Animal 18 (2024) 101102

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

# Predicting feed efficiency of Angus steers using the gastrointestinal microbiome

M. Congiu [a,b], J. Lourenco [b], A. Cesarani [a,b,*], U. Lamichhane [b], N.P.P. Macciotta [a], C. Dimauro [a]

[a] Dipartimento di Agraria, University of Sassari, Sassari 07100, Italy
[b] Department of Animal and Dairy Science, University of Georgia, Athens 30602, GA, USA

ARTICLE INFO

*Article history:*
Received 9 June 2023
Revised 29 January 2024
Accepted 30 January 2024
Available online 7 February 2024

*Keywords:*
Discriminant Analysis
Microbiota
Multivariate analysis
Residual Feed Intake
Rumen

ABSTRACT

Microbial composition of the gastrointestinal tracts is an important factor affecting the variation in feed efficiency in ruminants. Several studies have investigated the composition of the ruminal and fecal microbiotas, as well as their impacts on feed efficiency and digestion. In addition, next-generation DNA sequencing techniques have allowed us to gain a better understanding of such microbiomes. In this study, the beef cattle microbiome data were analyzed using both a multivariate and a univariate approach and the results were compared. Moreover, a statistical procedure to classify calves in two groups with extreme Residual Feed Intake (**RFI**) values, using their microbiota profile, was developed. Both fecal and ruminal samples were collected from 63 Angus steers at two different time points for evaluation of their microbiomes: at the beginning and at the end of the feedlot. An additional fecal sample was collected at weaning. A total of 149 and 119 bacterial families (**BFs**) were retrieved from the ruminal and fecal samples, respectively. A Canonical Discriminant Analysis (**CDA**) was used to investigate whether BFs were able to distinguish between rumen and fecal samples. A sub-sample of 28 steers was divided in two groups based on their feed efficiency status: positive or negative for RFI. Fecal samples collected at weaning were used to assign the positive and negative RFI animals to their corresponding groups using both Stepwise Discriminant Analysis and CDA. Results revealed that CDA was able to distinguish between rumen and fecal samples. *Peptostreptococcaceae* was the family most associated with the fecal samples, whereas *Prevotellaceae* the most associated with the ruminal samples. The CDA using 19 BFs selected from the stepwise was able to correctly assign all animals to the proper RFI groups (negative or positive). *Rhizobiaceae* was the family most associated with negative RFI, whereas *Comamonadacea* was the family most linked with positive RFI. The results from this study showed that the multivariate approach can be used to improve microbiome data analysis, as well as to predict feed efficiency in beef cattle using information derived from the fecal microbiome.



## Implications

Microbiome composition represents one of the biological factors associated with feed efficiency in ruminants. In recent years, this association has become a major research topic. This study demonstrated how to use a multivariate approach to highlight differences between rumen and fecal samples, as well as to predict feed efficiency. Our results, which need to be verified using both larger samples and different breeding situations, showed that it is possible to predict feed efficiency in steers based on the composition of their fecal microbiomes.

## Introduction

In recent years, the number of studies on livestock microbiome has rapidly increased due to the assessment of relevant associations between microbiota composition and production traits (Lourenco et al., 2019, Krause et al., 2020, Welch et al., 2020, Williamson et al., 2022). Several studies have also investigated the relationship between the microbiota and dietary changes (Correddu et al., 2019; Buffa et al., 2020; Monteiro et al., 2022). Significant technological achievements have contributed to improve the amount and quality of data available to study the microbiota. The development of next-generation DNA sequencing techniques and the wide utilization of metagenomics represents a fundamental step in microbiome investigation: they allow the identification of a larger number of microorganisms compared to previous

* Corresponding author at: Viale Italia 39, Sassari (SS) 07100, Italy.
*E-mail address:* acesarani@uniss.it (A. Cesarani).

https://doi.org/10.1016/j.animal.2024.101102

technologies based on *in vitro* methods (Escobar-Zepeda et al., 2015) and to identify microbes present in complex biological matrices through DNA sequencing (Matthews et al., 2019). More recently, most of the studies investigating the rumen or fecal microbiomes have been based on methods of high-throughput sequencing, particularly of the 16S rRNA gene, which allows the characterization of prokaryotes in complex biological mixtures such as gastrointestinal samples.

Knowing the microbial composition of the gastrointestinal tract is of great importance for the evaluation of feed efficiency and digestion. Many studies have reported a close relationship between feed efficiency and the bovine microbiota (Myer et al., 2015, Welch et al., 2020). Residual feed intake (**RFI**) is a measure of feed efficiency that is defined as the difference between actual and predicted feed intake based on body size and growth (Nkrumah et al., 2006; Arthur and Herd, 2008). Therefore, animals with low RFI, i.e., those that eat less than expected, are considered more efficient in their use of feed than those with high RFI. Selecting animals for feed efficiency could improve the sustainability of livestock farms. However, phenotyping of individual feed intake is expensive and time consuming, so, the identification of efficiency predictors is crucial, especially if an animal feed efficiency potential can be evaluated earlier in life, such as at weaning.

In cattle, differences in both ruminal and fecal microbiotas have been reported based on RFI classification (Shabat et al., 2016, Lourenco et al., 2022a, Zhou et al., 2023). While the collection of ruminal samples requires an invasive method such as fistulation or sampling through esophageal tubing to reach the rumen, fecal samples can be easily obtained from the rectum of the animal, simply by using a sterile glove. Several studies have highlighted the differences between the ruminal and fecal microbial taxonomic profile, in terms of both diversity and composition of bacterial families (**BFs**) (Welch et al., 2020; Lourenco et al., 2020; Zhou et al., 2023). This conclusion, however, was reached using univariate statistical approaches such as Student's *t*-test or ANOVA models, which do not consider the effects of other BFs when a single family is being individually analyzed.

In the present study, a multivariate statistical procedure was exploited to analyze BF differences in the rumen and feces of beef cattle. Unlike the univariate approach, the multivariate one is able, by definition, to capture multiple correlations among variables, the BF in this case, and highlight groups of BF whose presence or abundance are closely linked to each other. Then, a statistical procedure to classify calves into two groups with extreme RFI values, using the fecal microbiota profile, was developed.

## Material and methods

### *Animals*

The present study used animals that are part of an ongoing research in which the purpose is to investigate the performance of steers sired by Angus bulls that were divergently selected for residual average daily gain and marbling scores (Detweiler et al., 2019). For this study, 63 Angus steers, born in spring 2018 and belonging to a line selected for marbling and feed conversion, were enrolled (Lourenco et al., 2022b). All steers were raised together and were part of the same contemporary group. Animals were raised in a farm located in Calhoun (GA, US, 34°30′ N, 84°57′ W) of the University of Georgia's research and education center, and they were weaned at 7.5 months of age. Until weaning, the steers were raised without the addition of grains in their diets, which was composed of milk, forage, and free-choice minerals. The forage comprised of a mixture of Tall fescue (Festuca arundinacea) and Bermudagrass (Cynodon dactylon). The free-choice mineral contained 14% Mg, 12% Ca, 9% NaCl, 6% P, 3% K, 0.62% S, 0.06% F, 5 075 ppm Zn, 3 055 ppm Mn, 2 500 ppm Cu, 920 ppm Fe, 120 ppm I, 46 ppm Co, and 27 ppm Se (Godfrey's Feed, Madison, GA, USA). After weaning, the steers were backgrounded for approximately 5 months in a pasture-based system with the inclusion of some grains and co-products in their diet (3.6 kg/day of a mixture of 50% corn gluten feed and 50% soybean hulls). The steers were then transferred to a feedlot located in Brasstown (NC, US, 35°10′ N, 83°23′ W) at about 13 months of age. Following an adaptation period, feed intake of all steers was measured using a GrowSafe system (GrowSafe Systems Ltd., Calgary, Canada) during a 110-d finishing period. The use of this system allowed individual quantification of feed intake as well as the calculation of a feed efficiency metric, i.e., RFI; which was calculated as previously described by Hoque and Suzuki (2009). This calculation took into account the daily feed intake, metabolic BW at mid-test, average daily gain, and the regression coefficients of steers' daily feed intake on metabolic BW and average daily gain. More specifically, the equation used was:

$$\mathrm{RFIPhe} = \mathrm{FI} - \beta w(\mathrm{phe}) \times \mathrm{MWT} - \beta g(\mathrm{phe}) \times \mathrm{ADG}$$

where RFIphe = phenotypic residual feed intake, FI = daily feed intake, MWT = metabolic BW at mid-test, ADG = average daily gain, and βw(phe) and βg(phe) = partial regression coefficients of animal's FI on MWT and ADG, respectively (Hoque and Suzuki, 2009).

The feedlot-finishing diet was composed of corn (56.20%), dried distillers grains (19.54%), corn silage (8.05%), corn gluten feed (7.08%), vitamin/mineral mix (4.76%), and barley straw (4.37%).

Upon conclusion of their period in the feedlot, the steers were slaughtered at the University of Georgia's Meat Science Technology Center, located in Athens, GA, US (33°57′ N, 83°22′ W). Additional details about the animals, as well as the composition of their diets, are described in Lourenco et al. (2022b).

### *Collection and sample processing*

Samples for microbial evaluations were collected at three different time points: weaning (**WEA**), feedlot in (**FIN**) and feedlot out (**FOUT**). FIN represents the feedlot starting point when animals started receiving a high-grain diet, whereas FOUT represents the end of the feedlot period. At WEA, only fecal samples (28) were collected, whereas both ruminal and fecal samples (62 and 79, respectively) were collected in FIN (77) and FOUT (64). The ruminal contents were sampled by using an esophageal tubing introduced through the animal's mouth. Using a vacuum pump, about 300 ml of ruminal fluid was sampled from each animal and approximately 45 ml was transferred into a sterile conical tube (Lourenco et al., 2019; Lourenco et al., 2020). The fecal samples were collected from the rectum by digital palpation, and approximately 50 g of feces was put into a sterile conical tube for each animal (Welch et al., 2021). The gastrointestinal tract samples were immediately frozen by immersion in liquid nitrogen and subsequently stored at −80°C.

### *Microbial DNA extraction and taxonomic analysis*

All the procedures carried out to obtain the microbial compositions of the samples have been described in Lourenco et al. (2022b). In summary, the DNA extraction was performed following the methodology proposed by Rothrock et al (2014), with some minor modifications. The starting amount of sample was 350 mg, which was placed in a bead-beating tube along with 1 mL of InhibitEX Buffer and homogenized in a FastPrep-24 instrument for 40 seconds at 6.0 m/second. Next, tubes were incubated at 95°C for 5 minutes to maximize DNA recovery from Gram-positive organisms. After centrifugation, 200 µL of supernatant was mixed

with 15 μL of Proteinase K for the digestion of proteins (including DNases and RNases). Purification steps were concluded by centrifuging samples through a QIAamp spin column with the addition of 200 Proof ethanol, buffer AW1, and buffer AW2. Purified DNA was eluted in buffer ATE and checked by spectrophotometry using a Synergy H4 Hybrid Multi-Mode Microplate Reader (Agilent, Santa Clara, CA, USA). Samples with concentrations lower than 10 ng/μL, and/or A260/A280 ratios out of the 1.7–1.9 range were disqualified, and the DNA extraction process was repeated. The genomic DNA was analyzed at the Georgia Genomic and Bioinformatic Core (https://dna.uga.edu/) following the methodology described by Akerele et al. (2022), as follows. The V3-V4 region of the 16S rRNA gene was amplified using the primer pair S-D-Bact-0341-b-S-17 (CCTACGGGNGGCWGCAG); and S-D-Bact-0785-a-A-21 (GACTACHVGGGTATCTAATCC) (Klindworth et al., 2013). Each PCR reaction contained DNA, forward and reverse primers, and 2x Kapa HiFi Hotstart readyMix (Roche Diagnostics Corporation, Indianapolis, IN, USA). The DNA was amplified by 25 cycles of denaturation, annealing, and extension, followed by a cleaning step and another eight cycles of amplification to incorporate indexes (Illumina Nextera XT, Illumina, Inc., San Diego, CA, USA). All PCR products were cleaned using AMPure XP beads and ethanol. The libraries were pooled and sequenced on a MiSeq platform (Illumina, Inc., San Diego, CA, USA).

The resulting FASTQ files were demultiplexed and processed using QIIME 2 version 2021.11 (Bolyen et al., 2019). The DADA2 plugin (Callahan et al., 2016) was used to remove the primers, join sequences, denoise, and filter chimeras. Taxonomic classification was performed using a pretrained Naïve Bayes classifier (Pedregosa et al., 2011). The SILVA 138 SSU database (Quast et al., 2012) was used as the reference database. Amplicon sequence variants were classified at several taxonomic levels, and relative abundances at the family level were used in the current study. Although less biologically informative than the genus and species level, the family level was used in this study to evaluate if our models would still generate meaningful results even with less-than-ideal microbial data.

## Statistical analysis

### *Bacterial family differences in feces and rumen during the feedlot*

Since rumen samples were not collected at WEA; microbiome data obtained at this timepoint were excluded. Considering only the feedlot period, a total of 119 and 149 BFs were retrieved from the fecal and ruminal samples, respectively. Only the 100 BFs in common to the two compartments were used in the statistical analysis. In the univariate approach, an ANOVA was carried out to ascertain which BF were significantly different in rumen and feces using the following model:

$$y = \mu + \text{time} + \text{type} + \text{time} * \text{type} + \varepsilon \quad (1)$$

where y was one of the BFs, μ was the overall mean, time was the time point when BFs were quantified (FIN and FOUT), type was the gastrointestinal compartment where BFs were obtained (feces and rumen), time*type was the interaction factor and ε the random residuals. A BF was considered significantly associated with the type (i.e., feces or rumen) according to the Bonferroni correction for multiple tests ($P < 0.0005$).

In the multivariate approach, the canonical discriminant analysis (**CDA**) (Mardia and Jupp, 2000) was applied to test whether BFs were able to separate microbiome samples collected in rumen from those collected in feces. CDA is a multivariate statistical technique aimed to a) highlight differences among groups of individuals and b) improve understanding of the relationships among the involved variables. If p indicates the number of groups, the CDA derives p−1 equations (**CAN**) that are linear combinations of the original variables ($X_n$). In this study, since we had two groups (rumen and feces BF), only one CAN was obtained, which structure was:

$$CAN = C_1X_1 + C_2X_2 + \ldots\ldots + C_nX_n \quad (2)$$

where $C_i$ are the canonical coefficients (**CCs**) that provide information about the contribution of each $X_i$ to the CAN. To better decipher the role of an $X_i$ in the separation of groups, Rencher and Scott (1990) standardized all the CC values considering the multiple correlations among the original variables ($X_n$). The standardized CC and their corresponding absolute values can be used to characterize the CAN and to rank the variables according to their contribution to the function. On the other hand, the correlation between each original variable ($X_n$) and the newly extracted CAN can help to biologically interpret the CAN. However, the interpretation of this correlation is very different from that of standardized CC. As a CC provides information about the contribution of each $X_i$ to the CAN in the presence of the other variables (i.e., in a multivariate manner), the correlation between the original variable and the newly extracted CAN yield only a univariate information about the importance of each $X_i$ for the CAN, independent of the other variables. The significance of the separation between the two groups was tested by using Hotelling's T-square test (De Maesschalck et al., 2000). The CAN was then exploited to predict the group to which an individual belonged to. In practice, the CAN is applied to each individual producing a discriminant score. In our study, an animal was assigned to rumen or feces BF if its discriminant score was lower than the cutoff value obtained by calculating the weighted mean distance between the centroids of the two groups (Mardia and Jupp, 2000). Finally, the stepwise discriminant analysis was used to select the minimum number of BFs able to significantly separate feces and rumen samples and to correctly assign them to groups.

### *Fecal bacterial family differences at weaning between positive and negative residual feed intake classes*

BF data collected from feces at WEA were extracted from the complete dataset and the corresponding animals were divided into two groups according to their estimated RFI values at the feedlot: those with positive RFI (**PRFI**) and those with negative RFI (**NRFI**). An ANOVA model was applied to the data to ascertain which BFs were significantly different between PRFI and NRFI. The model was:

$$y = \mu + G + \varepsilon$$

where y was one of the fecal BF at weaning, μ was the overall mean, G was the fixed effect of the RFI group estimated at the feedlot (PRFI and NRFI), and ε was the random residual. A BF was considered significantly associated with the RFI class according to the Bonferroni correction ($P < 0.0005$).

The stepwise discriminant procedure was then applied to select among variables the minimum number of BFs able to significantly separate PRFI from NRFI. The obtained BFs were then submitted to CDA to test if the two RFI groups were significantly separated by the selected families and if the CAN was able to correctly assign animals to the true group of origin. Finally, to evaluate the ability of the procedure in correctly assigning new observations to one of the two groups, the leave-one-out cross-validation approach was adopted. Briefly, the dataset was split into a training set and a testing set, using all but one observation as part of the training set. The observation left from the dataset was used as a testing set. So, CDA was applied to the training set and the obtained CAN was used to assign the observation in the test set. The entire procedure was repeated until all observations were used in the test dataset.

## Results

### *Bacterial family differences in rumen and feces during the feedlot*

After applying all the quality-filtering steps, a total of 6 949 amplicon sequence variants were detected in the ruminal samples, with an average frequency per sample of 28 436 sequences. On average, each amplicon sequence variant appeared 393 times in the ruminal samples. In the feces, a total of 6 773 amplicon sequence variants were observed, with an average frequency per sample of 18 173, and each amplicon sequence variant was observed 287 times on average within the fecal samples.

In the ANOVA model, the difference in sampling time (FIN and FOUT) was significant only for four BFs (Table 1), whereas 35 families, whose least squares means and *P*-values are listed in Table 2, were significantly different between feces and rumen. No significant interaction between time and type was detected.

The CDA carried out on 100 BFs, separated significantly ($P < 0.0001$) rumen and feces samples. The univariate model, i.e., ANOVA, and the univariate information from the CDA (i.e., the correlation between the original variable and the CDA) led to very similar results. In fact, the 35 BFs selected by the ANOVA model showed the largest correlation values in the CDA analysis (Table 2). Moreover, the ranks of the BF according to the *P*-values from ANOVA and according to the correlation from CDA were almost the same. Fig. 1 shows the rank differences between CDA and ANOVA: 84% of the 35 BFs in CDA had the same or one difference in position compared to the rank of the ANOVA. Four BFs differed for two positions, and the remaining two BFs were differently ranked by three positions.

As expected, when the 35 BFs were ordered according to their standardized CC absolute value (instead of according to the correlation values), no close concordance between ANOVA and CDA ranks was observed. The stepwise discriminant analysis selected 19 BFs that were able to significantly ($P < 0.0001$) separate the two groups (i.e., feces and rumen samples). Over the 19 BFs, 13 were also significant in the ANOVA analysis (highlighted in bold in Table 3). BFs with negative values are more associated with fecal samples, whereas the BFs with positive values are mostly associated with rumen samples. *Peptostreptococcaceae* was the family that had the strongest association with the fecal samples, whereas *Prevotellaceae* was the BF that had the strongest association with the ruminal samples. These 2 BFs were also the first 2 selected by ANOVA.

Fig. 2 displays how microbiota samples are separated by the CDA developed using the 19 BFs as variables: fecal samples exhibited negative and rumen samples positive values, respectively.

### *Residual feed intake class discrimination with bacterial families in feces collected at weaning*

The two RFI groups, whose feces were collected at weaning, contained 14 animals each. The mean and the SD of the RFI values were −1.82 ± 0.99 for NRFI and 1.57 ± 0.73 for PNRFI. The ANOVA model, used to test which BF were different between negative and positive RFI groups, did not highlight any significant BF. However, the stepwise discriminant analysis selected 18 BFs (Table 4) that were able to significantly separate ($P < 0.0001$) the two groups in a new run of CDA. Moreover, the CDA developed by using the 18 selected BFs correctly assigned all animals to the proper group (i.e., NRFI and PRFI) in the leave-one-out procedure.

The Euclidean distance between the two group centroids was 88 ($P < 0.0001$), thus indicating a good separation between groups, as confirmed by Fig. 3 where the scores of involved individuals, in the space of the new CAN, are showed.

According to Table 4 and Fig. 3, BFs with negative CC values were more abundant in animals with NRFI. On the contrary, BF with positive CC prevailed in PRFI. In particular, *Rhizobiaceae* was the family most associated with the NRFI group, whereas *Comamonadaceae* was the family most associated with the PRFI group (Table 4), respectively.

**Table 1**
ANOVA least squares means and *P*-values for the differences of bacterial families collected at the feedlot starting point (FIN) and at the end of the feedlot period (FOUT) in Angus steers.

| Bacterial family | ANOVA | | |
|---|---|---|---|
| | FIN[1] | FOUT[1] | *P*-value |
| Bacillaceae | 0.098 | 0.016 | <0.001 |
| Planococcaceae | 0.027 | 0.002 | <0.001 |
| Pseudonocardiaceae | 0.086 | 0.028 | <0.001 |
| Thermoactinomycetaceae | 0.014 | 0.061 | <0.001 |

[1] Percentage abundance.

## Discussion

Differences between fecal and ruminal BF were analyzed both with univariate and multivariate statistical techniques. In the univariate approach, ANOVA tested if the abundance of a single BF was significantly different in rumen and feces, regardless of the presence of other families. Among the approximately 100 BFs analyzed, a total of 35 (Table 2) was significantly different between the two compartments. On the other hand, the CDA was able to successfully differentiate the ruminal sample from the fecal ones. Moreover, the CDA, although it is a multivariate technique, also produces some univariate results as the correlation between the single original variable (the BF) and the CAN. As displayed in Table 2, ANOVA and correlation between original variables and CAN gave very similar results because the lower ANOVA *P*-values were associated with the higher the correlation absolute values. This trend is confirmed by Fig. 1 that displays the number of BFs which ranking differed between ANOVA and CDA. Only 6 of 35 BFs had a different rank of 2 or 3 positions. The family *Prevotellaceae* had the largest positive (0.75) correlation with the extracted CAN. Welch et al. (2020) found that *Prevotellaceae* was the most abundant family in the rumen of Angus steers, regardless of their feed efficiency classification; however, this family had only a minor presence in the steers' hindgut (cecum and rectum). In a study involving creep-fed beef cattle, Lourenco et al. (2019) found a positive association between *Prevotella* and average daily gain. Seidel et al. (2022) observed that *Prevotellaceae* was the family with the greatest overall abundance in the rumen fluid of goats. Therefore, there is much evidence that the family *Prevotellaceae* plays a major biological role in the rumen, which may be explained by its broad substrate utilization capabilities (Seidel et al., 2022), and thus, this is one of the most typical bacterial family of the ruminal environment.

The *Peptostreptococcaceae* had the greatest negative correlation with the CAN and was the family more closely associated with the fecal samples. Welch et al. (2020) found that this family was the third most abundant in the fecal material of Angus steers collected during slaughter/evisceration process, with an average abundance varying between 11.7 and 12.7%; however, Peptostreptococcaceae was not among the top 10 families found in the rumen, indicating the strong association with the fecal material. Lourenco et al. (2022a) reported the average relative abundances of this family to be approximately 12% in the feces of feedlot steers during the feedlot phase, making it the third most abundant family in their fecal material. Moreover, the authors mentioned that many fermenters of structural carbohydrates prefer ammonia as their

**Table 2**
ANOVA least squares means and *P*-values for 35 bacterial families[1] significantly different in fecal and ruminal samples[2] and correlations between the single bacterial family and the canonical function from the canonical discriminant analysis (CDA) carried out on Angus steers.

| Bacterial family | ANOVA | | | CDA |
|---|---|---|---|---|
| | Feces[3] | Rumen[3] | *P*-value | Correlation |
| Prevotellaceae | 10.553 | 40.125 | 1.2E-26 | 0.748 |
| Peptostreptococcaceae | 9.355 | 0.008 | 1.8E-19 | −0.677 |
| Lachnospiraceae | 18.633 | 7.068 | 3.3E-17 | −0.645 |
| Erysipelotrichaceae | 2.516 | 0.157 | 6.4E-17 | −0.641 |
| Bacteroidales_UCG_001 | 0.000 | 0.123 | 2.5E-14 | 0.590 |
| Saccharimonadaceae | 0.128 | 0.994 | 2.6E-14 | 0.588 |
| Clostridiaceae | 5.419 | 0.039 | 3.4E-14 | −0.594 |
| Tannerellaceae | 0.918 | 0.000 | 3.6E-14 | −0.542 |
| Bacteroidales_RF16_group | 0.081 | 2.731 | 4.9E-14 | 0.588 |
| Hungateiclostridiaceae | 0.038 | 0.224 | 3.9E-13 | 0.566 |
| Gastranaerophilales | 0.036 | 0.302 | 1.7E-10 | 0.496 |
| Fibrobacteraceae | 0.000 | 0.105 | 2.3E-10 | 0.503 |
| F082 | 0.003 | 2.419 | 2.1E-09 | 0.469 |
| Bacteroidaceae | 2.539 | 0.005 | 1.1E-08 | −0.434 |
| Oscillospiraceae | 11.333 | 5.582 | 2.5E-08 | −0.463 |
| WCHB1_41 | 0.000 | 0.165 | 4.0E-08 | 0.449 |
| Monoglobaceae | 0.471 | 0.063 | 6.3E-08 | −0.430 |
| Pirellulaceae | 0.001 | 0.708 | 6.5E-08 | 0.442 |
| Butyricicoccaceae | 0.202 | 0.012 | 3.4E-07 | −0.424 |
| Peptococcaceae | 0.227 | 0.012 | 3.5E-07 | −0.399 |
| VadinBE97 | 0.000 | 0.100 | 3.9E-07 | 0.417 |
| Enterobacteriaceae | 0.514 | 0.011 | 4.4E-07 | −0.404 |
| Christensenellaceae | 1.046 | 2.082 | 7.9E-07 | 0.411 |
| Eubacterium coprostanoligenes group | 2.390 | 5.073 | 8.9E-07 | 0.396 |
| Desulfovibrionaceae | 0.000 | 0.022 | 3.4E-06 | 0.383 |
| Moraxellaceae | 0.001 | 0.062 | 4.1E-06 | 0.380 |
| Defluviitaleaceae | 0.026 | 0.075 | 5.6E-06 | 0.356 |
| Weeksellaceae | 0.000 | 0.011 | 3.6E-05 | 0.345 |
| Synergistaceae | 0.000 | 0.021 | 4.4E-05 | 0.342 |
| Eubacteriaceae | 0.002 | 0.013 | 9.9E-05 | 0.319 |
| Streptococcaceae | 1.260 | 0.301 | 2.1E-04 | −0.326 |
| Lactobacillaceae | 3.260 | 0.882 | 2.6E-04 | −0.316 |
| Mycoplasmataceae | 0.003 | 0.161 | 2.8E-04 | 0.298 |
| Paracaedibacteraceae | 0.000 | 0.006 | 3.1E-04 | 0.300 |
| Clostridium methylpentosum_group | 0.087 | 0.304 | 4.2E-04 | 0.288 |

[1] 100 bacterial families commonly detected in rumen and feces.
[2] 79 fecal and 62 ruminal samples.
[3] Percentage abundance.

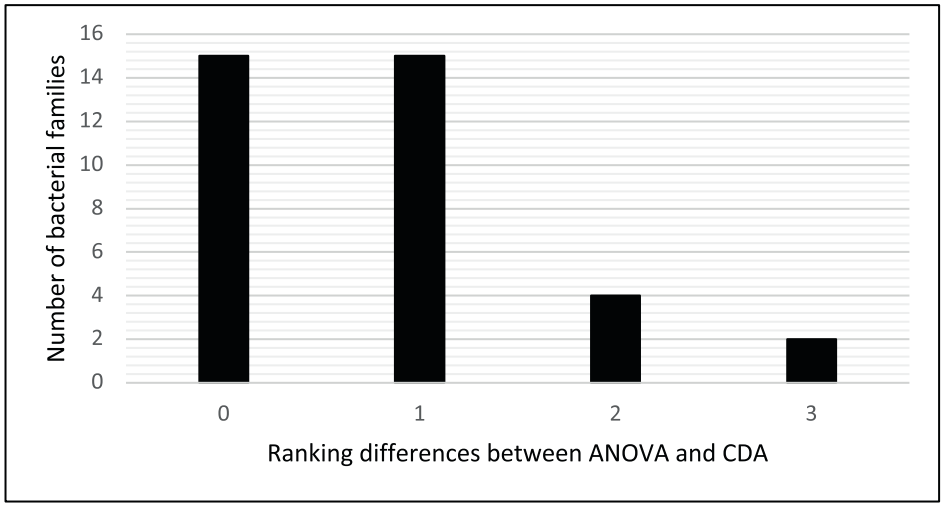


**Fig. 1.** Number of bacterial families whose ranking differed between ANOVA and Canonical Discriminant Analysis (CDA) carried out on Angus steers.

source of nitrogen, and members of Peptostreptococcaceae have the ability of producing large amounts of ammonia from different nitrogen sources, allowing this family to thrive in the large intestine. Therefore, our study, which found the Peptostreptococcaceae as the BF more strongly associated with the fecal samples, confirmed previous reports. The stepwise was applied to data to obtain the minimum number of BFs able to discriminate between fecal and ruminal samples. This procedure selected the 19 BFs listed in Table 3 with their standardized CC values. Using these families as predictors, the microbiota samples were correctly assigned to

**Table 3**
List of the 19 bacterial families, with their absolute standardized canonical coefficient (CC) and relative abundances, able to discriminate between fecal and ruminal samples collected from Angus steers.

| Bacterial families | | Relative abundances | |
|---|---|---|---|
| | Standardized CC | Feces | Rumen |
| Prevotellaceae[1] | 1.360 | 10.45 | 40.13 |
| Veillonellaceae | 0.783 | 0.03 | 0.27 |
| Bacteroidales_RF16_group[1] | 0.588 | 0.08 | 2.73 |
| Christensenellaceae[1] | 0.561 | 1.04 | 2.08 |
| Saccharimonadaceae[1] | 0.530 | 0.12 | 0.99 |
| Eubacterium coprostanoligenes group[1] | 0.519 | 2.41 | 5.07 |
| Fibrobacteraceae[1] | 0.440 | 0.00 | 0.10 |
| F082[1] | 0.341 | 0.00 | 2.42 |
| Defluviitaleaceae[1] | −0.017 | 0.03 | 0.07 |
| Erysipelatoclostridiaceae | −0.207 | 0.46 | 0.25 |
| Butyricicoccaceae[1] | −0.340 | 0.20 | 0.01 |
| Selenomonadaceae | −0.382 | 0.53 | 0.40 |
| Monoglobaceae[1] | −0.384 | 0.46 | 0.06 |
| Spirochaetaceae | −0.484 | 4.14 | 2.23 |
| Tannerellaceae[1] | −0.556 | 0.87 | 0.00 |
| Succinivibrionaceae | −0.585 | 2.89 | 1.60 |
| Lachnospiraceae[1] | −0.732 | 18.77 | 7.07 |
| Acidaminococcaceae | −0.733 | 3.90 | 1.68 |
| Peptostreptococcaceae[1] | −1.445 | 9.58 | 0.01 |

[1] Families significant also in the ANOVA. Bacterial families with negative values are more associated with fecal samples, whereas bacterial families with positive values are mostly associated with rumen samples.

**Table 4**
List of the 18 bacterial families, with their absolute standardized canonical coefficient (CC) and relative abundances, able to discriminate between Angus steers with negative (NRFI) and positive (PRFI) residual feed intake.

| Bacterial Families | | Relative abundances | |
|---|---|---|---|
| | Standardized CC | NRFI | PRFI |
| Rhizobiaceae | −403 | 0.01 | 0.00 |
| Erysipelatoclostridiaceae | −244 | 0.05 | 0.09 |
| Acidaminococcaceae | −170 | 1.04 | 0.31 |
| p_2534_18B5_gut_group | −161 | 0.70 | 0.09 |
| Fibrobacteraceae | −108 | 0.05 | 0.00 |
| Beijerinckiaceae | −107 | 0.02 | 0.01 |
| Lachnospiraceae | −61 | 6.64 | 10.78 |
| Enterobacteriaceae | −49 | 0.2 | 0.67 |
| Prevotellaceae | −34 | 8.92 | 5.53 |
| WCHB1_41 | −31 | 0.09 | 0.05 |
| Moraxellaceae | −25 | 0.01 | 0.02 |
| Campylobacteraceae | −25 | 0.28 | 0.07 |
| Oscillospiraceae | 31 | 17.17 | 20.05 |
| Acholeplasmataceae | 41 | 0.03 | 0.05 |
| RF39 | 61 | 0.02 | 0.09 |
| Succinivibrionaceae | 206 | 0.00 | 0.01 |
| Atopobiaceae | 294 | 0.05 | 0.19 |
| Comamonadaceae | 340 | 0.00 | 0.02 |

the feces or rumen (Fig. 2). Observing both Table 3 and Fig. 2, we can conclude that BF with negative CC values are associated with fecal samples, whereas those with positive CC are associated with rumen samples. However, among the 19 BFs, 13 were also significant in the univariate approach (in bold in Table 3), whereas the remaining six families had abundances not significantly different in the two compartments. As an example, the abundance of the *Veillonellaceae* family is not different between feces and rumen in the univariate approach; however, its CC is the second highest value among BFs able to separate rumen from feces samples. So, regardless of their abundance, our results suggest that, together, the 9 BFs with positive values are important in the rumen, whereas the remaining 11 BFs with negative values are relevant in the feces.

Besides being able to differentiate between ruminal and fecal samples, the classification models can be more useful whether they are able to differentiate and predict animal performance traits, such as feed efficiency. Differences in BF between animals with low and high RFI, both in rumen and feces, were found by Zhou et al. (2023) analyzing the microbiota of 10 Qinchuan cattle (5 with low and 5 with high RFI). The study by Lourenco et al. (2022a), which investigated the fecal microbiome of two groups of steers with low and high RFI, suggested that differences in feed efficiency might be at least partially ascribed to their intestinal microbial population, as several BFs were found to have different abundances in the two RFI groups. In the present study, the ANOVA model did not find any significant BF in the NRFI and PRFI groups. However, by using the stepwise, 18 highly discriminant BFs were identified and the subsequent CDA, analyzing their correlation structure, derived an equation (the CAN) that was able to highlight the differences among groups. The 18 BFs, which were not significant when they were individually studied with the AVOVA model, were able to clearly separate the two RFI groups when they were analyzed with the CDA which, being a multivariate technique, processes the involved variables simultaneously. Thus, the CDA approach has proven to be a valuable tool in evaluating one of the most critical points in animal production – animal feeding. Feeding cost is the single item that has the greatest impact on beef cattle operations (Mulliniks et al., 2020; Greenwood, 2021). Consequently, models that precisely assign animals to their correct feed efficiency groups based on fecal microbiome information can be very useful. The precise measurement of individual feed intake at population level is hampered by logistics and costs. Automatic feeding systems, i.e., mangers that recognize each animal and weigh their feed intake, are becoming available in experimental but also commercial farms (Romanzin et al., 2021; Ledda et al., 2023). The availability of these systems coupled with the microbiome data could help to increase the profitability of the farms. Thus, the prediction of the feed efficiency from the microbiota sampled from the feces could represent a valid, fast, and cheaper alternative to the individual mangers. The cost of DNA sequencing has been decreasing very rapidly in the last two decades, and even defying Moore's law

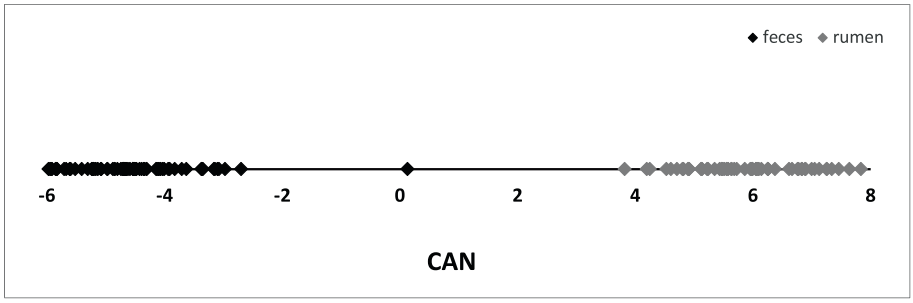


**Fig. 2.** Plot of the individual scores produced by the canonical discriminant equation (CAN) for fecal (black) or rumen (gray) samples collected from Angus steers.

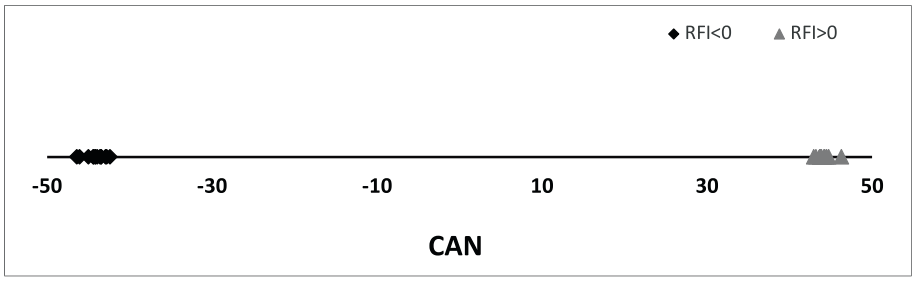


**Fig. 3.** Plot of the individual scores produced by the canonical discriminant equation (CAN) of Angus steers with negative (black) or positive (gray) residual feed intake (RFI) values.

(Lourenco and Welch, 2022c). As new sequencing technologies emerge, this cost is likely going to continue to decrease in the foreseeable future. Therefore, obtaining microbiome information from cattle fecal samples is likely going to become more affordable, which should encourage more producers to obtain it, especially if the microbiome information can provide insights into valuable traits such as feed efficiency. Moreover, if the microbiome test can be performed at earlier stages of the production cycle (e.g., at weaning, or even a few months before animals are slaughtered), this information becomes even more valuable since it can be used to change management decisions by producers. In addition, as the manipulation of microbiomes becomes a reality, producers can potentially alter unfavorable microbiomes of a certain group of animals through prebiotics, probiotics, and other products, and possibly change the fate of some animals by changing their gastrointestinal microbiomes.

Overall, CDA models can predict outcomes with a high level of accuracy, implying that the model successfully identified the most important variables in the dataset in order to make those accurate predictions. As described here, our CDA model using 18 BFs found in the feces of Angus steers did exactly that: it accurately classified all the steers as feed efficient and inefficient based on those families using the leave-one-out procedure. If these results are confirmed using larger samples in different breeding situations, given that (1) fecal samples are easy to obtain; and (2) microbiome evaluations are becoming more common and more affordable, this approach may have important practical implications and help cattle producers make management decisions.

## Conclusions

Results of the present study demonstrated that a multivariate approach can be used to select animals based on their residual feed intake and suggested the possibility to predict individual feed efficiency from the fecal microbiome. The recommended approach should be validated on a larger and independent dataset.

## Ethics approval

All procedures with live animals were approved by the University of Georgia's office of Animal Care and Use (Animal Use Protocol Number A2012 11-006-R1).

## Data and model availability statement

None of the data were deposited in an official repository. Data supporting this study could be available upon reasonable request from J. Lourenco (jefao@uga.edu).

## 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

**M. Congiu:** https://orcid.org/0009-0000-0351-9658.
**J. Lourenco:** https://orcid.org/0000-0003-1479-0528.
**A. Cesarani**: https://orcid.org/0000-0003-4637-8669.
**U. Lamichhane**: https://orcid.org/0000-0002-7984-3291.
**N.P.P. Macciotta**: https://orcid.org/0000-0001-5504-9459.
**C. Dimauro:** https://orcid.org/0000-0002-6588-923X.

## CRediT authorship contribution statement

**M. Congiu:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **J. Lourenco:** Writing – review & editing, Supervision, Project administration, Data curation. **A. Cesarani:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **U. Lamichhane:** Writing – review & editing, Data curation. **N.P.P. Macciotta:** Writing – review & editing, Supervision, Funding acquisition. **C. Dimauro:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization.

## Declaration of interest

None.

## Acknowledgements

None.

## Financial support statement

This research was funded by the grant "Fondo di Ateneo per la Ricerca 2020 – Una tantum per la Ricerca – Dimauro" and the research contract between the University of Georgia and the University of Sassari.

## References

Akerele, G., Al Hakeem, W.G., Lourenco, J., Selvaraj, R.K., 2022. The effect of necrotic enteritis challenge on production performance, cecal microbiome, and cecal tonsil transcriptome in broilers. Pathogens 11, 839.
Arthur, J.P., Herd, R.M., 2008. Residual feed intake in beef cattle. Revista Brasileira de Zootecnia 37, 269–279.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37, 852–857.
Buffa, G., Mangia, N.P., Cesarani, A., Licastro, D., Sorbolini, S., Pulina, G., Nudda, A., 2020. Agroindustrial by-products from tomato, grape and myrtle given at low dosage to lactating dairy ewes: effects on rumen parameters and microbiota. Italian Journal of Animal Science 19, 1462–1471.
Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13, 581–583.
Correddu, F., Fancello, F., Chessa, L., Atzori, A.S., Pulina, G., Nudda, A., 2019. Effects of supplementation with exhausted myrtle berries on rumen function of dairy sheep. Small Ruminant Research 170, 51–61.
De Maesschalck, R., Jouan-Rimbaud, D., Massart, D.L., 2000. The mahalanobis distance. Chemometrics and Intelligent Laboratory Systems 50, 1–18.
Detweiler, R.A., Dean Pringle, T., Rekaya, R., Wells, J.B., Segers, J.R., 2019. The impact of selection using residual average daily gain and marbling EPDs on growth, performance, and carcass traits in Angus steers. Journal of Animal Science 97, 2450–2459.
Escobar-Zepeda, A., Vera-Ponce de León, A., Sanchez-Flores, A., 2015. The road to metagenomics: from microbiology to DNA sequencing technologies and bioinformatics. Frontiers in Genetics 6, 348.
Greenwood, P.L., 2021. An overview of beef production from pasture and feedlot globally, as demand for beef and the need for sustainable practices increase. Animal 15, 100295.
Hoque, M.A., Suzuki, K., 2009. Genetics of residual feed intake in cattle and pigs: a review. Asian-Australasian Journal of Animal Sciences 22, 747–755.
Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research 41, e1-e1.
Krause, T.R., Lourenco, J.M., Welch, C.B., Rothrock, M.J., Callaway, T.R., Pringle, T.D., 2020. The relationship between the rumen microbiome and carcass merit in Angus steers. Journal of Animal Science 98, skaa287.
Ledda, A., Carta, S., Correddu, F., Cesarani, A., Atzori, A.S., Battacone, G., Macciotta, N. P.P., 2023. Dry matter intake prediction from milk spectra in sarda dairy sheep. Animals 13, 763.
Lourenco, J.M., Welch, C.B., 2022c. Using microbiome information to understand and improve animal performance. Italian Journal of Animal Science 21, 899–913.
Lourenco, J.M., Callaway, T.R., Kieran, T.J., Glenn, T.C., McCann, J.C., Stewart Jr, R.L., 2019. Analysis of the rumen microbiota of beef calves supplemented during the suckling phase. Frontiers in Microbiology 10, 1131.
Lourenco, J.M., Kieran, T.J., Seidel, D.S., Glenn, T.C., Silveira, M.F.D., Callaway, T.R., Stewart Jr, R.L., 2020. Comparison of the ruminal and fecal microbiotas in beef calves supplemented or not with concentrate. PloS One 15, e0231533.
Lourenco, J.M., Welch, C.B., Krause, T.R., Wieczorek, M.A., Fluharty, F.L., Rothrock, M. J., Pringle, T.D., Callaway, T.R., 2022a. Fecal microbiome differences in Angus steers with differing feed efficiencies during the feedlot-finishing phase. Microorganisms 10, 1128.
Lourenco, J.M., Krause, T.R., Welch, C.B., Callaway, T.R., Pringle, T.D., 2022b. Longitudinal changes of the ruminal microbiota in Angus Beef steers. Animals 12, 1066.
Mardia, K. V., Jupp, P. E., 2000. Directional statistics, vol. 2. Chichester, John Wiley & Sons, Inc., New York, NY, USA.
Matthews, C., Crispie, F., Lewis, E., Reid, M., O'Toole, P.W., Cotter, P.D., 2019. The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. Gut Microbes 10, 115–132.
Monteiro, H.F., Zhou, Z., Gomes, M.S., Peixoto, P.M.G., Bonsaglia, E.C.R., Canisso, I.F., Weimer, B.C., Lima, F.S., 2022. Rumen and lower gut microbiomes relationship with feed efficiency and production traits throughout the lactation of Holstein dairy cows. Scientific Reports 12, 1–14.
Mulliniks, J.T., Beard, J.K., King, T.M., 2020. Invited review: effects of selection for milk production on cow-calf productivity and profitability in beef production systems. Applied Animal Science 36, 70–77.
Myer, P.R., Smith, T.P., Wells, J.E., Kuehn, L.A., Freetly, H.C., 2015. Rumen microbiome from steers differing in feed efficiency. PloS One 10, e0129174.
Nkrumah, J.D., Okine, E.K., Mathison, G.W., Schmid, K., Li, C., Basarab, J.A., Price, M.A., Wang, Z., Moore, S.S., 2006. Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. Journal of Animal Science 84, 145–153.
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, E., 2011. Scikit-learn: machine learning in Python. The Journal of Machine Learning Research 12, 2825–2830.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41, D590–D596.
Rencher, A.C., Scott, D.T., 1990. Assessing the contribution of individual variables following rejection of a multivariate hypothesis. Communications in Statistics - Simulation and Computation 19, 535–553.
Romanzin, A., Degano, L., Vicario, D., Spanghero, M., 2021. Feeding efficiency and behavior of young Simmental bulls selected for high growth capacity: comparison of bulls with high vs. low residual feed intake. Livestock Science 249, 104525.
Rothrock Jr, M.J., Hiett, K.L., Gamble, J., Caudill, A.C., Cicconi-Hogan, K.M., Caporaso, J.G., 2014. A hybrid DNA extraction method for the qualitative and quantitative assessment of bacterial communities from poultry production samples. Journal of Visualized Experiments 94, 52161.
Seidel, D.S., Walker, J.W., Musser, J.M., Lourenco, J.M., Welch, C.B., Whitney, T.R., Callaway, T.R., 2022. Impact of concentrations of camphor on the in vitro mixed ruminal microorganism fermentation from goats selected for consumption of low and high levels of Juniperus spp.-2. translational. Animal Science 6, txac098.
Shabat, S.K.B., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Berg Miller, M.E., White, B.A., Shterzer, N., Mizrahi, I., 2016. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. The ISME Journal 10, 2958–2972.
Welch, C.B., Lourenco, J.M., Davis, D.B., Krause, T.R., Carmichael, M.N., Rothrock, M.J., Pringle, D.T., Callaway, T.R., 2020. The impact of feed efficiency selection on the ruminal, cecal, and fecal microbiomes of Angus steers from a commercial feedlot. Journal of Animal Science 98, skaa230.
Welch, C.B., Lourenco, J.M., Krause, T.R., Seidel, D.S., Fluharty, F.L., Pringle, T.D., Callaway, T.R., 2021. Evaluation of the fecal bacterial communities of angus steers with divergent feed efficiencies across the lifespan from weaning to slaughter. Frontiers in Veterinary Science 8, 597405.
Williamson, J.R., Callaway, T.R., Lourenco, J.M., Ryman, V.E., 2022. Characterization of rumen, fecal, and milk microbiota in lactating dairy cows. Frontiers in Microbiology 13, 984119.
Zhou, X., Ma, Y., Yang, C., Zhao, Z., Ding, Y., Zhang, Y., Wang, P., Zhao, L., Li, C., Su, Z., Wang, X., Ming, W., Zeng, L., Kang, X., 2023. Rumen and fecal microbiota characteristics of Qinchuan cattle with divergent residual feed intake. Microorganisms 11, 358.