

Fecal Microbiota and Their Association With Heat Stress in Bos Taurus

Bartosz Czech (✉ bartosz.czech@upwr.edu.pl)

Wroclaw University of Environmental and Life Sciences

Joanna Szyda

Wroclaw University of Environmental and Life Sciences

Kai Wang

China Agricultural University

Hanpeng Luo

China Agricultural University

Yachun Wang

China Agricultural University

Research Article

Keywords: 16S rRNA gene, heat stress, fecal microbiome, sequencing, V3-V4 regions, differential abundance

Posted Date: December 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1140713/v1>

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

[Read Full License](#)

RESEARCH

Fecal microbiota and their association with heat stress in *Bos taurus*

Bartosz Czech^{1*}, Joanna Szyda^{1,2}, Kai Wang³, Hanpeng Luo³ and Yachun Wang³

*Correspondence:

bartosz.czech@upwr.edu.pl

¹Biostatistics Group, Department of Genetics, Wrocław University of Environmental and Life Sciences, Wrocław, Poland

Full list of author information is available at the end of the article

Abstract

Background: Humans have been influencing climate changes by burning fossil fuels, farming livestock, and cutting down rainforests, which has led to global temperature rise. This problem of global warming affects animals by causing heat stress, which negatively affects their health, biological functions, and reproduction. On the molecular level, it has been proved that heat stress changes the expression level of genes and therefore causes changes in proteome and metabolome. The importance of a microbiome in many studies showed that it is considered as individuals' "second genome". Physiological changes caused by heat stress may impact the microbiome composition.

Results: In this study, we identified fecal microbiota associated with heat stress that was quantified by three metrics – rectal temperature, drooling, and respiratory scores and represented by their Estimated Breeding Values. For this purpose, the 16S rRNA sequencing technique was used. We analyzed the microbiota from 136 fecal samples of Chinese Holstein cows through a 16S rRNA gene sequencing approach. Sequence data were processed using a pipeline involving QIIME2 software together with SILVA database. Statistical modeling was performed using a negative binomial regression. The analysis revealed the total number of 24 genera and 12 phyla associated with heat stress metrics. *Rhizobium* and *Pseudobutyrvibrio* turned out to be the most significant genera, while *Acidobacteria* and *Gemmatimonadetes* were the most significant phyla. Phylogenetic analysis revealed that that three heat stress indicators quantify different metabolic ways of animals' reaction to heat stress. Other studies already identified that those genera had significantly increased abundance in mice exposed to stressor-induced changes.

Conclusions: This study provides insights into the analysis of microbiome composition in cattle using heat stress measured as a continuous variable. The bacteria highly associated with heat stress were highlighted and can be used as biomarkers in further microbiological studies.

Keywords: 16S rRNA gene; heat stress; fecal microbiome; sequencing; V3-V4 regions; differential abundance

Introduction

Global warming and the resulting long-term increase in temperatures are the main cause of heat stress in mammals [1]. Moreover, selection towards high production yield in livestock and associated with its high metabolic load is an additional factor that makes livestock especially prone to overheating. Heat stress negatively affects health, reproduction, and other biological functions [2]. Specifically, in dairy cattle, heat stress impedes milk production, welfare, and growth [3]. Unfortunately, the phenomenon of heat stress is common in current ages, and we should understand how its long-term susceptibility affects organisms. On the genomic level, heat stress is manifested by transcriptional and post-transcriptional regulation of heat stress-associated genes [4]. It is known that *Bos indicus* has greater heat tolerance than *Bos taurus* [5], which indicates a genetic component of heat resistance. A few genes responsible for thermotolerance in dairy cattle – HSF1, MAPK8IP1, and CDKN1B have been recently identified [6]. However, the effect of heat stress on animal-associated microbiotas is not well known. In cattle genomics, bacteria are the main cause of mastitis – one of the most prevalent diseases of dairy cattle [7]. In general, the composition of gut microbiota depends on multiple factors – genetic [8], dietary [9] and environmental [10].

Heat stress belongs to environmental factors that may change the composition of the microbiota. Studies in cattle reported microbial species which abundance depends on heat stress conditions. Chen and colleagues [11] reported the effect of heat stress on physiological characteristics and circulation levels of immune activity and the microbiome. In an experimental study Zhao and colleagues [12] identified bacterial species in the rumen microbiome associated with heat stress. In their study it was found that heat stress has no effect on both alpha and beta diversity, however the effect on the richness of microbiota was identified, especially significant increase in the abundance of *Streptococcus*, *Enterobacteriaceae*, *Ruminobacter*, *Treponema* and *Bacteroidaceae*. Sales and colleagues [13] in his study also reported that heat stress influenced microbiota in beef cattle rumen. Particularly, they found genera *Flavonifractor*, *Treponema*, *Ruminococcus*, and *Carnobacterium* significantly associated with heat stress. However, assessing the composition of microbiota in farm animals' environments is important to study its association with heat stress under breeding conditions. Moreover, the categorization between heat stress and normal conditions is a simplification. The level of an animal's heat stress is a continuous variable and thus can be assessed using quantitative metrics. This however implies non-standard statistical modeling of the association between heat stress traits

and microbiome as compared to the experimental-based, case-control setup. In our study, we focused on the identification of bacteria associated with heat stress measured by drooling score, rectal temperature, and respiratory score, expressed by the estimated breeding values.

Material and methods

The material comprises 136 fecal samples of 136 Chinese Holstein cows, which were collected in 2017, 2018, and 2019 directly in herds belonging to Beijing Shouong Livestock Development Co., Ltd. The cows from the same year had been fed with exactly the same total mixed ration diet for over 1 month and the cows from different years were fed with different total mixed ration diets with small change; however, all diets were based on corn silage and concentrate, and all the cows were fed *ad libitum*.

Fecal samples were collected directly from the cow's rectum using a method a bit similar to rectal inspection. Around 7 AM, before the new feed is provided to the cow, is the time point we selected to take fecal samples, as cows are calmer after a good rest during the night, samples are easier to keep during the morning cooler period in summer, also we can be more sure that cows are in similar digestion stage without stimulation from feed for a relatively long time, and feces accumulated in the cow's rectum. By wearing a disposable plastic long-armed glove, the sampler inserts his hand and arm into the cow's rectum, first removed the outer part of feces accumulated in the rectum, and then grabbed a certain amount of feces from the inner part by hand, after a few feces mixing actions in the rectum. A disposable plastic long-armed glove can be used once for each cow. After the sampler's hand holding feces moves out of the cow's rectum, one can turn the glove outside in. Feces will naturally accumulate into the finger parts of the glove. By cutting a small hole at the tip of the finger parts of the glove, a fecal sample can be easily transferred into a properly labeled sterile 5 ml cryopreservation tube. Since big particles within feces may precipitate at the bottom of the finger parts of the glove, the very first part of the fecal sample can be discarded. Fecal samples are normally then placed on dry ice maximum for 3-4 h before they stored at -80°C at the laboratory.

The thermal environment during the sampling process was measured by temperature, humidity, and Temperature Humidity Index (THI) presented in Table 1.

The procedures of DNA extraction, amplification, and sequencing were completed by Wekemo Tech Co., Ltd. (Shenzhen, China). Microbial DNA was extracted from fecal samples using the E.Z.N.A. soil DNA Kit (Omega Bio-tek, Norcross, GA,

Table 1 Characteristic of the thermal environment.

sampling date	Temperature (Td)	Humidity (RH)	THI
15 August, 2017	26.58	0.81	77.46
14 August, 2018	27.08	0.82	78.52
27 July, 2019	31.64	0.60	82.01

U.S.) according to manufacturer's protocols. The final DNA concentration and purification were determined by NanoDrop 2000 UV-vis spectrophotometer (Thermo Scientific, Wilmington, USA), and DNA quality was checked by 1% agarose gel electrophoresis. The V3-V4 hypervariable regions of the bacteria 16S rRNA gene were amplified with primers 338F(5'-ACTCCTACGGGAGGCAGCAG-3') and 806R(5'-GGACTACHVGGGTWTCTAAT-3') (for samples picked in 2017) as well as 341F(5'-CCTAYGGGRBGCASCAG-3') and 806R(5'-GGACTACNNGGGTATCTAAT-3') (for samples picked in 2018 and 2019) by thermocycler PCR system (GeneAmp 9700, ABI, USA). The PCR reactions were conducted using the following program: 3 min of denaturation at 95 °C, 27 cycles of 30 s at 95 °C, 30s for annealing at 55 °C, and 45s for elongation at 72 °C, and a final extension at 72 °C for 10 min. PCR reactions were performed in triplicate 20 μ L mixture containing 4 μ L of 5 \times FastPfu Buffer, 2 μ L of 2.5 mM dNTPs, 0.8 μ L of each primer (5 μ M), 0.4 μ L of FastPfu Polymerase and 10 ng of template DNA. The resulted PCR products were extracted from a 2% agarose gel and further purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, USA) and quantified using QuantiFluorTM-ST (Promega, USA) according to the manufacturer's protocol. Amplicons were sequenced using the HiSeq-PE250 (samples picked in 2017) and MiSeq-PE300 (samples picked in 2018 and 2019) Illumina platforms in paired-end modes. Part of the sequence data analyzed previously by Zhang and colleagues [14] was used in this study.

All the heat stress phenotypes were measured as it was described by Luo and colleagues [15]. In order to correct for the environmental effects that may affect phenotype values, cows' response to heat was expressed by breeding values for rectal temperature (RT), drooling (DS), and respiratory scores (RS) estimated using the following model:

$$y_{ijklqno} = \mu + fy m_i + p_j + s_k + m_l + t_q + thi + a_n + pe_n + \epsilon_{ijklqno}, \quad (1)$$

where $y_{ijklqno}$ refers to phenotype (RS, DS or RT), μ is the population mean, fym_i is the fixed effect for farm-year, p_j is the fixed effect of parity, s_k is the fixed effect of lactation stage, m_l is the fixed effect of the indication if the animal is before or after milking, t_q is the fixed effect of testing time (morning or afternoon), thi is the fixed effect of temperature-humidity index, a_n is the animal additive genetic effect, pe_n is the permanent environmental effect, and $\epsilon_{ijklqno}$ is the random residual. The covariance matrix of random effects has the following structure:

$$\text{var}\left(\begin{bmatrix} a \\ pe \\ \epsilon \end{bmatrix}\right) = \begin{bmatrix} A \otimes \sigma_a^2 & 0 & 0 \\ 0 & I \otimes \sigma_{pe}^2 & 0 \\ 0 & 0 & I \otimes \sigma_\epsilon^2 \end{bmatrix}. \quad (2)$$

The total number of 155 cows were used to estimate the breeding values. The reliability of calculate EBVs was presented in Table 2.

Table 2 Descriptive statistics of the reliability of the estimated breeding values.

phenotype	mean	median	standard deviation
rectal temperature	0.41	0.44	0.10
respiratory score	0.40	0.43	0.10
drooling score	0.33	0.35	0.09

Furthermore, the breeding values were additionally corrected by deregression [16] in order to remove the ancestral information from the EBVs.

Processing of sequencing data

The first step of the analysis included quality control of sequenced data. For this purpose, the FastQC [17] software was used. Then, poor quality reads and adapter sequences were removed using Trim Galore [18]. Followingly, cleaned reads were processed using the QIIME 2 [19] software. First of all, data were dereplicated – reads that are 100% the same will be pooled together. Next, reads are were de-noised – reads that occur very rarely are considered as PCR errors are removed as well as chimeric sequences and singletons. Those steps were done using DADA2 algorithm [20], implemented in QIIME 2. Afterward, the Amplicon Sequence Variants (ASVs) table that represents counts of occurrence of a given sequence in a sample was created. Diversity within samples (α -diversity) was calculated using Simpson's evenness and Shannon's diversity indices using the phyloseq [21] R package. The association of microbes composition with heat stress factors was tested using aGLMM-MiRKAT test implemented in GLMM-MiRKAT R package [22].

The SILVA database [23] was used to classify ASVs taxonomically. For the classification, the naive Bayes algorithm implemented in scikit-learn Python package was used [24].

Since taxa originally assigned by the SILVA database represent different levels of taxonomy, they were aggregated to genera and phyla levels. Genera and phyla with a variance below one and that occurred in less than three samples were excluded from downstream analyses. Such filtered genera and phyla tables were used for the further differential abundance analysis. Additionally, for organoleptic testing of batch effect occurrence, the Uniform Manifold Approximation and Projection (UMAP; [25]) dimensional reduction technique was used to find potential sources of unwanted variability. The phylogenetic tree was generated using the align-to-tree-mafft-fasttree pipeline implemented in QIIME2 software.

Differential abundance analysis

The edgeR [26] R package was used for the normalization of the processed ASV table as well as for statistical modeling of the association between the abundance of microbiota and heat stress indicators. In particular, the Trimmed Mean of M-values (TMM) based normalization [27] was applied. It identified and excluded highly abundant and highly variable genera and phyla, whereupon weighted mean of an abundance of remaining groups were used for the actual normalization [28]. The association between genera/phyla abundance and heat stress indicators was modeled using the negative binomial distribution:

$$K = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e \quad (3)$$

where K represents the counts of reads for a given genus/phylum, β_0 is the intercept, β_1 is the effect of the DRP (expressed by log fold change), X_1 is the design matrix for DRP, β_2 is the effect of the sampling year class, X_2 is the incidence matrix for sampling year, e is the random error.

$$K \sim NB(\mu_k, \phi) \quad (4)$$

where μ_k represents the mean of counts reads, and ϕ is the dispersion parameter such that $Var(K) = \mu_k + \phi\mu_k^2$ calculated using Cox-Reid approximate conditional inference moderated towards the mean [29].

The significance of the effect of DRP on the relative abundance of the genus/phylum was tested using the Likelihood Ratio Test [30]. The test statistic is as follows:

$$LR = -2\ln\left(\frac{L(m_1)}{L(m_2)}\right) \sim \chi^2(1) \quad (5)$$

where m_1 is the reduced model (i.e. the formula (5) without the effect of DRP), and m_2 is the full model (5).

Since each genus/phylum is tested separately, multiple testing correction method was applied using a False Discovery Rate (FDR) [31].

Results

Processing and classification of sequence variants

Table 3 Amplicon Sequence Variants classification results.

Taxonomic level	Number of unique features	Percent of classified reads
Domain	2	100.00
Phylum	29	97.94
Class	72	97.81
Order	114	97.50
Family	156	70.16
Genus	235	20.93
Species	152	2.35

46,825 unique sequences of V3 and V4 regions with a total of 6,486,706 reads were identified and classified. Table 3 summarizes the classification of ASVs based on the different taxonomic levels. In general, reads were classified into two domains – archaea (0.01%) and bacteria (99.99%). Almost all reads could be taxonomically assigned up to order, but species could be assigned only to 2.35% of reads. Further analysis was carried out using genus-level and phylum-level resolution.

Microbiota composition

The general composition of microbiota presented in all samples was presented on bar plot using genus-level and phylum-level resolution. Figure 1 presents the relative abundance of genera with average proportions of more than 0.5%. We can see that *Clostridium* is a genus with the highest relative abundance (15.14%). There were 209 genera with a relative abundance of less than 0.5%. Figure 2 presents the analogous visualization for phyla. *Firmicutes* is a phylum with the highest relative abundance (63.66%). Regarding the less abundant phyla, there were identified 22 phyla with less than 0.5% of the relative abundance.

Figure 1 The relative abundance of genera with average proportions of more than 0.5%.

Figure 2 The relative abundance of phyla with average proportions of more than 0.5%.

Clustering

Genera table was then clustered using UMAP algorithm. The projection of the UMAP coordinates calculated from the ASVs counts matrix on the genus level demonstrates three distinct clusters (Figure 3), which reflect the sampling year. In further analysis, the effect of the sampling year was corrected.

Figure 3 UMAP projection of the ASVs counts matrix on the genus level.

Correlation analysis of diversity metrics

In order to check whether the general diversity of microbiota within samples is correlated with the DRPs, a correlation analysis was performed. Correlations were generally positive, but non-significant (Table 4) with the highest correlation estimated between DRPs for the respiratory score and the Simpson's evenness index (0.27). Overall, non-significant correlations indicate that there is no linear dependence between DRPs and sample diversity calculated based on the abundance of genera.

Relationship of EBVs with microbiomes composition

aGLMM-MiRKAT test was calculated to test the association between the microbial community composition and EBVs. None of the analyzed EBVs showed statistically significant association with the microbial composition. It means that individual genera and phyla should be considered in a statistical model.

Differential abundance analysis

Based on the results of the negative binomial model and considering $FDR \leq 0.05$ 22 genera were significantly associated with rectal temperature with all but one (*Helococcus*) of them showing decreased abundance with the increase of rectal temperature. *Rhizobium* – that represents soil bacteria – was the most associated genus with the rectal temperature. The occurrence of this bacteria might be observed perhaps due to the specific metabolism or the specific plant diet.

Succinivibrio was the only genus associated with respiratory score and *Pseudobutyrvibrio* – with the drooling score. There was no overlap between genera signifi-

Table 4 Pearson correlation coefficients between DRPs and alpha diversity measures expressed by Simpson's evenness and Shannon diversity.

DRP	Simpson's evenness	Shannon diversity
Rectal temperature	0.25	-0.04
Drooling score	0.13	0.23
Respiratory score	0.27	0.11

cant for the three heat stress indicators (Table 5). Differential abundance analysis of phylum (Table 6) showed that 6 phyla were significantly associated with rectal temperature. All of them showed decreased abundance with the increase of rectal temperature. *Fibrobacteres* was the only phylum associated with respiratory score. Surprisingly, for drooling score, 5 differentially abundant phyla were identified. Five of them showed increase abundance with the increase of rectal temperature. Only *Fibrobacteres* showed the decrease abundance. *Fibrobacteres* was significantly associated with both drooling and respiratory scores, while *Nitrospirae*, *Gemmatimonadetes*, *Acidobacteria*, and *Planctomycetes* were significantly associated with both rectal temperature and drooling score.

In order to check the genetic relationship between those associated genera, the phylogenetic tree was created based on the 16S rRNA sequences. The genetic relationship of the associated genera was shown in Figure 4. Colors indicate the association between the genera and the phenotype. We can see, that *Succinivibrio* that was associated with the respiratory score phenotypes created the single clade with all the genera associated with the rectal temperature. Only *Pseudobutyrvibrio* that was associated with the drooling score creates a single, separate clade.

Figure 4 Phylogenetic tree of genera identified in fecal samples. Color indicates significantly associated genera with a given phenotype.

Discussion

This study aimed to identify genera that are associated with the rectal temperature, drooling score, and respiratory score, and in the consequences, associated with heat stress. The quantitative pseudophenotypes were used in order to model animals' microbiomes under conventional production conditions, without setting up a case (heat stress conditions) – control (standard conditions) experiment. Such an approach allows for the estimation of genera effect on heat stress under real conditions underlying dairy herd management.

The general composition of microbiota was not altered by heat stress. Therefore we focussed on single genera as potentially involved in heat stress response. Most

Table 5 Significant differentially abundant genera.

Genus	logFC	FDR
Rectal temperature		
<i>Rhizobium</i>	-16.97	3.88×10^{-5}
<i>Rhodoplanes</i>	-16.36	1.64×10^{-4}
<i>Kaistobacter</i>	-15.95	1.10×10^{-4}
<i>Streptomyces</i>	-15.82	1.10×10^{-4}
<i>Sphingomonas</i>	-15.60	1.37×10^{-4}
<i>Acidovorax</i>	-15.54	2.26×10^{-4}
<i>Nocardia</i>	-15.35	1.10×10^{-4}
<i>Cupriavidus</i>	-15.09	1.64×10^{-4}
<i>Candidatus Solibacter</i>	-14.27	1.15×10^{-3}
<i>Nocardioides</i>	-13.82	2.43×10^{-4}
<i>Brevundimonas</i>	-12.95	5.59×10^{-4}
<i>Kribbella</i>	-12.95	2.47×10^{-3}
<i>Amycolatopsis</i>	-12.34	2.03×10^{-3}
<i>DA101</i>	-12.25	1.75×10^{-3}
<i>Azospira</i>	-11.84	2.03×10^{-3}
<i>Catellatospora</i>	-11.33	8.41×10^{-3}
<i>Reyranella</i>	-11.12	5.69×10^{-3}
<i>Pseudonocardia</i>	-10.89	6.89×10^{-3}
<i>Devosia</i>	-10.64	2.28×10^{-3}
<i>Rhodococcus</i>	-10.23	1.11×10^{-2}
<i>Helcococcus</i>	8.31	8.22×10^{-3}
<i>YRC22</i>	-4.58	2.99×10^{-2}
Respiratory score		
<i>Succinivibrio</i>	6.24	3.33×10^{-2}
Drooling score		
<i>Pseudobutyrvibrio</i>	-16.64	1.68×10^{-3}

of the genera were significantly associated with rectal temperature which might be caused by the fact that samples and measurement came from the same environment (rectum). Since most of the significantly associated genera showed decreased abundance with the increase of heat stress, we can assume, that heat stress favors the inhibition of growth of some microbial populations.

Based on the current literature, Bailey [32] observed a reduced abundance of bacteria in genus *Pseudobutyrvibrio* in mice exposed to stressor-induced changes. Such reduced abundance was also observed by us the association with a drooling score. Baek [33] in his study observed that *Succinivibrio* shows increased abundance in cows under heat stress. In our study, this genus was also associated with the respiratory score metric. Interestingly, *Helcococcus*, the only genus that abundance increased with increasing rectal temperature, has not been reported in studies

Table 6 Significant differentially abundant phyla.

Phylum	logFC	FDR
Rectal temperature		
<i>Acidobacteria</i>	-26.99	1.18×10^{-16}
<i>Gemmatimonadetes</i>	-22.40	5.03×10^{-13}
<i>Chloroflexi</i>	-21.07	2.98×10^{-11}
<i>Nitrospirae</i>	-15.19	1.18×10^{-7}
<i>Planctomycetes</i>	-11.19	5.52×10^{-5}
<i>Euryarchaeota</i>	-6.80	2.99×10^{-2}
Respiratory score		
<i>Fibrobacteres</i>	-8.67	2.58×10^{-2}
Drooling score		
<i>Fibrobacteres</i>	-16.72	5.22×10^{-4}
<i>Nitrospirae</i>	15.73	4.22×10^{-4}
<i>Gemmatimonadetes</i>	15.17	4.69×10^{-4}
<i>Acidobacteria</i>	14.46	4.71×10^{-4}
<i>Planctomycetes</i>	12.50	1.13×10^{-2}

focused on heat stress and any stress-induced conditions, but it was reported as associated with postpartum endometritis by Miranda CasoLuengo [34]. Moreover, [35] showed that *Streptomyces* was reported as a genus with enriched relative abundance in Jersey cows in the normal condition compared to the heat stress condition. It is worthwhile to mention that many genera reported with the association to the rectal temperatures show the high fold change, suggesting that increased rectal temperature has a high impact on microbiota composition. *Proteobacteria* phylum that represents most of the associated genera in our study seems to be the most important phylum in heat stress conditions. Yu [36] already reported that *Proteobacteria* and *Firmicutes* are the most common phyla associated with heat stress conditions. Interestingly, analysis based on the phylum resolution showed that there were overlapping phyla. *Fibrobacteres* turned out to be the significantly associated phylum with respiratory and drooling scores. This phylum was already reported as significant in heat stress analysis of pigs reported by He [37]. *Chloroflexi* and *Planctomycetes* significantly associated with rectal temperature were also reported as a significant phyla in the analysis of short-term acute heat stress on the rumen microbiome of Hanwoo steers [33].

Differences found in microbial compositions and in genera/phyla abundance suggest that those changes might occur due to adapting to climate change. In this study, the abundance of *Fibrobacteres* was decreased due to heat stress. The role of this bacteria is the degradation of plant-based cellulose in ruminants and acetate

production. Ransom-Jones and colleagues [38] reported that glycosyl hydrolases of *Fibrobacteres* may produce carbohydrate activators, including cellulose enzymes and in consequence, cows may produce more energy with acetate in the rumen that can be associated with heat production. Some bacteria (e.g. *Pseudobutyrvibrio*) were described as a part of the microbiome, but their impact on host physiology is not yet known.

Heat stress modeled as a binary variable (i.e. normal vs. stress conditions) provides valuable insights into the understanding of the microbiome association to heat stress, however, it should be beard in mind that the real, production environment of a dairy cow markedly deviates from the experimental conditions. The most obvious differences comprise duration, intensity, and variation in ambient temperatures, which are typically not modeled in experiments. Therefore, our study, despite being more challenging from the analytical perspective, provided an attempt to analyze the microbiome dynamics directly in a production herd. In such a situation, an important aspect of the analysis is the heat stress “phenotype”. In order to pre-correct for a whole series of genetic (i.e. familial relationship) and environmental effects (such as parity or lactation stage) possibly affecting the heat stress indicator measurements, prior to the actual heat stress modeling, we decided to use breeding values as pseudophenotypes, which were then deregressed in order to remove ancestral and familiar contributions. Such an approach provided a novel approach for the investigation of bacteria in dairy cattle under heat stress condition.

Declarations

Ethics approval and consent to participate

The data collection process was carried out in strict accordance with the protocol approved by the Animal Welfare Committee of the China Agricultural University. All experimental protocols were approved by the Animal Welfare Committee of the China Agricultural University. All methods are reported in accordance with ARRIVE guidelines (<https://arriveguidelines.org>) for the reporting of animal experiments.

Consent for publication

Not provided.

Availability of data and materials

The sequence data are available in the NCBI Sequence Read Archive with accession number SRP202074 (<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP202074>).

Other datasets generated and/or analyzed during the current study are not publicly available due to institutional constraints but are available from Yachun Wang on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Wroclaw University of Environmental and Life Sciences (Poland) as the Ph.D. research program "Bon doktoranta SD UPWr".

The publication is financed under the Leading Research Groups support project from the subsidy increased for the period 2020–2025 in the amount of 2% of the subsidy referred to Art. 387 (3) of the Law of 20 July 2018 on Higher Education and Science, obtained in 2019.

Supported by China Agriculture Research System of MOF and MARA; The Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62); National Agricultural Genetic Improvement Program (2130135).

Authors' contributions

B.C., Y.W. and J.S conceived and conducted the experiment, B.C. analyzed the results and wrote the manuscript in consultation with J.S., K.W., H.L. and Y.W. All authors reviewed the manuscript.

Acknowledgements

Calculations have been carried out using resources provided by Wroclaw Centre for Networking and Supercomputing (<http://wcss.pl>), grant No. 509.

Computations were carried out at the Poznan Supercomputing and Networking Centre.

Author details

¹Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland. ²National Research Institute of Animal Production, Balice, Poland. ³College of Animal Science and Technology, China Agricultural University, 100193, Beijing, P.R. China; Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA; National Engineering Laboratory for Animal Breeding.

References

1. Brivio, F., Zurmühl, M., Grignolio, S., von Hardenberg, J., Apollonio, M., Ciuti, S.: Forecasting the response to global warming in a heat-sensitive species. *Scientific Reports* **9**(1) (2019). doi:[10.1038/s41598-019-39450-5](https://doi.org/10.1038/s41598-019-39450-5)
2. Polsky, L., von Keyserlingk, M.A.G.: Invited review: Effects of heat stress on dairy cattle welfare. *Journal of Dairy Science* **100**(11), 8645–8657 (2017). doi:[10.3168/jds.2017-12651](https://doi.org/10.3168/jds.2017-12651)
3. Nardone, A., Ronchi, B., Lacetera, N., Ranieri, M.S., Bernabucci, U.: Effects of climate changes on animal production and sustainability of livestock systems. *Livestock Science* **130**(1-3), 57–69 (2010). doi:[10.1016/j.livsci.2010.02.011](https://doi.org/10.1016/j.livsci.2010.02.011)

4. Li, Q., Yang, C., Du, J., Zhang, B., He, Y., Hu, Q., Li, M., Zhang, Y., Wang, C., Zhong, J.: Characterization of miRNA profiles in the mammary tissue of dairy cattle in response to heat stress. *BMC Genomics* **19**(1) (2018). doi:[10.1186/s12864-018-5298-1](https://doi.org/10.1186/s12864-018-5298-1)
5. Lees, A.M., Sejian, V., Wallage, A.L., Steel, C.C., Mader, T.L., Lees, J.C., Gaughan, J.B.: The impact of heat load on cattle. *Animals* **9**(6), 322 (2019). doi:[10.3390/ani9060322](https://doi.org/10.3390/ani9060322)
6. Sigdel, A., Abdollahi-Arpanahi, R., Aguilar, I., Peñagaricano, F.: Whole genome mapping reveals novel genes and pathways involved in milk production under heat stress in US holstein cows. *Frontiers in Genetics* **10** (2019). doi:[10.3389/fgene.2019.00928](https://doi.org/10.3389/fgene.2019.00928)
7. Hoque, M.N., Istiaq, A., Clement, R.A., Sultana, M., Crandall, K.A., Siddiki, A.Z., Hossain, M.A.: Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. *Scientific Reports* **9**(1) (2019). doi:[10.1038/s41598-019-49468-4](https://doi.org/10.1038/s41598-019-49468-4)
8. Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., Fitzsimmons, C., Plastow, G., Guan, L.L.: Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. *Microbiome* **7**(1) (2019). doi:[10.1186/s40168-019-0699-1](https://doi.org/10.1186/s40168-019-0699-1)
9. Scott, K.P., Gratz, S.W., Sheridan, P.O., Flint, H.J., Duncan, S.H.: The influence of diet on the gut microbiota. *Pharmacological Research* **69**(1), 52–60 (2013). doi:[10.1016/j.phrs.2012.10.020](https://doi.org/10.1016/j.phrs.2012.10.020)
10. Karl, J.P., Hatch, A.M., Arcidiacono, S.M., Pearce, S.C., Pantoja-Feliciano, I.G., Doherty, L.A., Soares, J.W.: Effects of psychological, environmental and physical stressors on the gut microbiota. *Frontiers in Microbiology* **9** (2018). doi:[10.3389/fmicb.2018.02013](https://doi.org/10.3389/fmicb.2018.02013)
11. Chen, S., Wang, J., Peng, D., Li, G., Chen, J., Gu, X.: Exposure to heat-stress environment affects the physiology, circulation levels of cytokines, and microbiome in dairy cows. *Scientific Reports* **8**(1) (2018). doi:[10.1038/s41598-018-32886-1](https://doi.org/10.1038/s41598-018-32886-1)
12. Zhao, Min, Zheng, Wang: Effect of heat stress on bacterial composition and metabolism in the rumen of lactating dairy cows. *Animals* **9**(11), 925 (2019). doi:[10.3390/ani9110925](https://doi.org/10.3390/ani9110925)
13. Sales, G.F.C., Carvalho, B.F., Schwan, R.F., de Figueiredo Vilela, L., Meneses, J.A.M., Gionbelli, M.P., da Silva Ávila, C.L.: Heat stress influence the microbiota and organic acids concentration in beef cattle rumen. *Journal of Thermal Biology* **97**, 102897 (2021). doi:[10.1016/j.jtherbio.2021.102897](https://doi.org/10.1016/j.jtherbio.2021.102897)
14. Zhang, G., Wang, Y., Luo, H., Qiu, W., Zhang, H., Hu, L., Wang, Y., Dong, G., Guo, G.: The association between inflammaging and age-related changes in the ruminal and fecal microbiota among lactating holstein cows. *Frontiers in Microbiology* **10** (2019). doi:[10.3389/fmicb.2019.01803](https://doi.org/10.3389/fmicb.2019.01803)
15. Luo, H., Brito, L.F., Li, X., Su, G., Dou, J., Xu, W., Yan, X., Zhang, H., Guo, G., Liu, L., Wang, Y.: Genetic parameters for rectal temperature, respiration rate, and drooling score in holstein cattle and their relationships with various fertility, production, body conformation, and health traits. *Journal of Dairy Science* **104**(4), 4390–4403 (2021). doi:[10.3168/jds.2020-19192](https://doi.org/10.3168/jds.2020-19192)
16. Jairath, L., Dekkers, J.C.M., Schaeffer, L.R., Liu, Z., Burnside, E.B., Kolstad, B.: Genetic evaluation for herd life in Canada. *Journal of Dairy Science* **81**(2), 550–562 (1998). doi:[10.3168/jds.s0022-0302\(98\)75607-3](https://doi.org/10.3168/jds.s0022-0302(98)75607-3)
17. Andrews, S.: FASTQC. A quality control tool for high throughput sequence data (2010)
18. Krueger, F.: Trim Galore! (2015)
19. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al.: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**(8), 852–857 (2019). doi:[10.1038/s41587-019-0209-9](https://doi.org/10.1038/s41587-019-0209-9)
20. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P.: DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods* **13**(7), 581–583 (2016). doi:[10.1038/nmeth.3869](https://doi.org/10.1038/nmeth.3869)
21. McMurdie, P.J., Holmes, S.: phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**(4), 61217 (2013). doi:[10.1371/journal.pone.0061217](https://doi.org/10.1371/journal.pone.0061217)
22. Koh, H., Li, Y., Zhan, X., Chen, J., Zhao, N.: A distance-based kernel association test based on the generalized linear mixed model for correlated microbiome studies. *Frontiers in Genetics* **10** (2019). doi:[10.3389/fgene.2019.00458](https://doi.org/10.3389/fgene.2019.00458)
23. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O.: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**(D1), 590–596 (2012). doi:[10.1093/nar/gks1219](https://doi.org/10.1093/nar/gks1219)
24. 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.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011)
25. McInnes, L., Healy, J., Melville, J.: UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. cite arxiv:1802.03426 (2018). <https://arxiv.org/pdf/1802.03426.pdf>

26. Robinson, M.D., McCarthy, D.J., Smyth, G.K.: edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1), 139–140 (2009). doi:[10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616)
27. Robinson, M.D., Oshlack, A.: A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**(3), 25 (2010). doi:[10.1186/gb-2010-11-3-r25](https://doi.org/10.1186/gb-2010-11-3-r25)
28. Smid, M., van den Braak, R.R.J.C., van de Werken, H.J.G., van Riet, J., van Galen, A., de Weerd, V., van der Vlugt-Daane, M., Bril, S.I., Lalmahomed, Z.S., Kloosterman, W.P., Wilting, S.M., Foekens, J.A., IJzermans, J.N.M., Martens, J.W.M., Sieuwerts, A.M.: Gene length corrected trimmed mean of m-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* **19**(1) (2018). doi:[10.1186/s12859-018-2246-7](https://doi.org/10.1186/s12859-018-2246-7)
29. Cox, D.R., Reid, N.: Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society: Series B (Methodological)* **49**(1), 1–18 (1987). doi:[10.1111/j.2517-6161.1987.tb01422.x](https://doi.org/10.1111/j.2517-6161.1987.tb01422.x)
30. Solomon, D.L.: A note on the non-equivalence of the neyman-pearson and generalized likelihood ratio tests for testing a simple null versus a simple alternative hypothesis. *The American Statistician* **29**(2), 101–102 (1975). doi:[10.1080/00031305.1975.10477383](https://doi.org/10.1080/00031305.1975.10477383)
31. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**(1), 289–300 (1995). doi:[10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)
32. Bailey, M.T., Dowd, S.E., Galley, J.D., Hufnagle, A.R., Allen, R.G., Lyte, M.: Exposure to a social stressor alters the structure of the intestinal microbiota: Implications for stressor-induced immunomodulation. *Brain, Behavior, and Immunity* **25**(3), 397–407 (2011). doi:[10.1016/j.bbi.2010.10.023](https://doi.org/10.1016/j.bbi.2010.10.023)
33. Baek, Y.C., Choi, H., Jeong, J.-Y., Lee, S.D., Kim, M.J., Lee, S., Ji, S.-Y., Kim, M.: The impact of short-term acute heat stress on the rumen microbiome of hanwoo steers. *Journal of Animal Science and Technology* **62**(2), 208–217 (2020). doi:[10.5187/jast.2020.62.2.208](https://doi.org/10.5187/jast.2020.62.2.208)
34. Luengo, R.M.-C., Lu, J., Williams, E.J., Miranda-CasoLuengo, A.A., Carrington, S.D., Evans, A.C.O., Meijer, W.G.: Delayed differentiation of vaginal and uterine microbiomes in dairy cows developing postpartum endometritis. *PLOS ONE* **14**(1), 0200974 (2019). doi:[10.1371/journal.pone.0200974](https://doi.org/10.1371/journal.pone.0200974)
35. Kim, D.-H., Kim, M.-H., Kim, S.-B., Son, J.-K., Lee, J.-H., Joo, S.-S., Gu, B.-H., Park, T., Park, B.-Y., Kim, E.-T.: Differential dynamics of the ruminal microbiome of jersey cows in a heat stress environment. *Animals* **10**(7), 1127 (2020). doi:[10.3390/ani10071127](https://doi.org/10.3390/ani10071127)
36. Yu, M.-F., Zhao, X.-M., Cai, H., Yi, J.-M., Hua, G.-H.: Dihydropyridine enhances the antioxidant capacities of lactating dairy cows under heat stress condition. *Animals* **10**(10), 1812 (2020). doi:[10.3390/ani10101812](https://doi.org/10.3390/ani10101812)
37. He, J., Guo, H., Zheng, W., Xue, Y., Zhao, R., Yao, W.: Heat stress affects fecal microbial and metabolic alterations of primiparous sows during late gestation. *Journal of Animal Science and Biotechnology* **10**(1) (2019). doi:[10.1186/s40104-019-0391-0](https://doi.org/10.1186/s40104-019-0391-0)
38. Ransom-Jones, E., Jones, D.L., McCarthy, A.J., McDonald, J.E.: The fibrobacteres: an important phylum of cellulose-degrading bacteria. *Microbial Ecology* **63**(2), 267–281 (2012). doi:[10.1007/s00248-011-9998-1](https://doi.org/10.1007/s00248-011-9998-1)

Figures

Figure 1

The relative abundance of genera with average proportions of more than 0.5%.

Figure 2

The relative abundance of phyla with average proportions of more than 0.5%.

Figure 3

UMAP projection of the ASVs counts matrix on the genus level.

Figure 4

Phylogenetic tree of genera identified in fecal samples. Color indicates significantly associated genera with a given phenotype.