Using nontargeted LC-MS metabolomics to identify the association of biomarkers in pig feces with feed efficiency

Jie Wu
South China Agricultural University
Yong Ye
South China Agricultural University
Jianping Quan
South China Agricultural University
Rongrong Ding
South China Agricultural University
Xingwang Wang
South China Agricultural University
Zhanwei Zhuang
South China Agricultural University
Shenping Zhou
South China Agricultural University
Qian Geng
South China Agricultural University
Cineng Xu
South China Agricultural University
Linjun Hong
South China Agricultural University
Zheng Xu
South China Agricultural University
Enqin Zheng
South China Agricultural University
Gengyuan Cai
South China Agricultural University
Jie Yang (✉ jieyang2012@hotmail.com)
South China Agricultural University https://orcid.org/0000-0002-7031-2160
Zhenfang Wu
South China Agricultural University

Research

Keywords: feed efficiency, pig, LC-MS, WGCNA

Posted Date: November 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-101175/v1

Abstract

Background: Improving feed efficiency is economically and environmentally beneficial in the pig industry. A deeper understanding of feed efficiency is essential on many levels for its highly complex nature. The aim of this project is to explore the relationship between fecal metabolites and feed efficiency-related traits, thereby identifying metabolites that may assist in the screening of the feed efficiency of pigs.

Result: We performed fecal metabolomics analysis on 50 individuals selected from 225 Duroc x (Landrace x Yorkshire) (DLY) commercial pigs, 25 with an extremely high feed efficiency and 25 with an extremely low feed efficiency. A total of 6,749 and 5,644 m/z features were detected in positive and negative ionization modes by liquid chromatography-mass spectrometry (LC/MS). Through weighted coexpression network analysis (WGCNA), we found that one module in each positive and negative mode was related to residual feed intake (RFI), and six and three metabolites were further identified. The nine metabolites were found to be involved in multiple metabolic pathways, including lipid metabolism (primary bile acid synthesis, linoleic acid metabolism), vitamin D, glucose metabolism, and others. In parallel, the expression patterns of seven genes involved in lipid metabolism were illuminated by real-time qPCR. Then, Lasso regression analysis was used to evaluate the importance of nine metabolites obtained by the annotation process.

Conclusions: Altogether, this study can not only provide new insights for the subsequent evaluation of commercial pig feed efficiency through small molecule metabolites, but also provide a reference for the development of new feed additives.

Background

Feed remains the main input cost in all animal production. Additionally, pork is one of the most important sources of meat for human beings. To cope with increasing market demand and breeding costs, improving the feed efficiency (FE) of pigs has always been a concern of breeders. Since FE cannot be measured directly for its complex pleiotropy, the feed conversion ratio (FCR) and residual feed intake (RFI) are often used as alternative selection conditions for FE [1–4]. However, although RFI traits have a strong genetic correlation with FCR (0.76 to 0.99) [5], a unified indicator for evaluating feed efficiency has not yet been determined.

At present, there are various omics methods to explore the molecular mechanisms that affect the feed efficiency of pigs, including transcriptomics [6, 7], genomics [5] and 16S rRNA gene sequencing [8]. However, metabolomics is rarely used to study pig feed efficiency phenotypes. Metabolites produced by the intestinal microbiota are increasingly recognized as an important part of human physiology [9]. As the downstream of the gene regulation network and protein interaction network, the metabolite can provide more detailed biological terminal information. By analyzing the changes in the expression of metabolites, it can help researchers find novel biomarkers and further understand the currently known metabolic pathways so that they can be applied to the study of various apparent traits. Furthermore, fecal

metabolites are the final products of the metabolism of cells and intestinal microbiota, which can help to reflect the absorption and digestion of nutrients by the intestinal flora and digestive tract more comprehensively. Finally, the analysis of fecal metabolomics provides a noninvasive way to study the correlation between biological traits and metabolites. Therefore, the fecal metabolome can not only partially explain the composition of the gut microbiota but also be used as biomarkers to investigate the relationship between gut microbial metabolism and host phenotypes [10].

Duroc · (Landrace · Yorkshire) (DLY) commercial pigs with good carcass and meat quality traits currently account for the largest sales share in the Chinese pork market [11]. Therefore, exploring the factors that affect the feed efficiency of DLY pigs has great significance for improving the economic benefits of the swine industry.

Liquid chromatography-mass spectrometry (LC/MS) is a relatively high-resolution separation analysis technique. Because of its high sensitivity, wide dynamic range, and lack of derivatization, LC-MS analysis has become a common technique for metabolomics research [12, 13]. However, few studies have used LC/MS technology to analyze the relationship between the fecal metabolome and feed efficiency. Hence, in this project, we analyzed the metabolomics of feces based on high and low feed efficiency groupings, aiming to unearth some small molecular metabolites that can evaluate feed efficiency and provide a new perspective for feed management in the pig industry.

Materials And Methods

Animals and sample collection

A total of 225 female DLY pigs in this study were provided by Guangdong Wen's Foodstuffs Group Co., Ltd. (Yunfu, China). The experimental animals were randomly divided into 30 pens, and each pen had 6 to 8 pigs. All pigs were arranged in a controlled environment with free access to water and food. Phenotypic data such as feed intake and weight per meal were recorded by the Osborne Feed Intake Recording Equipment (FIRE) Pig Performance Testing System (Osborne, KS, United States). The entire experimental process lasted approximately 12 weeks from an initial weight of approximately 30 kg to 100 kg. The calculation methods for FCR and RFI are based on previously described studies [14]. After ranking the FCR values of 225 pigs, the top 25 highest FCR and 25 lowest FCR were selected as the Low-FE and High-FE groups, respectively. Fecal samples of 50 sows were collected after rectal stimulation. From each pig 2–3 tubes of fecal samples were collected, and the tubes were immediately transferred to liquid nitrogen for temporary storage. Then, to minimize the effects of microbial fermentation, the samples were stored at − 80 °C until analysis. In the entire monitored experiment, except for the sampling required by the research, no behaviors that caused animal stress reactions were carried out. The experimental protocol used in this study is in accordance with the Animal Protection and Use Committee of South China Agricultural University (SCAU, Guangzhou, China) (Approval number SCAU#0017).

Real-time Quantitative Polymerase Chain Reaction

Gene expression levels were quantified using real-time qPCR to verify our analysis. Total RNA was extracted from pig liver tissues using the Trizol reagent, as described by the manufacturer (Invitrogen, Waltham, MA, USA). The liver tissue samples were collected from three animals in each of the high and low FE groups. Quality and quantity of extracted RNA samples were evaluated using Nanodrop 2000 (Thermo Scientific, Waltham, MA, USA). The cDNA was synthesized from mRNA using the PrimeScript RT reagent Kit (Takara Bio Inc., Kusatsu, Shiga, Japan). After reverse transcription, the RT-qPCR reactions were carried out in triplicate using the Power-Up SYBR green master mix (Thermo Fisher Scientific, Waltham, MA, USA) in QuantStudio 7 Flex real-time PCR system (Applied Biosystems, Life Technologies, USA). Relative expression was calculated based on $2^{-\Delta\Delta Ct}$ method [15]. The sequence of the primers for sterol 27-Hydroxylase (*CYP27A1*), cholesterol 7α-hydroxylase (*CYP7A1*), fibroblast growth factor receptor substrate 2 (*FRS2*), fatty acid desaturase-2 (*FADS2*), fatty acid desaturase-1 (*FADS1*), squalene epoxidase (*SQLE*), and nuclear receptor subfamily 1 group H member 4 (*NR1H4*) were listed in Supplementary Table S1. The β-actin gene was used as an internal control for quantification of all other proteins.

Fecal sample pretreatment

Fifty milligrams of fecal sample was accurately weighed, and 400 μl of extraction solution (methanol: water = 4:1) was added to the samples. Then, the high-throughput tissue crusher was used to crush at low temperature (60 Hz, -20 ℃). After vortex mixing and ultrasonic extraction on ice for 10 min · 3 (300 W), the extracted samples were placed at -20 ℃ for 30 min. The solution was then centrifuged at 13,000 g for 15 min (4 °C), and the supernatant was extracted and injected into the LC-MS system for analysis.

LC-MS analysis and quality control

The instrument platform for LC-MS analysis in this study was the AB SCIEX UPLC-TripleTOF system. The chromatographic conditions were as follows: the column was a BEH C18 column (100 mm × 2.1 mm i.d., 1.7 μm; Waters, Milford, USA); mobile phase A was water (containing 0.1% formic acid), and mobile phase B was acetonitrile/isopropanol (1/1) (containing 0.1% formic acid); the flow rate was 0.40 mL/min, the injection volume was 20 μL, and the column temperature was 40 °C. In addition, the sample mass spectrometer signals were collected using positive and negative ion scanning modes. The mass spectrometry conditions were as follows: electrospray capillary voltage, injection voltage and collision voltage: 1.0 kV, 40 V and 6 eV; ion source temperature and desolvation temperature: 120 ℃ and 500 ℃; carrier gas flow rate: 900 L/h; mass spectrometry scan range: 50-1000 m/z; resolution: 30,000.

To evaluate the stability of the analysis system and find the variables with large variations in the analysis system during analysis, all test samples were mixed as quality control (QC) samples. In the process of instrument analysis, a QC sample was inserted every 8–10 samples.

Data analysis

Before conducting statistical analysis, the raw data were imported into the metabolomics software ProgenesisQI (Waters Corporation, Milford, USA) to generate the matrix of retention time, mass-to-charge ratio, and peak intensity for baseline filtering, peak identification, integration, retention time correction,

and peak alignment. Furthermore, to obtain the final data matrix for subsequent analysis, the preprocessing process was as follows: (i) only variables with nonzero values above 80% in all samples were retained; (ii) missing values were filled using the k-nearest neighbors (KNN) approach in the R DMwR package; (iii) standardized values were obtained via the Z-score method, and the variables with relative standard deviation (RSD) ≥ 30% of the QC samples were deleted.

Principal component analysis (PCA) and (Orthogonal) partial least squares discrimination analysis (OPLS-DA) models were built using the ropls package in R [16]. PCA was used to observe the overall distribution between samples and the dispersion degree between groups. OPLS-DA was used to distinguish the different metabolites between groups. The goodness of fit (R^2) and goodness of prediction (Q^2) in cross-validation were used to evaluate the performance of the OPLS-DA model, and 500 permutation tests were performed.

Weighted Gene Correlation Network Analysis

Network and clustering analyses were performed using the R package Weighted Gene Coexpression Network Analysis (WGCNA) [17]. The Pearson correlation coefficient was calculated to obtain a coexpression similarity measure and used to subsequently construct an adjacency matrix using soft threshold combined with topological overlap matrix (TOM). Then, hierarchical clustering was performed based on the TOM. Briefly, the soft thresholds of the positive and negative ion modes were set to 3 and 8, respectively, to achieve the approximate scale-free topology of the signed network ($R^2 > 0.9$) (Supplementary Figure S1). In the dynamic tree cutting algorithm, deepSplit was set to 2 and minModuleSize was set to 50. The first principal component of the metabolite module was used as the feature vector of the module (including most of the variation information of all metabolites in the module), used to calculate the correlation coefficient between the metabolite module and feed efficiency, and then the most relevant module for subsequent analysis was selected. Subsequently, the gene significance (GS) and module membership (MM) of the most relevant module were calculated. Among these, GS can represent the correlation between metabolic characteristics and phenotype, and MM can represent the correlation between metabolic characteristics and module feature vectors. GS > 0.2 and MM > 0.8 were set as the threshold to screen the hub genes. Since WGCNA was first used for transcriptome data, we followed the term hub gene to represent the important metabolites identified. Subsequently, hub genes were identified by using the online Human Metabolome Database (HMDB, http://www.hmdb.ca/) and the METLIN public database (https://metlin.scripps.edu/). The *p*-values of the hub genes were computed using the Wilcoxon test. The pathways in which hub genes participated were identified in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/).

Lasso-penalized Linear Regression

We performed the Lasso regression in R using the glmnet [18] and caret packages. The sample data were randomly divided into a training set and a test set at a 1:1 ratio. Ten cross-validations were performed to calculate the lambda value (lambda = 0.08678594). Receiver operating characteristic (ROC) curves were

generated using the pROC curve, predictions were made on the training set and the test set, and the importance of the variables was evaluated by the varimp function of the caret package.

Result

Difference analysis of FE-related traits

A summary of FE-related phenotypic data, such as average feed intake (AFI), body weight (BW) and age, is shown in Supplementary Table S2. The calculated FCR and RFI data were used to compare the differences between the high- and low-FE groups. There were significant differences in the FCR phenotype between these two groups (*p-value* = $1.41 \cdot 10^{-9}$, Wilcoxon) (Fig. 1a). Similarly, the RFI of the low-FE group was significantly higher than that of the high-FE group (*p-value* = $3.67 \cdot 10^{-4}$, Wilcoxon) (Fig. 1b).

Statistics and Analysis of Different Metabolites

By liquid chromatography-mass spectrometry, a total of 5,644 and 6,749 m/z features were detected in negative and positive ionization modes, respectively. Our results showed that there was no clear separation between the high-FE and low-FE groups by PCA analysis (Supplementary Figure S2). Furthermore, we used the OPLS-DA model, which is a powerful statistical modeling tool, to identify differential metabolites between the two groups (Fig. 2). The ellipses in score plots for OPLS-DA were defined as the 95% critical limit of the Hotelling T2. For OPLS-DA analysis in both positive and negative ionization modes, the results showed that there was a complete separation between the high-FE group and the low-FE group ($R^2X = 0.357$, $R^2Y = 0.966$ and $Q^2Y = 0.292$ in positive mode; $R^2X = 0.330$, $R^2Y = 0.948$ and $Q^2Y = 0.178$ in negative mode) (Fig. 2a, b). From the results of OPLS-DA, it can be found that although there was a clear differentiation between the two groups in the positive and negative modes, a lower Q^2Y value indicated poor predictability and low quality of the model.

Metabolism network analysis

Referring to the results of the above difference analysis, we tried to use another approach to explore the association between metabolome networks and phenotypes. WGCNA is a method of grouping genes with similar gene expression patterns into a module and finding the hub gene in the module. This analysis was then performed in the positive and negative ion modes. By calculating the adjacency between metabolic features and merging the closer modules, a total of 14 (14) modules were obtained in positive (negative) modes (Supplemental Figure S3). Then, by constructing the correlation matrix between the modules and the phenotypic data related to feed efficiency, the module with the closest correlation to the sample traits was identified. Our results showed that the metabolites of the MEtan module had the strongest correlation with RFI traits in positive modes (r = 0.42, *p-value* = 0.004) (Fig. 3a). Similarly, the MEgreenyellow module in negative mode was positively correlated with RFI (r = 0.44, *p-value* = 0.002) (Fig. 3c). In the MEtan (MEgreenyellow) module, a total of 31 (20) metabolic features with important contributions were screened from 65 (145) features (Module Membership = 0.8, Gene significance = 0.2)

(Fig. 3b, d). After annotation, we obtained six and three metabolites in the positive and negative modes, respectively (Supplemental Table S3). The pathways involved in the nine metabolites can be roughly divided into four categories, including lipid metabolism (primary bile acid synthesis, linoleic acid metabolism), vitamin D, glucose metabolism, and others. Separately, three metabolites, 3a,7a,12a-trihydroxy-5b-cholestan-26-al (THC26), 3alpha,7alpha- dihydroxycoprostanic acid (DHCA) and 5β-cholestane-3α,7α,12α,22-tetrol (22-OH-THC), which were negatively correlated with RFI, were related to the synthesis of primary bile acids (KEGG ID: M00120). C24:5n-6,9,12,15,18 (C24:5n-6) was involved in the metabolism of linoleic acid (KEGG ID: M00592). There are two metabolites belonging to the vitamin D category, including (10S)-1α,19,25-trihydroxy-10,19-dihydrovitamin D3 and (22E)-1α-hydroxy-22,23-didehydrovitamin D3. The above six metabolites in the negative ion mode were all negatively correlated with RFI. The metabolite 2-Keto-3-deoxy-D-gluconic acid (KDG) is involved in glucose metabolism (KEGG ID: M00052). The remaining two metabolites are 6-hydroxyhexanoic acid and m-coumaric acid. To further validate our analysis, seven genes related to lipid metabolism were quantified by real-time qPCR experiment, including *CYP27A1*, *CYP7A1*, *FADS2*, *FADS2*, *FRS2*, *SQLE*, *NR1H4*. Five of the these genes were related to primary bile acid synthesis [19–21], and the remaining two were involved in the process of linoleic acid metabolism [22]. Since the expression level of mRNA changes dynamically in different stages of the liver, and protein plays a direct regulatory role in biological processes, genes that conformed to the trend of expression changes were identified as possibly affecting the metabolic process. The results of real-time qPCR showed that the expression trends of five genes were consistent with our analysis (Supplemental Table S4).

Lasso Regression

After identifying hub genes, we performed Lasso regression analysis on the nine annotated metabolites for feature selection. From the analysis results of WGCNA, we regrouped the samples into high- and low-FE groups based on the RFI values of individuals. To detect the quality of the model, the ROC curve was used to reflect its specificity and sensitivity. The results showed that the model has excellent prognostic effectiveness, with AUCs of the training set and the test set of 0.8817 and 0.8542, respectively (Fig. 4a). Subsequently, to evaluate the most valuable potential biomarkers, we ranked the importance of each metabolite according to the absolute value of the regression coefficient in the regression model (Fig. 4b).

Discussion

Improving the feed efficiency traits of livestock is of great significance, but it is not easy to estimate. Therefore, any measure that can effectively predict feed efficiency is meaningful for production. Although there is currently much work to study FE at the genetic level, few studies have linked metabolites to feed efficiency phenotypic traits. In this study, we analyzed and compared the metabolites in the feces of pigs in the high-FE and low-FE groups by LC-MC technology and interpretation tools, including WGCNA and Lasso regression. To the best of our knowledge, this is the first report combining these methods to study the metabonomic profile related to feed efficiency and related traits in DLY pigs.

At present, FCR and RFI are commonly used to evaluate FE traits, and it is believed that RFI can better represent feed efficiency [2, 3, 23], which is consistent with our WGCNA analysis results. In our data, we found that the use of powerful tools such as PCA and OPLS-DA were not sufficient to distinguish the different features between the high- and low-FE animals. There are many possible explanations for the unsatisfactory results of PCA and OPLS-DA, including but not limited to (1) the sampling process was carried out after the individual growth indicators were measured. When the pig reaches the slaughter weight (approximately 100 kg), its metabolic activity is often not as active as before, and the increase in weight has little effect on the growth performance of pigs after 100 kg [24]; (2) throughout the experiment, all test subjects were clinically healthy. Therefore, there is no physiological interference between the FE groups that could cause large metabolome differences. (3) The number of animal individuals in our study (25 individuals per group) may not have high statistical power, so more animal groups and more targeted experimental designs may be needed to evaluate feed efficiency in the future. Because the results of the PCA and OPLS-DA models were not ideal, we then adopted WGCNA analysis to select the modules and metabolites most closely related to RFI and FCR. After screening and annotation, we obtained nine metabolites in these models. Based on these metabolites, we identified four pathways from the KEGG database that were also significantly related to feed efficiency, including lipid metabolism (primary bile acid synthesis, linoleic acid metabolism), vitamin D, and glucose metabolism. Moreover, the Lasso regression model showed that all nine annotated metabolites contribute to feed efficiency.

The metabolite 22-OH-THC is a kind of bile alcohol, which is the end product of catabolism of cholestanoic acids [25–27]. Bile alcohol may be regarded as an intermediate and side product from the normal pathways in bile acid biosynthesis [28]. Notably, THC26 and DHCA were mainly involved in the biosynthesis of primary bile acids. The specific synthesis process is that cholesterol 7α-hydroxylase (CYP27A1) catalyzes the oxidation of steroid side chains to form THC26 or DHCA in the mitochondria of liver cells and then obtains the primary bile acid cholic acid (CD) or chenodeoxycholic acid (CDCA) under the catalysis of various enzymes [19, 29–32]. Interestingly, although the synthesis of bile acids is determined by a variety of cytochrome P450 enzymes (CYPs), both THC26 and DHCA are intermediate products catalyzed by CYP27A1 [33]. Bile acids start from the catabolism of cholesterol and are the final product of cholesterol catabolism; they play a critical role in food digestion and nutrient absorption, helping the absorption of lipids and fat-soluble vitamins in the intestine [32, 34–36]. After passing down the intestine with bile, approximately 95% of bile acids are reabsorbed in the terminal ileum and circulate back to the liver through the portal vein [29, 35, 37]. The performance of these functions of bile acid mainly depends on its enterohepatic circulation process, which is of great significance for nutrient absorption and distribution, metabolic regulation and homeostasis [29, 35, 37–39]. The results of metabolite network analysis showed that three metabolites related to bile acid synthesis were significantly negatively correlated with RFI traits, which means that they were positively correlated with feed efficiency. This result was consistent with our finding that these metabolites had higher levels in the high feed efficiency group. At present, little is known about the transport of bile acids and intermediates between different chambers, which may provide some references for understanding the important factors in the synthesis of bile acids.

In addition, (10S)-1α,19,25–trihydroxy-10,19-dihydrovitamin D3 and (22E)-1α-hydroxy–22,23-didehydrovitamin D3 belong to vitamin D, which is a steroid derivative [40, 41]. Vitamin D has various effects on lipid metabolism and immune system function through its effects on nuclear hormone receptors (such as vitamin D receptor and PPARγ) [42, 43]. Similarly, our results were consistent with the effect of vitamin D on lipid metabolism mediated through these receptors. Previous studies have shown that the CYP27A1 enzyme can catalyze the hydroxylation of compounds both in the biosynthesis of bile acids and the bioactivation of vitamin D3 [44–46]. The acidic pathway (or alternative pathway) of bile acid synthesis is initiated by CYP27A1, which is a mitochondrial cytochrome P450 enzyme widely distributed in most tissues and macrophages [29, 47]. CYP27A1 can not only catalyze the 25-hydroxylation of vitamin D3, which is required for the conversion of vitamin D3 into a functionally active form, but may also regulate cholesterol homeostasis by promoting the synthesis of bile acids or producing active oxysterols [46–49]. Although there are no current reports on the effect of adding this enzyme, this may warrant further concern. Additionally, metabolite C24:5n-6 was involved in the alpha linolenic acid and linoleic acid metabolism pathways [50]. Linoleic acid is the main dietary n-6 polyunsaturated fatty acid (PUFA), and livestock mainly obtain it from diets such as vegetable oil, soybeans, and corn [51]. Previous studies reported that higher n-6 PUFA intake can reduce liver fat in overweight individuals, improve liver metabolism, and regulate the balance between fatty acid oxidation and lipid synthesis [52, 53]. In the process of linoleic acid metabolism, linoleic acid is catalyzed by the rate-limiting enzyme fatty acid desaturase-2 (FADS2), and after a series of extensions, C24:5n-6 is produced by FADS2 catalyzed C24:4n-6 [54]. Our results showed that C24:5n-6 correlated negatively and significantly with RFI traits and was significantly higher in the high-FE group than in the low-FE group (*p-value* = 0.002). Notably, in the process of linoleic acid synthesis, there was no significant difference between the upstream metabolite linoleic acid in the high and low FE groups, while the downstream C24:5n-6 was extremely different in the two groups. Similarly, there is currently no research on adding FASD2 enzyme to feed, but the potential for it to lead to differences in the feed efficiency of livestock cannot be ignored. Based on this, these findings can also provide some references for further analysis of linoleic acid metabolism. Improving feed efficiency is a concern, but these two enzymes and potential metabolic markers deserve further evaluation and research to clarify their biological significance.

Conclusion

In this study, through WGCNA coexpression analysis, we found two important modules and 210 potential metabolic features related to feed efficiency in the positive and negative modes. The nine potential metabolic markers obtained from screening and annotation were assessed by the Lasso model, which offered possibilities of collecting metabolic hallmarks in feces under minimizing stress reactions in the future. Based on the identified several important metabolites, from the perspectives of pathways, biological processes, and functions, we have revealed two potential key enzymes that may affect the feed efficiency of pigs. In short, our integrated method using WGCNA analysis, data annotation, and Lasso regression enriched the molecular markers of the biological processes related to feed efficiency. This study identifies important pathways and candidate biomarkers closely related to RFI, that contribute to

understanding the molecular mechanism of feed efficiency more comprehensively and provides an important reference for further verifying the application of metabolite signatures to identify pig feed efficiency traits.

Abbreviations

AFI: Average feed intake; BW: Body weight; C24:5n-6: C24:5n-6,9,12,15,18; CA: Cholic Acid; CDCA: Chenodeoxycholic acid; CYP27A1: Cholesterol 7α-hydroxylase; DHCA: 3alpha,7alpha-Dihydroxycoprostanic acid; FADS2: Fatty acid desaturase-2; FCR: Feed conversion ratio; FE: Feed efficiency; GS: Gene significance; H-FE: High feed efficiency; KDG: 2-Keto-3-deoxy-D-gluconic acid; L-FE: Low feed efficiency; MM: Module membership; OPLS-DA: Orthogonal partial least squares discriminant analysis; PCA: principal component analysis; PUFA: Polyunsaturated fatty acid; RFI: Residual feed intake; THC26: 3a,7a,12a-Trihydroxy-5b-cholestan-26-al; WGCNA: Weighted gene co-expression network analysis; 22-OH-THC: 5β-Cholestane-3α,7α,12α,22-tetrol.

Declarations

Ethics approval and consent to participate

The experimental protocol used in this study was approved by the Animal Protection and Use Committee of South China Agricultural University (SCAU, Guangzhou, China).

Consent for publication

Not applicable.

Availability of data and material

All data generated and analyzed during this study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (31972540)、the National High-quality lean-type Pig Breeding United Research Program of China、the Guangdong YangFan Innovative and Entrepreneurial Research Team Program (2016YT03H062), and the Key-Area Research and Development Program of Guangdong Province (2018B020203002).

Authors' Contributions

JY and ZW: conceived and designed the experiments. JW, YY, JQ, RD, XW: collected the samples and recorded the phenotypes. RD: calculated the FCR and RFI of experimental pigs. JW and YY: data analyzed. JW, JY, and ZW: wrote and revised the manuscript. All authors reviewed and approved the manuscript.

Acknowledgements

We thank Guangdong Wen's Foodstuffs Group for constant support.

References

1. Carmelo VAO, Banerjee P, da Silva Diniz WJ, Kadarmideen HN. Metabolomic networks and pathways associated with feed efficiency and related-traits in Duroc and Landrace pigs. Scientific reports. 2020;10(1):1-14.
2. Do DN, Strathe AB, Ostersen T, Pant SD, Kadarmideen HN. Genome-wide association and pathway analysis of feed efficiency in pigs reveal candidate genes and pathways for residual feed intake. Front Genet. 2014;5:307.
3. McCormack UM, Curião T, Metzler-Zebeli BU, Magowan E, Berry DP, Reyer H, et al. Porcine feed efficiency-associated intestinal microbiota and physiological traits: Finding consistent cross-locational biomarkers for residual feed intake. Msystems. 2019;4(4).
4. Saintilan R, Mérour I, Schwob S, Sellier P, Bidanel J, Gilbert H. Genetic parameters and halothane genotype effect for residual feed intake in Piétrain growing pigs. Livestock Science. 2011;142(1-3):203-209.
5. Do DN, Strathe AB, Jensen J, Mark T, Kadarmideen HN. Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds. Journal of animal science. 2013;91(9):4069-4079.
6. Vigors S, O'Doherty JV, Bryan K, Sweeney T. A comparative analysis of the transcriptome profiles of liver and muscle tissue in pigs divergent for feed efficiency. BMC genomics. 2019;20(1):461.
7. Ramayo-Caldas Y, Ballester M, Sánchez JP, González-Rodríguez O, Revilla M, Reyer H, et al. Integrative approach using liver and duodenum RNA-Seq data identifies candidate genes and pathways associated with feed efficiency in pigs. entific Reports. 2018;8(1):558.
8. Quan J, Cai G, Ye J, Yang M, Ding R, Wang X, et al. A global comparison of the microbiome compositions of three gut locations in commercial pigs with extreme feed conversion ratios. Scientific reports. 2018;8(1):1-10.
9. De Vadder F, Kovatcheva-Datchary P, Goncalves D, Vinera J, Zitoun C, Duchampt A, et al. Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. Cell. 2014;156(1-2):84-96.
10. Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. Nature genetics. 2018;50(6):790-795.

11. Jiang Y, Zhu L, Tang G, Li M, Jiang A, Cen W, et al. Carcass and meat quality traits of four commercial pig crossbreeds in China. Genetics and Molecular Research. 2012;11(4):4447-4455.
12. Novais FJ, Pires PRL, Alexandre PA, Dromms RA, Iglesias AH, Ferraz JBS, et al. Identification of a metabolomic signature associated with feed efficiency in beef cattle. BMC Genomics. 2019;20(1).
13. Gika HG, Theodoridis GA, Plumb RS, Wilson ID. Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics. Journal of pharmaceutical and biomedical analysis. 2014;87:12-25.
14. Wu J, Wang X, Ding R, Quan J, Ye Y, Gu T, et al. Identification of Important Proteins and Pathways Affecting Feed Efficiency in DLY Pigs by iTRAQ-Based Proteomic Analysis. Animals. 2020;10(2):189.
15. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-408.
16. Thévenot EA, Roux A, Xu Y, Ezan E, Junot C. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. Journal of proteome research. 2015;14(8):3322-3335.
17. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9(1):559.
18. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996;58(1):267-288.
19. Li T, Apte U. Bile Acid Metabolism and Signaling in Cholestasis, Inflammation, and Cancer. Advances in Pharmacology. 2015;74:263.
20. Dong Y, Li C, Zhang Y, He Q, Daud MK, Chen J, Zhu S. Glutathione S-Transferase Gene Family in Gossypium raimondii and G. arboreum: Comparative Genomic Study and their Expression under Salt Stress. Front Plant Sci. 2016;7:139.
21. Reyer H, Oster M, Magowan E, Dannenberger D, Ponsuksili S, Wimmers K. Strategies towards Improved Feed Efficiency in Pigs Comprise Molecular Shifts in Hepatic Lipid and Carbohydrate Metabolism. International journal of molecular sciences. 2017;18(8).
22. Gol S, Pena RN, Rothschild MF, Tor M, Estany J. A polymorphism in the fatty acid desaturase-2 gene is associated with the arachidonic acid metabolism in pigs. Scientific reports. 2018;8(1):14336.
23. Ding R, Yang M, Wang X, Quan J, Zhuang Z, Zhou S, et al. Genetic architecture of feeding behavior and feed efficiency in a Duroc pig population. Front Genet. 2018;9:220.
24. Oliveira E, Bertol T, Coldebela A, Santos Filho J, Scandolera A, Warpechowski M. Live performance, carcass quality, and economic assessment of over 100kg slaughtered pigs. Arquivo Brasileiro de Medicina Veterinária e Zootecnia. 2015;67(6):1743-1750.
25. Iguchi Y, Yamaguchi M, Sato H, Kihira K, Nishimaki-Mogami T, Une M. Bile alcohols function as the ligands of membrane-type bile acid-activated G protein-coupled receptor. Journal of Lipid Research. 2010;51(6):1432-1441.

26. Iguchi Y, Kihira K, Nishimaki-Mogami T, Une M. Structure-activity relationship of bile alcohols as human farnesoid X receptor agonist. Steroids. 2010;75(1):95-100.
27. Kuramoto T, Matsumoto N, Hoshita T. Syntheses of 22- and 23-hydroxylated bile alcohols. Chemical & Pharmaceutical Bulletin. 1978;26(9):2788-2792.
28. Kuroki S, Shimazu K, Kuwabara M, Une M, Hoshita T. Identification of bile alcohols in human bile. Journal of Lipid Research. 1985;26(2):230-240.
29. Chiang JY. Bile acid metabolism and signaling. Comprehensive Physiology. 2013;3(3):1191-1212.
30. Cali JJ, Russell DW. Characterization of human sterol 27-hydroxylase. A mitochondrial cytochrome P-450 that catalyzes multiple oxidation reaction in bile acid biosynthesis. Journal of Biological Chemistry. 1991;266(12):7774.
31. Pikuleva IA, Babiker A, Waterman MR, Bjorkhem I. Activities of recombinant human cytochrome P450c27 (CYP27) which produce intermediates of alternative bile acid biosynthetic pathways. Journal of Biological Chemistry. 1998;273(29):18153.
32. Shapiro H, Kolodziejczyk AA, Halstuch D, Elinav E. Bile acids in glucose metabolism in health and disease. Journal of Experimental Medicine. 2018;215(2):jem.20171965.
33. Gómez C, Stücheli S, Kratschmar DV, Bouitbir J, Odermatt A. Development and Validation of a Highly Sensitive LC-MS/MS Method for the Analysis of Bile Acids in Serum, Plasma, and Liver Tissue Samples. Metabolites. 2020;10(7).
34. Halilbasic E, Claudel T, Trauner M. Bile acid transporters and regulatory nuclear receptors in the liver and beyond. J Hepatol. 2013;58(1):155-168.
35. Schneider KM, Albers S, Trautwein C. Role of bile acids in the gut-liver axis. J Hepatol. 2018:1083-1085.
36. de Aguiar Vallim TQ, Tarling EJ, Edwards PA. Pleiotropic roles of bile acids in metabolism. Cell Metab. 2013;17(5):657-669.
37. Garcia M, Thirouard L, Sedès L, Monrose M, Holota H, Caira F, et al. Nuclear Receptor Metabolism of Bile Acids and Xenobiotics: A Coordinated Detoxification System with Impact on Health and Diseases. International Journal of Molecular Sciences. 2018;19(11).
38. Roberts DMS, Magnusson BM, Burczynski FJ, Weiss M. Enterohepatic Circulation. Clinical Pharmacokinetics. 2002;41(10):751-790.
39. Hofmann AF. The enterohepatic circulation of bile acids in mammals: form and functions. Frontiers in Bioence. 2009;14(7):2584-2598.
40. Sicinski RR, Prahl JM, Smith CM, DeLuca HF. New 1alpha,25-dihydroxy-19-norvitamin D3 compounds of high biological activity: synthesis and biological evaluation of 2-hydroxymethyl, 2-methyl, and 2-methylene analogues. J Med Chem. 1998;41(23):4662-4674.
41. Mullard G: **Metabolomic investigations into human apocrine sweat secretions**. University of Nottingham; 2012.

42. Salamon H, Bruiners N, Lakehal K, Shi L, Ravi J, Yamaguchi KD, et al. Cutting edge: Vitamin D regulates lipid metabolism in Mycobacterium tuberculosis infection. J Immunol. 2014;193(1):30-34.
43. White JH. Vitamin D metabolism and signaling in the immune system. Reviews in endocrine and metabolic disorders. 2012;13(1):21-29.
44. Su P, RENNERT H, SHAYIQ RM, YAMAMOTO R, ZHENG Y-M, ADDYA S, et al. A cDNA encoding a rat mitochondrial cytochrome P450 catalyzing both the 26-hydroxylation of cholesterol and 25-hydroxylation of vitamin D3: gonadotropic regulation of the cognate mRNA in ovaries. DNA and cell biology. 1990;9(9):657-665.
45. Usui E, Noshiro M, Ohyama Y, Okuda K. Unique property of liver mitochondrial P450 to catalyze the two physiologically important reactions involved in both cholesterol catabolism and vitamin D activation. FEBS letters. 1990;274(1-2):175-177.
46. Jones G, Prosser DE, Kaufmann M. Cytochrome P450-mediated metabolism of vitamin D. Journal of lipid research. 2014;55(1):13-31.
47. Pikuleva IA, Babiker A, Waterman MR, Björkhem I. Activities of recombinant human cytochrome P450c27 (CYP27) which produce intermediates of alternative bile acid biosynthetic pathways. J Biol Chem. 1998;273(29):18153-18160.
48. Shinkyo R, Sakaki T, Kamakura M, Ohta M, Inouye K. Metabolism of vitamin D by human microsomal CYP2R1. Biochem Biophys Res Commun. 2004;324(1):451-457.
49. Bodin K, Andersson U, Rystedt E, Ellis E, Norlin M, Pikuleva I, et al. Metabolism of 4 beta - hydroxycholesterol in humans. J Biol Chem. 2002;277(35):31534-31540.
50. Martin J, Stapleton RD. Parenteral and Enteral Nutrition with Omega-3 Fatty Acids.
51. Pérez MM, Gonçalves ECS, Salgado JCS, Rocha MdS, Almeida PZd, Vici AC, et al. Production of omegas-6 and 9 from the hydrolysis of açaí and buriti oils by lipase immobilized on a hydrophobic support. Molecules. 2018;23(11):3015.
52. Bjermo H, Iggman D, Kullberg J, Dahlman I, Johansson L, Persson L, et al. Effects of n-6 PUFAs compared with SFAs on liver fat, lipoproteins, and inflammation in abdominal obesity: a randomized controlled trial. Am J Clin Nutr. 2012;95(5):1003-1012.
53. Clarke SD. The multi-dimensional regulation of gene expression by fatty acids: polyunsaturated fats as nutrient sensors. Curr Opin Lipidol. 2004;15(1):13-18.
54. Glaser C, Heinrich J, Koletzko B. Role of FADS1 and FADS2 polymorphisms in polyunsaturated fatty acid metabolism. Metab Clin Exp. 2010;59(7):993-999.

Figures

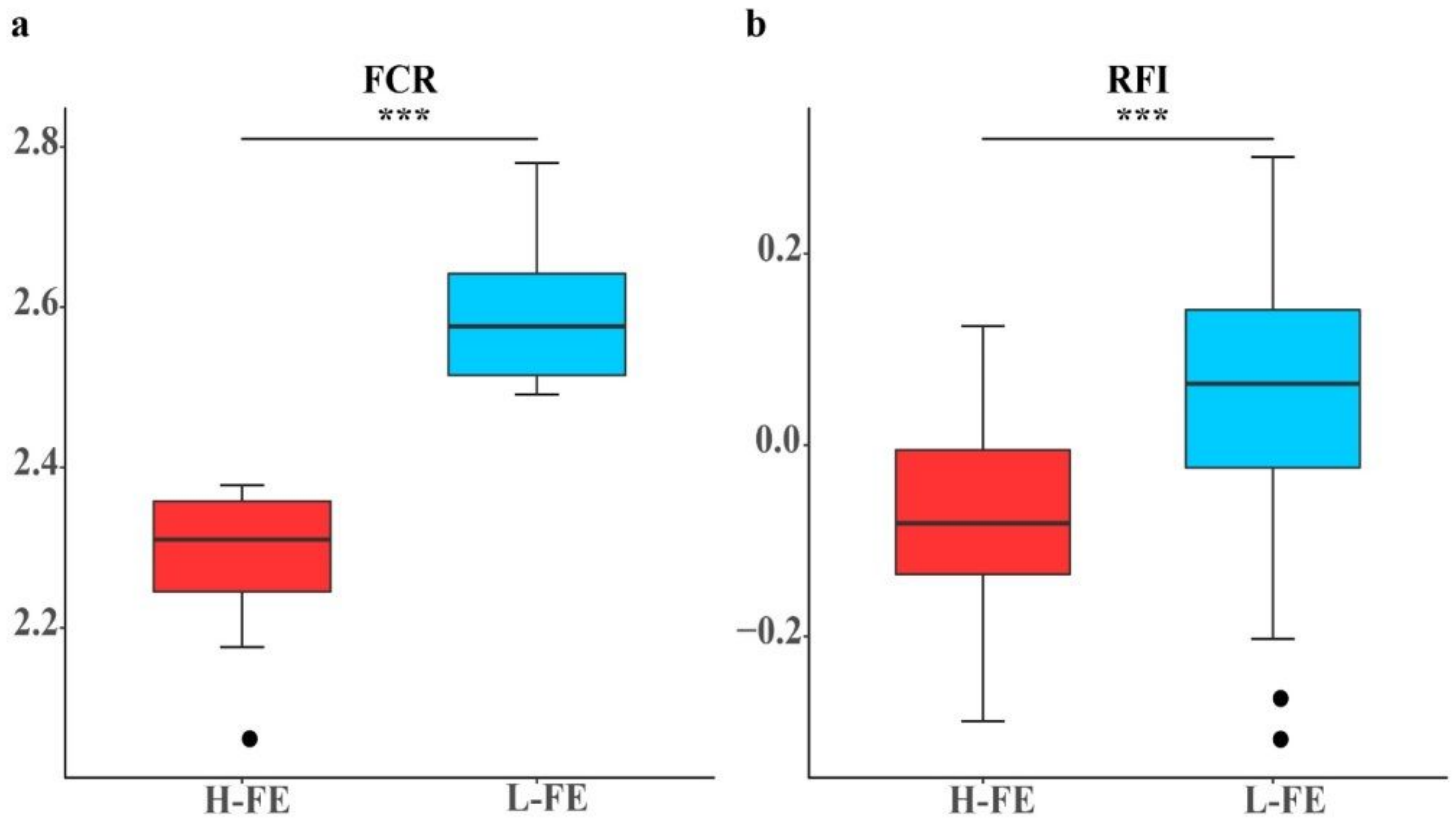

Figure 1

Differences in FCR trait between high- and low-FE groups. H-FE, high-feed efficiency; L-FE, low-feed efficiency; FCR, feed conversion ratio; RFI, residual feed intake. ***A significant difference of p-value < 0.001.

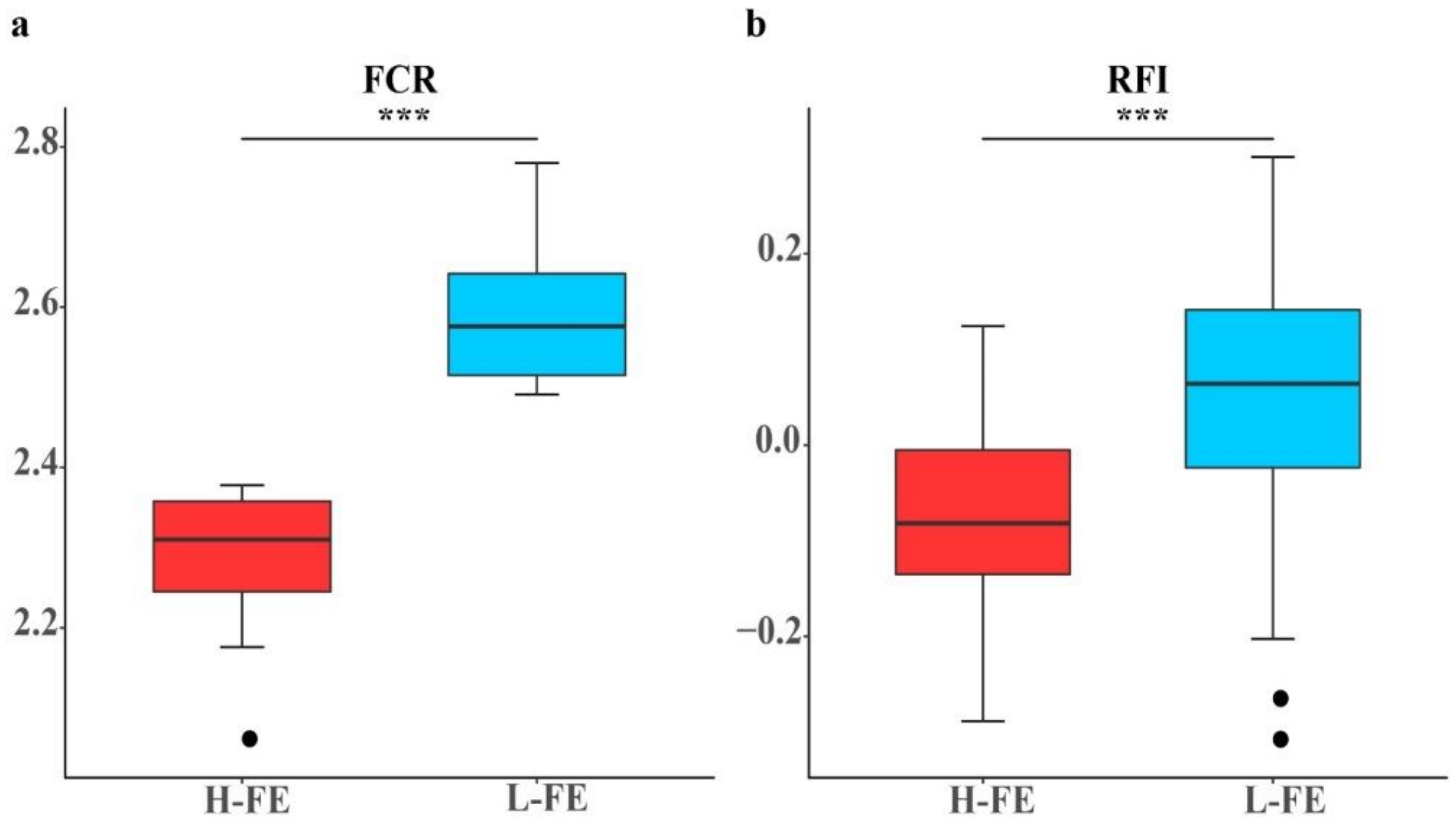

Figure 1

Differences in FCR trait between high- and low-FE groups. H-FE, high-feed efficiency; L-FE, low-feed efficiency; FCR, feed conversion ratio; RFI, residual feed intake. ***A significant difference of p-value < 0.001.

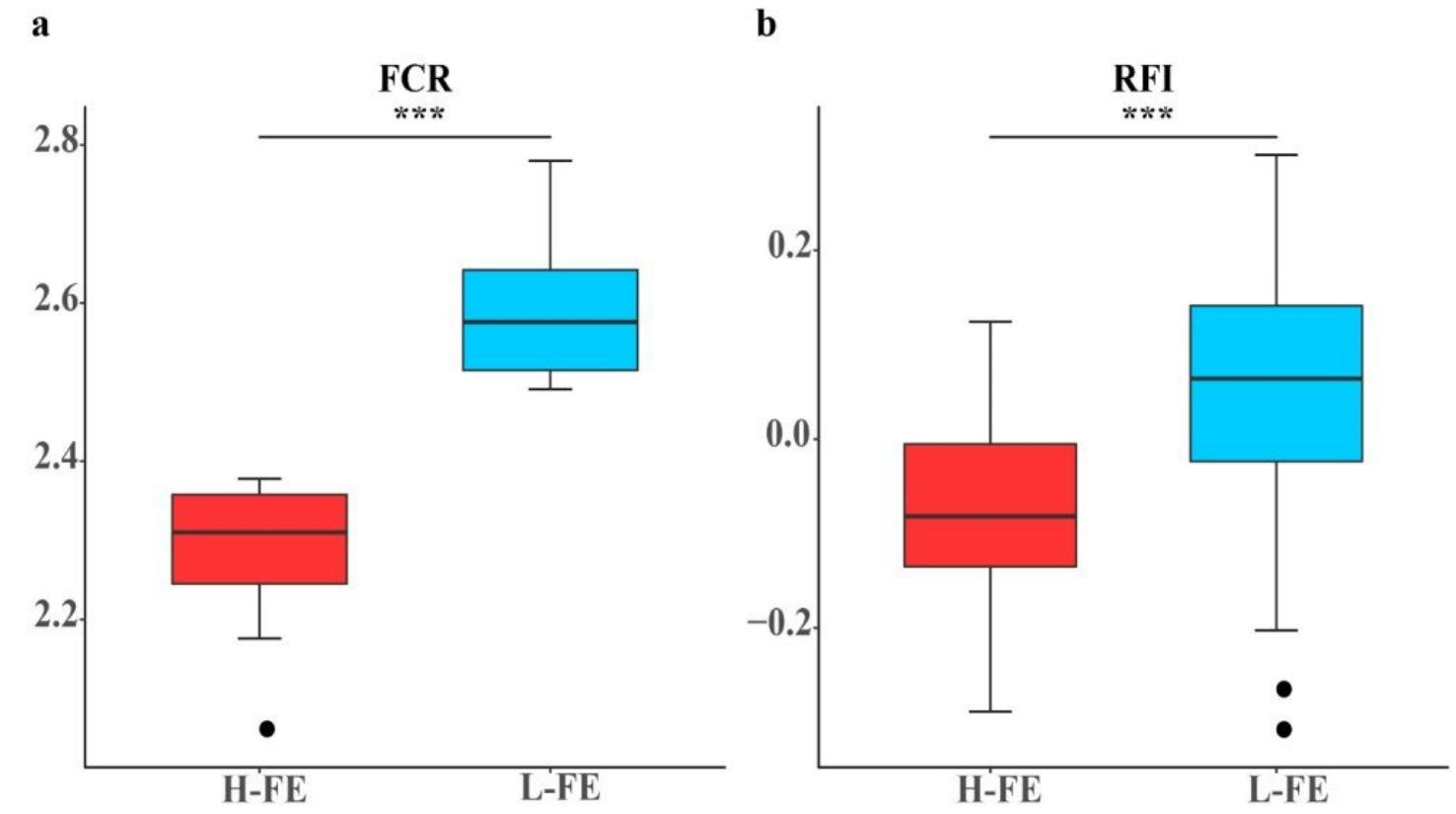

Figure 1

Differences in FCR trait between high- and low-FE groups. H-FE, high-feed efficiency; L-FE, low-feed efficiency; FCR, feed conversion ratio; RFI, residual feed intake. ***A significant difference of p-value < 0.001.

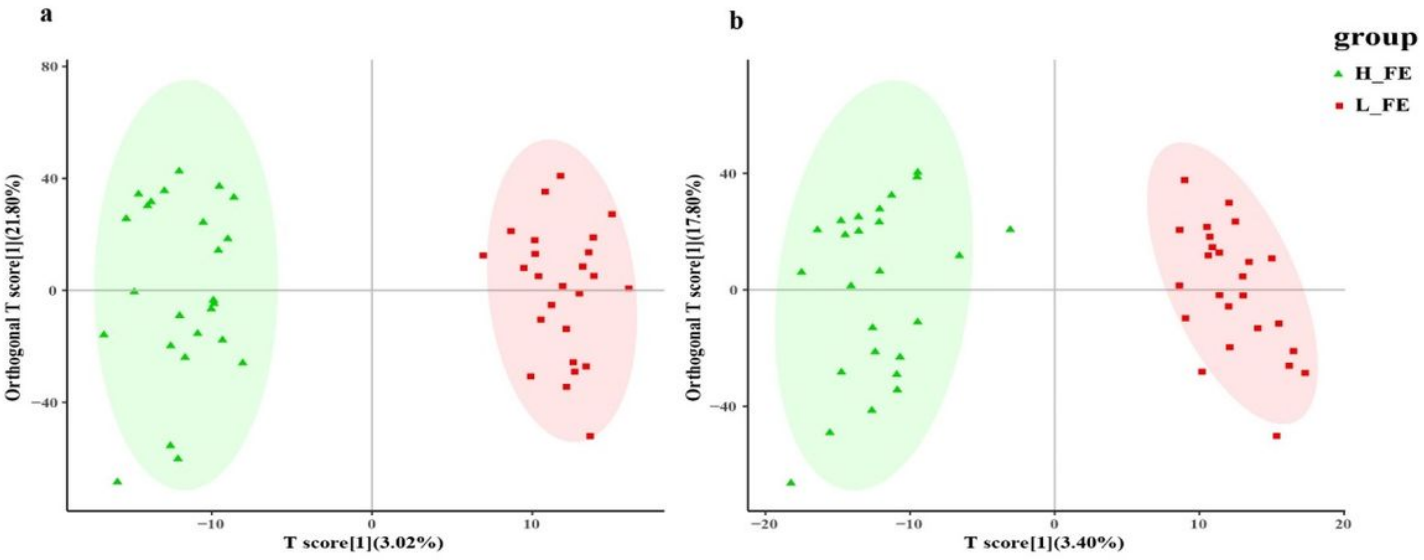

Figure 2

(Orthogonal) Partial Least Squares Discrimination Analysis ((O) PLS-DA) score plots based on LC/MS data of fecal samples from H-FE (green) and L-FE (red) of (a) positive and (b) negative modes.

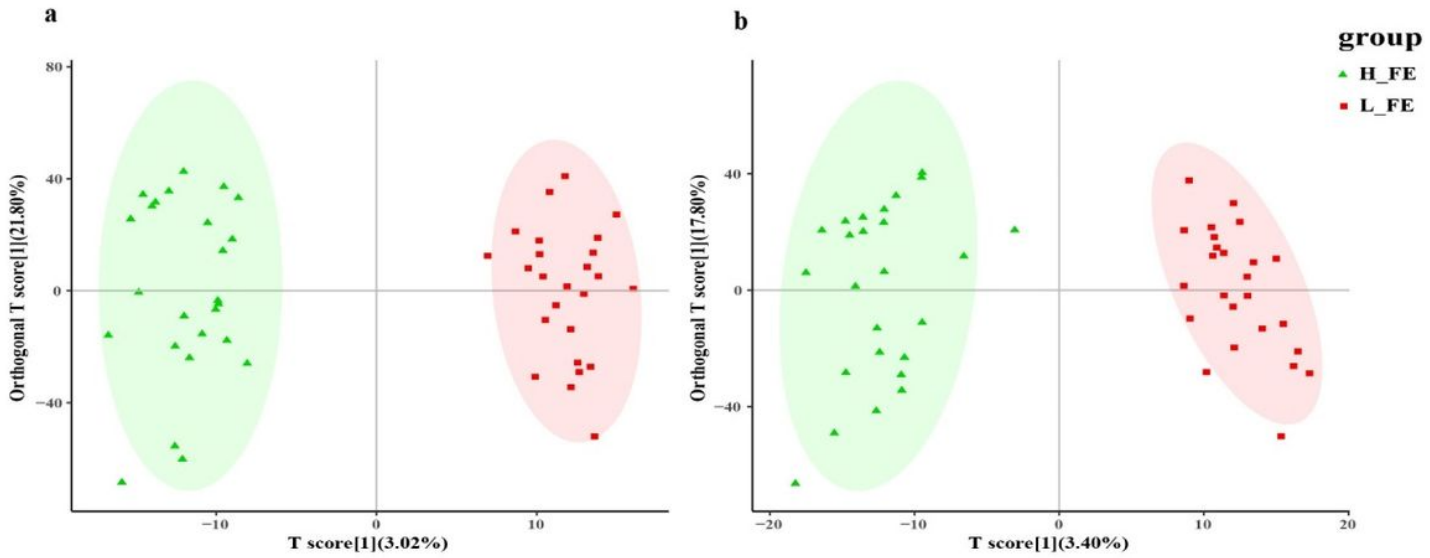

Figure 2

(Orthogonal) Partial Least Squares Discrimination Analysis ((O) PLS-DA) score plots based on LC/MS data of fecal samples from H-FE (green) and L-FE (red) of (a) positive and (b) negative modes.

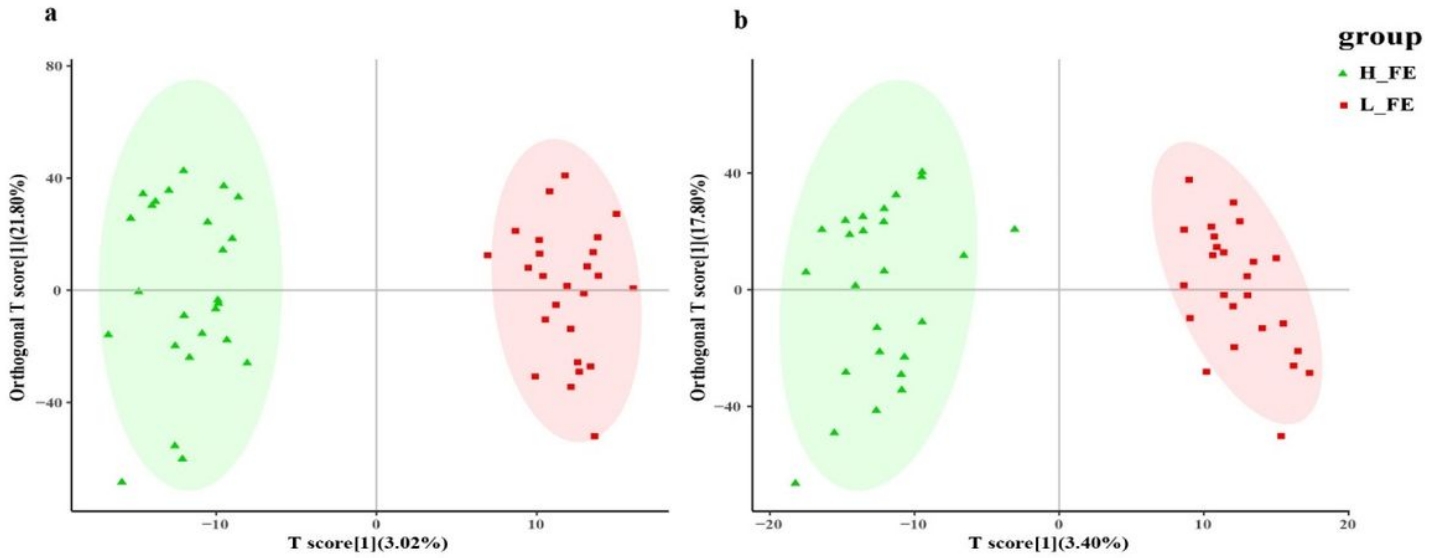

Figure 2

(Orthogonal) Partial Least Squares Discrimination Analysis ((O) PLS-DA) score plots based on LC/MS data of fecal samples from H-FE (green) and L-FE (red) of (a) positive and (b) negative modes.

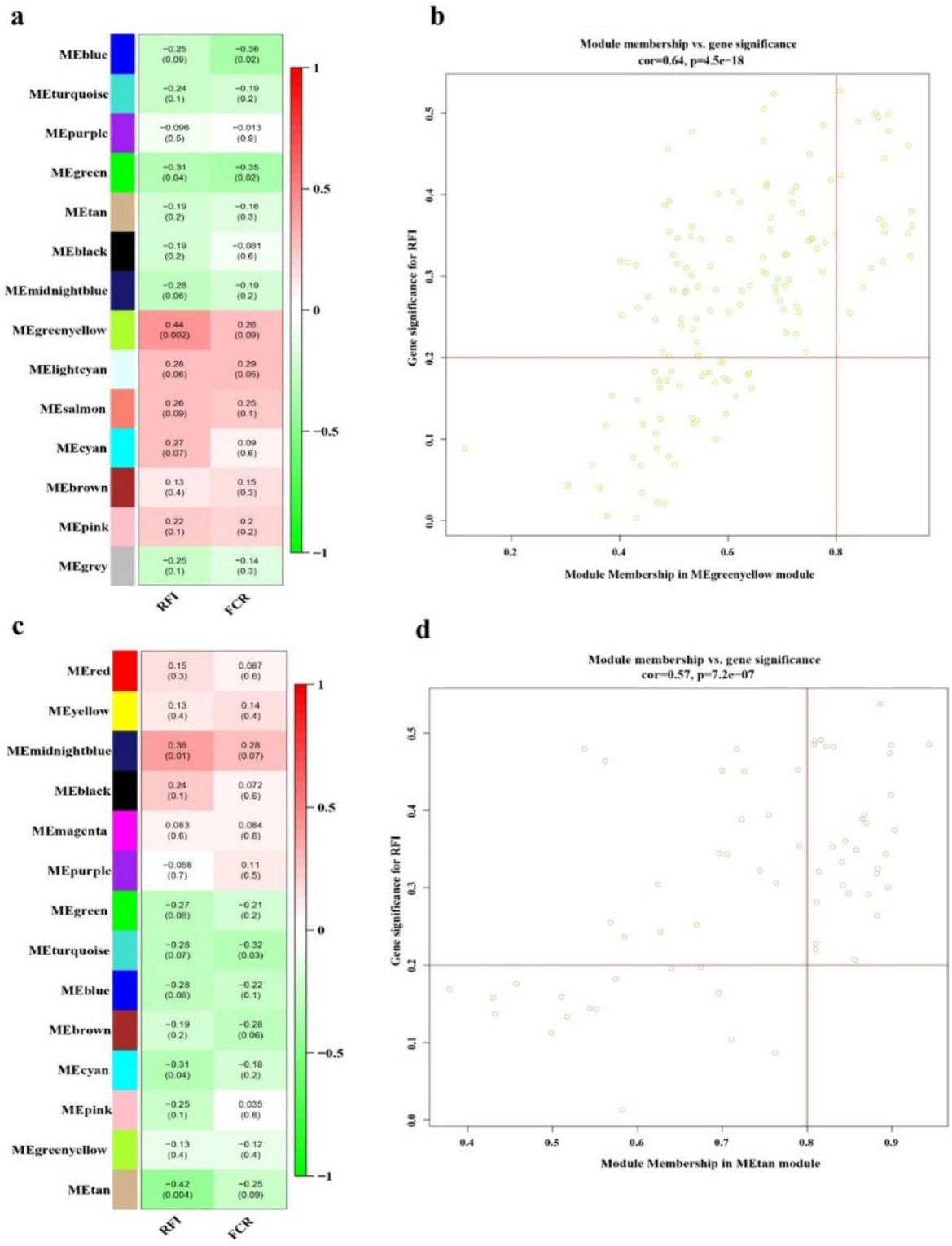

Figure 3

Coexpression network analysis of metabolic features. The left panel of the figure shows the correlation between the module and RFI or FCR in (a) negative and (c) positive models. The right panel of the figure shows the scatter plot of module membership and the gene significance in (b) MEgreenyellow or (d) MEtan module. Each row corresponds to ME, and each column corresponds to traits; the number in each

module represents the Pearson correlation between the module and RFI or FCR; the number in parentheses represents the p-value of the correlation.

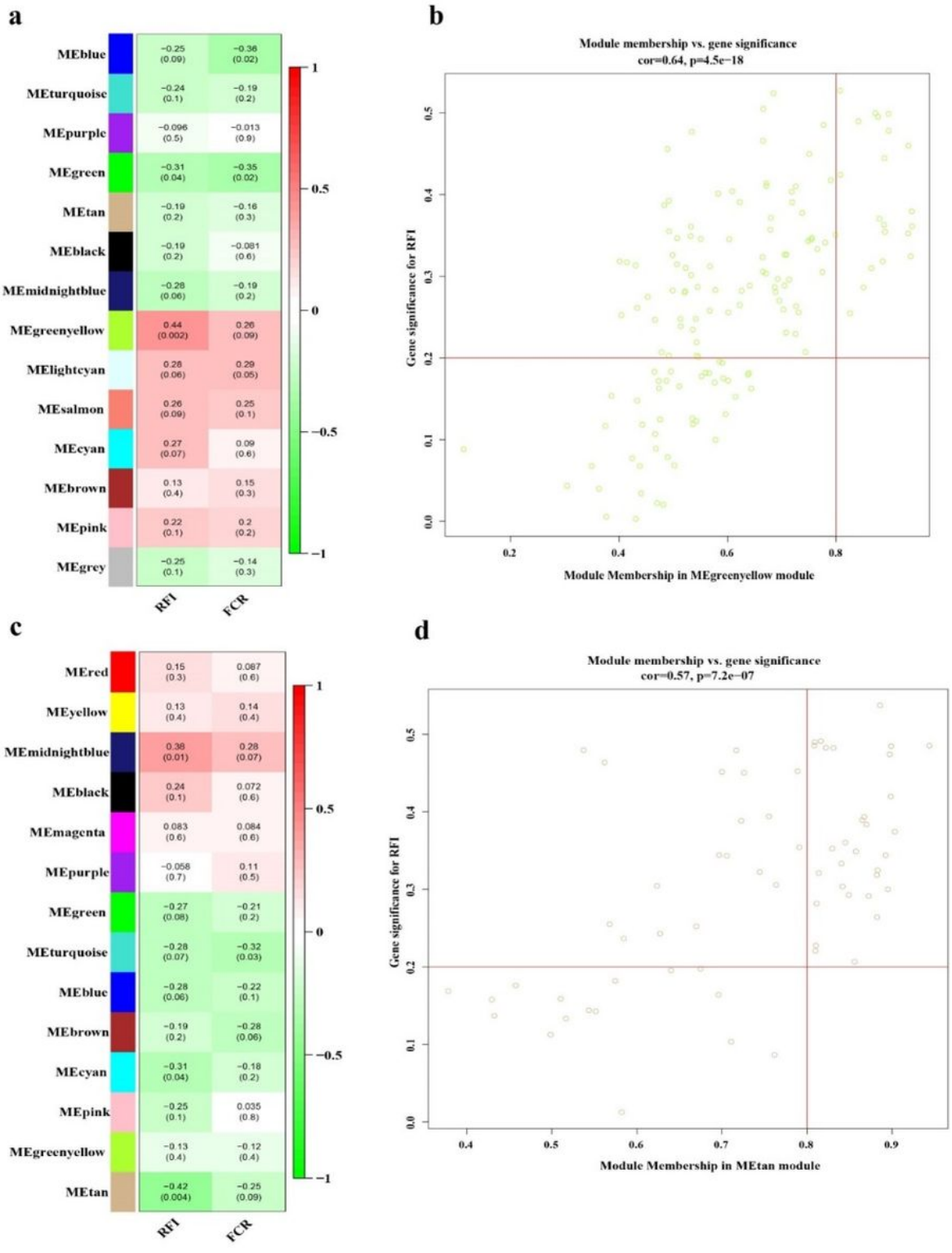

Figure 3

Coexpression network analysis of metabolic features. The left panel of the figure shows the correlation between the module and RFI or FCR in (a) negative and (c) positive models. The right panel of the figure shows the scatter plot of module membership and the gene significance in (b) MEgreenyellow or (d)

MEtan module. Each row corresponds to ME, and each column corresponds to traits; the number in each module represents the Pearson correlation between the module and RFI or FCR; the number in parentheses represents the p-value of the correlation.

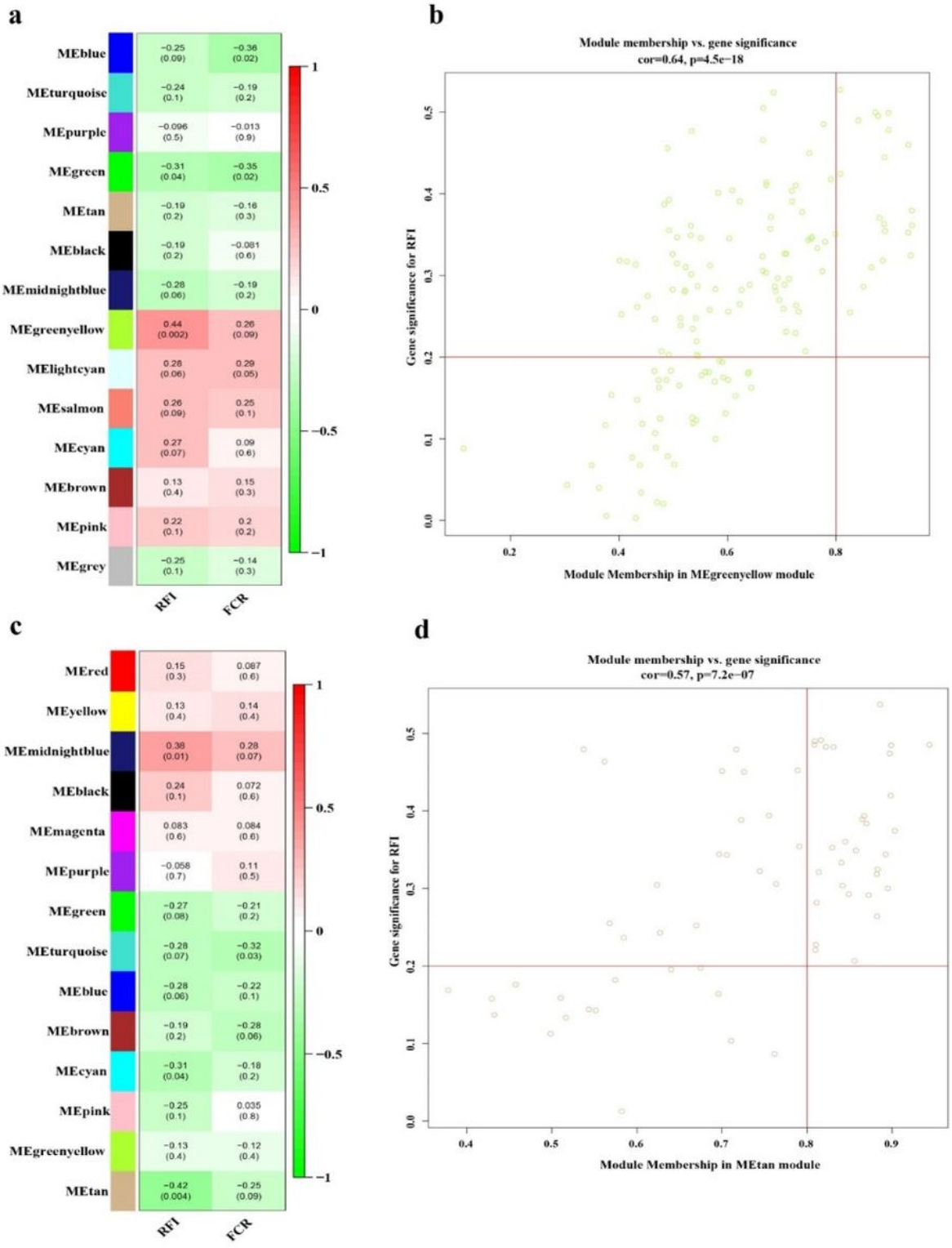

Figure 3

Coexpression network analysis of metabolic features. The left panel of the figure shows the correlation between the module and RFI or FCR in (a) negative and (c) positive models. The right panel of the figure

shows the scatter plot of module membership and the gene significance in (b) MEgreenyellow or (d) MEtan module. Each row corresponds to ME, and each column corresponds to traits; the number in each module represents the Pearson correlation between the module and RFI or FCR; the number in parentheses represents the p-value of the correlation.

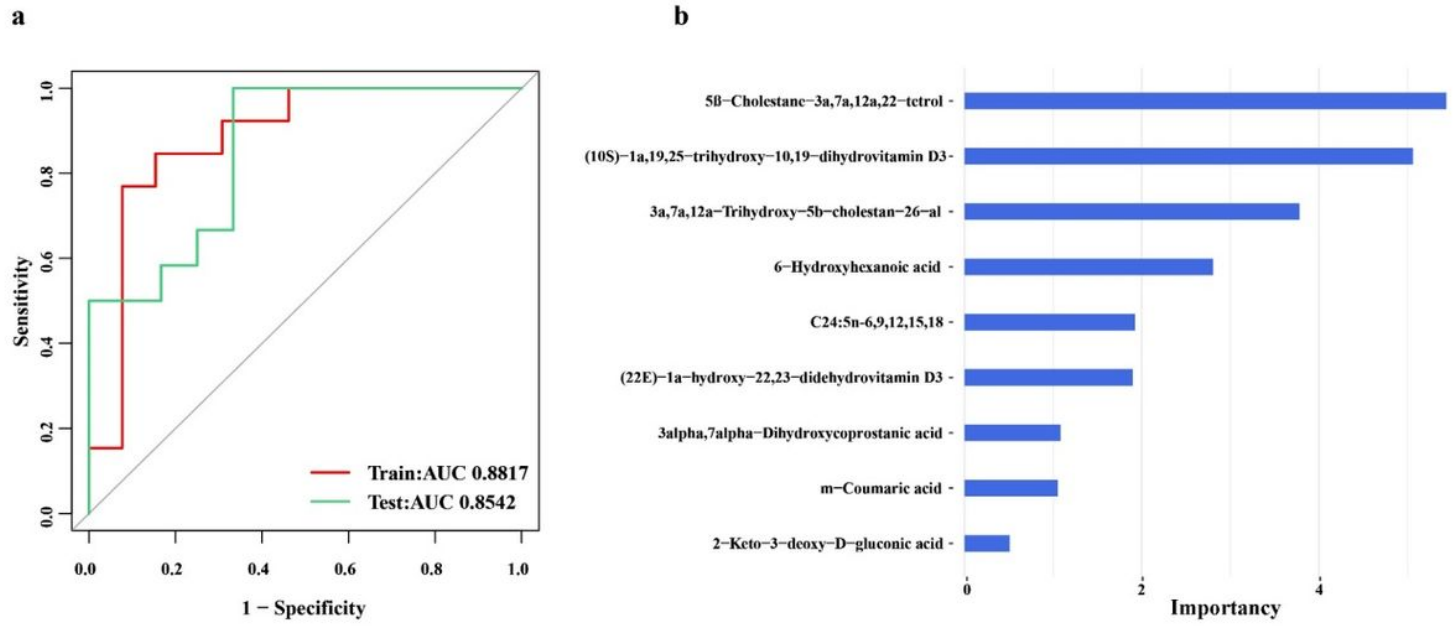

Figure 4

Assessing the weight of nine metabolites using Lasso regression analysis. (a) ROC curve of the training set (red) and the test set (green). (b) Regression coefficients of nine metabolites in the Lasso model. The y-axis of the graph on the right represents metabolites, and the x-axis represents the regression coefficient of metabolites.

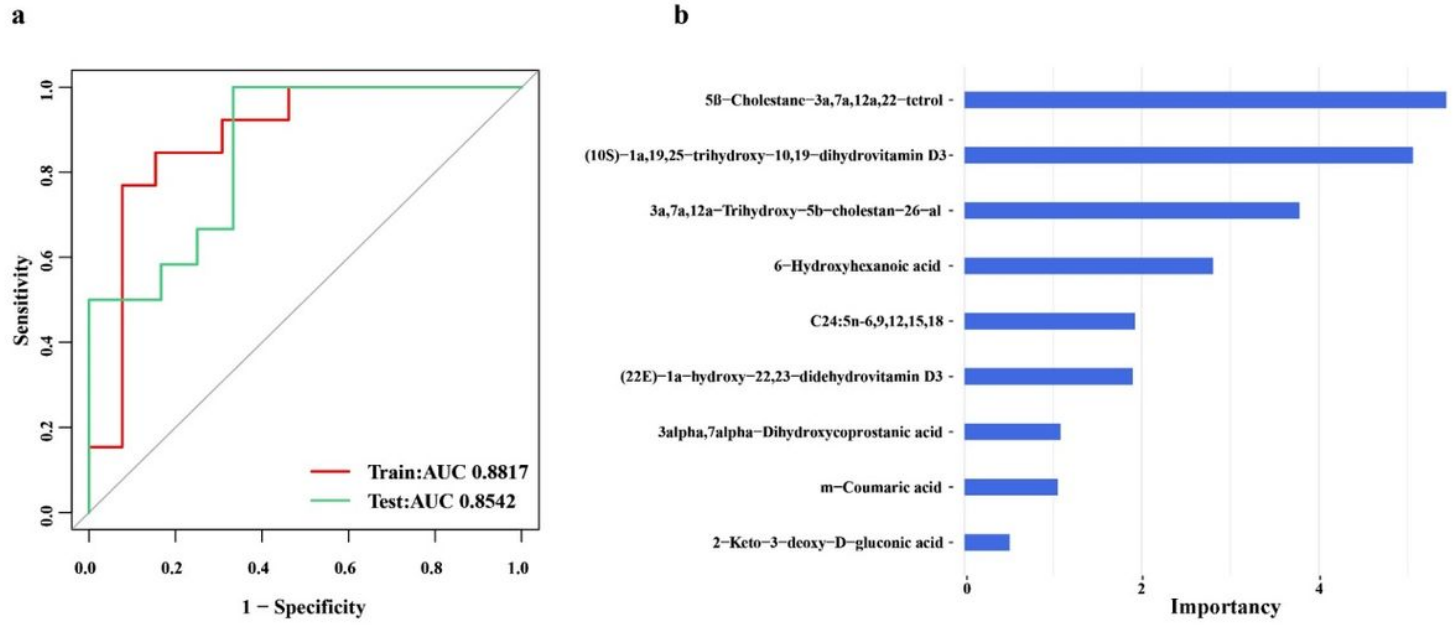

Figure 4

Assessing the weight of nine metabolites using Lasso regression analysis. (a) ROC curve of the training set (red) and the test set (green). (b) Regression coefficients of nine metabolites in the Lasso model. The y-axis of the graph on the right represents metabolites, and the x-axis represents the regression coefficient of metabolites.

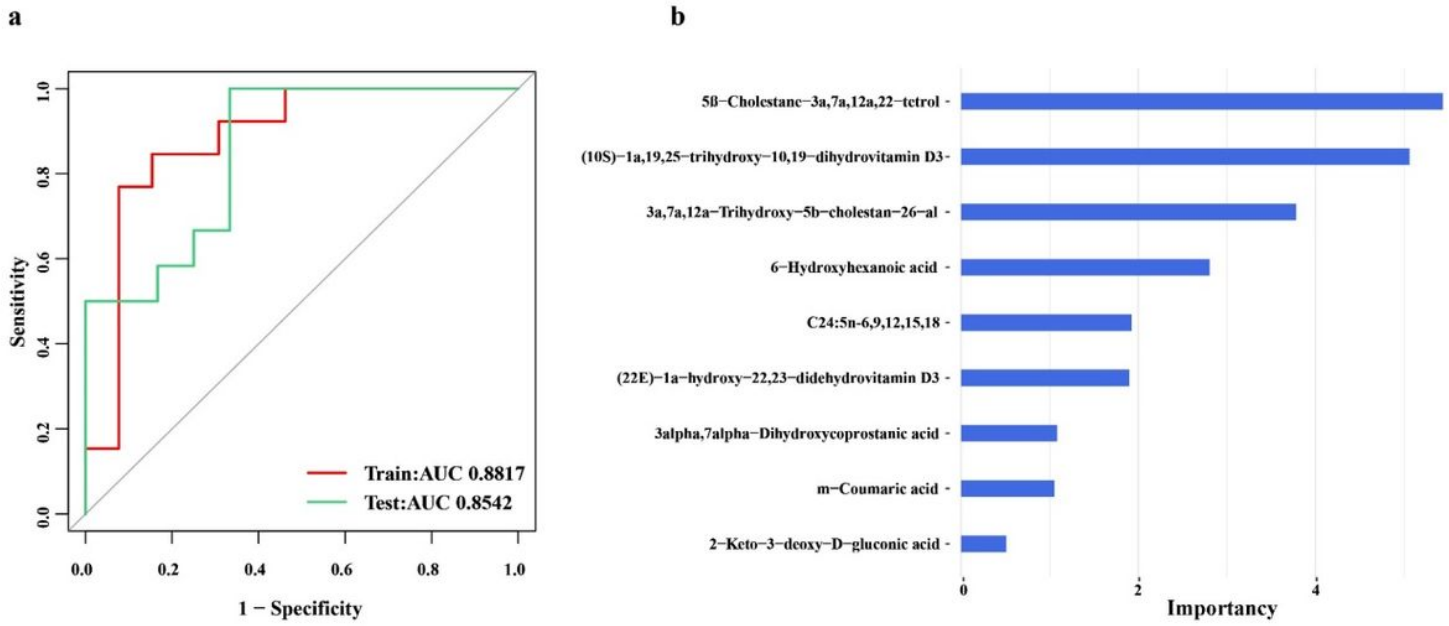

Figure 4

Assessing the weight of nine metabolites using Lasso regression analysis. (a) ROC curve of the training set (red) and the test set (green). (b) Regression coefficients of nine metabolites in the Lasso model. The y-axis of the graph on the right represents metabolites, and the x-axis represents the regression coefficient of metabolites.

Supplementary Files

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

- SupplementaryFigure.docx
- SupplementaryFigure.docx
- SupplementaryFigure.docx
- SupplementaryTableS1.xlsx
- SupplementaryTableS1.xlsx
- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx
- SupplementaryTableS2.xlsx
- SupplementaryTableS2.xlsx

- SupplementaryTableS3.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS4.xlsx