expected, using ADG-derived metabolites in the FCR sample set led to lower OOB values (Fig. 4B). In the ADG sample set, FCR-derived metabolites generally underperformed compared to the whole metabolite set (OOB score = 0.67), with values around 0.56 (Fig. 4C). When considering the ADG-derived metabolites,

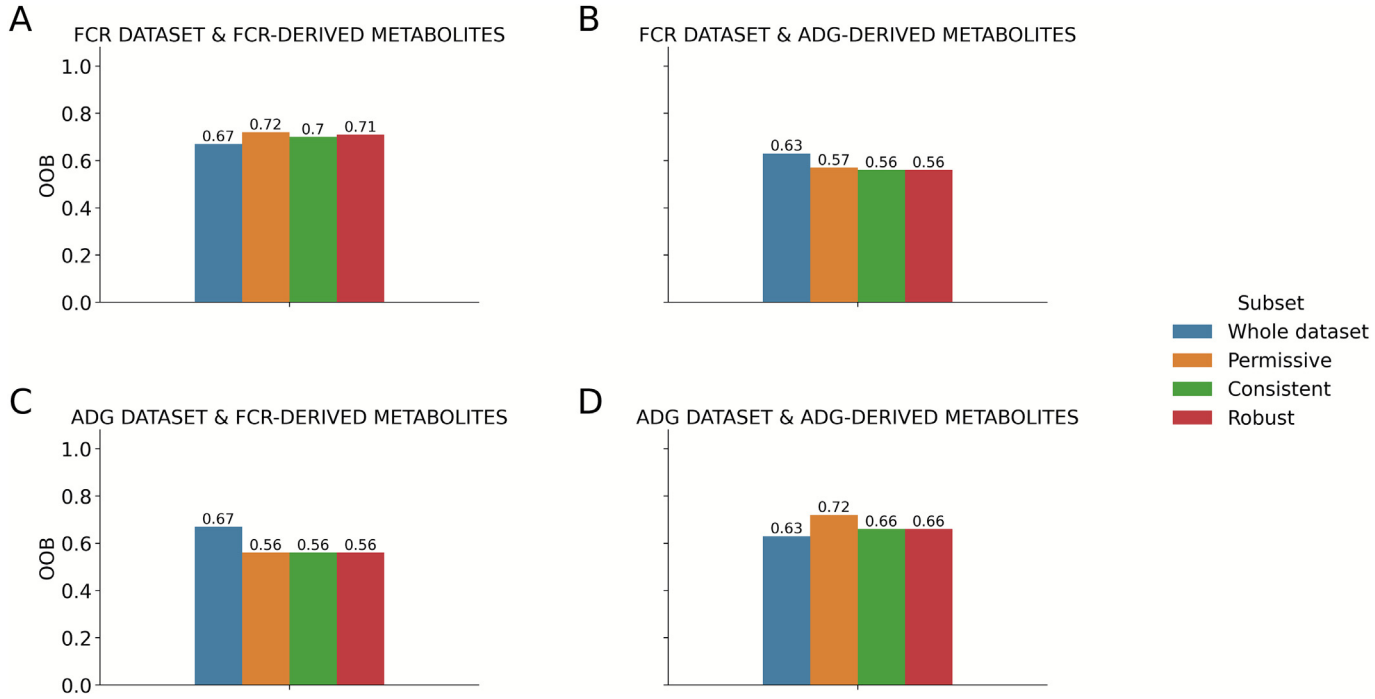

Fig. 4. Out Of Bag (OOB) scores for the different models trained with different metabolite subsets for Italian Large White pigs. (A) FCR dataset with FCR-derived metabolites; (B) FCR dataset with ADG-derived metabolites (cross-test analysis); (C) ADG dataset with FCR-derived metabolites (cross-test analysis); (D) ADG dataset with ADG-derived metabolites. Abbreviations: FCR: Feed Conversion Ratio; ADG: Average Daily Gain.

the OOB score was lower for the whole metabolite set (OOB score = 0.63) than when using the FCR-derived metabolite sets (OOB score = 0.72) (Fig. 4D).

Discussion

In the heavy pig production system, typically associated with PDO products, it is crucial to optimise feed efficiency and growth to achieve economic and environmental sustainability, while also maintaining high product quality (Rocadembosch et al., 2016; Soleimani and Gilbert, 2021). Inadequate feed efficiency in pigs can lead to increased production costs and heightened environmental impact, as feed is needed for maintenance. At the same time, it is important to manage growth rate carefully to prevent pigs from reaching the required slaughter weight before they have reached the minimum age required for slaughter (Disciplinare di produzione della Denominazione di Origine Protetta Prosciutto di Parma, 2023). This balancing act highlights the importance of understanding the complex relationship between feed efficiency and growth. While these two traits are partially related, they are in fact different traits influenced by separate biological and metabolic mechanisms (Arthur et al., 2009; Patience et al., 2015). In this study, we characterised the metabolomic plasma profiles in Italian Large White heavy pigs associated with feed efficiency (FCR) and growth (ADG), by analysing over 700 metabolites in an untargeted approach. We examined two separate cohorts of Italian Large White pigs that were extreme and divergent in terms of their FCR and ADG traits, with a 52% overlap between the two groups of animals. This overlap is a result of the moderate-to-high negative correlation observed between FCR and ADG in our datasets.

This extreme-phenotype approach is widely applied in biomarker discovery studies as it maximises the biological contrast between groups, thereby increasing the power to detect metabolite differences associated with the traits of interest (e.g. Liu et al., 2021; Seyres et al., 2022; Zubiri-Gaitán et al., 2023).

Despite the high correlation and overlap in pigs between the two datasets, distinct metabolite patterns were observed to differentiate the two groups. Boruta analyses conducted on the datasets revealed only one common metabolite out of the 18 detected. As an all-relevant feature selection algorithm, Boruta is designed to capture the full set of variables contributing to group differences, making it well−suited for our objective of highlighting all metabolites associated with the observed patterns rather than reducing them to a minimal predictive subset. Moreover, Boruta can overcome limitations of other approaches, such as LASSO, particularly with regard to multicollinearity (Sztepanacz and Houle, 2024). Among the remaining detected metabolites, 10 were unique to FCR, and 7 were unique to ADG, while 1 was shared. When considered separately, these metabolites exhibited medium−to−high AUC values and highly significant MWU *P*, confirming their ability to discriminate between extreme and divergent values. Individually, most metabolites selected in the FCR dataset still displayed significant discriminatory abilities between high and low ADG, even if they were not selected in the Boruta analysis. The opposite was not true. These results align with the biological background of these two traits. FCR is a composite measure that considers feed intake relative to weight gain, encompassing aspects of both growth and feed utilisation efficiency (Patience et al., 2015). In contrast, ADG is a more narrowly defined trait that directly measures growth performance, without considering feed utilisation. Therefore, metabolites associated with FCR may reflect processes that influence both feed metabolism and growth-related pathways, even if they were not directly identified as the most important contributors to ADG in Boruta feature selection.

As far as we know, none of the metabolites detected in this study have been mentioned in the previous works related to pig feed efficiency and growth. However, the main classes of bile acid metabolism and ceramides were detected by previous studies, aligning with our own work (Carmelo et al., 2020; Wu et al., 2021; Ye et al., 2021). These differences may be due to several factors, primarily the use of different breeds and genetics in the studies, leading to varying selection goals and metabolomic patterns (Bovo et al., 2025a,b,c). In our study, L-carnitine and O-adipoylcarnitine were selected using the Boruta analysis in the FCR dataset. These two metabolites were found to be present at significantly higher concentrations in the high FCR group. Although they were not selected in the Boruta analysis for ADG, both carnitine and O-adipoylcarnitine exhibited significant opposite trends in the ADG dataset (as indicated by MWU *P*), with higher concentrations observed in the low ADG group, particularly for O-adipoylcarnitine. L-carnitine is a key molecule in fatty acid metabolism, serving as a cofactor in mitochondrial fatty acid oxidation by transporting long-chain fatty acids across the inner mitochondrial membrane. It is also involved in other functions such as acyl moiety transport from peroxisomes, and regulation of the intramitochondrial acyl-CoA/CoA ratio (reviewed in Eder, 2009). Carnitine is mainly absorbed from the diet but can also be formed through biosynthesis (Vaz and Wanders, 2002). In livestock, this molecule can be used as an additive to improve the growth performance of both monogastric and ruminant livestock species with contrasting results (Martín et al., 2022; Ringseis et al., 2018a, 2018b). In our study, higher levels in the plasma of the high FCR group suggest that inefficient animals may need to mobilise a higher level of lipids for oxidation to compensate for an elevated energy demand. Despite the lack of a direct correlation in our dataset, L-carnitine and O-adipoylcarnitine belong to the same fatty acid metabolic pathway. In fact, O-adipoylcarnitine is an adipic acid ester of carnitine, which transports acyl-groups (i.e. organic acids and fatty acids) from the cytoplasm into the mitochondria so that they can be broken down to produce energy (Indiveri et al., 2011; Jennings et al., 2023; Schooneman et al., 2013). L-carnitine and O-adipoylcarnitine may accumulate in plasma and urine in the case of fatty acid oxidation disorders (Merritt et al., 2018). The elevated concentrations of these metabolites in the high FCR group suggest an accumulation of fatty acid intermediates. It can be hypothesised that pigs in the high FCR group have higher rates of fatty acid oxidation, leading to increased production of adipoylcarnitine as a byproduct. Animals with high FCR values typically have higher energy demands and lower feed efficiency, consistent with their metabolic profile characterised by increased substrate turnover and incomplete oxidation. Animals with slower growth may rely on less efficient and more energy-consuming routes to meet their energy needs. The higher levels of both metabolites, found in the low ADG group, further support this hypothesis.

In the FCR dataset, isoleucylhydroxyproline and prolylhydroxyproline, which showed a high correlation (ρ = 0.94) in our analyses, are significantly higher in the low FCR group and not significant in the ADG dataset. Both isoleucylhydroxyproline and prolylhydroxyproline are dipeptides derived from the breakdown of collagen or collagen-like proteins, indicating a connection to connective tissue turnover and remodelling (Asai et al., 2020). While no direct link with feed efficiency has been established yet, the increased abundance of these dipeptides in low FCR pigs may suggest higher rates of efficient protein turnover, potentially enhancing tissue maintenance and repair without excessive metabolic cost (Sarri et al., 2024).

Another cluster of correlated metabolites in the FCR dataset included p-cresol glucuronide (a peptide), phenylacetylglycine (an amino acid) and one unclassified metabolite (X-16580). While phenylacetylglycine and p-cresol glucuronide showed significantly higher abundance in the low FCR group, the unclassified metabolite exhibited an opposite trend. In the ADG dataset, only pheny-

lacetylglycine and p-cresol glucuronide displayed significant differences, being more abundant in the high ADG group, while the unclassified X-16580 metabolite did not show significant differences. Phenylacetylglycine and p-cresol glucuronide are bioproducts of gut microbial metabolism of phenylalanine and tyrosine, respectively, and have long been known to circulate in the blood (Stachulski et al., 2023; Wikoff et al., 2009). While direct associations between phenylacetylglycine and p-cresol glucuronide with feed efficiency are not yet established, several studies have linked changes in gut microbiome with feed efficiency-related traits. Changes in gut microbiota affect nutrient utilisation of the animal and growth (Maltecca et al., 2020). The high correlation of the unnamed metabolite X-16580 may suggest its possible common origin (i.e. gut microbiota activity). It is interesting to observe that this metabolite may target nutrient optimisation rather than growth; however, further analyses will be needed to clarify its role.

As expected, and previously discussed, the seven metabolites retrieved from the ADG dataset did not show differences in the FCR dataset. Moreover, no high correlations have been detected among any of them. Among those metabolites, 6β-hydroxylithocholate, N6-methyllysine, X-12216, and vanillactate were more abundant in the low ADG group, while trimethylamine N-oxide, 2-hydroxydecanoate, and 1-oleoyl-GPE (18:1) were higher in the high ADG group. 6β-hydroxylithocholate, a bile acid derivative, and trimethylamine N-oxide, a product of choline metabolism, are both primarily influenced by gut microbial activity. Vanillactate, a secondary metabolite of tyrosine metabolism, is also derived from microbial fermentation (Ufnal et al., 2015; Wilson et al., 2019; Winston and Theriot, 2020). With the current knowledge, it is not possible to determine the interplay of these compounds that can modify ADG performance in pigs or other mammals. However, it is known that specific gut microbial profiles are associated with enhanced weight gain in several livestock species, supporting our observations from FCR-related metabolites regarding feed efficiency (Fang et al., 2020; Quan et al., 2019; Wang et al., 2021).

N6-methyllysine, 2-hydroxydecanoate, and 1-oleoyl-GPE (18:1), despite being involved in different metabolic pathways and still not well investigated in livestock, have all been shown to have a role in influencing growth patterns and weight gain (Axelrod et al., 2024; Olarini et al., 2022; Zhang et al., 2022). 2-Hydroxydecanoate has also been associated with sexual differentiation in Italian Large White pigs (Bovo et al., 2025a). In our study, however, metabolite levels were corrected for sex, and the datasets were balanced by sex ratio, indicating that the detected metabolites reflect efficiency status rather than sex effects.

Metabolites do not function in isolation; instead, they are integral components of biological systems, interconnected within complex metabolic pathways that are not yet fully characterised. In this study, we expanded our analyses by statistically examining metabolite clusters using PCA and Random Forest, two complementary approaches (Biau and Scornet, 2016; Jollife and Cadima, 2016). PCA was used to assess how Principal Components generated by the selected metabolites were able to discriminate between groups in the two datasets, while Random Forest was used to assess the power of classification of the metabolite clusters, exploiting the ability of this approach to self-generate internal test sets. Random Forest is considered one of the most stable machine learning algorithms that does not require a separate testing set (Breiman, 2001; Liaw and Wiener, 2002). These methods provided further assessment of the metabolites identified through Boruta analysis, confirming their relevance to the two targeted traits. In our study, metabolite sets corresponding to their respective datasets confirmed a higher predictive performance in their own dataset, with both traits achieving the highest OOB score with their respective permissive metabolite subsets. In the PCA approach, some variability was detected among the different metabolite subsets which showed best clustering performances when utilising only the metabolites that belonged to the "robust" class, while for Random Forest, this difference was not detected. However, when metabolites were switched, results declined for both datasets in both approaches, confirming the higher specificity of the selected biomarkers over the traits they were selected for.

Conclusions

This study provides the first characterisation of metabolites linked to feed efficiency and growth-related traits in an Italian Large White heavy pig cohort, which is under selection to optimise PDO dry-cured ham production. The Boruta machine learning algorithm allowed us to identify, from approximately 700 analysed metabolites, those most strongly associated with the two targeted traits: FCR and ADG. The metabolic profiles that emerged were linked to fatty acid oxidation, bile acid metabolism, collagen synthesis, and microbiome activity. While some of the main activities were common to both traits, the metabolomic profile reflects unique patterns that characterise them. Specifically, these analyses clarify the role of FCR, which partially includes ADG (growth), but not vice versa. In fact, some metabolites linked to FCR were shown to discriminate ADG, but this was not observed in the other direction. The importance of these metabolites in characterising high– and low–performing pigs for these traits was also confirmed by PCA and random forest analyses. This highlights the importance of considering these traits independently when identifying biomarkers for breeding programmes or nutritional interventions. It also provides candidate biomarkers to be integrated into breeding programmes for the improvement of feed efficiency and growth in Italian Large White pigs, thereby enhancing the sustainability of the heavy pig breeding system. Although our untargeted metabolomic approach offers valuable insights into metabolic signatures associated with feed efficiency and growth, we acknowledge that these complex traits are polygenic and multifactorial. Integrating metabolomics with other omics layers, such as genomics, transcriptomics, proteomics, and microbiomics, would be essential future step to provide a deeper mechanistic understanding and to validate the biomarkers across broader phenotypic spectra. Future studies should also include pigs monitored over the entire growth phase and cohorts representing intermediate phenotypes to complete the metabolomic picture of these traits.

Supplementary material

Supplementary Material for this article (https://doi.org/10.1016/j.animal.2025.101725) can be found at the foot of the online page, in the Appendix section.

Ethics approval

Not applicable.

Data and model availability statement

The datasets generated and/or analysed during the current study are owned by a third party. However, they can be obtained and are available from the corresponding author upon reasonable request.

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

During the preparation of this work the author(s) did not use any AI and AI-assisted technologies.

Author ORCIDs

Francesca Bertolini: https://orcid.org/0000-0003-4181-3895.
Samuele Bovo: https://orcid.org/0000-0002-5712-8211.
Matteo Bolner: https://orcid.org/0000-0002-4985-0191.
Giuseppina Schiavo: https://orcid.org/0000-0002-3497-1337.
Anisa Ribani: https://orcid.org/0000-0001-6778-1938.
Paolo Zambonelli: https://orcid.org/0000-0002-2532-5528.
Stefania Dall'Olio: https://orcid.org/0000-0003-1384-3771.
Maurizio Gallo: NA.
Luca Fontanesi: https://orcid.org/0000-0001-7050-3760.

CRediT authorship contribution statement

F. Bertolini: Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualisation. **S. Bovo:** Writing – review & editing, Methodology, Data curation, Conceptualisation. **M. Bolner:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **G. Schiavo:** Writing – review & editing, Funding acquisition, Data curation. **A. Ribani:** Writing – review & editing, Data curation. **P. Zambonelli:** Writing – review & editing. **S. Dall'Olio:** Writing – review & editing. **M. Gallo:** Writing – review & editing, Resources. **L. Fontanesi:** Writing – review & editing, Funding acquisition, Conceptualisation.

Declaration of interest

None.

Acknowledgements

We acknowledge Associazione Nazionale Allevatori Suini (ANAS) for providing the biological samples.

Financial support statement

This work was financially supported by (1) the Horizon Europe Re-Livestock project that has received funding from: the European Union's Horizon Europe research and innovation programme under the grant agreement No. 101059609 (for the elaboration and analysis of the data); (2) the Italian PRIN2022 project FEEDTHEPIG, Funded by the European Union - NextGenerationEU under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research - Component 2 From research to business - Investment 1.1 Notice PRIN 2022 PNRR (DD N. 1409 del 14/09/2022), proposal code P2022FZMJ9 - CUP J53D23018310001 (for the analysis of part of the data); (3) the Italian PRIN2022 HamCapture, funded by the European Union - NextGenerationEU under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research - Component 2 From research to business - Investment 1.1 Notice Prin 2022 - DD N. 104 del 2/2/2022, proposal code 202238NP9N – CUP J53D23009570001 (for the analysis of part of the data); (4) the Programma di Sviluppo Rurale Nazionale (PSRN) SUIS-2, co-funded by the European Agricultural Fund for Rural Development of the European Union and by the MASAF (for pig genotyping data); (5) and was carried out within the Agritech National Research Center (Agritech Spoke 1) and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022) - This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

Arthur, P.F., Giles, L.R., Eamens, G.J., Barchia, I.M., James, K.J., 2009. Measures of growth and feed efficiency and their relationships with body composition and carcass traits of growing pigs. Animal Production Science 49, 1105–1112.
Asai, T.T., Oikawa, F., Yoshikawa, K., Inoue, N., Sato, K., 2020. Food-derived collagen peptides, prolyl-hydroxyproline (Pro-Hyp), and hydroxyprolyl-glycine (Hyp-Hly) enhance growth of primary cultured mouse skin fibroblast using fetal bovine serum free from hydroxyprolyl peptide. International Journal of Molecular Sciences 21, 229.
Axelrod, C.L., Hari, A., Dantas, W.S., Kashyap, S.R., Schauer, P.R., Kirwan, J.P., 2024. Metabolomic fingerprints of medical therapy versus bariatric surgery in patients with obesity and type 2 diabetes: the STAMPEDE trial. Diabetes Care 47, 2024–2032.
Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J., 2011. Multiple imputation by chained equations: what is it and how does it work? International Journal of Methods in Psychiatric Research 20, 40–49.
Benjamin, D.J., Berger, J.O., Johannesson, M., Nosek, B.A., Wagenmakers, E.J., Berk, R., Bollen, K.A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C.D., Clyde, M., Cook, T.D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., Johnson, V.E., 2018. Redefine statistical significance. Nature Human Behaviour 2, 6–10.
Bertolini, F., Schiavo, G., Galimberti, G., Bovo, S., D'Andrea, M., Gallo, M., Buttazzoni, L., Rothschild, M.F., Fontanesi, L., 2018. Genome-wide association studies for seven production traits highlight genomic regions useful to dissect dry-cured ham quality and production traits in Duroc heavy pigs. Animal 12, 1777–1784.
Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227.
Bosi, P., Russo, V., 2004. The production of the heavy pig for high quality processed products. Italian Journal of Animal Science 3, 309–321.
Bovo, S., Mazzoni, G., Calò, D.G., Galimberti, G., Fanelli, F., Mezzullo, M., Schiavo, G., Scotti, E., Manisi, A., Samoré, A.B., Bertolini, F., Trevisi, P., Bosi, P., Dall'Olio, S., Pagotto, U., Fontanesi, L., 2015. Deconstructing the pig sex metabolome: targeted metabolomics in heavy pigs revealed sexual dimorphisms in plasma biomarkers and metabolic pathways. Journal of Animal Science 93, 5681–5693.
Bovo, S., Mazzoni, G., Galimberti, G., Calò, D.G., Fanelli, F., Mezzullo, M., Schiavo, G., Manisi, A., Trevisi, P., Bosi, P., Dall'Olio, S., Pagotto, U., Fontanesi, L., 2016. Metabolomics evidences plasma and serum biomarkers differentiating two heavy pig breeds. Animal 10, 1741–1748.
Bovo, S., Schiavo, G., Galimberti, G., Fanelli, F., Bertolini, F., Dall'Olio, S., Pagotto, U., Fontanesi, L., 2023. Comparative targeted metabolomic profiles of porcine plasma and serum. Animal 17, 101029.
Bovo, S., Bolner, M., Schiavo, G., Galimberti, G., Bertolini, F., Dall'Olio, S., Ribani, A., Zambonelli, P., Gallo, M., Fontanesi, L., 2025a. Description of metabolic differences between castrated males and intact gilts obtained from high-throughput metabolomics of porcine plasma. Journal of Animal Science 103, skaf178.
Bovo, S., Bolner, M., Schiavo, G., Galimberti, G., Bertolini, F., Dall'Olio, S., Ribani, A., Zambonelli, P., Gallo, M., Fontanesi, L., 2025b. High-throughput untargeted metabolomics reveals metabolites and metabolic pathways that differentiate two divergent pig breeds. Animal 19, 101393.
Bovo, S., Ribani, A., Fanelli, F., Galimberti, G., Martelli, P.L., Trevisi, P., Bertolini, F., Bolner, M., Casadio, R., Dall'Olio, S., Gallo, M., Luise, D., Mazzoni, G., Schiavo, G., Taurisano, V., Zambonelli, P., Bosi, P., Pagotto, U., Fontanesi, L., 2025c. Merging metabolomics and genomics provides a catalog of genetic factors that influence molecular phenotypes in pigs linking relevant metabolic pathways. Genetics Selection Evolution 57, 11.
Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.
Cai, W., Casey, D.S., Dekkers, J.C.M., 2008. Selection response and genetic parameters for residual feed intake in Yorkshire swine. Journal of Animal Science 86, 287–298.
Camp Montoro, J., Manzanilla, E.G., Solà-Oriol, D., Muns, R., Gasa, J., Clear, O., Díaz, J. A.C., 2020. Predicting productive performance in grow-finisher pigs using birth and weaning body weight. Animals 10, 1017.
Carmelo, V.A.O., Banerjee, P., da Silva Diniz, W.J., Kadarmideen, H.N., 2020. Metabolomic networks and pathways associated with feed efficiency and related-traits in Duroc and Landrace pigs. Scientific Reports 10, 255.
de Lima, A.O., Koltes, J.E., Diniz, W.J.S., de Oliveira, P.S.N., Cesar, A.S.M., Tizioto, P.C., Afonso, J., de Souza, M.M., Petrini, J., Rocha, M.I.P., Cardoso, T.F., Neto, A.Z., Coutinho, L.L., Mourão, G.B., Regitano, L.C.A., 2020. Potential biomarkers for feed efficiency-related traits in Nelore cattle identified by co-expression network and integrative genomics analyses. Frontiers in Genetics 11, 189.
Dervishi, E., Reimert, I., van der Zande, L.E., Mathur, P., Knol, E.F., Plastow, G.S., 2021. Relationship between indirect genetic effects for growth, environmental enrichment, coping style and sex with the serum metabolome profile of pigs. Scientific Reports 11, 23377.
Dervishi, E., Bai, X., Dyck, M.K., Harding, J.C.S., Fortin, F., Dekkers, J.C.M., Plastow, G., 2023. GWAS and genetic and phenotypic correlations of plasma metabolites with complete blood count traits in healthy young pigs reveal implications for pig immune response. Frontiers in Molecular Biosciences 10, 1140375.
Disciplinare di produzione della Denominazione di Origine Protetta Prosciutto di Parma. Retrieved on 21 February 2025 from: www.prosciuttodiparma.com/wp-content/uploads/2023/09/Disciplinare_Prosciutto_di_Parma_06_03_2023_ITA.pdf.

Dudzik, D., Barbas-Bernardos, C., García, A., Barbas, C., 2018. Quality assurance procedures for mass spectrometry untargeted metabolomics. A review. Journal of Pharmaceutical and Biomedical Analysis 147, 149–173.
Eder, K., 2009. Influence of L-carnitine on metabolism and performance of sows. British Journal of Nutrition 102, 645–654.
Fang, S., Chen, X., Pan, J., Chen, Q., Zhou, L., Wang, C., Xiao, T., Gan, Q.F., 2020. Dynamic distribution of gut microbiota in meat rabbits at different growth stages and relationship with average daily gain (ADG). BMC Microbiology 20, 1–13.
Faquih, T., van Smeden, M., Luo, J., Le Cessie, S., Kastenmüller, G., Krumsiek, J., Noordam, R., van Heemst, D., Rosendaal, F.R., van H. Vlieg, A., van Dijk, K.W., Mook-Kanamori, D.O., 2020. A workflow for missing values imputation of untargeted metabolomics data. Metabolites 10, 486.
Fontanesi, L., 2016. Metabolomics and livestock genomics: insights into a phenotyping frontier and its applications in animal breeding. Animal Frontiers 6, 73–79.
Fontanesi, L., Speroni, C., Buttazzoni, L., Scotti, E., Dall'Olio, S., Costa, L.N., Davoli, R., Russo, V., 2010. The insulin-like growth factor 2 (IGF2) gene intron3-g.3072g>A polymorphism is not the only sus scrofa chromosome 2p mutation affecting meat production and carcass traits in pigs: evidence from the effects of a cathepsin D (CTSD) gene polymorphism. Journal of Animal Science 88, 2235–2245.
Fu, L., Jiang, Y., Wang, C., Mei, M., Zhou, Z., Jiang, Y., Song, H., Ding, X., 2020. A genome-wide association study on feed efficiency related traits in landrace pigs. Frontiers in Genetics 11, 692.
Gardiner, G.E., Metzler-Zebeli, B.U., Lawlor, P.G., 2020. Impact of intestinal microbiota on growth and feed efficiency in pigs: a review. Microorganisms 8, 1886.
Gertsman, I., Barshop, B.A., 2018. Promises and pitfalls of untargeted metabolomics. Journal of Inherited Metabolic Disease 41, 355–366.
Gilbert, H., Bidanel, J.P., Gruand, J., Caritez, J.C., Billon, Y., Guillouet, P., Lagant, H., Noblet, J., Sellier, P., 2007. Genetic parameters for residual feed intake in growing pigs, with emphasis on genetic relationships with carcass and meat quality traits. Journal of Animal Science 85, 3182–3188.
Goldansaz, S.A., Guo, A.C., Sajed, T., Steele, M.A., Plastow, G.S., Wishart, D.S., 2017. Livestock metabolomics and the livestock metabolome: a systematic review. PLoS One 12, e0177675.
Gowda, G.N., Djukovic, D., 2014. Overview of mass spectrometry-based metabolomics: opportunities and challenges. Methods in Molecular Biology 1198, 3–12.
Grubbs, J.K., Dekkers, J.C.M., Huff-Lonergan, E., Tuggle, C.K., Lonergan, S.M., 2016. Identification of potential serum biomarkers to predict feed efficiency in young pigs. Journal of Animal Science 94, 1482–1492.
Hao, D., Bai, J., Du, J., Wu, X., Thomsen, B., Gao, H., Su, G., Wang, X., 2021. Overview of metabolomic analysis and the integration with multi-omics for economic traits in cattle. Metabolites 11, 753.
Homma, C., Hirose, K., Ito, T., Kamikawa, M., Toma, S., Nikaido, S., Satoh, M., Uemoto, Y., 2021. Estimation of genetic parameter for feed efficiency and resilience traits in three pig breeds. Animal 15, 100384.
Imaz, J.A., García, S., González, L.A., 2022. The metabolomics profile of growth rate in grazing beef cattle. Scientific Reports 12, 2554.
Indiveri, C., Iacobazzi, V., Tonazzi, A., Giangregorio, N., Infantino, V., Convertini, P., Console, L., Palmieri, F., 2011. The mitochondrial carnitine/acylcarnitine carrier: function, structure and physiopathology. Molecular Aspects of Medicine 32, 223–233.
Jennings, E.A., Abi-Rached, Z.H., Jones, D.E., Ryan, R.O., 2023. 3-Methylglutarylcarnitine: a biomarker of mitochondrial dysfunction. Clinica Chimica Acta 551, 117629.
Jollife, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374, 20150202.
Kavlak, A.T., Uimari, P., 2019. Estimation of heritability of feeding behaviour traits and their correlation with production traits in Finnish Yorkshire pigs. Journal of Animal Breeding and Genetics 136, 484–494.
Keel, B.N., Snelling, W.M., Lindholm-Perry, A.K., Oliver, W.T., Kuehn, L.A., Rohrer, G. A., 2020. Using SNP weights derived from gene expression modules to improve GWAS power for feed efficiency in pigs. Frontiers in Genetics 10, 1339.
Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010. Boruta - a system for feature selection. Fundamenta Informaticae 101, 271–285.
Li, J., Wang, Y., Mukiibi, R., Karisa, B., Plastow, G.S., Li, C., 2022. Integrative analyses of genomic and metabolomic data reveal genetic mechanisms associated with carcass merit traits in beef cattle. Scientific Reports 12, 3389.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2, 18–22.
Liu, M., Xie, Z., Costello, C.A., Zhang, W., Chen, L., Qi, D., Furey, A., Randell, E.W., Rahman, P., Zhai, G., 2021. Metabolomic analysis coupled with extreme phenotype sampling identified that lysophosphatidylcholines are associated with multisite musculoskeletal pain. Pain 162, 600–608.
Luise, D., Bovo, S., Bosi, P., Fanelli, F., Pagotto, U., Galimberti, G., Mazzoni, G., Dall'Olio, S., Fontanesi, L., 2020. Targeted metabolomic profiles of piglet plasma reveal physiological changes over the suckling period. Livestock Science 23, 103890.
Makkar, H.P.S., 2018. Feed demand landscape and implications of food-not feed strategy for food security and climate change. Animal 12, 1744–1754.
Maltecca, C., Bergamaschi, M., Tiezzi, F., 2020. The interaction between microbiome and pig efficiency: a review. Journal of Animal Breeding and Genetics 137, 4–13.
Martín, A., Giráldez, F.J., Cremonesi, P., Castiglioni, B., Biscarini, F., Ceciliani, F., Santos, N., Andrés, S., 2022. Dietary administration of L-carnitine during the fattening period of early feed restricted lambs modifies ruminal fermentation but does not improve feed efficiency. Frontiers in Physiology 13, 840065.
Merritt 2nd, J.L., Norris, M., Kanungo, S., 2018. Fatty acid oxidation disorders. Annals of Translational Medicine 6, 2018473.
Messad, F., Louveau, I., Renaudeau, D., Gilbert, H., Gondret, F., 2021. Analysis of merged whole blood transcriptomic datasets to identify circulating molecular biomarkers of feed efficiency in growing pigs. BMC Genomics 22, 1–14.
Metzler-Zebeli, B.U., Lerch, F., Yosi, F., Vötterl, J.C., Koger, S., Aigensberger, M., Rennhofer, P.M., Berthiller, F., Schwartz-Zimmermann, H.E., 2023. Creep feeding and weaning influence the postnatal evolution of the plasma metabolome in neonatal piglets. Metabolites 13, 214.
Noblet, J., van Milgen, J., 2004. Energy value of pig feeds: effect of pig body weight and energy evaluation system. Journal of Animal Science 82, E229–E238.
Olarini, A., Ernst, M., Gürdeniz, G., Kim, M., Brustad, N., Bønnelykke, K., Cohen, A., Hougaard, D., Lasky-Su, J., Bisgaard, H., Chawes, B., Rasmussen, M.A., 2022. Vertical transfer of metabolites detectable from newborn's dried blood spot samples using UPLC-MS: a chemometric study. Metabolites 12, 94.
Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., Xia, J., 2022. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nature Protocols 17, 1735–1761.
Patience, J.F., Rossoni-Serão, M.C., Gutiérrez, N.A., 2015. A review of feed efficiency in swine: biology and application. Journal of Animal Science and Biotechnology 6, 1–9.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.
Picone, G., Zappaterra, M., Luise, D., Trimigno, A., Capozzi, F., Motta, V., Davoli, R., Nanni Costa, L., Bosi, P., Trevisi, P., 2018. Metabolomics characterization of colostrum in three sow breeds and its influences on piglets' survival and litter growth rates. Journal of Animal Science and Biotechnology 9, 1–12.
Quan, J., Cai, G., Yang, M., Zeng, Z., Ding, R., Wang, X., Zhuang, Z., Zhou, S., Li, S., Yang, H., Li, Z., Zheng, E., Huang, W., Yang, J., Wu, Z., 2019. Exploring the fecal microbial composition and metagenomic functional capacities associated with feed efficiency in commercial dly pigs. Frontiers in Microbiology 10, 52.
R Core Team, 2022. R: a language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.
Ringseis, R., Keller, J., Eder, K., 2018a. Basic mechanisms of the regulation of L-carnitine status in monogastrics and efficacy of L-carnitine as a feed additive in pigs and poultry. Journal of Animal Physiology and Animal Nutrition 102, 1686–1719.
Ringseis, R., Keller, J., Eder, K., 2018b. Regulation of carnitine status in ruminants and efficacy of carnitine supplementation on performance and health aspects of ruminant livestock: a review. Archives of Animal Nutrition 72, 1–30.
Rocadembosch, J., Amador, J., Bernaus, J., Font, J., Fraile, L.J., 2016. Production parameters and pig production cost: temporal evolution 2010-2014. Porcine Health Management 2, 11.
Ruggeri, R., Bee, G., Correa, F., Trevisi, P., Ollagnier, C., 2025. Intrauterine growth restriction, defined by an elevated brain-to-liver weight ratio, affects faecal microbiota composition and, to a lesser extent, plasma metabolome profile at different ages in pig. Animal Microbiome 7, 17.
Sarri, L., Balcells, J., Seradj, A.R., de la Fuente, G., 2024. Protein turnover in pigs: a review of interacting factors. Journal of Animal Physiology and Animal Nutrition 108, 451–469.
Schooneman, M.G., Vaz, F.M., Houten, S.M., Soeters, M.R., 2013. Acylcarnitines: reflecting or inflicting insulin resistance? Diabetes 62, 1–8.
Seyres, D., Cabassi, A., Lambourne, J.J., Burden, Farrow, S., McKinney, H., Batista, J., Kempster, C., Pietzner, M., Slingsby, O., Cao, T.H., Quinn, P.A., Stefanucci, L., Sims, M.C., Rehnstrom, K., Adams, C.L., Frary, A., Ergüener, B., Kreuzhuber, R., Mocciaro, G., D'Amore, S., Koulman, A., Grassi, L., Griffin, J.L., Ng, L.L., Park, A., Savage, D.B., Langenberg, C., Bock, C., Downes, K., Wareham, N.J., Allison, M., Vacca, M., Kirk, P.D.W., Frontini, M., 2022. Transcriptional, epigenetic and metabolic signatures in cardiometabolic syndrome defined by extreme phenotypes. Clinical Epigenetics 14, 39.
Silva, F., Lopes, M.S., Lopes, P.S., Gasparino, E., 2019. A genome-wide association study for feed efficiency-related traits in a crossbred pig population. Animal 13, 2447–2456.
Soleimani, T., Gilbert, H., 2021. An approach to achieve overall farm feed efficiency in pig production: environmental evaluation through individual life cycle assessment. International Journal of Life Cycle Assessment 26, 455–469.
Stachulski, A.V., Knausenberger, T.B.A., Shah, S.N., Hoyles, L., McArthur, S., 2023. A host–gut microbial amino acid co-metabolite, p-cresol glucuronide, promotes blood–brain barrier integrity in vivo. Tissue Barriers 11, 2073175.
Sztepanacz, J.L., Houle, D., 2024. Regularized regression can improve estimates of multivariate selection in the face of multicollinearity and limited data. Evolution Letters 8, 361–373.

Ufnal, M., Zadlo, A., Ostaszewski, R., 2015. TMAO: a small molecule of great expectations. Nutrition 31, 1317–1323.
van Buuren, S., 2018. Flexible Imputation of Missing Data. CRC Press, Boca Raton, FL, USA.
Vaz, F.M., Wanders, R.J.A., 2002. Carnitine biosynthesis in mammals. Biochemical Journal 361, 417–429.
Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272.
Wang, Z., He, Y., Wang, C., Ao, H., Tan, Z., Xing, K., 2021. Variations in microbial diversity and metabolite profiles of female landrace finishing pigs with distinct feed efficiency. Frontiers in Veterinary Science 8, 702931.
Wikoff, W.R., Anfora, A.T., Liu, J., Schultz, P.G., Lesley, S.A., Peters, E.C., Siuzdak, G., 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. Proceedings of the National Academy of Sciences of the United States of America 106, 3698–3703.
Wilson, M.P., Plecko, B., Mills, P.B., Clayton, P.T., 2019. Disorders affecting vitamin B6 metabolism. Journal of Inherited Metabolic Disease 42, 629–646.
Winston, J.A., Theriot, C.M., 2020. Diversification of host bile acids by members of the gut microbiota. Gut Microbes 11, 158–171.
Wu, J., Ye, Y., Quan, J., Ding, R., Wang, X., Zhuang, Z., Zhou, S., Geng, Q., Xu, C., Hong, L., Xu, Z., Zheng, E., Cai, G., Wu, Z., Yang, J., 2021. Using nontargeted LC-MS metabolomics to identify the Association of Biomarkers in pig feces with feed efficiency. Porcine Health Management 7, 39.
Ye, Y., Wu, J., Quan, J., Ding, R., Yang, M., Wang, X., Zhou, S., Zhuang, Z., Huang, S., Gu, T., Hong, L., Zheng, E., Wu, Z., Yang, J., 2021. Lipids and organic acids in three gut locations affect feed efficiency of commercial pigs as revealed by LC–MS-based metabolomics. Scientific Reports 11, 7746.
Zhang, B., Hu, S., Baskin, E., Patt, A., Siddiqui, J.K., Mathé, E.A., 2018. RaMP: a comprehensive relational database of metabolomics pathways for pathway enrichment analysis of genes and metabolites. Metabolites 8, 16.
Zhang, R., Sun, X., Huang, Z., Pan, Y., Westbrook, A., Li, S., Bazzano, L., Chen, W., He, J., Kelly, T., Li, C., 2022. Examination of serum metabolome altered by cigarette smoking identifies novel metabolites mediating smoking-BMI association. Obesity 30, 943–952.
Zubiri-Gaitán, A., Blasco, A., Hernández, P., 2023. Plasma metabolomic profiling in two rabbit lines divergently selected for intramuscular fat content. Communication Biology 6, 893.