Journal of Animal Science, 2025, **103**, skaf206
https://doi.org/10.1093/jas/skaf206
Advance access publication 17 June 2025
Animal Genetics and Genomics

Differential intestinal microbiome response to heat stress in two rabbit maternal lines: a comparative analysis using Random Forest, BayesC, and PLS-DA

Iliyass Biada,[†] Francesco Tiezzi,[‡] Noelia Ibáñez-Escriche,[†,1] María Luz García,[‖] María José Argente,[‖] and María Antonia Santacreu[†]

[†]Instituto de Ciencia y Tecnología Animal, Universitat Politècnica de València, València, Spain
[‡]Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali, Università degli Studi di Firenze, Firenze, Italy
[‖]Departamento de Tecnología Agroalimentaria, Universidad Miguel Hernández de Elche, Orihuela, Spain
[1]Corresponding author: noeibes@dca.upv.es

Abstract

Heat stress affects livestock productivity and health, particularly in rabbits, due to their physiological vulnerabilities. This study explores the relationship between environmental conditions, genetic lines backgrounds, and soft fecal microbiota. A 2 × 2 factorial design was used, involving 2 maternal rabbit lines: A (standard longevity line) and LP (high longevity line), exposed to heat stress and thermal comfort. Samples were analyzed with multiple models to assess the impact of heat stress on microbiota by comparing microbial diversity and evaluating the classification performance of Random Forest, Partial Least Squares Discriminant Analysis (PLS-DA), and Bayesian Regression (BayesC). Heat stress influenced microbial diversity in both lines, increasing alpha diversity and driving significant beta-diversity shifts (2.3% variance, $P < 0.001$). This could be due to intestinal barrier disruption, which facilitate pathogen proliferation. The high longevity line LP exhibited higher richness under thermal comfort, whereas heat stress equalized these differences between lines, possibly due to increased pathogen proliferation in the low longevity line A. These differences in response to heat stress may be influenced by the crosstalk between microbiota and host genetics, shaping distinct adaptive mechanisms in each line. Prediction accuracy and key selected variables distinguishing between lines A and LP varied across thermal conditions, with the area under the curve exceeding 0.92 under heat stress and 0.87 in thermal comfort. This reflects different microbiome regulations between the 2 lines under heat stress. Potential stress-associated taxa such as *Erysipelatoclostridium* and *Monoglobus* were more abundant in the low longevity line A. These results highlight LP's higher longevity and expected resilience, while line A's susceptibility is reflected in a higher abundance of heat stress-associated taxa in the latter. This underscores soft fecal microbiota as a potential biomarker for heat stress resilience and emphasizes the role of host–microbiota interactions in mediating genetic-environmental responses. Additionally, this study highlights the value of combining modeling approaches, which enhance accuracy and reveal key taxa driving heat stress responses. Among the models tested, PLS-DA achieved the highest accuracy, while Random Forest identified a smaller yet biologically relevant subset of taxa, providing valuable phylogenetic and taxonomic insights.

Lay Summary

This study highlights the significant impact of heat stress on the microbiome of maternal rabbit lines with differing genetic backgrounds. By revealing how genetic and environmental factors interact to shape the microbiota, it underscores the potential of soft feces microbiome composition as a biomarker for heat stress resilience. These findings are relevant to rabbit breeding and provide valuable insights applicable to other animal production species facing similar climate challenges. Additionally, comparing machine learning and Bayesian methods offers robust tools for understanding microbial architecture, guiding the scientific community in microbiome data analysis. This dual contribution supports sustainable livestock production and advances microbiome research methodologies.

Keywords: Longevity, microbiota, classification models, rabbits, thermal conditions, resilience.

Abbreviations: A, rabbit maternal genetic line A; AUC, area under the curve; ASV, amplicon sequence variant; BayesC, Bayesian regression model C; BER, balanced error rate; CLR, centered log ratio; GIT, gastrointestinal tract; HPD95%, highest posterior density interval at 95%; LP, rabbit maternal genetic line LP; MCMC, Markov chain Monte Carlo; ML, machine learning; OOB, out-of-bag; PCA, principal component analysis; PERMANOVA, permutational multivariate analysis of variance; PLS-DA, partial least squares discriminant analysis; PCR, polymerase chain reaction; SCFAs, short-chain fatty acids; THI, temperature-humidity index; VIP, variable importance in projection.

Introduction

Rabbit farmers, especially in the Mediterranean region, seek animals that are highly productive, more resilient, and long-lived (Cullere and Dalle Zotte, 2018). This is becoming increasingly important as rising temperatures due to climate change pose a major environmental challenge (Marai et al., 2002; Giorgi and Lionello, 2008; Ebeid et al., 2023). Rabbits are particularly sensitive to heat stress due to their dense fur and limited number of sweat glands, which makes natural

Received March 10, 2025 Accepted June 16, 2025.

heat dissipation through sweating and respiration less efficient (Dalmau et al., 2014). While many commercial rabbitries in southern Europe employ ventilation and cooling systems, the intensification of summer heatwaves (particularly in Mediterranean regions like Spain) has exposed the limitations of these technologies (Egea, 2024). Factors such as humidity, equipment aging, or limited economic capacity to modernize systems can compromise their effectiveness. Consequently, heat stress negatively affects reproductive traits, decreasing fertility, embryonic development, litter size, litter weight, and milk production (Ebeid et al., 2023). It also impacts growth traits such as body weight, daily gain, meat quality, and growth rate (Zeferino et al., 2013). These adverse effects underscore the importance of addressing heat stress for the sustainability of rabbit farming under changing climatic conditions.

One of the critical yet underexplored effects of heat stress on rabbits is its impact on the intestinal microbiota. Recent studies have established links between intestinal microbiota, life expectancy, and overall health in rabbits (Gerber, 2014; Funosas et al., 2021; Biada et al., 2024). In healthy animals, the intestinal microbiota maintains intestinal integrity by producing short-chain fatty acids (**SCFAs**), which help the gut barrier and support immune function (Priyadarshini et al., 2018). However, under heat stress conditions, the integrity of the gut barrier is compromised due to increased permeability, which allows harmful substances to enter the bloodstream, leading to inflammation and illness (He et al., 2021). Heat stress has also been shown to directly alter intestinal microbiota composition and diversity in monogastric animals (He et al., 2021) and in rabbits (Bai et al., 2021; Liu et al., 2022). This disruption can lead to dysbiosis, further worsening the situation by reducing the production of SCFAs, weakening the gut's protective barrier, and ultimately compromising overall animal health. Specific microbial taxa are particularly responsive to heat stress and may serve as valuable biomarkers for monitoring its effects on health, growth, and productivity (He et al., 2021; Bai et al., 2021; Liu et al., 2022).

This study hypothesizes that heat stress differentially affects soft fecal microbiota composition and diversity based on the maternal genetic line. This results in distinct microbial profiles that reflect variations in physiological and microbial mechanisms for regulating and responding to heat stress between the genetic lines. By identifying these microbial markers, this study aims to offer new insights into how rabbits and other livestock cope with environmental stressors, enabling farmers to manage heat stress better and enhance the sustainability of animal production systems in a changing climate.

Microbiome data is high dimensional, sparse, and compositional (Gloor et al., 2017). In this study, we apply different machine learning (**ML**) techniques and Bayesian regression to address these complexities and compare the predictive capabilities of different models. ML methods have gained significant traction in microbiome research because they can handle complex, high-dimensional datasets (Namkung, 2020; Papoutsoglou et al., 2023). Among these methods, Random Forest is among the most frequently employed for microbiome classification and biomarker discovery (Marcos-Zambrano et al., 2021). It is an ensemble learning method that builds multiple decision trees and aggregates their outputs to improve prediction accuracy. Partial Least Squares Discriminant Analysis (PLS-DA) represents a different class of ML that uses latent variables. It maximizes the covariance between predictor variables and the response variable, making it particularly effective for reducing dimensionality while retaining the ability to differentiate between groups (Papoutsoglou et al., 2023). Lastly, BayesC is a model from the Bayesian alphabet family used in genomic prediction, especially with high-dimensional data like single nucleotide polymorphisms. BayesC can be suited for microbiome prediction because it assumes that only a subset of predictors (such as microbial taxa) has non-zero effects, allowing the model to focus on the most informative variables (Habier et al., 2011). Also, BayesC can be adapted to handle non-continuous distributions of the outcome variable (Pérez and de los Campos 2014). These 3 methods (Random Forest, PLS-DA, and BayesC) were chosen for their distinct strengths and approaches in prediction and variable selection. This will allow for a thorough comparison of model performance and biological interpretation to determine which approach best suits microbiome data analysis.

This study has the overarching objective of examining the effect of heat stress on the soft fecal microbiome of 2 maternal rabbit lines with differing longevities. First, we aim to explore the relationship between environmental conditions, genetic lines, and microbiota. This is done by comparing the microbiota's alpha- and beta-diversity measures. Second, we aim to offer insights into microbial architecture by comparing the 3 modeling approaches' predictive performance and biological interpretability. These are Random Forest, PLS-DA, and BayesC. They are compared in terms of their ability to classify genetic lines based on microbiome information and identify key microbial taxa associated with heat stress. This dual focus allows us to deepen our understanding of the biological effects of heat stress and assess which computational methods are most effective in predicting and interpreting microbiome data.

Material and Methods

Ethics approval

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

Animals and experimental design

A total of 199 nulliparous adult female rabbits (24 to 27 wk) from 2 highly productive commercial maternal lines with different longevity (A and LP) were employed in this experiment. All animals were born and raised at the experimental farm of the *Universitat Politècnica de València*. Data collection was conducted across 2 distinct periods: a heat stress period during summer (June 19, 2023, to September 25, 2023) with a temperature-humidity index (**THI**) of 25.7 ± 1.2, and a thermal comfort period during winter (November 27, 2023, to February 13, 2024) with a THI of 17.5 ± 1.7. Ambient temperature and humidity were recorded inside the experimental farm, and THI was calculated following the formula adapted for rabbits by Marai et al. (2002). In both thermal conditions, animals were exposed to the respective environment during at least the final stage of gestation, the day of parity, and the first week postpartum.

The dataset comprised 98 females from line A (47 in heat stress and 51 in thermal comfort) and 101 females from line LP (50 in heat stress and 51 in thermal comfort). Line LP was founded using longevity and productivity criteria, specifically from females with an average number of parities of

30, which is 6 times higher than the average longevity (Sánchez et al., 2008). Line A is a standard commercial maternal line that was founded by sampling New Zealand White rabbits reared by farmers near Valencia, Spain (Sánchez et al., 2008). All females were housed in individual cages (extractable nest boxes with isolated plastic) under a photoperiod of 16-h light: 8-h dark at the *Universitat Politècnica de València* (Spain) farm. The animals were fed ad libitum with a standard commercial diet throughout the experimental period, formulated to meet their nutritional requirements. The feed composition included 16.8% crude protein, 3.4% crude fat, 15.5% crude fiber, 7.9% crude ash, 1.00% calcium, 0.58% phosphorus, and 0.19% sodium. Additional details on ingredients and additives are provided in Table S1.

Sample collection and DNA extraction

All samples and physiological data were recorded during the first week postpartum in both thermal conditions. Specifically, on the first morning following the first parity of each female, litter size and body temperature were recorded. Body temperature was assessed using infrared thermography of the eyeball with a "Testo Irsoft" thermograph, and data analysis was performed using Irsoft software (version 1.7). Two consecutive measurements were taken for each rabbit at a consistent distance, with the average maximum eyeball temperature calculated. Measurements showing a difference greater than 0.5 °C between readings were excluded. Under thermal comfort conditions, the mean eyeball temperature was 35.4 ± 0.6 °C, which increased to 37.43 ± 0.6 °C during heat stress. The average litter size was 10.5 ± 2.7 under thermal comfort conditions and 10.3 ± 2.2 under heat stress.

Soft fecal samples were collected from females during the first postpartum week. Sampling was conducted twice daily by gently stimulating the perianal region to facilitate fecal release. Immediately after successful collection, the samples were stored at −80 °C to maintain their integrity until DNA extraction. HigherPurity Stool DNA Isolation Kit was used for bacterial DNA extraction (Canvax Reagents SLU, Valladolid, Spain). The following modifications to the standard protocol were applied: Approximately 0.2 g of fecal matter was disrupted using two 4-mm glass beads in a bead homogenizer at maximum speed (6 m/s) for 1 min, combined with the lysis buffer. The samples were then heated to 95 °C for 5 min, a process that was repeated twice. Following this, centrifugation at 10,000 × *g* for 30 s was performed, and the supernatant was carefully transferred to a fresh tube, adhering to the manufacturer's guidelines. The final elution step yielded DNA in a 100-µL volume. DNA concentration and quality were assessed using a Nanodrop ND-1000 spectrophotometer and a Qubit 4 fluorometer with the dsDNA HS DNA assay kit.

PCR amplification, barcoding, and DNA sequencing

The V3 and V4 regions of the bacterial 16S rRNA gene were amplified and purified following Illumina's 16S Metagenomic Sequencing Library Preparation protocol, as detailed in Biada et al. (2024). PCR amplification was carried out using primers recommended for this protocol under specific thermal cycling conditions: an initial denaturation at 95 °C for 3 minutes, followed by 25 cycles, each consisting of 30 s at 95 °C, 30 s at 55 °C, and 30 s at 72 °C, with a final extension step at 72 °C for 5 minutes. For multiplexing, dual indices were added to the PCR products using the Nextera XT Index Kit (FC-131-2001). The amplified products, approximately 550 bp in length, were verified using a Bioanalyzer DNA 1000 chip. Final sequencing was performed using a 2 × 300 bp paired-end run on a MiSeq Sequencer, adhering to the instructions provided by Illumina.

Bioinformatic analyses

Primary processing was carried out on the raw sequencing reads, including quality control filtering using fastp (Chen et al., 2018) and removal of primers using Cutadapt (Martin, 2011). Then, the sequences were processed using the DADA2 pipeline (Callahan et al., 2016) in R version 4.3.2. Forward and reverse reads were trimmed to remove the low-quality portion of the sequences (minimum of 25 Phred). Removal of chimeras was performed. Amplicon sequence variants (**ASVs**) were inferred after denoising, and the result was the construction of ASV tables. The ASV table was imported into QIIME2 software, version 2021.11 (Bolyen et al., 2019). Taxonomic classification was performed using a trained Silva-based 16S classifier (the classifier was trained on the primers used for amplification and the length of the sequence reads). Also, to consider evolutionary relatedness between the ASVs identified, a phylogenetic tree was constructed using QIIME2. Final ASVs were filtered to remove potential artifacts; sequences not classified under the kingdom Bacteria and those with a relative abundance below 0.01% were excluded (Bokulich et al., 2013).

Alpha-diversity analysis

Three alpha-diversity metrics, Observed richness, Shannon, and Pielou evenness, were computed using the phyloseq package (McMurdie and Holmes, 2013). Each alpha-diversity index was modeled using a linear model. The model considered fixed effects, including the animal's genetic line (A or LP), thermal condition effect (heat stress and comfort), and the interaction between the line and thermal condition. Sequencing depth per sample (N) was included to control variation in sequencing depth across samples. The model equation was

$$y_{ijk} = Line_j + Thermal\ condition_k + Line : Thermal\ condition_{jk} + \log(N_{ijk}) + e_{ijk} \quad (1)$$

Where y_{ijk} represents the alpha-diversity measurement for the sample i of line j (A or LP), thermal condition k (heat stress or comfort) with N_{ijk} being the total sampling depth and e_{ijk} representing the error term of each sample. To ensure robustness in the analysis, the same model was repeated using rarefied tables instead of including N_{ijk} for sequencing depth control.

The traits were analyzed using Bayesian methodology in R using the brms package (Bürkner, 2017). Default priors provided by brms were used (Bürkner, 2017), and they were as follows: the intercept was assigned a weakly informative Student's *t*-distribution with 3 °C of freedom and a scale parameter that depends on the standard deviation of the response (with a minimum of 2.5). The rest of the fixed effects coefficients were assigned flat priors. The residuals were assumed to follow a normal distribution $N(0, \sigma_e^2)$, and the error standard deviation σ_e was assigned weakly informative prior as with the same parameters as the intercept. Each model was run by 4 chains of 50,000 iterations and a 10% burn-in of 5000 iterations. Inferences were made from the estimated marginal posterior distributions of the differences between the 2 lines A and LP, heat stress and thermal comfort conditions and the

interaction. The probability P0 that the difference is higher (if positive) or lower (if negative) than 0 and HPD95% (the highest posterior density interval at 95%) were calculated. The relevance threshold is P0 greater than or equal to 80%, and the mean difference in standard deviation (sd) units is higher than 1/3.

Beta-diversity analysis

To account for the compositionality of microbiome data prior to computing the Aitchison dissimilarity matrix, a pseudo count was added, and the data was transformed using the centered log ratio (**CLR**) (Greenacre et al., 2021) . Then, principal component analysis (**PCA**) was computed using Aitchison dissimilarity to visually represent between-sample dissimilarity. Finally, the Permutational Multivariate Analysis of Variance (PERMANOVA) test was used to assess the effect of the predictors on the Aitchison distance between samples (10,000 permutations). The model was computed using the vegan package in R (Oksanen et al., 2024) and included the effect of line, thermal condition, and interaction.

Data processing, modeling and variable selection

The dataset consists of 199 samples, with 99 females from line A (47 in heat stress and 51 in thermal comfort) and 101 females from line LP (50 in heat stress and 51 in thermal comfort). To assess the performance of the models under different environmental conditions, we tested 3 data subsets. Subset 1 included samples from the heat stress condition only (97 samples), subset 2 comprised samples from the thermal comfort condition (102 samples). Subset 3 consisted of a balanced dataset of 100 samples, comprising 50 randomly selected animals from the heat stress group and 50 from the thermal comfort group. This subset was designed to avoid biases related to unequal sample sizes and to serve as a reference for evaluating model performance under mixed environmental conditions.

We followed the framework for microbiome-based classification problems suggested by (Topçuoğlu et al., 2020), with some modifications. We split the data into 50% training and 50% testing for each subset while maintaining the overall class distribution. Repeated 5-fold cross-validation (100 times) was performed on the training set to tune the hyperparameters. Once tuned, the models were trained on the whole training set and evaluated on the testing set. Data splitting, hyperparameter selection, training, and testing were repeated 100 times to minimize bias from random splits. The area under the receiver operating characteristic curve (**AUC**) was used to evaluate model performance.

All models (BayesC, Random Forest, and PLS-DA) were employed for the 3 data subsets, where the CLR transformed ASVs were used as predictors (1,416 ASVs) to predict the genetic line of the female rabbits (Line A or LP).

BayesC: We employed the BayesC regression model, implemented in the BGLR R package (Pérez and de los Campos, 2014), to predict the genetic line of female rabbits (A or LP). Given the categorical nature of the dependent variable, we used a probit link function to handle its binary nature. In this model, the observed binary variable (y, genetic line: A or LP) is linked to an unobserved continuous variable called liability (λ) through the following relationship:

$$y = \begin{cases} 1 \; if \; \lambda \; > 0 \\ 0 \; if \; \lambda \; \leq 0 \end{cases}$$

The liability variable λ is modeled as:

$$\lambda = \quad \mu + Ma + e \quad . \tag{2}$$

where λ is the response variable on the liability scale, which is assumed to follow a normal distribution $\lambda \sim N(\eta, 1)$ with a mean η equal to the linear predictor and variance of one (Albert and Chib, 1993). Here μ is a population mean, a is a vector of ASV effects (1,416 levels), **M** is a matrix of centered and scaled ASV counts, and e is a vector of random residuals assumed $N(0, \; \sigma_e^2)$.

CLR transformed ASVs were fitted to the model using a scaled-t mixture prior distribution, which uses a mixture of a point of mass at zero and a Gaussian slab. Regression coefficients were assigned Gaussian priors with a mean of zero and variance σ_β^2. The variance parameter σ_β^2 and residual variance σ_e^2 were assigned scaled-inverse Chi-square densities, indexed by a scale and a degree of freedom parameter, using BGLR's default values. The proportion of non-zero effects was modeled with the parameter π, which followed a Beta prior $\pi \sim Beta(\quad p_0, \quad \pi_0)$. Here, the expected proportion of non-zero effects is $E(\pi) = \pi_0$, while p_0 determined prior strength with higher values leading to tighter distributions around π_0. For this model, p_0 was set to 1,000,000 to provide a strongly informative prior, while π_0 was tuned across different values to optimize the model through a grid search. Markov Chain Monte Carlo (MCMC) sampling was used to estimate the effects of the predictors and provide posterior distributions for these effects. Each model was run by 4 chains with a length of 100,000 iterations, a lag of 5, and a burn-in of 10,000 iterations.

Random forest: The classification model was implemented using the randomForest package in R (Liaw and Wiener, 2002), based on Breiman's original random forest algorithm (Breiman, 2001). This approach builds multiple decision trees using bootstrap samples drawn with replacements from the original dataset. At each node, a randomly selected subset of features is chosen for the splitting criterion, with the *mtry* parameter set to specify the number of predictors considered at each split. For this model, *mtry* was tuned to optimize classification performance while other parameters were kept at their default values. A total of 1,500 trees (*ntree* = 1,500) were grown to ensure robust classification accuracy, with each tree reaching its maximum extent until the terminal nodes became homogeneous. The out-of-bag (**OOB**) error was used to evaluate the model's performance, which also provided a measure of feature importance and class-specific error rates across the predictions.

The general form of the Random Forest model is given by:

$$\hat{y} = \; \mu + \frac{1}{T} \sum_{t=1}^{T} h_t(y : M) \tag{3}$$

Where $\hat{y}$ is the predicted probability for the genetic line (A or LP), μ is the population mean, T represents the total number of decision trees in the forest (T = 1,500), and $h_t(y : M)$ denotes the prediction for the tth tree, where $t \in (1, \; T)$. M is the matrix of centered and scaled ASV counts. Each tree $(h_t(y : M) \;\; t \in (1, \; T))$ is distinct from any other as it is constructed from a bootstrap sample of the original data. By design, the model minimizes overfitting and improves predictive performance by aggregating predictions of different

decision trees. Additionally, Random Forest quantified feature importance using 2 standard metrics: Mean Decrease Accuracy, which evaluates the drop in model accuracy when a feature is permuted, and Mean Decrease Gini, which assesses a feature's contribution to node impurity reduction. These measures allowed identification of the most influential ASVs for classifying rabbit lines.

Partial least square discriminant analysis: The PLS-DA model was constructed using the mixOmics package in R (Rohart et al., 2017). This model identifies latent variables (or components) that maximize covariance between the predictor matrix **M** and the response **y**. The relationship can be expressed as:

$$y = \sum_{h=1}^{m} (\mathbf{M}\omega_h^*)c_h + e \quad (4)$$

Where **y** is the response variable (genetic line A or LP), **M** is the matrix of centered and scaled ASVs, ω_h^* represents the weight vectors for each component h, which capture the contribution of each predictor variable to the component while maximizing the covariance with the response variable, c_h is the regression coefficient for the hth component, and e is the error term.

Variable selection

To identify the most relevant variables for prediction and compare across models, variable selection was performed within each training set over 100 repetitions. Variables consistently selected in at least 50% of repetitions were considered the most relevant. BayesC: Variable selection is inherently part of the BayesC method, as it calculates posterior probabilities of inclusion for each variable. Different thresholds were tried, and we selected the 150 variables with the highest posterior probabilities to align with the approximate number of features chosen in the Random Forest and PLS-DA models. Random Forest: Variable selection followed the 2-step process outlined by Genuer et al. (2010) for their interpretation method. First, variables with low importance (assessed by the Mean Decrease Accuracy repeated over 50 iterations) were eliminated. Next, a nested collection of models was constructed, selecting variables leading to the smallest OOB error. PLS-DA: We applied the Variable Importance in Projection (**VIP**) criterion, iteratively eliminating variables with VIP scores less than 1.5 until the model achieved the lowest Balanced Error Rate (BER).

Differential abundance

To complement the predictive models (PLS-DA, BayesC, and Random Forest), which identified key taxa differentiating the 2 genetic lines, we performed a differential abundance analysis to determine which line these predictive taxa were more abundant. This step was an additional discovery to validate the features already identified as predictive. To determine which taxa were more abundant in each genetic line, we followed the recommendations from Nearing et al. (2022), who systematically evaluated various differential abundance methods for microbiome studies. Their findings suggested that ALDEx2 and ANCOM-II provided the most consistent results across different studies. We used ALDEx2 for this analysis, as it aligns with our use of centered log-ratio (CLR) transformation. Using the ALDEx2 package in R, we identified differentially abundant taxa between the 2 rabbit lines (A and LP) in each data subset.

Results

Diversity results

At the level of alpha diversity (within sample diversity), line LP had higher alpha diversity for all 3 indices (Table 1). These differences were relevant for Shannon and observed richness (P0 >= 80% and mean difference in sd units higher than 1/3). Concerning the thermal condition effect, alpha diversity increases during heat stress in comparison to thermal comfort, with

Table 1. Summary of the main effect estimates for alpha-diversity indices Shannon, Pielou evenness and observed richness obtained with the linear model

Contrast	Index	PM[1]	HPD95%[2]	P0(%)[3]
A[4] to LP[5]	Shannon	−0.9	[−0.11, 0.04]	82
	Pielou evenness	−0.6	[−0.009, 0.005]	74
	Observed richness	−1.0	[−44, 15]	83
HS[6] to TC[7]	Shannon	1.5	[−0.02, 0.14]	93
	Pielou evenness	0.8	[−0.004, 0.010]	80
	Observed richness	2.4	[7, 68]	99
A (HS) to LP (HS)	Shannon	−0.6	[−0.14, 0.07]	72
	Pielou evenness	−0.4	[−0.012, 0.007]	70
	Observed richness	−0.5	[−55, 31]	68
A (TC) to LP (TC)	Shannon	−0.7	[−0.14, 0.07]	76
	Pielou evenness	−0.5	[−0.012, 0.008]	66
	Observed richness	−0.9	[−59, 23]	81

[1]Marginal posterior mean in standard deviation (sd) units.
[2]Highest posterior density intervals at 95% probability.
[3]Probability of the difference being higher (if the difference is positive) or lower (if negative) than zero.
[4]Rabbit's maternal genetic line A.
[5]Rabbit's maternal genetic line LP.
[6]Heat stress.
[7]Thermal comfort.

differences being relevant for all indices. Finally, no relevant differences were observed regarding the interaction effects, with similar trends within each thermal condition (Line LP with higher alpha diversity values than A). However, the increase of diversity during heat stress was more pronounced in line A, since the contrasts between lines were lower, and none of the 3 indices showed relevant differences. The analysis was repeated using rarefaction, and the same results were found, with the exception of no relevant differences between heat stress and thermal comfort for evenness diversity (Table S2).

Concerning beta diversity, PERMANOVA results showed that thermal condition significantly (P-value = 0.000) explained the highest proportion of the total dissimilarity between samples, 2.3%, followed by the line effect (P-value = 0.000), which explained 1.7%. Finally, the interaction significantly explained 0.8% (P-value = 0.001). Concerning the results of the PCA based on Aitchison dissimilarity, the first and second principal components explained 32% and 10% of the total variance, respectively. While these components accounted for the majority of the variation, the third and fourth components, together explaining an additional 9% revealed a clearer separation of bacterial communities according to thermal conditions (Fig. 1). Concerning line effect, the fourth and fifth components (5% of the total variance) had the better separation (Fig. 1).

Predictive performance of models

We assessed the predictive power of microbiome data in distinguishing between maternal genetic lines A and LP in this population of first-parity females. Predictions were evaluated using 3 subsets: females under heat stress, females under thermal comfort, and a combined dataset of both conditions, adjusted to match the sample sizes of the other subsets. Three distinct models were employed to analyze the data: ML approaches using Random Forest, PLS-DA, and BayesC. Fig. 2 reports the AUC values over 100 repetitions for each model and data subset. The initial results indicate that all models performed well, with AUC scores (average ± sd) of 0.90 ± 0.05 for Random Forest, 0.94 ± 0.05 for BayesC, and 0.96 ± 0.03 for PLS-DA.

We employed a permutation test to ensure that the models were not randomly discriminating between the classes. The permuted AUC scores were 0.51 ± 0.08 for Random Forest, 0.55 ± 0.06 for BayesC, and 0.50 ± 0.07 for PLS-DA, indicating that predictions were close to random chance under permutation. These initial results indicate that the microbiome of the 2 maternal lines, A and LP, is different.

Influence of model choice on prediction performance and variable selection

While all models demonstrated strong predictive power across the different datasets (Fig. 2), the choice of model impacted both the prediction performance and the variable selection process. In prediction performance, PLS-DA consistently outperformed the other models, achieving the highest AUC values across all subsets of the data. Specifically, PLS-DA achieved AUC values of 0.99 ± 0.01 under heat stress, 0.95 ± 0.03 in thermal comfort, and 0.95 ± 0.03 in mixed data (Combination of heat stress and thermal condition). BayesC followed closely behind PLS-DA regarding prediction performance; it demonstrated a strong mean AUC of 0.96 ± 0.03 under heat stress, 0.92 ± 0.05 in thermal comfort, and 0.92 ± 0.06 for mixed data. Although still effective, Random Forest performed somewhat worse than both PLS-DA and BayesC with mean AUC of 0.92 ± 0.05 under heat stress, 0.87 ± 0.06 in thermal comfort, and 0.92 ± 0.03 for mixed data. Moreover, PLS-DA demonstrated lower variability across all data subsets than BayesC and Random Forest. The AUC values for PLS-DA were highly stable, with a sd of 0.01 under heat stress and 0.03 in thermal comfort and mixed data. This indicates that PLS-DA not only provides higher AUC but also exhibits a strong level of consistency, suggesting it is less prone to fluctuations in prediction performance across different training sets.

We applied a variable selection process to identify the most important features that discriminated between the 2 rabbit genetic lines (A and LP). This aims to highlight the most discriminatory taxa and compare consistency in selected variables across the models and thermal conditions. The number of selected variables for each model and data subset is

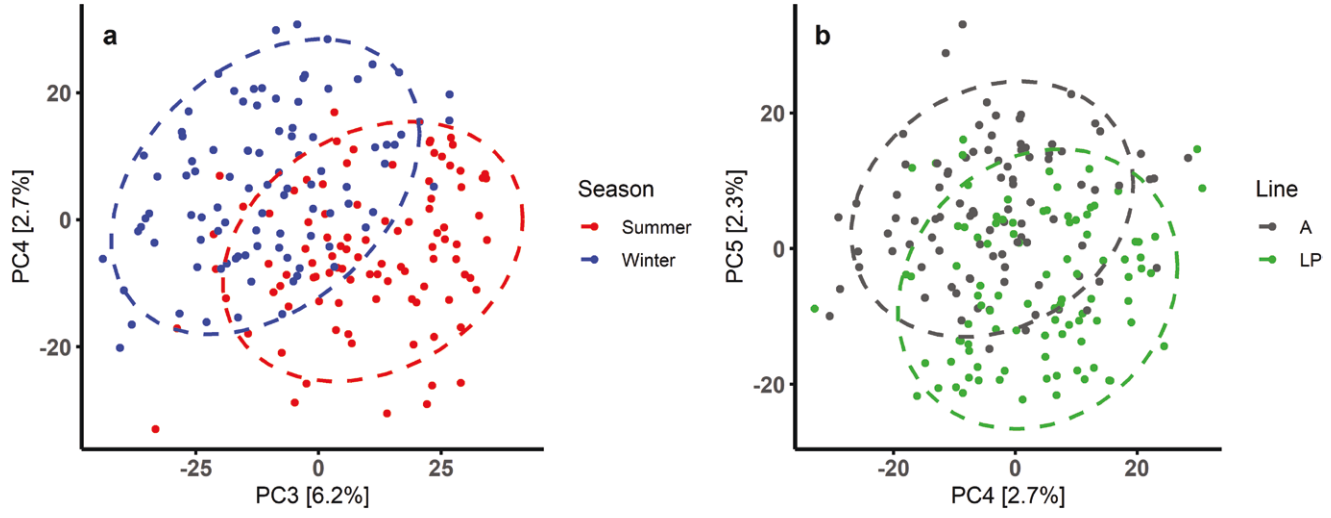

Figure 1. (A) Visualization of third and fourth principal components (PC3 and PC4) representing the dissimilarity between samples (based on Aitchison distance), with samples colored by thermal condition. (B) Visualization of fourth and fifth principal components (PC4 and PC5) with samples colored by Line.

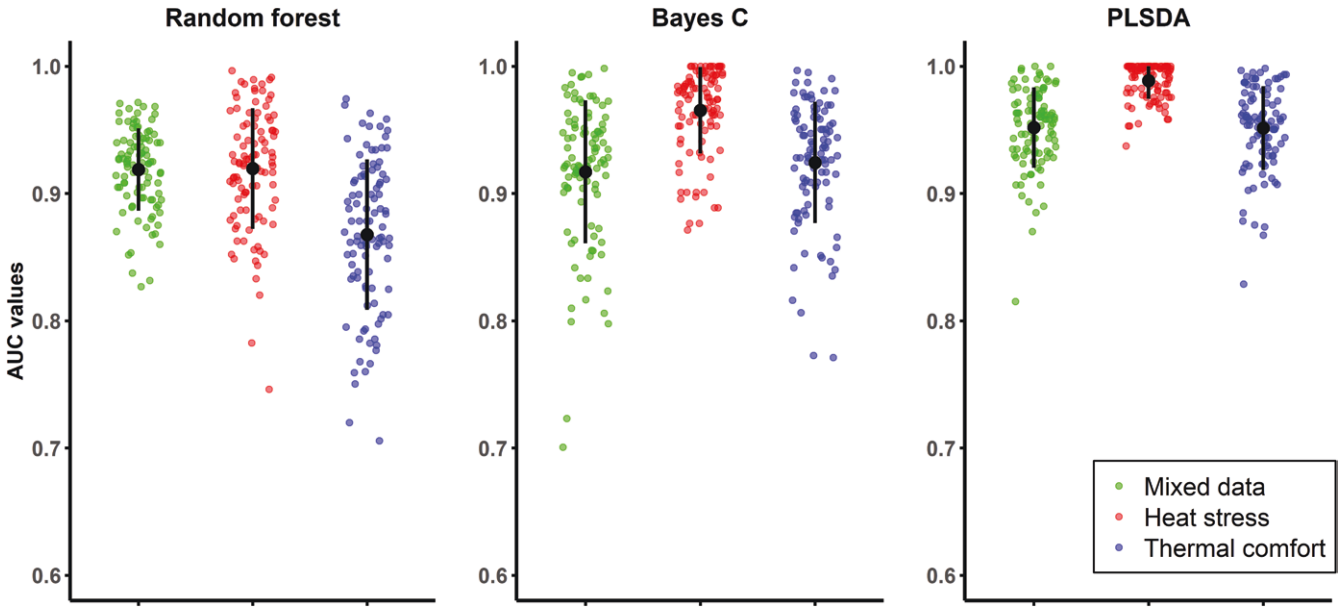

Figure 2. AUC values for Random Forest, BayesC, and PLS-DA models prediction of genetic line A or LP based on microbiome data across mixed, heat stress, and thermal comfort conditions. Each dot represents the AUC value from one repetition, with black error bars indicating the mean ± standard deviation (sd) for each group.

Table 2. Number of important variables identified by Random Forest, BayesC, and PLS-DA models under heat stress, thermal comfort, and mixed conditions, along with the shared variables across models

Subset	Random Forest	BayesC	PLS-DA	Intersect
Heat stress	15	71	96	14
Thermal comfort	17	56	86	15
Mixed	21	25	50	16

summarized in Table 2, and the overlap between the selected variables is illustrated in Fig. 3. PLS-DA consistently selected a higher number of variables across all subsets, especially under heat stress (96 variables) and thermal comfort (86 variables), whereas BayesC selected a moderate number. Random Forest selected fewer variables overall (between 15–21 variables across subsets), but its selections showed a high degree of overlap with both BayesC and PLS-DA (Fig. 3). This overlap suggests that Random Forest, which evaluates features independently and additively while capturing interactions, leads to a more conservative selection of variables. Consequently, the selected variables represent core taxa consistently identified as important across other models.

To compare the biological relevance of the variables selected by each model, we examined the taxonomic information and phylogenetic trees of the selected ASVs. For PLS-DA, the selected ASVs collapsed into 23 and 34 taxonomic units for heat stress and thermal comfort (Table S3). Notably, 14 taxonomic units (10 genera and 4 families) were common between heat stress and thermal comfort, while the remaining taxa were unique to each condition. The phylogenetic tree (Fig. S1) showed similar results, with ASVs from both conditions clustering closely on adjacent branches. However, some ASVs representing the unique taxa to each condition were more separated and appeared on distant branches. BayesC followed a similar trend to PLS-DA, with selected ASVs collapsing into 24 distinct taxonomic units for both heat stress and thermal comfort (Table S4). From which, 13 (9 genera and 4 families) were shared between the 2 thermal conditions. The phylogenetic analysis mirrored that of PLS-DA, with many ASVs clustering together, while a few (belonging to the unique taxonomic units of each condition) remained dispersed across separate branches (Fig. S2). In contrast, Random Forest selected fewer ASVs, collapsing into 7 and 10 taxonomic units for heat stress and thermal comfort (Table S5). From which, only 2 (families *Eubacteriaceae* and *Lachnospiraceae*) were shared between heat stress and thermal comfort. The phylogenetic tree revealed a different pattern: ASVs belonging to different thermal conditions were spread across distant branches, though those from the same condition tended to cluster together (Fig. S3). Even the ASVs of the shared families between conditions were placed on separate branches, suggesting they likely belong to distinct genera or species.

In conclusion, PLS-DA selected more variables and achieved better prediction performance compared to Random Forest and BayesC. Random Forest, while having lower prediction accuracy, selected fewer variables that showed clearer biological differences between thermal conditions at both the taxonomic and phylogenetic levels. BayesC performed in between the 2, balancing variable selection and prediction accuracy. These results suggest that selecting a larger number of variables may be necessary for achieving better class discrimination.

Line prediction in different thermal conditions

When comparing the 3 scenarios (heat stress, thermal comfort, and mixed) we observed that the mean AUC values were consistently higher during the heat stress period across all 3 models. The average AUC during heat stress was 0.96 ± 0.03, higher than the 0.91 ± 0.06 observed under thermal comfort. When both periods were combined (Mixed), the models achieved intermediate results with an AUC of 0.93 ± 0.04.

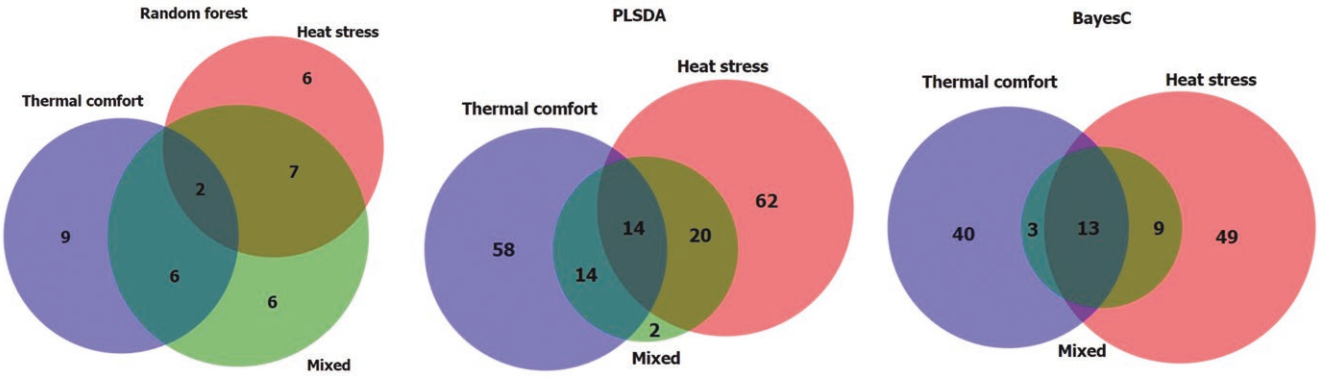

Figure 3. Venn diagrams showing the overlap of variables selected across different models Random Forest, BayesC and PLS-DA for different data subsets (heat stress, thermal comfort and mixed). The size of circles is proportional to number of values within it.

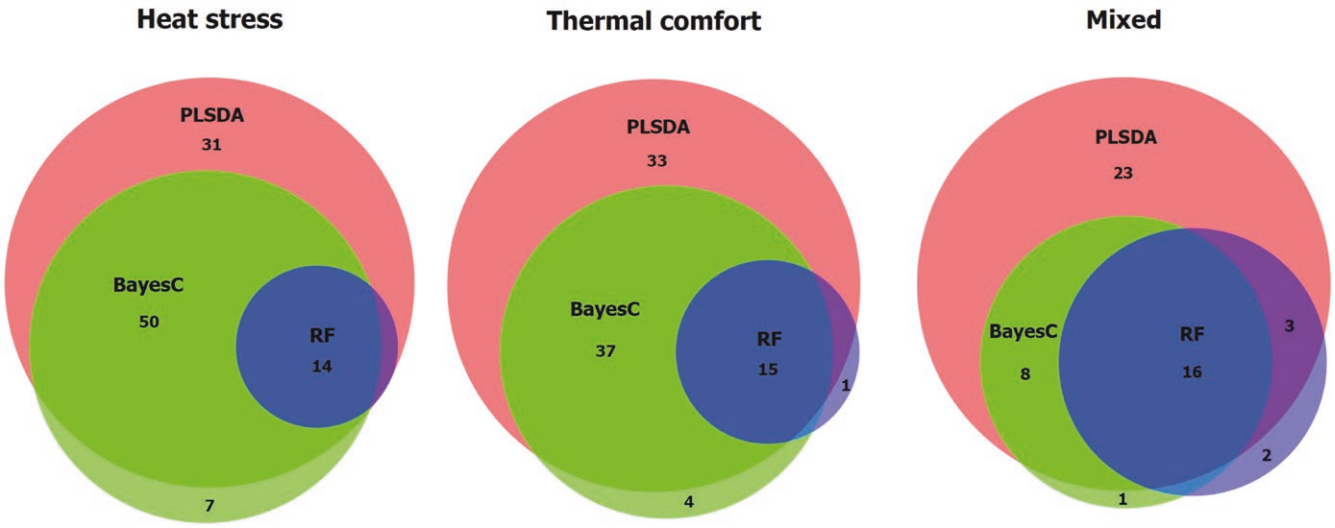

Figure 4. Venn diagrams showing the overlap of variables selected across thermal conditions (heat stress, thermal comfort, and mixed) for Random Forest, BayesC and PLS-DA models. The size of circles is proportional to number of values within it.

These results suggest that predicting genetic lines based on microbiome data was most effective during the heat stress period, indicating that the microbiome of the 2 rabbit lines diverged more under heat stress conditions.

Variable selection further validated these findings, revealing that distinct taxa were identified as the key variables discriminating between the 2 lines, depending on whether the conditions were heat stress or thermal comfort (Fig. 4). Specifically, the Venn diagrams show that, across all 3 models, most of the important variables under heat stress and thermal comfort are unique to each condition, with only a small portion shared. The combined subset (mixed) showed few or no unique important variables, with most of its important variables overlapping with the 2 other subsets. These results indicate that key variables selected with high discriminatory power between the 2 lines A and LP are different depending on the thermal conditions.

This pattern suggests that the differences in the microbiomes between the genetic lines are conditional on environmental conditions, such as heat stress. The more pronounced differences observed during heat stress, along with the distinct key variables selected compared to thermal comfort, indicate that the microbiome undergoes more significant and varied changes in response to heat stress.

Potential taxa as biomarkers for heat stress

From this point onward, we focus on the intersection of variables selected by the 3 models (Random Forest, PLS-DA, and BayesC) which identified 14 variables under heat stress and 15 under thermal comfort (Table 2). Differential abundance analysis was performed using ALDEx2, aligning with our use of CLR transformation. As anticipated, all taxa identified as key predictors were also found to be differentially abundant (Supplementary Material S1 and Fig. 5).

Similar to the previous section, we explored taxonomic and phylogenetic differences across thermal conditions. The final ASVs were grouped into 4 and 9 taxonomic units for heat stress and thermal comfort, respectively. From which, only 2 taxonomic units (families *Eubacteriaceae* and *Lachnospiraceae*) were shared between the 2 thermal conditions. The unique taxonomic units for heat stress included the genera *Erysipelatoclostridium* and *Monoglobus*, while thermal comfort specific taxa comprised the genera *RF39, Clostridia UCG-014, Olsenella, Parvibacter, Lachnospiraceae NK4A136*

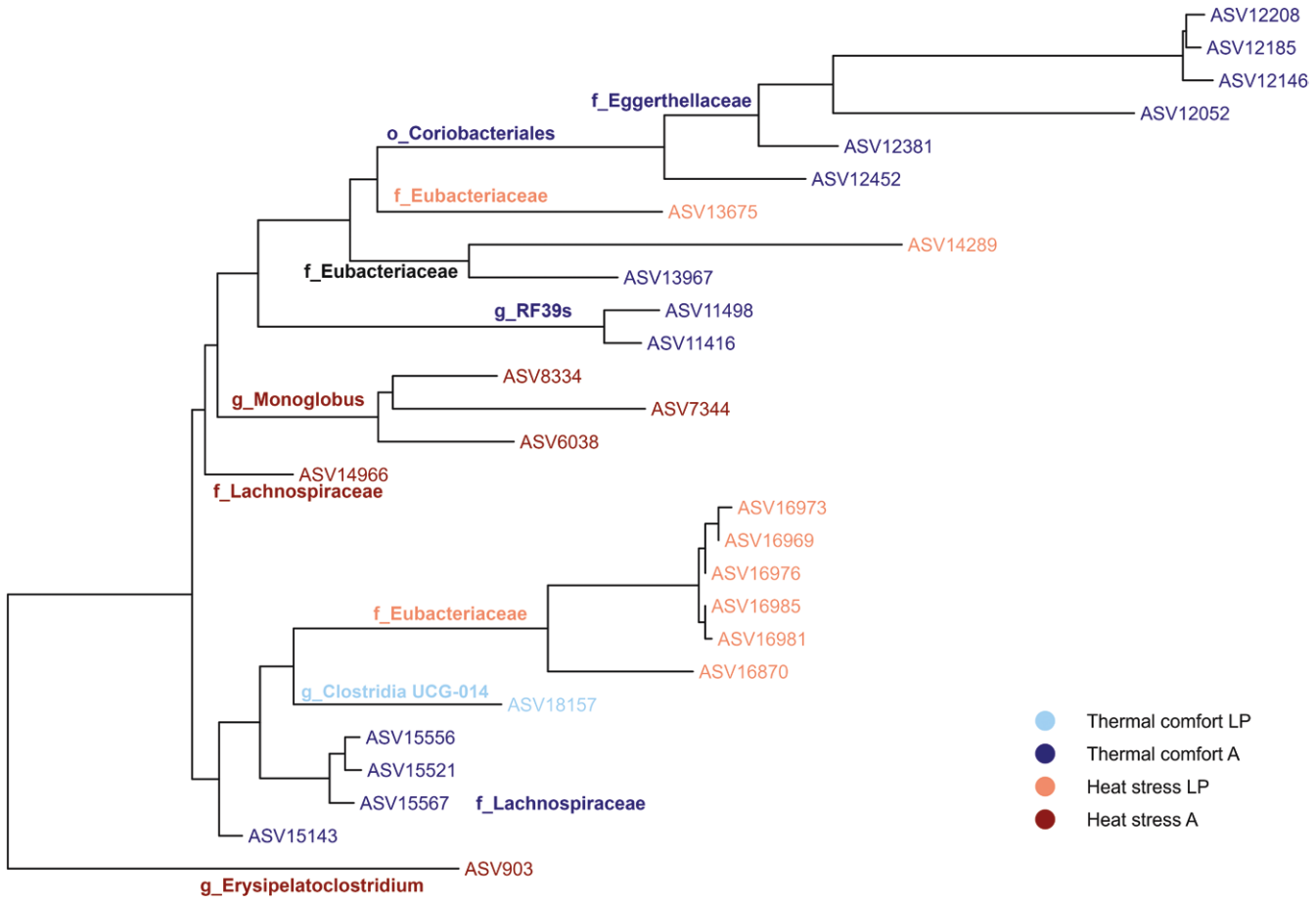

Figure 5. Phylogenetic tree of selected Amplicon Sequence Variants (ASVs) by random forest, BayesC, and PLS-DA for discriminating rabbit genetic lines (A and LP) under heat stress and thermal comfort. Red/orange and blue/sky blue indicate in which line (A and LP) the ASV is more abundant.

group, and the family *Eggerthellaceae*. In the phylogenetic tree, ASVs from the same thermal conditions generally clustered together, in contrast ASVs from different thermal conditions, were dispersed across distant branches (Fig. 5). The different shared taxa and the clustering of ASVs within the same thermal conditions on the phylogenetic tree support the idea that heat stress and thermal comfort drive divergent microbiome profiles, with potential functional implications for rabbits' health and resilience.

Discussion

Microbiome diversity

The 2 maternal genetic lines, A and LP, exhibited differences in both alpha and beta diversity. Line LP demonstrated higher richness compared to line A, particularly under thermal comfort conditions. This aligns with previous findings in the same genetic lines (Biada et al., 2024), which also reported higher alpha diversity in line LP and distinct beta-diversity patterns between the lines. During heat stress, differences were not relevant between the 2 lines. These results suggest that genetic background plays a role in shaping microbiome diversity, but heat stress may override some of these genetic influences on microbiota diversity.

Heat stress induced important changes in both alpha and beta diversity across both rabbit lines, indicating a stress-induced microbiome shift. Specifically, alpha diversity (richness and evenness) increased during heat stress for both lines, with a more pronounced increase in the less longeval line A compared to line LP. This finding diverges from earlier studies in rabbits, which reported no changes in alpha diversity under heat stress compared to thermal comfort (Bai et al., 2021; Liu et al., 2022; Li et al., 2023). However, they did observe compositional changes, such as increased abundances of majority phylum *Firmicutes* (Bai et al., 2021) and *Proteobacteria* (Liu et al., 2022). The observed increase in richness during heat stress in this study aligns with reports in other monogastric animals, including sows (He et al., 2019), pigs (Le Sciellour et al., 2019), and broilers (Shi et al., 2019; Liu et al., 2020). Concerning beta diversity, the study revealed small but significant shifts in microbiome composition between thermal comfort and heat stress conditions, accounting for a small but significant portion of the variance (2.3%). Similar to the findings of this study, differences in microbiome composition visualized using PCA/PCoA have been reported in other rabbit studies (Bai et al., 2021; Liu et al., 2022; Li et al., 2023) and in various monogastric species (Le Sciellour et al., 2019; Xiong et al., 2019; Liu et al., 2020). These results indicate the sensitivity of microbiota diversity to environmental stressors. Heat stress likely disrupts gut integrity, increasing oxidative stress and facilitating pathogen proliferation, which collectively results in a more diverse microbial community and elevated alpha diversity (Patra, 2021). Based on our findings, the disruption of gut barrier and pathogen proliferation appears to be more pronounced in line A, which has lower longevity and is expected to be less resilient compared to line LP.

Additionally, the divergence in microbial community composition among individuals, reflected in altered beta diversity, highlights the heightened responsiveness of microbiota to environmental stress (Patra, 2021).

Microbiome differences between lines

The model results reveal that all 3 models, across both thermal conditions, achieved high accuracies in predicting the 2 genetic lines, A and LP, based on microbiome data. This strongly indicates that the microbiome composition differs between these 2 rabbit lines. Since the animals compared were reared in the same exact environmental conditions, such differences suggest that genetic factors tied to the origin of these lines may shape their microbiomes. These 2 lines were founded using distinct criteria, line A, a standard commercial maternal line, originated from New Zealand White rabbits reared near Valencia, Spain. In contrast, line LP, a robust maternal line established in 2002, was founded from females selected for exceptional reproductive longevity, achieving a minimum of 25 parities with an average prolificacy of 9 or more kits per parity, consistent with commercial averages in Spain (Sánchez et al., 2008). Since their foundation, these lines have been maintained genetically separated and have both been selected for litter size at weaning. Genetic differences between lines can significantly influence microbiome composition. The microbiome is known to be shaped, at least in part, by the host genome (Blekhman et al., 2015; Bonder et al., 2016). Through the concept of the holobiont and its associated hologenome, the host and its microbiome form an interconnected unit where genetic variations in the host affect microbiome structure (Rosenberg and Zilber-Rosenberg, 2018). Moreover, in another study, they found differences in the microbiome sampled at first parity of these 2 lines A and LP (Biada et al., 2024). In this context, the genetic divergence between lines A and LP likely underpins the microbiome differences identified in this study.

Prediction accuracy and thermal conditions

As mentioned above, the microbiome profiles of lines A and LP exhibited differences under both heat stress and thermal comfort conditions. However, prediction accuracies (as line classification ability) were consistently higher under heat stress across all models, suggesting that the microbiome of these genetic lines diverges more distinctly when subjected to heat stress. This indicates a stress-induced response that highlights differences in the adaptive capacities of the 2 lines. These findings were further validated by examining the key variables separating the 2 lines during heat stress and thermal comfort. The final ASVs selected differed between the 2 thermal conditions, with no shared genera. The only overlapping taxonomic units across conditions were the families *Eubacteriaceae* and *Lachnospiraceae*, both large families encompassing diverse species and subspecies (Cotozzolo et al., 2021). Additionally, phylogenetic analysis revealed significant distances between the ASVs selected under each condition. Under thermal comfort, the observed differences between lines A and LP reflect baseline microbiome variation in non-stressful environments, consistent with findings from previous studies (Biada et al., 2024). In contrast, the differences observed during heat stress highlight how the 2 lines, A and LP, respond differently to environmental challenges, with the key selected taxa offering insights into the distinct microbiome dynamics within each line under heat stress. Specifically, the stress-responsive taxa identified included ASVs from the genera *Erysipelatoclostridium*, *Monoglobus*, and the family *Lachnospiraceae* that were more abundant in line A. Conversely, all more abundant ASVs in line LP were exclusively from the family *Eubacteriaceae*.

The genus *Erysipelatoclostridium* has been associated with elevated serum cortisol levels (Feng et al., 2022), indicating a potential link to stress responses. Elevated cortisol can disrupt homeostasis, leading to adverse effects on health and longevity. Additionally, the family *Erysipelotrichaceae*, which includes *Erysipelatoclostridium*, has been implicated in inflammation-related gastrointestinal (GIT) conditions and metabolic disorders (Kaakoush, 2015; Bermingham et al., 2017). On the other hand, *Monoglobus* presents a dual perspective. While its abundance has been positively correlated with beneficial factors such as catalase activity, CD4+ T-cell counts, and butyrate production, key indicators of a healthy immune response (Li et al., 2022; Wang et al., 2022), its overrepresentation might not always be advantageous. Elevated levels of *Monoglobus* have been linked to increased blood ammonia levels, which disrupt intestinal barrier function, increase permeability, and trigger systemic inflammatory responses (Sakurai et al., 2022).

The families *Lachnospiraceae* and *Eubacteriaceae* are among the most abundant families in the rabbit intestinal microbiota (Cotozzolo et al., 2021). They include a diverse range of genera and species with varying metabolic roles (Parte et al., 2011), which complicates the interpretation of their differential abundance between genetic lines. Therefore, we will focus on general trends rather than specific functional interpretations. Members of *Lachnospiraceae* are known to ferment dietary fiber into SCFAs, yet different taxa within this family are also linked to various diseases (Vacca et al., 2020). On the other hand, the role of *Eubacteriaceae* in rabbits remains poorly understood. However, *Eubacterium*, the primary genus in this family (Read et al., 2019), is associated with butyrate production, a critical process for energy homeostasis, colonic motility, immunomodulation, and the suppression of gut inflammation (Mukherjee et al., 2020).

In summary, under heat stress, line A exhibited higher abundance of potentially detrimental taxa such as *Erysipelatoclostridium* and *Monoglobus*, suggesting an altered or imbalanced immune and metabolic state. As mentioned earlier, line A is a standard commercial maternal line. In contrast, line LP, a robust maternal line established in 2002, was founded using longevity criteria (Sánchez et al., 2008). Since its foundation, line LP has consistently demonstrated superior longevity and is expected to exhibit greater resilience. Supporting this, does from line A are almost 2 times more likely to die or be culled compared to those from Line LP (El Nagar et al., 2021). These microbiota differences may reflect physiological disparities between the lines and offer insight into their differing resilience to heat stress. Such differences may also help explain the observed variation in longevity, with these taxa potentially serving as biomarkers for stress susceptibility and contributors to reduced longevity in line A.

Comparative of PLS-DA, Random Forest, and BayesC models

In evaluating the predictive performance of Random Forest, PLS-DA, and BayesC, our results indicate that all 3 models performed well across thermal conditions. Among these, PLS-DA exhibited slightly higher predictive accuracy, as

evidenced by consistently higher AUC values. Ruiz-Perez et al. (2020) through simulation and real-life microbiome datasets, demonstrated the effectiveness of PLS-DA for both classification and variable selection. Their findings highlighted that PLS-DA works well when signal features exhibit a clustered distribution, even key features are hidden among a large number of noise. Moreover, in a comparative study, Troll et al. (2020) found that PLS outperformed Random Forest in regression problems involving gut microbiome data. BayesC also performed closely to PLS-DA in our study, but to our knowledge, BayesC has not been applied to microbiome classification tasks. However, related Bayesian Alphabet models, such as Bayesian lasso, have been used for microbiome data analysis. For instance, Maltecca et al. (2019) applied Bayesian lasso to predict growth and carcass traits in microbiome data, demonstrating comparable performance to Random Forest in terms of prediction accuracy. Despite our results, Random Forest remains one of the most widely used ML techniques in microbiome research. According to Iadanza et al. (2020), it often generates the highest prediction accuracies compared to other ML and deep learning methods. Contrary to our findings, Papoutsoglou et al. (2023) showed that while preprocessing techniques such as filtering low-abundance taxa and performing CLR transformation improved the accuracy of PLS-DA, Random Forest still outperformed PLS-DA based on F1 scores.

Regarding variable selection, the majority of selected variables were shared across the 3 models, indicating consistency in their ability to identify key taxa. However, there were differences in the patterns of variable selection. Random Forest, for instance, tended to select fewer taxa. Notably, the variables it selected to predict the genetic lines (A and LP) in heat stress were more distinct, both taxonomically and phylogenetically, from those selected in thermal comfort. This distinction provided a clearer differentiation of how different microbiota reacts differently within each line in face of heat stress. In contrast, PLS-DA followed by BayesC identified a broader range of taxa. While these additional variables may have contributed to their improved predictive performance in comparison to Random Forest, many of them were not as biologically or taxonomically distinct. This suggests that Random Forest focuses on a narrower but more meaningful subset of taxa, while PLS-DA and BayesC capture a more extensive array of features, potentially enhancing prediction but at the cost of including less relevant variables. In the case of PLS-DA, this broader selection can be attributed to its ability to capture latent variables (linear combinations of original features that maximize covariance with the dependent variable). Although this approach can enhance prediction performance, it may also include variables that are less biologically relevant. By comparison, both BayesC and Random Forest assume that each feature contributes independently and additively to the prediction, but Random Forest additionally captures interactions, both conditionally within individual trees and through its ensemble approach. This approach leads to more conservative variable selection by identifying a smaller subset of highly relevant features. As a result, Random Forest is widely recognized as one of the most effective methods for feature selection in microbiome analysis, outperforming other ML techniques, including PLS (Statnikov et al. 2013; Liu et al. 2017; Namkung, 2020). In a review study, It consistently selected fewer number of variables across 8 datasets of microbiome data (Statnikov et al., 2013).

In summary, while PLS-DA and BayesC demonstrated slightly better predictive performance, Random Forest selected a smaller but more taxonomically and phylogenetically relevant subset of variables. This distinction highlights the complementary strengths of these methods: PLS-DA and BayesC captured a broader range of features, enhancing predictive accuracy, while Random Forest focused on core taxa critical for distinguishing between conditions, making it highly effective for targeted feature selection. The differing variable selection patterns among the models provide valuable insights into the microbial architecture of the genetic lines and their responses to heat stress. Specifically, PLS-DA's ability to identify latent variables offers a comprehensive view of microbiome dynamics, whereas Random Forest's conservative selection emphasizes the most biologically relevant taxa.

Conclusion

Heat stress altered the microbiome in maternal rabbit lines by increasing alpha diversity and shifting beta diversity, likely due to gut barrier disruption that increased pathogen proliferation. While line LP had higher microbiome richness under normal conditions, heat stress caused a greater diversity increase in the less robust line A, narrowing the differences. Line A also showed higher levels of potential stress-associated taxa like *Erysipelatoclostridium* and *Monoglobus*, indicating a line-specific microbial response. Predictive models revealed that microbiome responses differed between lines under thermal stress, with PLS-DA providing the highest accuracy and capturing more features, while Random Forest identified fewer but biologically relevant taxa. These complementary methods enhance our understanding of how microbiomes mediate stress responses.

These findings support the potential of soft fecal microbiota as biomarkers of heat-stress resilience in livestock, encouraging further validation across populations and integration into resilience-oriented breeding programs. They also underscore the importance of genetic and environmental interactions in shaping the microbiome and guiding strategies to safeguard animal health under climate stress.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Acknowledgements

Iliyass BIADA acknowledges a CIACIF/2021/005 scholarship from Generalitat Valenciana. The authors also acknowledge AGROALNEXT program and Tamara Romero Rueda and Rafael Torres Lozano for their collaboration in the collection of samples and handling of the animals. Funding for open access charge: CRUE-Universitat Politècnica de València. This research is part of the AGROALNEXT program (AGROALNEXT/2022/037) and has been financed by MCIN with funds from the European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat Valenciana. Additionally, this work was supported by COST Action CA22112, which provided funding for a Short-Term Scientific Mission (STSM) awarded to Iliyass Biada.

Conflict of interest statement. The authors declare that they have no competing interests.

Availability of data and materials

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB83677.

Author contributions

Iliyass Biada (Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review & editing), Francesco Tiezzi (Methodology, Writing—review & editing), Noelia Ibañez-Escriche (Conceptualization, Methodology, Writing—review & editing), Maria Luz Garcia (Funding acquisition, Project administration, Writing—review & editing), María Jose Argente (Funding acquisition, Project administration), and María Antonia Santacreu (Conceptualization, Funding acquisition, Methodology, Writing—review & editing).

Literature Cited

Albert, J. H., and S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88:669–679. doi:10.1080/01621459.1993.10476321

Bai, X., Y. Shi, L. Tang, L. Chen, H. Fan, H. Wang, J. Wang, X. Jia, S. Chen, and S. Lai. 2021. Heat stress affects faecal microbial and metabolic alterations of rabbits. Front. Microbiol. 12:817615. doi:10.3389/fmicb.2021.817615

Bermingham, E. N., P. Maclean, D. G. Thomas, N. J. Cave, and W. Young. 2017. Key bacterial families (Clostridiaceae, Erysipelotrichaceae and Bacteroidaceae) are related to the digestion of protein and energy in dogs. PeerJ. 5:e3019. doi:10.7717/peerj.3019

Biada, I., N. Ibáñez-Escriche, A. Blasco, C. Casto-Rebollo, and M. A. Santacreu. 2024. Microbiome composition as a potential predictor of longevity in rabbits. Genet. Sel. Evol. 56:25. doi:10.1186/s12711-024-00895-6

Blekhman, R., J. K. Goodrich, K. Huang, Q. Sun, R. Bukowski, J. T. Bell, T. D. Spector, A. Keinan, R. E. Ley, D. Gevers, et al. 2015. Host genetic variation impacts microbiome composition across human body sites. Genome Biol. 16:191. doi:10.1186/s13059-015-0759-1

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, and F. Asnicar, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37:852–857. doi:10.1016/j.patrec.2010.1038/s41587-019-0209-9

Bokulich, N. A., S. Subramanian, J. J. Faith, D. Gevers, J. I. Gordon, R. Knight, D. A. Mills, and J. G. Caporaso. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat. Methods. 10:57–59. doi:10.1038/nmeth.2276

Bonder, M. J., A. Kurilshikov, E. F. Tigchelaar, Z. Mujagic, F. Imhann, A. V. Vila, P. Deelen, T. Vatanen, M. Schirmer, S. P. Smeekens, et al. 2016. The effect of host genetics on the gut microbiome. Nat. Genet. 48:1407–1412. doi:10.1038/ng.3663

Breiman, L. 2001. Random forests. Mach. Learn. 45:5–32. doi:10.1023/A:1010933404324

Bürkner, P. -C. 2017. brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw. 80:1–28. doi:10.18637/jss.v080.i01

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583. doi:10.1038/nmeth.3869

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34:i884–i890. doi:10.1093/bioinformatics/bty560

Cotozzolo, E., P. Cremonesi, G. Curone, L. Menchetti, F. Riva, F. Biscarini, M. L. Marongiu, M. Castrica, B. Castiglioni, D. Miraglia, et al. 2021. Characterization of bacterial microbiota composition along the gastrointestinal tract in rabbits. Animals (Basel). 11:31. doi:10.3390/ani11010031

Cullere, M., and A. Dalle Zotte. 2018. Rabbit meat production and consumption: state of knowledge and future perspectives. Meat Sci. 143:137–146. doi:10.1016/j.meatsci.2018.04.029

Dalmau, A., B. Catanese, O. Rafel, P. Rodriguez, C. Fuentes, P. Llonch, E. Mainau, A. Velarde, J. Ramón, E. Taberner, et al. 2014. Effect of high temperatures on breeding rabbit behaviour. Anim. Prod. Sci. 55:1207–1214. doi:10.1071/an13440

Ebeid, T. A., H. S. Aljabeili, I. H. Al-Homidan, Z. Volek, and H. Barakat. 2023. Ramifications of heat stress on rabbit production and role of nutraceuticals in alleviating its negative impacts: an updated review. Antioxidants (Basel). 12:1407. doi:10.3390/antiox12071407

Egea, C. 2024. Estrés por calor, ¿es posible minimizarlo? Bol. Cunicult. 212:6–8. https://www.navarra.es/nr/rdonlyres/af2bd786-0a6d-4564-9076-be42220b4843/225685/calorytrabajoprofesional.pdf

El Nagar, A. G., M. Baselga, C. Mínguez, M. A. Santacreu, and J. P. Sánchez. 2021. Functional longevity in five rabbit lines founded on different criteria: comparison at foundation and at fixed times after selection. J. Anim. Breed. Genet. 138:508–517. doi:10.1111/jbg.12543

Feng, Y., H. Fan, X. Liang, X. Wang, G. Gao, and S. Gun. 2022. Environmental enrichment changes rabbits' behavior, serum hormone level and further affects cecal microbiota. PeerJ. 10:e13068. doi:10.7717/peerj.13068

Funosas, G., X. Triadó-Margarit, F. Castro, R. Villafuerte, M. Delibes-Mateos, C. Rouco, and E. O. Casamayor. 2021. Individual fate and gut microbiome composition in the European wild rabbit (Oryctolagus cuniculus). Sci. Rep. 11:766. doi:10.1038/s41598-020-80782-4

Genuer, R., J. -M. Poggi, and C. Tuleau-Malot. 2010. Variable selection using random forests. Pattern Recognit. Lett. 31:2225–2236. doi:10.1016/j.patrec.2010.03.014

Gerber, G. K. 2014. The dynamic microbiome. FEBS Lett. 588:4131–4139. doi:10.1016/j.febslet.2014.02.037

Giorgi, F., and P. Lionello. 2008. Climate change projections for the Mediterranean region. Glob. Planet. Change. 63:90–104. doi:10.1016/j.gloplacha.2007.09.005

Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: and this is not optional. Front. Microbiol. 8:2224. doi:10.3389/fmicb.2017.02224

Greenacre, M., M. Martínez-Álvaro, and A. Blasco. 2021. Compositional Data Analysis of Microbiome and Any-Omics Datasets: A Validation of the Additive Logratio Transformation. Front Microbiol. 12. doi:10.1016/j.patrec.2010.3389/fmicb.2021.727398

Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. BMC Bioinf. 12:186. doi:10.1186/1471-2105-12-186

He, J., H. Guo, W. Zheng, Y. Xue, and R. Zhao. 2019. Heat stress affects fecal microbial and metabolic alterations of primiparous sows during late gestation. J. Anim. Sci. Biotechnol. 10(1). doi:10.1016/j.patrec.2010.1186/s40104-019-0391-0

He, Y., C. Maltecca, and F. Tiezzi. 2021. Potential use of gut microbiota composition as a biomarker of heat stress in monogastric species: a review. Animals (Basel). 11:1833. doi:10.3390/ani11061833

Iadanza, E., R. Fabbri, D. Bašić-ČiČak, A. Amedei, and J. H. Telalovic. 2020. Gut microbiota and artificial intelligence approaches: a scoping review. Health Technol. 10:1343–1358. doi:10.1007/s12553-020-00486-7

Kaakoush, N. O. 2015. Insights into the role of Erysipelotrichaceae in the human host. Front. Cell. Infect. Microbiol. 5:84. doi:10.3389/fcimb.2015.00084

Le Sciellour, M., O. Zemb, I. Hochu, J. Riquet, H. Gilbert, M. Giorgi, Y. Billon, and J. -L. Gourdine. 2019. Effect of chronic and acute heat challenges on fecal microbiota composition, production, and thermoregulation traits in growing pigs1,2. J. Anim. Sci. 97:3845–3858. doi:10.1016/j.patrec.2010.1093/jas/skz222

Li, R., X. Yi, J. Yang, Z. Zhu, Y. Wang, X. Liu, X. Huang, Y. Wan, X. Fu, W. Shu, et al. 2022. Gut microbiome signatures in the progression

of Hepatitis B virus-induced liver disease. Front. Microbiol. 13:916061. doi:10.3389/fmicb.2022.916061

Li, K., M. M. Abdelsattar, M. Gu, W. Zhao, H. Liu, Y. Li, P. Guo, C. Huang, and S. Fang. 2023. The Effects of Temperature and Humidity Index on Growth Performance, Colon Microbiota, and Serum Metabolome of Ira Rabbits. Animals 13(12):1971. doi:10.1016/j.patrec.2010.3390/ani13121971

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News, 2, 18-22. http://CRAN.R-project.org/doc/Rnews/

Liu, Y., S. Tang, C. Fernandez-Lozano, C. R. Munteanu, A. Pazos, Y. Yu, Z. Tan, and H. González-Díaz. 2017. Experimental study and Random Forest prediction model of microbiome cell surface hydrophobicity. Expert Syst. Appl. 72:306–316. doi:10.1016/j.eswa.2016.10.058

Liu, G., H. Zhu, T. Ma, Z. Yan, Y. Zhang, Y. Geng, and Y. Zhu. 2020. Effect of chronic cyclic heat stress on the intestinal morphology, oxidative status and cecal bacterial communities in broilers. J. Therm. Biol. 91:102619. doi:10.1016/j.patrec.2010.1016/j.jthermbio.2020.102619

Liu, H., B. Zhang, Li Fan, L. Liu, T. Yang, H. Zhang, et al. 2022. Effects of heat stress on growth performance, carcass traits, serum metabolism, and intestinal microflora of meat rabbits. Front. Microbiol. 13:998095. doi:10.3389/fmicb.2022.998095

Maltecca, C., D. Lu, C. Schillebeeckx, N. P. McNulty, C. Schwab, C. Shull, and F. Tiezzi. 2019. Predicting growth and carcass traits in swine using microbiome data and machine learning algorithms. Sci. Rep. 9:6574. doi:10.1038/s41598-019-43031-x

Marai, I. F. M., A. A. M. Habeeb, and A. E. Gad. 2002. Rabbits' productive, reproductive and physiological performance traits as affected by heat stress: a review. Livest. Prod. Sci. 78:71–90. doi:10.1016/s0301-6226(02)00091-x

Marcos-Zambrano, L. J., K. Karaduzovic-Hadziabdic, T. Loncar Turukalo, P. Przymus, V. Trajkovik, O. Aasmets, M. Berland, A. Gruca, J. Hasic, K. Hron, et al. 2021. Applications of machine learning in human microbiome studies: a review on feature selection, biomarker identification, disease prediction and treatment. Front. Microbiol. 12:634511. doi:10.3389/fmicb.2021.634511

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 17:10–12. doi:10.14806/ej.17.1.200

McMurdie, P. J., and S. Holmes. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 8:e61217. doi:10.1371/journal.pone.0061217

Mukherjee, A., C. Lordan, R. P. Ross, and P. D. Cotter. 2020. Gut microbes from the phylogenetically diverse genus Eubacterium and their various contributions to gut health. Gut Microbes. 12:1802866. doi:10.1080/19490976.2020.1802866

Namkung, J. 2020. Machine learning methods for microbiome studies. J. Microbiol. 58:206–216. doi:10.1007/s12275-020-0066-8

Nearing, J. T., G. M. Douglas, M. G. Hayes, J. MacDonald, D. K. Desai, N. Allward, C. M. A. Jones, R. J. Wright, A. S. Dhanani, A. M. Comeau, et al. 2022. Microbiome differential abundance methods produce different results across 38 datasets. Nat. Commun. 13:342. doi:10.1038/s41467-022-28034-z

Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, et al. 2024. vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/index.html

Papoutsoglou, G., S. Tarazona, M. B. Lopes, T. Klammsteiner, E. Ibrahimi, J. Eckenberger, P. Novielli, A. Tonda, A. Simeon, R. Shigdel, et al. 2023. Machine learning approaches in microbiome research: challenges and best practices. Front. Microbiol. 14. doi:10.3389/fmicb.2023.1261889

Parte, A., N. R. Krieg, W. Ludwig, W. B. Whitman, B. P. Hedlund, B. J. Paster, et al. 2011. Bergey's manual of systematic bacteriology: volume 4: the Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes. Springer Science & Business Media.

Patra, A. K. 2021. Heat stress on microbiota composition, barrier integrity, and nutrient transport in gut, production performance, and its amelioration in farm animals. J. Anim. Sci. Technol. 63:211–247. doi:10.1016/j.patrec.2010.5187/jast.2021.e48

Pérez, P., and G. de los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics. 198:483–495. doi:10.1534/genetics.114.164442

Priyadarshini, M., K. U. Kotlo, P. K. Dudeja, and B. T. Layden. 2018. Role of short chain fatty acid receptors in intestinal physiology and pathophysiology. Compr. Physiol. 8:1091–1115. doi:10.1002/cphy.c170050

Read, T., L. Fortun-Lamothe, G. Pascal, M. Le Boulch, L. Cauquil, B. Gabinaud, et al. 2019. Diversity and co-occurrence pattern analysis of cecal microbiota establishment at the onset of solid feeding in young rabbits. Front. Microbiol. 10:973. doi:10.3389/fmicb.2019.00973

Rohart, F., B. Gautier, A. Singh, and K. -A. L. Cao. 2017. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13:e1005752. doi:10.1371/journal.pcbi.1005752

Rosenberg, E., and I. Zilber-Rosenberg. 2018. The hologenome concept of evolution after 10 years. Microbiome. 6:78. doi:10.1186/s40168-018-0457-9

Ruiz-Perez, D., H. Guan, P. Madhivanan, K. Mathee, and G. Narasimhan. 2020. So you think you can PLS-DA? BMC Bioinf. 21:2. doi:10.1186/s12859-019-3310-7

Sakurai, K., T. Toshimitsu, E. Okada, S. Anzai, I. Shiraishi, N. Inamura, S. Kobayashi, T. Sashihara, and T. Hisatsune. 2022. Effects of Lactiplantibacillus plantarum OLL2712 on memory function in older adults with declining memory: a randomized placebo-controlled trial. Nutrients. 14:4300. doi:10.3390/nu14204300

Sánchez, J. P., P. Theilgaard, C. Mínguez, and M. Baselga. 2008. Constitution and evaluation of a long-lived productive rabbit line. J. Anim. Sci. 86:515–525. doi:10.2527/jas.2007-0217

Shi, D., L. Bai, Q. Qu, S. Zhou, M. Yang, S. Guo, and Q. Li. 2019. Impact of gut microbiota structure in heat-stressed broilers. Poultry Science 98(6):2405–2413. doi:10.1016/j.patrec.2010.3382/ps/pez026

Statnikov, A., M. Henaff, V. Narendra, K. Konganti, Z. Li, L. Yang, Z. Pei, M. J. Blaser, C. F. Aliferis, and A. V. Alekseyenko. 2013. A comprehensive evaluation of multicategory classification methods for microbiomic data. Microbiome. 1:11. doi:10.1186/2049-2618-1-11

Topçuoğlu, B. D., N. A. Lesniak, M. T. Ruffin, J. Wiens, and P. D. Schloss. 2020. A framework for effective application of machine learning to microbiome-based classification problems. mBio. 11. doi:10.1128/mbio.00434-20

Troll, M., S. Brandmaier, S. Reitmeier, J. Adam, S. Sharma, A. Sommer, M. -A. Bind, K. Neuhaus, T. Clavel, J. Adamski, et al. 2020. Investigation of adiposity measures and Operational Taxonomic unit (OTU) data transformation procedures in stool samples from a German Cohort Study using machine learning algorithms. Microorganisms. 8:547. doi:10.3390/microorganisms8040547

Vacca, M., G. Celano, F. M. Calabrese, P. Portincasa, M. Gobbetti, and M. De Angelis. 2020. The controversial role of human gut Lachnospiraceae. Microorganisms. 8:573. doi:10.3390/microorganisms8040573

Xiong, Y., H. Yi, Q. Wu, and Z. Jiang. 2019. Effects of acute heat stress on intestinal microbiota in grow-finishing pigs, and associations with feed intake and serum profile. J. Appl. Microbiol. 128 :840–852. doi:10.1016/j.patrec.2010.1111/jam.14504

Wang, Q., S. -X. Zhang, M. -J. Chang, J. Qiao, C. -H. Wang, X. -F. Li, Q. Yu, and P. -F. He. 2022. Characteristics of the gut microbiome and its relationship with peripheral CD4+ T cell subpopulations and cytokines in rheumatoid arthritis. Front. Microbiol. 13:799602. doi:10.3389/fmicb.2022.799602

Zeferino, C. P., C. M. Komiyama, S. Fernandes, J. R. Sartori, P. S. S. Teixeira, and A. S. A. M. T. Moura. 2013. Carcass and meat quality traits of rabbits under heat stress. Animal. 7:518–523. doi:10.1017/S1751731112001838