WROCŁAW UNIVERSITY OF ENVIRONMENTAL
AND LIFE SCIENCES
FACULTY OF BIOLOGY AND ANIMAL SCIENCE

DOCTORAL DISSERTATION

The application of bioinformatics tools for the analysis of the genetic background of heat stress resistance in cattle

Wykorzystanie narzędzi bioinformatycznych do analizy genetycznego podłoża odporności na stres cieplny na przykładzie bydła

---

Author:
Bartosz Czech, MSc, Eng.

Supervisor:
prof. Joanna Szyda, PhD
Department of Genetics,
Wrocław University of Environmental
and Life Sciences

Auxiliary supervisor:
prof. Gustavo de los Campos, PhD
Department of Epidemiology,
Michigan State University

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

Scientific field: natural sciences
Scientific discipline: biological sciences

Wrocław 2023

This dissertation is dedicated to my Mom Stefania and Dad Józef. Thank you for supporting my long academic adventure, pieces of advice, patience, and faith in me.

In memory of my Dad Józef.

Acknowledgements

I would like to express my heartfelt gratitude to my Supervisor, **Prof. Joanna Szyda**, for her invaluable guidance, unwavering support, and insightful mentorship throughout the course of my research. Her expertise, dedication, and encouragement have been instrumental in shaping this dissertation and my academic journey as a whole. I am truly fortunate to have had the privilege of working under her supervision. Thank you for your patience, encouragement, and belief in my potential. Your mentorship has been a source of inspiration, and I am deeply thankful for your contributions to my academic and personal growth.

I am deeply grateful to my colleague and companion, **Dr. Magda Mielczarek** for her unwavering support and camaraderie throughout my academic journey. Her warm smiles, infectious laughter, and light-hearted jokes have been a constant source of positivity, making both conferences and everyday interactions a joyous experience. Her genuine friendship and encouragement have been instrumental in shaping my scientific path.

I would like to express my deep gratitude to all my friends and supporters who have played an instrumental role in my PhD journey. Your diverse forms of support, encouragement, and contributions have been invaluable in shaping the course of my research and personal growth. Your unwavering belief in my abilities, the insightful discussions we've shared, and the various ways you've stood by me have made this journey meaningful and fulfilling. This thesis is a testament to the collective impact of your kindness and assistance, and I am truly thankful for the role each of you has played in my success.

"All models are wrong, but some are useful."

George E. P. Box

Contents

List of publications constituting the Doctoral Dissertation

1. **Czech, B.**, Szyda, J., Wang, K., Luo, H., & Wang, Y. (2022). Fecal microbiota and their association with heat stress in Bos taurus. In BMC Microbiology (Vol. 22, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1186/s12866-022-02576-0

   Impact Factor = 4.465

2. **Czech, B.**, Wang, Y., Wang, K., Luo, H., Hu, L., & Szyda, J. (2022). Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition. In Frontiers in Microbiology (Vol. 13). Frontiers Media SA. https://doi.org/10.3389/fmicb.2022.998093

   Impact Factor = 6.064

3. **Czech, B.**, Wang, Y., & Szyda, J. (2023). Genome-wide association study of heat stress response in Bos taurus. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2023.06.05.543663

   Preprint (under review)

Abstract

Global warming and the associated rise in temperatures pose a significant threat to mammals, inducing heat stress and adversely impacting their health and biological functions. This issue is particularly pertinent in livestock farming, where animals bred for high production yields and increased metabolic loads are especially vulnerable to heat stress. In dairy cattle, heat stress leads to reduced milk production, compromised welfare, and stunted growth. This research addresses the pressing need to understand the long-term effects of heat stress susceptibility on organisms, focusing on the genomic, transcriptomic, and microbiota levels of Holstein cattle under heat stress conditions. Bioinformatics emerges as a pivotal tool in this research, aiding in the identification of genetic variants, candidate genes, and pathways associated with heat stress response. The integration of genomics, transcriptomics, and metagenomics data provides a holistic understanding of how these factors interplay in the face of heat stress. Three distinct studies are presented: the first identifies microbial markers indicative of heat stress in cattle; the second unravels gene expression regulation influenced by the microbiome; and the third identifies genetic markers associated with heat stress resilience. These findings collectively inform strategies to enhance animal welfare and productivity amidst climate-induced heat stress. In the face of climate change's global impact, this study emerges as a pivotal foundation, delving into the biological intricacies that underlie the effects of heat stress on cattle. By dissecting the interplay of microbiome, transcriptome, and genome, this research unveils the complex biological mechanisms shaping cattle's responses to environmental challenges. Moreover, this dissertation imparts invaluable biological insights that may refine livestock management and breeding strategies, ultimately strengthening agricultural sustainability and bolstering global food security.

Abstract in Polish

Globalne ocieplenie i związany z nim wzrost temperatur stanowią znaczące zagrożenie dla ssaków, prowadząc do stresu cieplnego i negatywnie wpływając na ich zdrowie oraz funkcje biologiczne. Problem ten jest szczególnie istotny w hodowli zwierząt, gdzie zwierzęta hodowane ze względu na wysoką wydajność produkcyjną i zwiększone obciążenie metaboliczne są szczególnie podatne na stres cieplny. W przypadku krów mlecznych, stres cieplny prowadzi do zmniejszenia produkcji mleka, pogorszenia warunków bytowania i zahamowania wzrostu. Niniejsze badania adresują pilną potrzebę zrozumienia długoterminowych skutków podatności na stres cieplny u organizmów, skupiając się na analizie genomu, transkryptomu i mikrobiomu krów rasy Holstein w warunkach stresu cieplnego. Bioinformatyka wyłania się jako kluczowe narzędzie w tej pracy badawczej, wspomagając identyfikację zmian genetycznych, genów kandydujących i szlaków związanych z reakcją na stres cieplny. Integracja danych genomicznych, transkryptomicznych oraz mikrobiomicznych pozwala na całościowe zrozumienie, jak stres cieplny wpływa na zmiany molekularne. Przedstawione są trzy odrębne badania: pierwsze identyfikuje mikrobiologiczne markery wskazujące na stres cieplny u bydła; drugie wskazuje zmiany na poziomie ekspresji genów oraz interakcje zachodzące między transkryptomem gospodarza a jego mikrobiomem; trzecie badanie identyfikuje genetyczne markery związane z odpornością na stres cieplny. Wyniki te łącznie dostarczają informacji na temat strategii poprawy dobrostanu zwierząt i produktywności w obliczu stresu cieplnego. Badania przedstawione w niniejszej pracy mogą stać się fundamentem do dalszych badań, zgłębiając biologiczne mechanizmy leżące u podstaw wpływu stresu cieplnego na bydło. Poprzez analizę współdziałania mikrobiomu, transkryptomu i genomu, badania te ujawniają złożone mechanizmy biologiczne kształtujące odpowiedzi bydła na wyzwania środowiskowe. Ponadto, niniejsza dysertacja dostarcza spostrzeżeń biologicznych, które mogą udoskonalić zarządzanie zwierzętami hodowlanymi i strategie hodowlane, ostatecznie wzmacniając zrównoważony rozwój rolnictwa i zwiększając globalne bezpieczeństwo żywnościowe.

List of Abbreviations

ASV	Amplicon Sequence Variant
BTA	Bos Taurus Autosome
cDNA	complementary DNA
DB	Database
DEG	Differentially Expressed Gene
DNA	Deoxyribonucleic Acid
DRP	Deregressed Proof
DS	Drooling Score
EBV	Estimated Breeding Value
F	Forward
FDR	False Discovery Rate
GO	Gene Ontology
GSEA	Gene-Set Enrichment Analysis
GWAS	Genome-Wide Association Study
HS	Heat Stress
HSP	Heat Shock Protein
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAF	Minor Allele Frequency
mRNA	messenger RNA
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
R	Reverse
RNA	Ribonucleic Acid
RNA-seq	RNA sequencing
ROS	Reactive Oxygen Species
rRNA	ribosomal RNA
RS	Respiratory Score
RT	Rectal Temperature
SCFA	Short-Chain Fatty Acid
SNP	Single-Nucleotide Polymorphism
THI	Temperature Humidity Index
TOM	Topological Overlap Matrix
UMAP	Uniform Manifold Approximation and Projection
VEP	Variant Effect Predictor
WGCNA	Weighted Correlation Network Analysis

Chapter 1

Introduction

Heat is a significant environmental stressor that can affect the physiological, metabolic, and behavioral responses of animals, including cattle (Gonzalez-Rivas et al., 2020). Heat stress in cattle occurs when the body's ability to dissipate heat is overwhelmed by the environmental heat load, resulting in an imbalance between heat production and heat dissipation. As a result, cattle are subjected to a range of negative consequences that can impact their health, welfare, and productivity (Roth, 2020). Cattle are particularly susceptible to heat stress due to their limited ability to regulate body temperature. Unlike humans, who can sweat to dissipate heat, cattle primarily rely on respiratory evaporative cooling to maintain their core body temperature. As a result, they are vulnerable to heat stress when the ambient temperature and humidity exceed their thermal comfort zone (Cartwright et al., 2023). Furthermore, as a result of the multi-generational process of selecting dairy cattle with a primary focus on enhancing milk production, a breeding objective that has persisted until very recently, the inherent physiology of these animals can induce metabolic stress. This stress bears resemblance to the effects of heat stress, as it emerges from the exceptionally elevated metabolic rates observed in the most productive cows. This overlap in stress factors highlights a noteworthy parallel between the challenges posed by intense metabolic activity and the challenges posed by elevated temperatures. The negative effects of heat stress in cattle have been well-documented, including decreased feed intake, reduced milk production, altered immune function, impaired reproductive performance, and even mortality in severe cases (Krishnan et al., 2017). Heat stress can also impact the quality of animal products, such as meat and milk, due to changes in the animal's metabolism and physiological processes (Dahl, Tao, and Laporta, 2020). Given the potential impact of heat stress on animal welfare and productivity, it is crucial to understand the underlying biological mechanisms that contribute to heat stress tolerance in cattle. One of the primary responses to heat stress in cattle is an increase in respiration rate and panting, which increases the rate of heat loss through evaporative cooling. The evaporation of moisture from the respiratory tract and skin surface helps to dissipate excess heat from the body (Becker, Collier, and Stone, 2020). In addition to respiratory evaporative cooling, heat stress also triggers changes in blood flow and metabolism. As body temperature increases, blood vessels in the skin dilate to facilitate heat loss, which can result in a decrease in blood flow to other organs, such as the gastrointestinal tract and mammary glands. This can lead to decreased feed intake, reduced milk production, and altered nutrient absorption (Rebez et al., 2023). Heat stress can also affect metabolic processes in the body, including changes in hormone secretion and glucose metabolism. In response to heat stress, cattle produce more cortisol, which can affect immune function and induce a state of inflammation in the body. It can also disrupt glucose homeostasis and lead to insulin resistance, which can impair

nutrient utilization and increase the risk of metabolic disorders such as ketosis (Abbas et al., 2020; Mellado et al., 2023). Furthermore, heat stress impacts cattle reproduction, as high temperatures can reduce sperm and oocyte quality, and decrease embryo survival rates. Heat stress can also interfere with the normal estrous cycle and result in delayed or irregular breeding patterns (Miętkiewska, Kordowitzki, and Pareek, 2022). Overall, the physiological responses to heat stress in cattle are complex and multifaceted, involving changes in respiration, blood flow, metabolism, and hormone secretion. Understanding these responses is essential for developing effective strategies to mitigate the negative effects of heat stress on cattle health and productivity (Idris et al., 2021).

Heat stress in cattle can also result in changes at the cellular level. The excessive heat load can induce a range of cellular stress responses, such as the activation of heat shock proteins (HSPs) and the production of reactive oxygen species (ROS) (Khan et al., 2023; Hassan et al., 2019). HSPs are a family of proteins that are involved in cellular stress responses, including those triggered by heat stress. They help to maintain protein folding and prevent the aggregation of damaged proteins, which can be harmful to the cell. Heat stress can induce the expression of HSPs in cattle, which can help protect cells and tissues from heat-induced damage (Archana et al., 2017). However, the production of ROS can also be induced by heat stress, which can lead to oxidative stress and damage to cellular components, such as DNA, lipids, and proteins. This can impair cellular function and contribute to a range of negative consequences, such as reduced immune function, impaired reproductive performance, and increased susceptibility to disease (Guo et al., 2021). Studies have shown that heat stress can alter the expression of genes involved in immune function, metabolism, and reproductive processes. For example, heat stress can increase the expression of genes involved in inflammation and decrease the expression of genes involved in nutrient absorption and metabolism (Garner et al., 2020; Bai et al., 2020). The molecular changes induced by heat stress in cattle can have a range of consequences for animal health and productivity. Understanding the molecular mechanisms underlying these changes is essential for developing targeted interventions to mitigate the negative effects of heat stress on cattle. By identifying key molecular targets, researchers can develop strategies to improve the resilience of cattle to heat stress and enhance their overall health and productivity.

The microbiome refers to the collection of microorganisms that live in and on an animal. The microbiome plays a crucial role in animal health and might be altered in response to heat stress. It can affect animal physiology, influencing a range of physiological processes, including nutrient absorption, immune function, and metabolism (Liu et al., 2023). Metagenomic sequencing can be used to identify the microbial species present in the cattle gut. Bioinformatics tools can be used to perform taxonomic profiling and functional annotation of the microbiome. Differential abundance analysis can be performed to identify microbial species that are enriched or depleted in response to heat stress. Additionally, network analysis can be used to identify groups of co-occurring microbial species that are associated with heat stress tolerance. Microbiome analysis can provide insights into how the intestinal microbiome responds to heat stress and identify potential microbiome-based interventions to improve heat stress tolerance in cattle (Zhang et al., 2022). Studies have shown that heat stress can alter the abundance of specific microbial species in the rumen and feces of cattle, with a reduction in beneficial bacteria such as cellulolytic bacteria and an increase in opportunistic pathogens. These changes in the microbiota may have a range of consequences for cattle health

and productivity. For example, a decrease in beneficial bacteria can impair the animal's ability to digest and utilize nutrients from feed, leading to reduced feed efficiency and production performance. Changes in the microbiota can also impact immune function, with a reduction in beneficial bacteria potentially increasing the risk of disease (Park et al., 2022; Kim et al., 2022). Moreover, microbiota can also play a role in mitigating the negative effects of heat stress on the animal. Certain microbial species have been shown to produce metabolites such as short-chain fatty acids (SCFAs) and antioxidants that can help improve nutrient utilization, reduce inflammation, and protect against oxidative stress (Ruiz-González, Rico, and Rico, 2022). Overall, the relationship between heat stress and microbiota in cattle is complex and multifaceted. Further research is needed to better understand the mechanisms underlying these interactions and to develop strategies to mitigate the negative effects of heat stress on microbiota and animal health.

The transcriptome refers to the set of all RNA transcripts in a cell or tissue. RNA sequencing (RNA-seq) can be used to quantify gene expression levels in response to heat stress. Bioinformatics tools can be used to identify differentially expressed genes (DEGs) between heat-stressed and non-stressed cattle. Functional enrichment analysis can be performed to identify overrepresented biological pathways and gene ontology terms among the DEGs. These pathways and terms can provide insights into the molecular mechanisms underlying heat stress response. Additionally, co-expression network analysis can be performed to identify groups of genes that are co-regulated in response to heat stress. Network analysis can provide information about how genes work together to respond to heat stress and identify novel pathways and gene candidates for further study (Luo et al., 2022).

The cattle genome was sequenced in 2009 (Burt, 2009), and since then, researchers have used bioinformatics tools to annotate and compare the genomes of heat-stressed and non-stressed cattle. One of the primary aims of genome analysis is to identify genetic variants associated with heat stress tolerance. Genome-wide association studies (GWAS) can be performed to identify single nucleotide polymorphisms (SNPs) associated with heat stress response. These SNPs hold the potential for genomic selection, facilitating the breeding of cattle with improved heat tolerance, among other breeding objectives. Gene ontology and pathway enrichment analysis can be performed to gain insights into the biological processes and molecular pathways involved in heat stress response (Bohlouli et al., 2022).

Multiomics integration, which involves integrating data from multiple omics platforms, including genomics, transcriptomics, and metagenomics, can provide a more comprehensive understanding of the biological systems involved in heat stress response in cattle. Multiomics integration has several advantages over singleomics in heat stress studies. Firstly, heat stress is a complex biological process that involves multiple biological systems, including the host genome, transcriptome, and microbiome. Singleomics approaches, which focus on only one of these systems, may provide limited insights into the biological processes involved in heat stress response. Multiomics integration allows for a more comprehensive understanding of the complex interactions between these systems and can identify key genes, pathways, and interactions involved in heat stress response (Subramanian et al., 2020). Secondly, multiomics integration can help overcome limitations associated with singleomics approaches. For example, genomics data can identify genetic variations associated with heat stress tolerance, but it does not provide information on the functional consequences of these variations. Transcriptomics data can provide information on changes in gene expression under heat stress conditions, but it does not provide information on the

genetic basis of these changes. Integrating genomics and transcriptomics data can help overcome these limitations and provide a more comprehensive understanding of the functional consequences of genetic variations associated with heat stress tolerance (Manzoni et al., 2016). Thirdly, multiomics integration can provide insights into host-microbe interactions involved in heat stress response. The intestinal microbiome plays a key role in host metabolism, immunity, and overall health, and is known to be affected by heat stress. Singleomics approaches may not capture the complex interactions between the host and the microbiome, which can limit our understanding of the biological processes involved in heat stress response. Integrating metagenomics and transcriptomics data can help identify key host-microbe interactions involved in heat stress response and provide insights into potential targets for interventions (Czech et al., 2022b).

Bioinformatics plays an important role in the analysis of genome, transcriptome, and microbiome data in heat stress research in cattle. By identifying genetic variations, candidate genes, and pathways involved in heat stress response, bioinformatics can provide insights into the molecular mechanisms underlying heat stress tolerance.

Multiomic data integration may pose challenges, but understanding these processes can yield results in improving the life and health of animals.

This dissertation will review the current scientific literature on the physiological, metabolic, and behavioral responses of cattle to heat stress and discuss the potential strategies for mitigating their negative effect. The dissertation is structured into three main segments, each shedding light on different facets of how heat stress influences cattle and their biological makeup. The first part of the dissertation centers on investigating the influence of heat stress on the microbial composition of the cattle microbiota. This involves delving into how elevated temperatures affect the balance and diversity of microorganisms residing within the cattle, which can subsequently influence their health and overall well-being. In the second segment, the focus shifts to a detailed analysis of differential gene expression in cattle experiencing heat stress. This involves examining how the heat stress condition triggers changes in the expression of genes within the cattle's genetic makeup. Furthermore, this research seeks to elucidate the intricate interplay between these changes in gene expression and the alterations observed in the microbiota composition. By exploring this interaction, the dissertation aims to provide a more holistic understanding of the physiological response to heat stress. The third major component of the research involves a genome-wide associated analysis of cattle subjected to heat stress. This entails investigating the genetic variations present within the cattle population and identifying specific genes or genetic markers that are associated with their response to heat stress. By pinpointing these genetic factors, researchers can gain insights into the genetic basis of the cattle's ability to cope with elevated temperatures.

Chapter 2

Objective

The primary objective of this dissertation was to investigate the interplay between microbiome composition, gene expression patterns, genetic diversity, and heat stress tolerance. The specific research objectives were as follows:

- identification and characterization of the fecal microbiota associated with heat stress;
- identification of genes differentially expressed under heat stress;
- exploration of the dynamics of interactions between host transcriptome (cattle) and fecal microbiome under heat stress;
- identification of associations between genetic variation and the level of heat stress response.

This research aimed to provide insights into the interplay between heat stress and the microbiome, transcriptome, and genomic variation within the context of Chinese Holstein cattle as the experimental subject. The findings will contribute to the development of biomarkers, breeding strategies, and management practices aimed at improving cattle resilience and productivity under heat stress conditions – an environmental stressor that has recently been gaining importance. Additionally, this research will enhance our understanding of the underlying mechanisms involved in heat stress responses and shed light on potential therapeutic targets for mitigating heat stress effects. This study's outcomes are anticipated to yield dual-fold contributions: firstly, they are poised to deliver substantial benefits to animal breeders and stakeholders within the cattle industry by furnishing evidence-based strategies that proactively address the multifaceted challenges precipitated by heat stress, thereby ameliorating both animal welfare and productivity. Secondly, the study will serve as an instrumental tool for advancing our comprehension of the intricate molecular mechanisms governing an organism's adaptive response to elevated temperature conditions. These insights are pivotal for the informed development of strategies aimed at mitigating the deleterious effects of heat stress and enhancing the overall resilience of cattle populations.

Chapter 3

Publications constituting the Doctoral Dissertation

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

This research is a comprehensive exploration of the impact of heat stress on the composition of the fecal microbiome in Chinese Holstein cattle, which serves as an apt model organism to investigate the ramifications of heat stress. Recognizing the multifaceted nature of heat stress, which defies simple quantification through a single metric, our investigation centers on three distinct indicators of heat stress: rectal temperature, drooling, and respiratory scores. By leveraging cutting-edge techniques, including 16S rRNA gene sequencing, bioinformatics, and statistical analyses, this study endeavors to unravel the intricate associations linking specific microbial genera and phyla with the diverse facets of heat stress. These findings constitute a crucial milestone in our understanding of heat stress in cattle, offering valuable insights into the potential identification of biomarkers indicative of this condition. This underscores the pivotal role played by the microbiome in responding to the intricate physiological alterations triggered by heat stress. Moreover, this research augments our comprehension of microbiome composition analysis in cattle, introducing the concept of utilizing heat stress as a continuous variable for investigation. The microbial taxa that exhibit strong associations with heat stress are prominently featured in this study, holding promise as biomarkers for subsequent microbiological investigations. As we conclude this study, it becomes evident that these discoveries not only advance our understanding of heat stress but also pave the way for further nuanced inquiries into microbiome dynamics in response to environmental stressors.

Czech *et al. BMC Microbiology* (2022) 22:171
https://doi.org/10.1186/s12866-022-02576-0

BMC Microbiology

RESEARCH Open Access

Fecal microbiota and their association with heat stress in Bos taurus

Bartosz Czech[1*], Joanna Szyda[1], Kai Wang[2], Hanpeng Luo[2] and Yachun Wang[2]

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 represented by their Estimated Breeding Values. We analyzed the microbiota from 136 fecal samples of Chinese Holstein cows through a 16S rRNA gene sequencing approach. 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 *Pseudobutyrivibrio* turned out to be the most significant genera, while *Acidobacteria* and *Gemmatimonadetes* were the most significant phyla. Phylogenetic analysis revealed 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 associated with 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

*Correspondence: bartosz.czech@upwr.edu.pl
[1]Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland
Full list of author information is available at the end of the article

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 Shounong 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. The experimental design deliberately did not involve formal case and control groups but was carried out on the production population.

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. Moreover, 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, 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 bacterial 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

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

performed in triplicate 20 μL mixture containing 4 μL of 5 μL FastPfu Buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of each primer (5 μM), 0.4 μL 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 QuantiFluor™-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 particular, each lactating cow was recorded twice a day for 2 consecutive days. 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:

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

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:

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

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

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

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

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 were pooled together. Next, reads were denoised – reads that occur very rarely were considered to be PCR errors and removed, as well as chimeric sequences and singletons. Those steps were done using DADA2 algorithm [20]. implemented in QIIME 2. All the sequencing runs were processed separately. 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 (SILVA SSU 138.1) [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. Filtered tables were used for the further differential abundance analysis. Additionally, for organoleptic testing of batch effect occurrent, 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 [26] implemented in QIIME2 software.

Differential abundance analysis

The edgeR [27] 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 [28] was applied. It identified and excluded highly abundant and highly variable genera and phyla, whereupon weighted mean of an abundance of remaining groups was used for the actual normalization [29]. The association between genera/phyla abundance and heat stress indicators was modeled using the negative binomial distribution:

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

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

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.

$$K \sim NB(\mu_k, \phi) \tag{4}$$

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 [30].

The significance of the effect of DRP on the relative abundance of the genus/phylum was tested using the Likelihood Ratio Test [31]. The test statistic is as follows:

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

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) [32].

Results

Processing and classification of sequence variants

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 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.

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 (Fig. 3), which reflect the sampling year. In further analysis, the effect of the sampling year was corrected.

Correlation analysis of diversity metrics

In order to check whether the general diversity of microbiota within samples is correlated with the DRPs, a

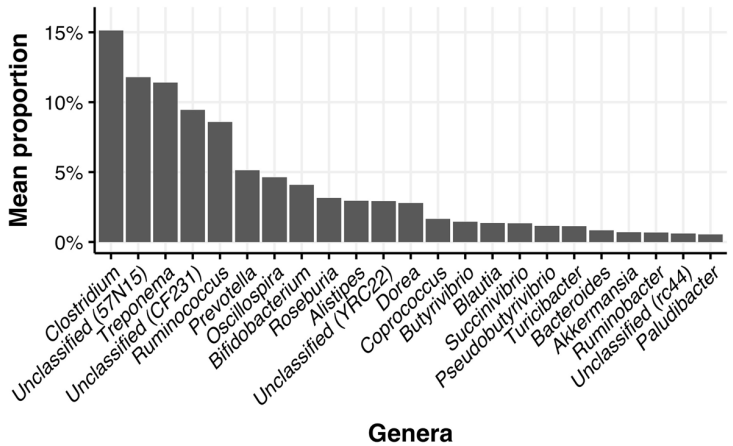

Fig. 1 The relative abundance of genera with average proportions of more than 0.5%

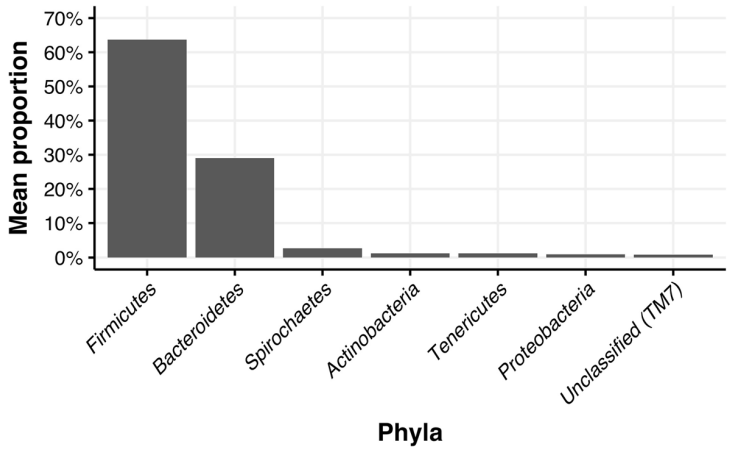

Fig. 2 The relative abundance of phyla with average proportions of more than 0.5%

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 performed 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 *Pseudobutyrivibrio* – with the drooling score. There was no overlap between genera significant for the three heat stress indicators (Table 5). Differential abundance analysis of phylum (Table 6) showed that 6 phyla were significiantly associated with rectal tempearture. 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

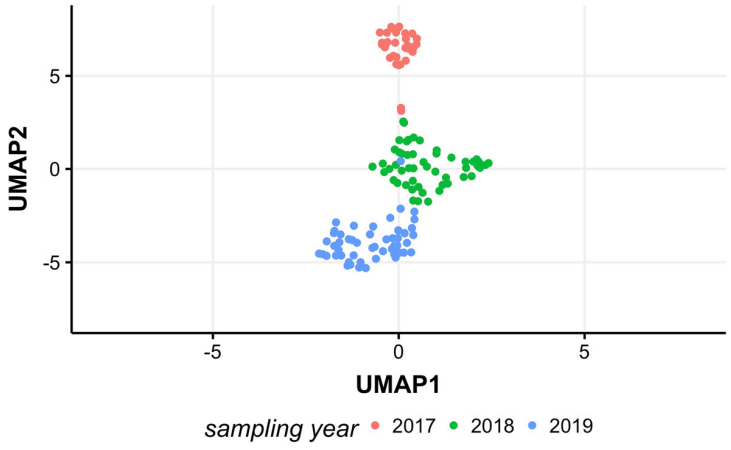

Fig. 3 UMAP projection of the ASVs counts matrix on the genus level

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

the descrease abundance. *Fibrobacteres* was significantly associated with both drooling and respiratory scores, while *Nitrospiarae*, *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 Fig. 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 *Pseudobutyrivibrio* that was associated with the drooling score creates a single, separate clade.

Table 5 Significant differentially abundant genera

Genus	logFC	FDR
Rectal temperature		
Rhizobium	-16.97	3.88×10^{05}
Rhodoplanes	-16.36	1.64×10^{04}
Kaistobacter	-15.95	1.10×10^{04}
Streptomyces	-15.82	1.10×10^{04}
Sphingomonas	-15.60	1.37×10^{04}
Acidovorax	-15.54	2.26×10^{04}
Nocardia	-15.35	1.10×10^{04}
Cupriavidus	-15.09	1.64×10^{04}
Candidatus Solibacter	-14.27	1.15×10^{03}
Nocardioides	-13.82	2.43×10^{04}
Brevundimonas	-12.95	5.59×10^{04}
Kribbella	-12.95	2.47×10^{03}
Amycolatopsis	-12.34	2.03×10^{03}
DA101	-12.25	1.75×10^{03}
Azospira	-11.84	2.03×10^{03}
Catellatospora	-11.33	8.41×10^{03}
Reyranella	-11.12	5.69×10^{03}
Pseudonocardia	-10.89	6.89×10^{03}
Devosia	-10.64	2.28×10^{03}
Rhodococcus	-10.23	1.11×10^{02}
Helcococcus	8.31	8.22×10^{03}
YRC22	-4.58	2.99×10^{02}
Respiratory score		
Succinivibrio	6.24	3.33×10^{02}
Drooling score		
Pseudobutyrivibrio	-16.64	1.68×10^{03}

Table 6 Significant differentially abundant phyla

Phylum	logFC	FDR
Rectal temperature		
Acidobacteria	-26.99	1.18×10^{16}
Gemmatimonadetes	-22.40	5.03×10^{13}
Chloroflexi	-21.07	2.98×10^{11}
Nitrospirae	-15.19	1.18×10^{07}
Planctomycetes	-11.19	5.52×10^{05}
Euryarchaeota	-6.80	2.99×10^{02}
Respiratory score		
Fibrobacteres	-8.67	2.58×10^{02}
Drooling score		
Fibrobacteres	-16.72	5.22×10^{04}
Nitrospirae	15.73	4.22×10^{04}
Gemmatimonadetes	15.17	4.69×10^{04}
Acidobacteria	14.46	4.71×10^{04}
Planctomycetes	12.50	1.13×10^{02}

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 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 [33] observed a reduced abundance of bacteria in genus *Pseudobutyrivibrio* in mice exposed to stressor-induced changes. Such reduced abundance was also observed by us the association with a drooling score. Baek [34] 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 focused on heat stress and any stress-induced conditions, but it was reported as associated with postpartum endometritis by Miranda CasoLuengo [35]. Moreover, [36] showed that *Streptomyces* was

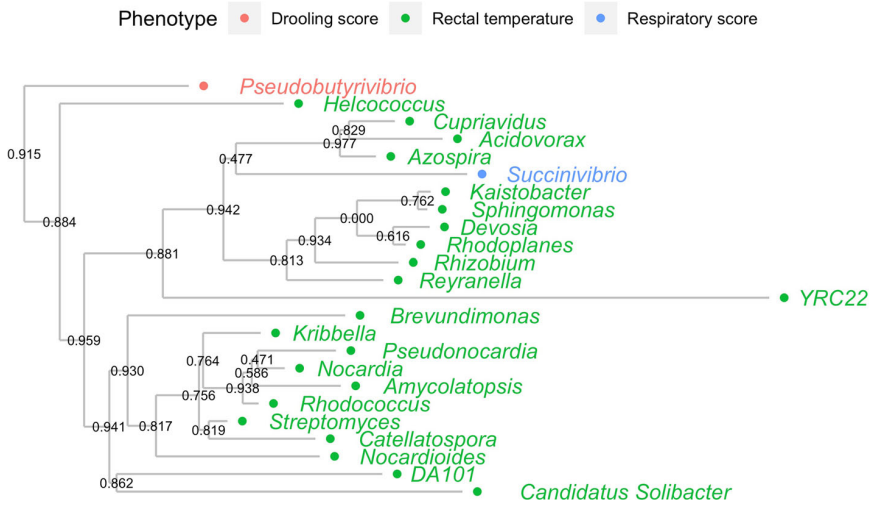

Fig. 4 Phylogenetic tree of genera identified in fecal samples. Color indicates significantly associated genera with a given phenotype

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 [37] 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 [38]. *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 [34].

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 [39] 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. *Pseudobutyrivibrio*) 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.

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.

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. All authors read and approved the final manuscript.

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). Special Item of Regional Collaborative Innovation in Tibet Autonomous Region (Grant No. QYXTZX-LS2021-01).

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.

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.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland. [2] Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA; National Engineering Laboratory for Animal Breeding; College of Animal Science and Technology, China Agricultural University, 100193 Beijing, People's Republic of China.

Received: 4 December 2021 Accepted: 7 June 2022
Published online: 05 July 2022

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. Sci Rep. 2019;9(1). https://doi.org/10.1038/s41598-019-39450-5.
2. Polsky L, von Keyserlingk MAG. Invited review: Effects of heat stress on dairy cattle welfare. J Dairy Sci. 2017;100(11):8645–57. https://doi.org/10.3168/jds.2017-12651.
3. Nardone A, Ronchi B, Lacetera N, Ranieri MS, Bernabucci U. Effects of climate changes on animal production and sustainability of livestock systems. Livest Sci. 2010;130(1-3):57–69. 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. 2018;19(1). https://doi.org/10.1186/s12864-018-5298-1.
5. Lees AM, Sejian V, Wallage AL, Steel CC, Mader TL, Lees JC, Gaughan JB. The impact of heat load on cattle. Animals. 2019;9(6):322. 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. Front Genet. 2019;10. https://doi.org/10.3389/fgene.2019.00928.
7. Hoque MN, Istiaq A, Clement RA, Sultana M, Crandall KA, Siddiki AZ, Hossain MA. Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. Sci Rep. 2019;9(1). 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 LL. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome. 2019;7(1). https://doi.org/10.1186/s40168-019-0699-1.
9. Scott KP, Gratz SW, Sheridan PO, Flint HJ, Duncan SH. The influence of diet on the gut microbiota. Pharmacol Res. 2013;69(1):52–60. https://doi.org/10.1016/j.phrs.2012.10.020.
10. Karl JP, Hatch AM, Arcidiacono SM, Pearce SC, Pantoja-Feliciano IG, Doherty LA, Soares JW. Effects of psychological, environmental and physical stressors on the gut microbiota. Front Microbiol. 2018;9. 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. Sci Rep. 2018;8(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. 2019;9(11):925. https://doi.org/10.3390/ani9110925.
13. Sales GFC, Carvalho BF, Schwan RF, de Figueiredo Vilela L, Meneses JAM, Gionbelli MP, da Silva Ávila CL. Heat stress influence the microbiota and organic acids concentration in beef cattle rumen. J Therm Biol. 2021;97:102897. 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. Front Microbiol. 2019;10. https://doi.org/10.3389/fmicb.2019.01803.
15. Luo H, Brito LF, 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. J Dairy Sci. 2021;104(4):4390–403. https://doi.org/10.3168/jds.2020-19192.
16. Jairath L, Dekkers JCM, Schaeffer LR, Liu Z, Burnside EB, Kolstad B. Genetic evaluation for herd life in canada. J Dairy Sci. 1998;81(2):550–62. 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. https://www.bibsonomy.org/bibtex/2b6052877491828ab53d3449be9b293b3/ozborn. Accessed 21 Nov 2021.
18. Krueger F. Trim Galore!. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 21 Nov 2021.
19. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852–7. https://doi.org/10.1038/s41587-019-0209-9.
20. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from illumina amplicon data. Nat Methods. 2016;13(7):581–3. https://doi.org/10.1038/nmeth.3869.
21. McMurdie PJ, Holmes S. phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE. 2013;8(4):61217. 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. Front Genet. 2019;10. https://doi.org/10.3389/fgene.2019.00458.
23. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2012;41(D1):590–6. 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. J Mach Learn Res. 2011;12:2825–30.
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. Accessed 21 Nov 2021.
26. Yamada KD, Tomii K, Katoh K. Application of the MAFFT sequence alignment program to large data—reexamination of the usefulness of chained guide trees. Bioinformatics. 2016;32(21):3246–51. https://doi.org/10.1093/bioinformatics/btw412.

27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616.
28. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):25. https://doi.org/10.1186/gb-2010-11-3-r25.
29. Smid M, van den Braak RRJC, van de Werken HJG, van Riet J, van Galen A, de Weerd V, van der Vlugt-Daane M, Bril SI, Lalmahomed ZS, Kloosterman WP, Wilting SM, Foekens JA, IJzermans JNM, Martens JWM, Sieuwerts AM. 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. 2018;19(1). https://doi.org/10.1186/s12859-018-2246-7.
30. Cox DR, Reid N. Parameter orthogonality and approximate conditional inference. J R Stat Soc Ser B (Methodol). 1987;49(1):1–18. https://doi.org/10.1111/j.2517-6161.1987.tb01422.x.
31. Solomon DL. 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. Am Stat. 1975;29(2):101–2. https://doi.org/10.1080/00031305.1975.10477383.
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodol). 1995;57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
33. Bailey MT, Dowd SE, Galley JD, Hufnagle AR, Allen RG, Lyte M. Exposure to a social stressor alters the structure of the intestinal microbiota: Implications for stressor-induced immunomodulation. Brain Behav Immun. 2011;25(3):397–407. https://doi.org/10.1016/j.bbi.2010.10.023.
34. Baek YC, Choi H, Jeong J-Y, Lee SD, Kim MJ, Lee S, Ji S-Y, Kim M. The impact of short-term acute heat stress on the rumen microbiome of hanwoo steers. J Anim Sci Technol. 2020;62(2):208–17. https://doi.org/10.5187/jast.2020.62.2.208.
35. Luengo RM-C, Lu J, Williams EJ, Miranda-CasoLuengo AA, Carrington SD, Evans ACO, Meijer WG. Delayed differentiation of vaginal and uterine microbiomes in dairy cows developing postpartum endometritis. PLoS ONE. 2019;14(1):0200974. https://doi.org/10.1371/journal.pone.0200974.
36. 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. 2020;10(7):1127. https://doi.org/10.3390/ani10071127.
37. 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. 2020;10(10):1812. https://doi.org/10.3390/ani10101812.
38. 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. J Anim Sci Biotechnol. 2019;10(1). https://doi.org/10.1186/s40104-019-0391-0.
39. Ransom-Jones E, Jones DL, McCarthy AJ, McDonald JE. The fibrobacteres: an important phylum of cellulose-degrading bacteria. Microb Ecol. 2012;63(2):267–81. https://doi.org/10.1007/s00248-011-9998-1.

Publisher's Note

3.2 Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition

The primary objective of this scientific inquiry was to comprehensively investigate the profound impact of heat stress on the transcriptome. This study was designed to address two crucial aspects: firstly, the meticulous analysis of gene expression patterns utilizing RNA-seq data, and secondly, the utilization of previously acquired findings from the analysis of microbiota, the diverse microbial community residing within an organism, to expand the scope of the investigation. By incorporating these additional findings, the study aimed to perform an interaction analysis, thereby unraveling the intricate interplay between gene expression and microbiota in response to heat stress. To comprehensively explore this interplay, several heat stress metrics were considered, including rectal temperature, drooling score, and respiratory score. Through the meticulous analysis of gene expression patterns and the utilization of 16S rRNA sequencing, a technique that enables the identification and quantification of bacterial species based on their ribosomal RNA, the research successfully identified a set of differentially expressed genes that are intricately associated with the response to heat stress across three distinct phenotypes. Moreover, the study revealed a significant and noteworthy relationship between gene expression and the abundance of microbiota. This finding sheds light on specific bacterial species that potentially exert influence on gene regulation during episodes of heat stress. By elucidating the impact of microbiota on gene expression, this study contributes to our understanding of the intricate molecular mechanisms underlying the response to heat stress, highlighting the potential role of specific bacterial species in modulating gene expression patterns during such physiological challenges.

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 23 November 2022
DOI 10.3389/fmicb.2022.998093

OPEN ACCESS

EDITED BY
Rachel Pilla,
Texas A&M University, United States

REVIEWED BY
Li Min,
Guangdong Academy of Agricultural Sciences (GDAAS), China
Bonnie Mallard,
University of Guelph, Canada

*CORRESPONDENCE
Joanna Szyda
joanna.szyda@upwr.edu.pl

SPECIALTY SECTION
This article was submitted to Microorganisms in Vertebrate Digestive Systems, a section of the journal Frontiers in Microbiology

RECEIVED 19 July 2022
ACCEPTED 09 November 2022
PUBLISHED 23 November 2022

CITATION
Czech B, Wang Y, Wang K, Luo H, Hu L and Szyda J (2022) Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition. *Front. Microbiol.* 13:998093. doi: 10.3389/fmicb.2022.998093

Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition

Bartosz Czech[1], Yachun Wang[2], Kai Wang[2], Hanpeng Luo[2], Lirong Hu[2] and Joanna Szyda[1]*

[1]Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland, [2]National Engineering Laboratory for Animal Breeding, Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, College of Animal Science and Technology, China Agricultural University, Beijing, China

Climate change affects animal physiology. In particular, rising ambient temperatures reduce animal vitality due to heat stress and this can be observed at various levels which included genome, transcriptome, and microbiome. In a previous study, microbiota highly associated with changes in cattle physiology, which included rectal temperature, drooling score and respiratory score, were identified under heat stress conditions. In the present study, genes differentially expressed between individuals were selected representing different additive genetic effects toward the heat stress response in cattle in their production condition. Moreover, a correlation network analysis was performed to identify interactions between the transcriptome and microbiome for 71 Chinese Holstein cows sequenced for mRNA from blood samples and for 16S rRNA genes from fecal samples. Bioinformatics analysis was performed comprising: i) clustering and classification of 16S rRNA sequence reads, ii) mapping cows' transcripts to the reference genome and their expression quantification, and iii) statistical analysis of both data types—including differential gene expression analysis and gene set enrichment analysis. A weighted co-expression network analysis was carried out to assess changes in the association between gene expression and microbiota abundance as well as to find hub genes/microbiota responsible for the regulation of gene expression under heat stress. Results showed 1,851 differentially expressed genes were found that were shared by three heat stress phenotypes. These genes were predominantly associated with the cytokine-cytokine receptor interaction pathway. The interaction analysis revealed three modules of genes and microbiota associated with rectal temperature with which two hubs of those modules were bacterial species, demonstrating the importance of the microbiome in the regulation of gene expression during heat stress. Genes and microbiota from the significant modules can be used as biomarkers of heat stress in cattle.

KEYWORDS

heat stress, cattle, 16S rRNA, RNA-seq, WGCNA, NGS, sequencing, multiomics

1. Introduction

Humans induce global warming that negatively influences all organisms on the Earth, e.g., by the occurrence of heat stress (HS) in livestock, which negatively affects its vitality, physiological responses, and behavior. Heat stress can inhibit milk production in dairy cattle which can result in significant losses to the industry (Garner et al., 2020). Moreover, HS in cows results in further economic losses by reducing reproduction (Macciotta et al., 2017). Currently, phenotypes such as rectal temperature, drooling score, and respiration rate score are standard physiological indicators of HS (Brito et al., 2020; Luo et al., 2021). Recently, due to genomic selection being targeted to increased milk production in cattle, cows tend to be more susceptible to HS (Biffani et al., 2016). The phenomenon of HS from the molecular perspective is a complex challenge and still, many mechanisms are unknown. A previous study focusing on the differential abundance of microbiota already demonstrated that HS affects the microbial composition of the colon (Czech et al., 2022). Only a few studies have demonstrated the effect of HS on the transcriptome profile of cattle, and have identified genes potentially associated with the HS response. Gao et al. (2019) pointed out that amino acid and glucose transport were downregulated by HS. In another study looking at the effect of HS on the transcriptome profile of mammary glands of cows (Yue et al., 2020), the authors indicated that HS affects dairy cows' immunity and thus has a potential impact on milk yield. Sigdel *et al.* also presented the association analysis of HS cattle using SNP markers from the Cooperative Dairy DNA Repository and the Council on Dairy Cattle Breeding, and identified genes *HSF1*, *MAPK8IP1*, and *CDKN1B* that were directly involved in the cellular response to HS (Sigdel et al., 2019).

In general, the impact of HS on cows is fairly difficult to assess due to the complexity of the metabolism and physiology of cows. However, the development of molecular techniques like next-generation sequencing, mass spectrometry, and other techniques allow us to look more deeply into these mechanisms by obtaining information about the entire biology of the system. Additionally, high-performance computers with new algorithms allow for considering more complex statistical models that allow for better insights into the complexity of organisms (Park et al., 2021b). Many other studies that focused on the integration analysis of host transcriptome and microbiota already demonstrated the importance of the multiomics approach to identify biomarkers underlying diseases and complex traits (Wang et al., 2019). In livestock, only a few studies have been focusing on the integration of host transcriptome and microbiome interactions. One of the studies showed the impact of the interaction of host transcriptome and microbiome on the physiology of full-sibs broilers with divergent feed conversion ratio (Shah et al., 2019). Ramayo-Caldas et al. (2021) investigated the joint effects of host genomic variation and the gut microbiome variation in the context of immune response in pigs. Also, Carillier-Jacquin et al. (2022) considered the importance of using both sources of information for the accuracy of prediction of pig digestibility coefficients, concluding that the incorporation of gut microbiome information is important for prediction and even outperforms the importance of host genetic variation. Another study in chickens showed the influence of nutrition on the interaction between transcriptome and microbiome which in turn influenced egg production in aged laying hens (Liu et al., 2022). Although those studies have revealed (and stressed) the importance of the incorporation of microbiome information into the evaluation of phenotypes that are important for livestock, the particular impact of the interaction of host transcriptome and microbiome on HS is still not well understood. Moreover, previous studies (Freitas et al., 2022) indicated a complex genetic architecture underlying the HS response, so it is expected that using all available sources of omic information is crucial for the modeling of this phenotype.

Therefore, it is worth mentioning that almost all HS studies used the case-control experimental design. In this study, we used a continuous variable to measure the HS response, to reflect the production environment which allowed us to study potential changes in gene expression level, microbiome abundance, and their interactions under standard conditions. This study aimed to identify genes differentially expressed between cows characterized by different additive genetic effects of HS response measured by the three HS indicators: rectal temperature, drooling score, and respiratory score, and to perform an integration of multiomics data of host gene expression levels in relation to its microbiome composition.

2. Materials and methods

2.1. Material

Fecal and blood samples from 71 Chinese Holstein cows were collected in 2017, 2018, and 2019. Cows were sampled once over the course of 3 years. In this study, no artificial heat stress challenge was imposed since the major goal was to assess the impact of heat stress that occurs during standard production conditions and is due to a combination of several climatic and production factors. Heat stress phenotypes used in this study were represented by the additive genetic effect of each cow, corrected for environmental factors such as lactation stage, age at calving, parity, and temperature-humidity index, that were expressed as deregressed estimated breeding value (DRP) for rectal temperature, respiratory score, and drooling score. We used a mixed linear model for the DRP estimation:

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

where $y_{ijklqno}$ refers to phenotype (RS, DS or RT), μ is the population mean, fym_i is the fixed herd-year effect, 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 whether the measurement was taken before or after milking, t_q is the fixed effect of testing time (morning or afternoon), *thi* is the fixed effect of the 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:

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

More information about the sampling procedure, the housing of animals, phenotypes, THI, and the statistical model used to estimate DRP and a description of the dataset were included in a previous study (Czech et al., 2022).

2.1.1. 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.

2.2. Methods

2.2.1. Microbiome

Deoxyribonucleic acid was isolated from fecal samples, which represent the microbiome composition of the colon, and was used for sequencing of V3 and V4 regions of the 16S rRNA gene using the Illumina MiSeq and HiSeq platforms. Sequenced reads were cleaned and processed using QIIME2 software (Bolyen et al., 2019) with the SILVA database (Quast et al., 2012) to cluster and classify them to taxonomical levels. The final output is represented by the amplicon sequence variants table with information about the frequency of a given taxon in a given fecal sample. The procedure is explained in detail by Czech et al. (2022).

2.2.2. mRNA-seq

Total RNA was isolated from leukocytes according to the instructions of the TRIzol Reagent method (Rio et al., 2010). The cDNA library was prepared using mRNA molecules and sequenced using the NovaSeq 6000 System Illumina platform. Ribonucleic acid concentration and quality were determined using Equalbit RNA BR Assay Kit (Invitrogen, California, USA) and the Nanodrop 2000 (Thermo, Massachusetts, USA). Ribonucleic acid integrity was assessed using 1% agarose gel electrophoresis and then used for library construction with 28S/18S >1. For the RNA-Seq library, 2 μg total RNA was firstly used for purification and fragmentation with NEBNext Poly(A) mRNA Magnetic Isolation Module (Cat No. E7490S, New England Biolabs (UK) Ltd., Hitchin, Herts, UK) and then followed by cDNA library with NEBNext Ultra RNA Library Prep Kit for Illumina (Cat No. E7530S, New England Biolabs (UK) Ltd., Hitchin, Herts, UK). All libraries were quantitated by the Equalbit DNA BR Assay Kit (Invitrogen, California, USA) and pooled to generate equimolarly, and finally submitted for sequencing by the NovaSeq 6000 System (Illumina, Inc., San Diego, California, USA) which generated 150 base paired-end reads.

Sequenced reads were evaluated in the context of their quality and cleaned using Fastp software (Chen et al., 2018). Filtered reads were mapped to the bovine genome (ARS-UCD1.2) using STAR software (Dobin et al., 2012) and Picard (Broad Institute, 2022) was applied to mark duplicates. Finally, RNA-SeQC (DeLuca et al., 2012) software was used to quantify the expression. Gene expression was analyzed using the DESeq2 R package (Love et al., 2014) to perform differential gene expression analysis fitting the negative binomial regression model adjusted for the sequencing year. The effect of HS was expressed as the average fold change per DRP increased by one unit. The Wald test was used to assess the significance of slope estimates. *P*-values obtained separately for each gene were corrected for multiple testing using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) for controlling the False Discovery Rate (FDR). Genes with the FDR < 0.05 were considered to be associated with HS. Next, we performed Gene-Set Enrichment Analysis (GSEA) based on Gene Ontology (GO) (Ashburner et al., 2000; Consortium, 2020) implemented in the goseq R package (Young et al., 2010) and metabolic pathways were defined by Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa, 2000) implemented in the clusterProfiler R package (Yu et al., 2012).

2.2.3. Omics integration

The final step of the analysis was the integration of microbiota abundance identified in the 16S rRNA data with the gene expression identified in the RNA-seq data. To study the transcriptome-microbiome interaction we applied the weighted co-expression network analysis implemented in the WGCNA R package (Langfelder and Horvath, 2008). The analysis was split into steps comprising: i. creating a correlation matrix using Pearson's correlation coefficient between all pairs of genes-genera; ii. creating adjacency matrix (matrix-based representation of a graph) using the formula: $a_{mn} = |c_{mn}|^\beta$, where a_{mn} is an adjacency between gene/genus m and gene/genus n, c_{mn} is a Pearson's correlation coefficient, and β is a soft-power threshold determined based on the standard scale-free topology network (Chen and Shi, 2004);

iii. transformation of the adjacency matrix into the topological overlap matrix (TOM) which is the matrix of the similarity in terms of the commonality of the connected nodes (Yip and Horvath, 2007); iv. the dynamic tree cutting algorithm was used for the hierarchical clustering of TOM into modules, as clusters of highly interconnected genes and genera; in order to obtain co-expressed modules, the parameters of the algorithm were set to minModuleSize = 20 for the gene/genus dendrogram and minimum height = 0.25 to cut the tree, in order to merge similar modules; v. identification of eigengenes for each module that is expressed by the first principal component of the expression matrix, vi. Pearson's correlation analysis of eigengenes with phenotypes (with *t*-test for testing the significance of the correlation coefficient), and finally, vii. identification of hub genes/genera – genes/genera that have the highest correlations with other genes/genera contained within each module. Genes contained within significantly associated modules were then subjected to GSEA.

3. Results

3.1. Microbiome

The analysis of the 16S rRNA gene allowed us to identify 232 unique genera. The genera, with abundance exceeding 10% in all the samples, were *Clostridium*, *57N15*, and *Treponema*. A detailed analysis of microbiota was described by Czech et al. (2022).

3.2. mRNA-seq

The analysis of the RNA-seq data identified 2,035 differentially expressed genes for rectal temperature, 1,886 for drooling score, and 1,958 for respiratory score. The expressions of the majority of those genes were down-regulated with increasing HS response, i.e., the higher value of phenotypes, the lower expression. This comprised 85% of down-regulated genes for rectal temperature, 78% for drooling score, and 80% for respiratory score. The most highly up-regulated genes were *ENSBTAG00000048590* (for rectal temperature), *ENSBTAG00000054209* (for respiratory score), and *SLC22A1* (for drooling score), while genes with the highest down-regulated expression were *ENSBTAG00000024272* (for rectal temperature), *ENSBTAG00000050067* (for respiratory score), and *ENSBTAG00000051290* (for drooling score). The 1,851 genes significantly associated with HS were common for all three phenotypes (Figure 1).

Next, we performed GSEA in which we identified seven KEGG pathways enriched among significantly differentially expressed genes that were shared between all three phenotypes: herpes simplex virus 1 infection (bta05168), viral protein interaction with cytokine and cytokine receptor (bta04061), chemokine signaling pathway (bta04062), cytokine-cytokine receptor interaction (bta04060), PI3K-Akt signaling pathway (bta04151), antifolate resistance (bta01523), and EGFR tyrosine kinase inhibitor resistance (bta01521). Results of GSEA for KEGGs were visualized in Figure 2. On the plot, we can see that Herpes simplex virus 1 infection is characterized with the lowest *P*-value of $2.73 \cdot 10^{-12}$ and also demonstrated the highest gene ratio of significantly associated genes that composed this pathway. Significantly enriched GO terms related to biological processes were identified only for respiratory score and were related to cell surface receptor signaling pathway (GO:0007166), cellular response to endogenous stimulus (GO:0071495), G protein-coupled receptor signaling pathway (GO:0007186), and metal ion transport (GO:0030001) (Figure 3).

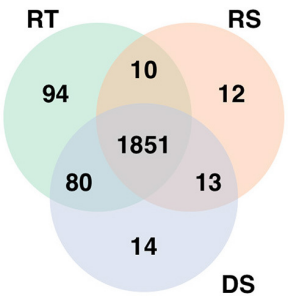

FIGURE 1
Genes significantly differentially expressed with increasing heat stress for each phenotype. RT denotes rectal temperature, RS respiratory score, and DS drooling score.

3.3. Omics integration

By applying steps described in the method section, the weighted co-expression network was generated. The adjacency matrix was created by raising the correlation matrix to the power of 4 (β parameter, Figure 4). In the next step, the TOM dissimilarity matrix was computed and used for the hierarchical clustering. Genes and bacteria were clustered into 20 modules, which ranged in size from 36 to 3015 genes/bacterial genera per module (Figure 5).

The effect of each gene/bacterial genera was expressed by the eigengene value, and the correlation of each eigengene with each HS phenotype was calculated (Figure 6). Three modules demonstrated significant correlations with rectal temperature (positive correlation for MEtan, and negative correlations for MElightycan, and MEroyalblue). Module MEtan consists of 129 genes but no bacterial genera, MElightycan of 26 genes and 26 bacterial genera, and module MEroyalblue of 2 genes and 34 bacterial genera.

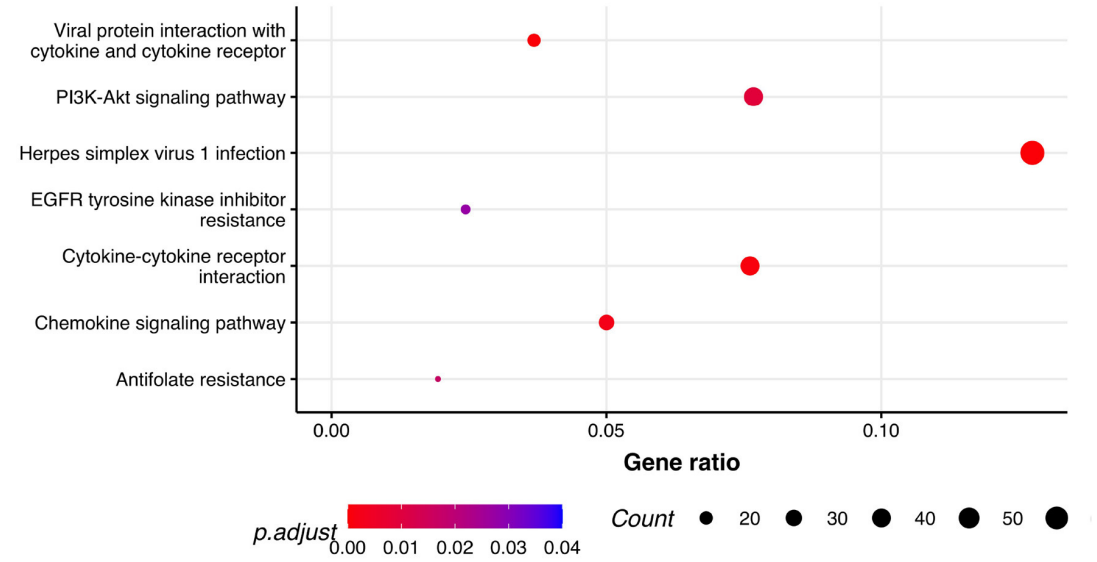

FIGURE 2
Pathway analysis. Dot plot of the statistically significant KEGG pathways shared between all the phenotypes. Gene ratio represents genes related to KEGG pathway/total number of significantly differentially expressed genes and count is the number of genes that belong to a given pathway.

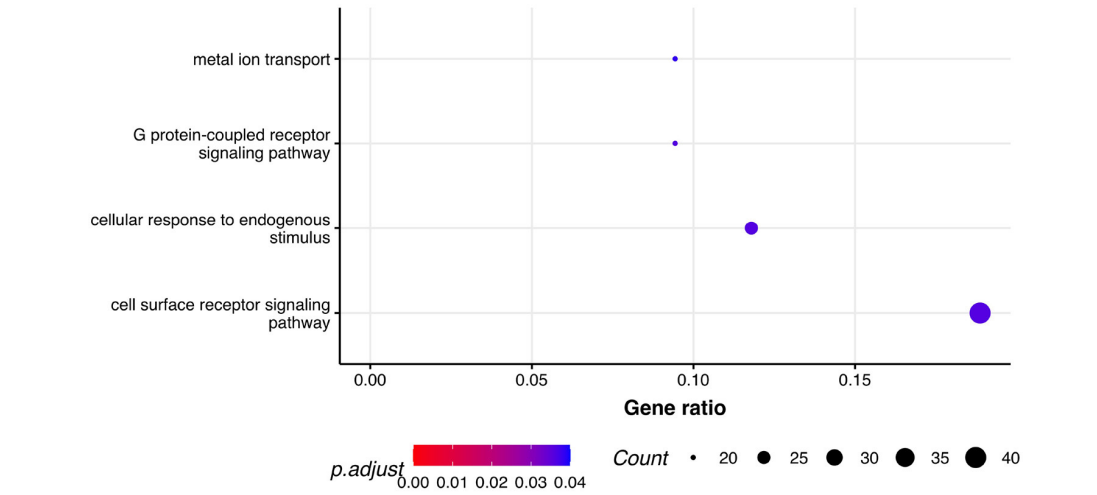

FIGURE 3
Pathway analysis. Dot plot of the statistically significant GO pathways related to biological processes for respiratory score phenotype. Gene ratio represents genes related to GO terms / total number of significantly differentially expressed genes and count is the number of genes that belong to a given pathway.

Further, we identified hub genes/bacterial genera representative for each of the three modules: *CSF3R* gene in MEtan, *Lactococcus* bacteria in MElightcyan, and *Rhizobium* bacteria in MEroyalblue. There was no overlap between genes contained within the significant modules and in the differential gene expression analysis. All genes from significant modules were annotated to GO terms and KEGG pathways. MEtan module was enriched in a pathway related to *Pertussis* and *Salmonella* infection (bta05133 and bta05132, respectively) and in GO terms related to the cellular response to organic substance (GO:0071310), response to oxygen-containing compound (GO:1901700), cellular response to lipid (GO:0071396), and cellular response to lipopolysaccharide (GO:0071222). The other two modules

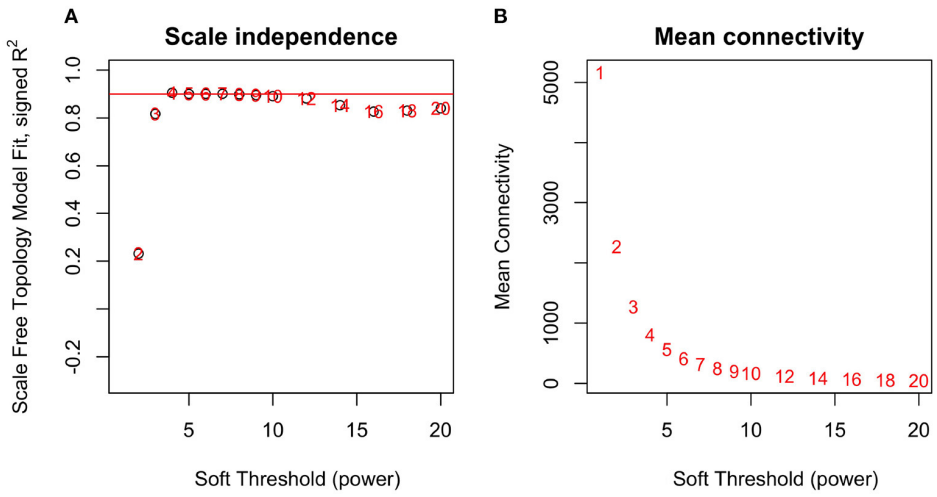

FIGURE 4
Network topology analysis for soft-thresholding powers in WGCNA-scale-free fit index for different powers **(A)** and mean connectivity analysis for different soft-thresholding powers **(B)**.

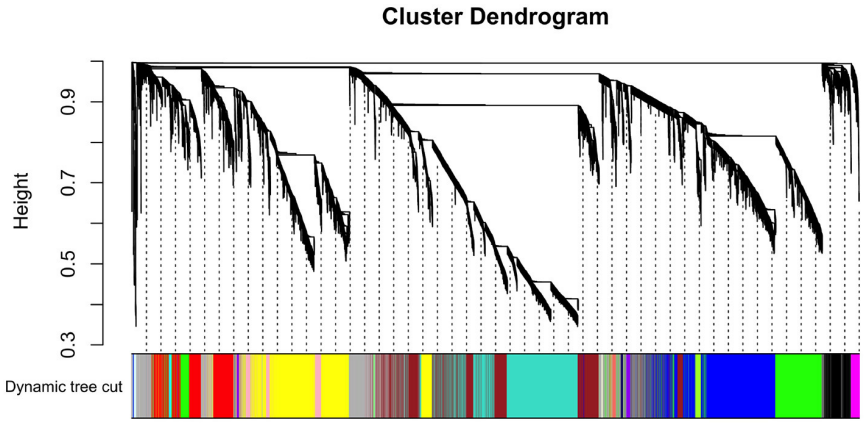

FIGURE 5
Hierarchical cluster tree of co-expressed genes and bacteria.

demonstrated no significant enrichment of KEGG and GO terms.

3.4. Discussion

This experiment is one of the first investigations in which the combined data of host transcriptome and microbiota were used together to study heat stress in cattle. The changes in cows' response to HS were identified on the level of gene expression alteration as well as on the level of the interaction with microbiota. Heat stress is undoubtedly a complex process that scientists today must face in order to protect animals. However, due to its physiological complexity, we are not able to assess in detail the changes in molecular mechanisms underlying HS response in livestock. The progressive development of molecular biology and bioinformatics allows for a broader look into changes in organisms, allowing simultaneous insight into the cell at virtually every stage of its life cycle. The RNA-seq technology has become a very powerful method for identifying candidate genes associated with complex traits. Already in Garner et al. (2020) identified *BDKRB1* and *SNORA19* as potential candidate genes related to HS. Sigdel et al. (2019) reported *HSF1*, *MAPK8IP1*, and *CDKN1B* as genes responsible

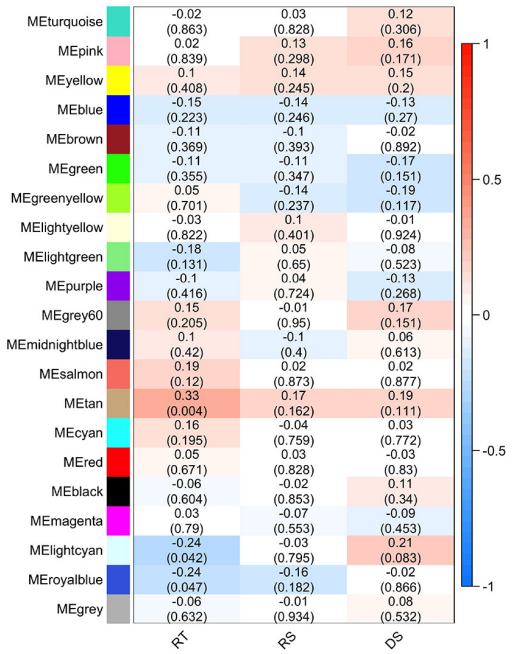

FIGURE 6
Module-trait associations. Each row corresponds to a module eigengene, column to a trait-rectal temperature (RT), respiratory score (RS), and drooling score (DS). Each cell contains the corresponding Pearson's correlation and *P*-value.

for thermotolerance in dairy cattle. Moreover, Diaz et al. (2021) identified five genes: *E2F8*, *GATAD2B*, *BHLHE41*, *FBXO44*, and *RAB39B* which were significantly associated with HS. In this study, it was found that the gene *RAB39B* was significantly associated with three phenotypes which included rectal temperature, drooling, and respiratory scores.

For the microbiome, Sales et al. (2021) identified four bacterial genera related to HS—*Flavonifractor*, *Treponema*, *Ruminococcus*, and *Carnobacterium*. In this study, it was identified that HS inhibits gene expression of several genes that might be related to the reduction of energy during overheating. Because HS appears to be a physiologically complex phenomenon, the multiomics approach that accounts not only for alteration in gene expression and changes in the microbiome composition but also for the interaction between them is an important approach. Recently also Martínez-Álvaro et al. (2022) demonstrated that in cattle the host genome affects not only the composition of the rumen microbiome but also the level of expression of microbial genes related to methane emissions. In this study, which is a follow-up analysis, genes and pathways were identified that are significantly associated with HS phenotypes. Additionally, interactions involving mRNA levels and microbiota in cattle were analyzed. Although the overlap between these findings and the microbiome and genes related to HS reported in the literature is constrained to only *RAB39B*, therefore it is hypothesized that this approach which is focused on the interaction between microbiome and host genetics was able to identify new components of the HS response that have been missed in the single omics analyzes. A loss of interaction under increased HS was observed. In two out of three significant modules, bacteria played a key role in the regulation of gene expression and controlled the abundance of other bacteria, while *CSF3R* gene was identified as the only hub gene in all significantly associated coexpression modules. Currently, the importance of this gene in the context of HS in cattle has not been reported yet. However, in human genetics, this gene is associated with congenital neutropenia (Triot et al., 2014). Park et al. (2021a) already reported that HS may affect neutrophil phagocytosis. Therefore, these results may indicate that gene *CSF3R* might be strictly associated with both neutrophils and HS response in cattle. Other significant hubs were represented by bacteria. *Lactococcus* bacteria that was identified as the hub of MElightcyan module was already indicated in the literature as a genus associated with bovine mastitis (Rodrigues et al., 2016). This observation stressed the important role of the gut microbiome in the regulation of gene expression. Our analysis indicates that for such physiologically complex phenomena like HS not only the effect of particular omics-based sources of information is important, but also the consideration of interactions between them.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The sequence data is available in the NCBI Sequence Read Archive with accession number SRP202074. Other datasets generated and/or analyzed during the current study are not publicly available due to institutional constraints. Requests to access these datasets should be directed to wangyachun@cau.edu.cn.

Author contributions

BC, YW, and JS conceived and conducted the experiment. BC analyzed the results and wrote the manuscript in consultation with JS, YW, KW, HL, and LH. All authors reviewed the manuscript read and approved the final manuscript.

Funding

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

Acknowledgments

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.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Biffani, S., Bernabucci, U., Vitali, A., Lacetera, N., and Nardone, A. (2016). Short communication: effect of heat stress on nonreturn rate of italian holstein cows. *J. Dairy Sci.* 99, 5837–5843. doi: 10.3168/jds.2015-10491

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Brito, L. F., Oliveira, H. R., McConn, B. R., Schinckel, A. P., Arrazola, A., Marchant-Forde, J. N., et al. (2020). Large-scale phenotyping of livestock welfare in commercial production systems: a new frontier in animal breeding. *Front. Genet.* 11, 793. doi: 10.3389/fgene.2020.00793

Broad Institute (2022). *Picard Tools*. Available online at: http://broadinstitute.github.io/picard/ (accessed July 15, 2022).

Carillier-Jacquin, C., Deru, V., Tusell, L., Bouquet, A., Jacquin, L., and Gilbert, H. (2022). Predicting pig digestibility coefficients with microbial and genomic data using machine learning prediction algorithms. Available online at: https://hal.inrae.fr/hal-03739243/document

Chen, Q., and Shi, D. (2004). The modeling of scale-free networks. *Physica A* 335, 240–248. doi: 10.1016/j.physa.2003.12.014

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

Consortium, T. G. O. (2020). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* 49, D325-D334. doi: 10.1093/nar/gkaa1113

Czech, B., Szyda, J., Wang, K., Luo, H., and Wang, Y. (2022). Fecal microbiota and their association with heat stress in bos taurus. *BMC Microbiol.* 22, 171. doi: 10.1186/s12866-022-02576-0

DeLuca, D. S., Levin, J. Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., et al. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* 28, 1530–1532. doi: 10.1093/bioinformatics/bts196

Diaz, F. A., Gutierrez-Castillo, E. J., Foster, B. A., Hardin, P. T., Bondioli, K. R., and Jiang, Z. (2021). Evaluation of seasonal heat stress on transcriptomic profiles and global DNA methylation of bovine oocytes. *Front. Genet.* 12, 699920. doi: 10.3389/fgene.2021.699920

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2012). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Freitas, P., Johnson, J., Wen, H., Maskal, J., Byrd, M., Tiezzi, F., et al. (2022). Genome-wide association analyses for various thermotolerance indicators measured during lactation in maternal-line pigs. Available online at: https://www.researchgate.net/publication/362322441_Genome-wide_association_analyses_for_various_thermotolerance_indicators_measured_during_lactation_in_maternal-line_pigs

Gao, S. T., Ma, L., Zhou, Z., Zhou, Z. K., Baumgard, L. H., Jiang, D., et al. (2019). Heat stress negatively affects the transcriptome related to overall metabolism and milk protein synthesis in mammary tissue of midlactating dairy cows. *Physiol. Genomics* 51, 400–409. doi: 10.1152/physiolgenomics.00039.2019

Garner, J. B., Chamberlain, A. J., Jagt, C. V., Nguyen, T. T. T., Mason, B. A., Marett, L. C., et al. (2020). Gene expression of the heat stress response in bovine peripheral white blood cells and milk somatic cells *in vivo*. *Sci. Rep.* 10, 19181. doi: 10.1038/s41598-020-75438-2

Kanehisa, M. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinform.* 9, 559. doi: 10.1186/1471-2105-9-559

Liu, Z., Cao, Y., Ai, Y., Yin, X., Wang, L., Wang, M., et al. (2022). Microbiome and ileum transcriptome revealed the boosting effects of selenium yeast on egg production in aged laying hens. *Anim. Nutr.* 10, 124–136. doi: 10.1016/j.aninu.2022.04.001

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Luo, H., Brito, L. F., Li, X., Su, G., Dou, J., Xu, W., et al. (2021). 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. *J. Dairy Sci.* 104, 4390–4403. doi: 10.3168/jds.2020-19192

Macciotta, N., Biffani, S., Bernabucci, U., Lacetera, N., Vitali, A., Ajmone-Marsan, P., et al. (2017). Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle. *J. Dairy Sci.* 100, 4683–4697. doi: 10.3168/jds.2016-12249

Martínez-Álvaro, M., Auffret, M. D., Duthie, C.-A., Dewhurst, R. J., Cleveland, M. A., Watson, M., et al. (2022). Bovine host genome acts on rumen microbiome function linked to methane emissions. *Commun. Biol.* 5, 350. doi: 10.1038/s42003-022-03293-0

Park, D. S., Gu, B.-H., Park, Y. J., Joo, S. S., Lee, S.-S., Kim, S.-H., et al. (2021a). Dynamic changes in blood immune cell composition and function in holstein and jersey steers in response to heat stress. *Cell Stress Chaperones* 26, 705–720. doi: 10.1007/s12192-021-01216-2

Park, Y., Heider, D., and Hauschild, A.-C. (2021b). Integrative analysis of next-generation sequencing for next-generation cancer research toward artificial intelligence. *Cancers* 13, 3148. doi: 10.3390/cancers13133148

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 41, D590-D596. doi: 10.1093/nar/gks1219

Ramayo-Caldas, Y., Zingaretti, L. M., Pérez-Pascual, D., Alexandre, P. A., Reverter, A., Dalmau, A., et al. (2021). Leveraging host-genetics and gut microbiota to determine immunocompetence in pigs. *Anim. Microbiome* 3, 74. doi: 10.1186/s42523-021-00138-9

Rio, D. C., Ares, M., Hannon, G. J., and Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*,2010, pdb.prot5439. doi: 10.1101/pdb.prot5439

Rodrigues, M., Lima, S., Higgins, C., Canniatti-Brazaca, S., and Bicalho, R. (2016). The lactococcus genus as a potential emerging mastitis pathogen group: A report on an outbreak investigation. *J. Dairy Sci*. 99, 9864–9874. doi: 10.3168/jds.2016-11143

Sales, G. F. C., Carvalho, B. F., Schwan, R. F., de Figueiredo Vilela, L., Meneses, J. A. M., Gionbelli, M. P., et al. (2021). Heat stress influence the microbiota and organic acids concentration in beef cattle rumen. *J. Therm. Biol*. 97, 102897. doi: 10.1016/j.jtherbio.2021.102897

Shah, T. M., Patel, J. G., Gohil, T. P., Blake, D. P., and Joshi, C. G. (2019). Host transcriptome and microbiome interaction modulates physiology of full-sibs broilers with divergent feed conversion ratio. *npj Biofilms Microbiomes* 5, 24. doi: 10.1038/s41522-019-0096-3

Sigdel, A., Abdollahi-Arpanahi, R., Aguilar, I., and Pe nagaricano, F. (2019). Whole genome mapping reveals novel genes and pathways involved in milk production under heat stress in US holstein cows. *Front. Genet*. 10, 928. doi: 10.3389/fgene.2019.00928

Triot, A., Järvinen, P. M., Arostegui, J. I., Murugan, D., Kohistani, N., Díaz, J. L. D., et al. (2014). Inherited biallelic CSF3r mutations in severe congenital neutropenia. *Blood* 123, 3811–3817. doi: 10.1182/blood-2013-11-535419

Wang, Q., Wang, K., Wu, W., Giannoulatou, E., Ho, J. W. K., and Li, L. (2019). Host and microbiome multi-omics integration: applications and methodologies. *Biophys. Rev*. 11, 55–65. doi: 10.1007/s12551-018-0491-7

Yip, A. M., and Horvath, S. (2007). Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinform*. 8, 22. doi: 10.1186/1471-2105-8-22

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol*. 11, R14. doi: 10.1186/gb-2010-11-2-r14

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an r package for comparing biological themes among gene clusters. *OMICS* 6, 284–287. doi: 10.1089/omi.2011.0118

Yue, S., Wang, Z., Wang, L., Peng, Q., and Xue, B. (2020). Transcriptome functional analysis of mammary gland of cows in heat stress and thermoneutral condition. *Animals* 10, 1015. doi: 10.3390/ani10061015

3.3 Genome-wide association study of heat stress response in Bos taurus

By analyzing DNA variation expressed by single nucleotide polymorphisms (SNPs) from 68 cows identified by using oligonucleotide microarrays this study revealed 17 SNPs distributed across three chromosomes that are strongly associated with heat stress. Notably, these discerned SNPs are intricately linked to the genes *PDZRN4* and *PRKG1*, which are known to participate in fundamental protein degradation pathways and the regulation of blood vessel dilation, respectively. These significant findings underscore the potential importance of *PDZRN4* and *PRKG1* in the heat stress tolerance of cattle, thereby providing valuable genetic markers for prospective research investigations and breeding programs. Nevertheless, it is worth underscoring that this study signifies the necessity of doing more in-depth exploration to elucidate the intricate biological pathways that influence heat stress tolerance. Furthermore, it is noteworthy to mention that the low statistical power observed in this study is attributed to the relatively modest sample size utilized, underscoring the necessity for subsequent investigations to corroborate and extend the scope of our current findings within this pivotal realm of research. Consequently, further research endeavors are warranted to refine our understanding of this complex biological phenomenon.

Genome-wide association study of heat stress response in *Bos taurus*

Bartosz Czech[1, *], Yachun Wang[2], Joanna Szyda[1]

1 Biostatistics Group, Department of Genetics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland
2 National Engineering Laboratory for Animal Breeding, Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, College of Animal Science and Technology, China Agricultural University, Beijing, China

* bartosz.czech@upwr.edu.pl

Abstract

Heat stress is a major challenge in cattle production, affecting animal welfare, productivity, and economic viability of the industry. In this study, we conducted a genome-wide association study (GWAS) to identify genetic markers associated with tolerance to heat stress in Chinese Holstein cattle. We genotyped 68 cows using Illumina 150K Bovine BeadChip microarray and analysed 112,081 single nucleotide polymorphisms using a linear model-based GWAS approach. We identified 17 SNPs distributed on three chromosomes that showed statistically significant associations with tolerance to heat stress in Chinese Holstein cattle. Five of them were located in introns of two genes, *PDZRN4* and *PRKG1*. *PDZRN4* is involved in protein degradation pathways, while *PRKG1* encodes a protein kinase involved in smooth muscle relaxation and blood vessel dilation. Our findings highlight the potential importance of *PDZRN4* and *PRKG1* in heat stress tolerance in cattle and provide valuable genetic markers for further research and breeding programmes aimed at improving the tolerance to heat stress in Holstein cattle. However, more studies are needed to elucidate the exact mechanisms by which these SNPs contribute to tolerance to heat stress and their potential implications for practical cattle breeding strategies.

Author summary

Heat stress is a critical challenge in cattle production, leading to reduced productivity and increased mortality rates. In our study, we conducted a genome-wide association study (GWAS) to identify genetic markers associated with indicators of tolerance to heat stress in cattle. We found significant associations between indicators of heat stress tolerance and specific single nucleotide polymorphisms (SNPs) located in two genes, *PDZRN4* and *PRKG1*. These genes are known to play roles in protein degradation pathways and smooth muscle relaxation, respectively, and have previously been implicated in physiological responses to heat stress in other species. Our findings provide insight into the genetic mechanisms underlying heat stress tolerance in cattle and could potentially be used in genomic selection programmes aimed at improving heat stress tolerance in cattle populations. More research is needed to elucidate the functional importance of these SNPs and their potential applications in cattle breeding programmes.

Introduction

Heat stress in cattle occurs when the animal's body temperature rises above physiologically normal levels due to exposure to high temperatures, humidity and solar radiation [1]. This can occur in both dairy and beef cattle and is a significant problem for the livestock industry, particularly in regions with hot and humid climates [2].

Heat stress can affect cattle in several ways. First, it can cause a decrease in feed intake and, later, reduce weight gain or milk production [3]. Second, heat stress can result in respiratory distress, panting, and increased water consumption, which can put additional strain on the animal's cardiovascular system [4]. Finally, severe heat stress can lead to dehydration, electrolyte imbalances, and dramatic changes in animal physiology [5]. In addition to the direct impact on animal health, heat stress also results in economic losses due to reduced milk production and a lower reproduction rate [6].

An exposition of cows to prolonged periods of heat stress changes in gene expression and epigenetic modifications, which can ultimately affect the animal's health, productivity and even the genetics of their offspring [7].

Studies have shown that heat stress can lead to changes in the expression of genes related to immune function, metabolism, and reproduction. For example, heat stress can cause a decrease in the expression of genes involved in milk production and an increase in the expression of genes involved in stress responses [8].

In the case of heat stress in cattle, GWAS can be used to identify genetic variants that are associated with body temperature, drooling score, and respiratory score, for example. This will allow breeding strategies to be developed to select animals that are more tolerant to heat stress and maintain productivity under hot and humid conditions [9].

Moreover, from the scientific perspective, GWAS allows understanding of the genetic basis of heat stress, including the biological pathways and mechanisms involved in the response to heat stress. In general, the combination of NGS, genotyping microarrays, and GWAS can provide a powerful approach to the identification of genetic variants and even candidate genes associated with the response to heat stress in cattle. This knowledge can be used to develop new management practises, breeding strategies, and therapeutics to improve animal welfare and productivity in a changing environment.

The purpose of this study was to identify genetic variants and metabolic pathways associated with the response to heat stress in cattle that lead to a better understanding of the functional basis of tolerance to heat stress in cattle.

Materials and methods

The material consists of 68 cows representing Chinese Holstein cattle. These individuals were genotyped using the Illumina 150K Bovine BeadChip (Illumina Inc., San Diego, CA, USA), which consists of 123,268 single nucleotide polymorphisms (SNPs). For all animals, the responses to heat stress were expressed by rectal temperature (RT), drooling score (DS), and respiratory score (RS). These phenotypes were represented by deregressed proofs (DRP) of estimated breeding values (EBVs) predicted as previously described in Czech *et al.* [10] [11].

The preprocessing of genotype data consisted of retaining: i) individuals with call rate greater than 0.9, ii) SNPs with minor allele frequency (MAF) greater than 0.05, and iii) SNPs that were in the Hardy-Weinberg equilibrium (P-value > 0.05). The filtration process was performed using PLINK software (v1.90b6.21) [12]. Subsequently, GWAS was performed separately for each phenotype, using the following model:

$$y = \mu + \beta X + \epsilon \tag{1}$$

where y is a vector of DRPs, X contains SNP genotype coded as 0, 1, or 2, representing the number of reference alleles, β is the SNP additive effect, and ϵ represents residuals. The significance of a SNP effect was tested using the likelihood ratio test with the reduced model represented by model 1 without the SNP effect. The estimation of the model parameter and testing of the significance of the SNP effect were performed using the GEMMA software [13]. To control for multiple testing P-values were adjusted using the Bonferroni correction. Significant SNPs were considered based on the adjusted P-values lower than 0.05. All significant SNPs were annotated using the Variant Effect Predictor (VEP) implemented in the ensemblVEP R package with Ensembl Release 109 (Feb 2023) [14]. Additionally, the Animal QTL database was used (QTLdb) to explain the genetic basis of variation in heat stress phenotypes [15]. Next, the Gene-Set Enrichment Analysis (GSEA) was performed to detect potential functional pathways underlying the heat stress response by applying one-sided version of Fisher's exact test. The Gene Ontology (GO) [16] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [17] were considered in GSEA implemented in the clusterProfiler R package [18].

Ethics approval and consent to participate

The data collection process was carried out strictly according to 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 reporting animal experiments.

Results

Genome-wide association study

The filtration process retained 112 081 out of 123 268 SNPs (91%) for GWAS for all 68 individuals. As a result of rectal temperature, 17 significant SNPs were identified, while no significant hits were observed for drooling and respiratory scores. Significant SNPs associated with rectal temperature were located on chromosomes 5, 17, and 26. On the *Bos taurus* autosome (BTA) 5 there were three significant SNPs, on BTA17 there were 12 significant SNPs, while on BTA26 there were only two significant SNPs. Manhattan plots were presented in Figure 1 (for rectal temperature), in Figure 2 (for drooling score) and in Figure 3 (for respiratory score). Table 1 shows detailed information on the significantly associated SNPs with rectal temperature. All SNPs were further annotated and processed through GSEA.

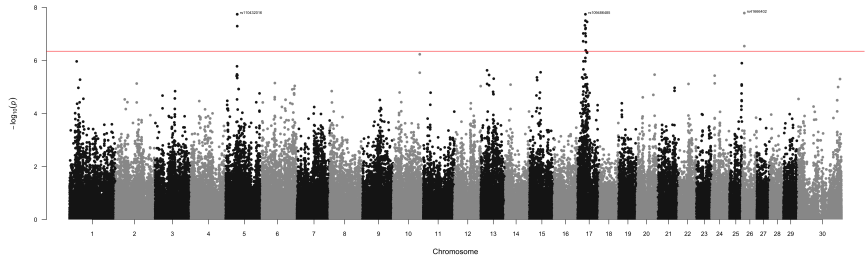

Fig 1. Manhattan plot for associations of SNPs with the rectal temperature. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

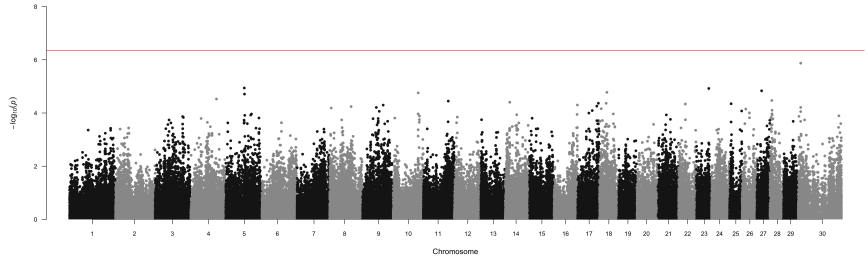

Fig 2. Manhattan plot for associations of SNPs with the drooling score. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

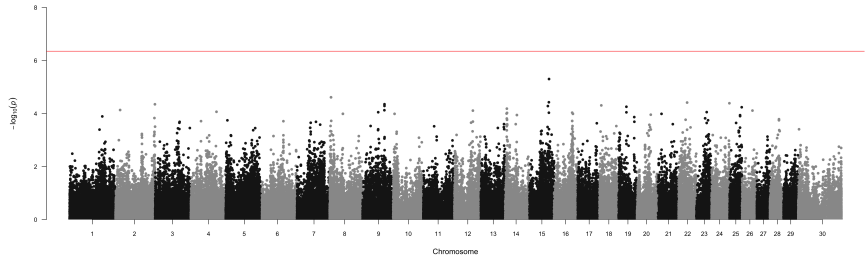

Fig 3. Manhattan plot for associations of SNPs with the respiratory score. X-axis: SNPs positions on chromosomes, Y-axis: -log10 p-value. The red line indicates the 0.05 significance threshold corrected for multiple testing. Labels show the top significant SNPs.

Table 1. The genomic annotation of associated SNPs with rectal temperature.

Chr.	Position	Alternative allele	MAF	Genomic annotation	Gene	p-adj
5	39 389 648	T	0.478	intron variant	*PDZRN4*	0.002
5	39 451 743	C	0.478	intron variant	*PDZRN4*	0.002
5	39 470 772	A	0.463	intron variant	*PDZRN4*	0.006
17	18 454 906	G	0.191	downstream gene variant	ENSBTAG00000015811	0.021
17	18 911 911	C	0.265	intergenic variant	-	0.011
17	24 328 037	C	0.309	intergenic variant	-	0.005
17	25 977 762	G	0.265	intergenic variant	-	0.011
17	26 111 051	T	0.235	intergenic variant	-	0.007
17	26 151 841	C	0.250	intergenic variant	-	0.002
17	26 207 829	C	0.235	intergenic variant	-	0.004
17	26 946 234	C	0.243	intergenic variant	-	0.007
17	27 068 312	C	0.243	intergenic variant	-	0.013
17	27 100 624	G	0.287	intergenic variant	-	0.047
17	27 837 317	T	0.221	intergenic variant	-	0.023
17	31 952 291	G	0.243	intergenic variant	-	0.004
26	8 148 418	G	0.353	intron variant	*PRKG1*	0.002
26	8 194 110	T	0.353	intron variant	*PRKG1*	0.032

Variants annotation and Gene-Set Enrichment Analysis

Results of the annotation process performed using VEP were summarised in Table 1 which shows that all significantly associated SNPs with rectal temperature on BTA5 and BTA26 were located in introns of *PDZRN4* (BTA5) and *PRKG1* (BTA25) genes. On BTA17, 11 out of 12 SNPs were located in the intergenic regions, while one SNP was located in the downstream part of the ENSBTAG00000015811 gene. The annotation of the phenotype based on QTLdb demonstrated that SNPs 17:27837317 (rs110432016) and 17:25977762 (rs109962820) were related to a maternal component of calving ease, dairy form, daughter pregnancy rate, foot angle, milk fat percentage, milk fat yield, net merit, length of productive life, milk protein percentage, milk protein yield, rear leg placement, and teat length. GSEA based on the GO indicated that genes related to the associated SNPs were enriched in the following ontologies: GO:0003682:chromatin binding, GO:0004672:protein kinase activity, GO:0004674:protein serine/threonine kinase activity, GO:0004692:cGMP-dependent protein kinase activity, GO:0005524:ATP binding, GO:0006468:protein phosphorylation, GO:0042802:identical protein binding, GO:0005515:protein binding, and GO:0046872:metal ion binding. GSEA for the KEGG pathways showed only enrichment of the cGMP-PKG signaling pathway (bta04022).

Discussion

Heat stress is an important environmental challenge for livestock production, including cattle, as it can negatively affect animal health, welfare, and productivity. In this study, we performed a GWAS to identify genetic markers associated with heat stress in cattle. Our findings revealed significant associations between heat stress and single nucleotide polymorphisms (SNPs) located in the *PDZRN4* and *PRKG1* genes, shedding light on the mechanisms underlying the response to heat stress in cattle. This GWAS study serves as a follow-up to previous analyses of differential gene expression and differential abundance of the microbiota in the context of heat stress, providing further insight into the complex interactions between genetics, gene expression, microbiota, and response to heat stress. Previous studies have already demonstrated the importance of rectal temperature as a main indicator of heat stress in cattle. It has been shown that all of the three phenotypes (rectal temperature, drooling score, and respiratory score), rectal temperature showed a major association with gene expression and abundance of microbiota in cattle under heat stress conditions [10] [11]. The heat stress phenotype is difficult to quantify and out of the three measurements that were available in this study, only rectal temperature appeared to be the most representative of heat stress.

There are many publications on GWAS related to heat stress in cattle, however, almost all of them focused on the standard case-control experimental design in which one cannot identify potential candidate genes responsible for the heat stress response. It is due to the complex nature of heat stress and the involvement of multiple genes and environmental factors. However, unlike previous studies that have focused on controlled experimental environments, our study examined animals in their production environment, providing valuable information on the genetic factors that influence the tolerance of heat stress in cattle under production conditions. Although this approach allows for capturing the genetic variation present in the population, it also has limitations, including potential confounding factors and the lack of control over environmental variables that may interact with the heat stress phenotype in real production systems.

The identification of SNPs in *PDZRN4* and *PRKG1* associated with heat stress in cattle suggests that these genes may play a role in the physiological response of cattle to heat stress. *PDZRN4* gene (PDZ domain containing the ring finger 4) also known as

LNX4 (Ligand of Numb Protein-X 4) plays a potential role as a tumour suppressor gene and may have an antiproliferative effect on hepatocellular carcinoma cell proliferation [19]. Another study showed that *PDZRN4* is a functional suppressor of prostate cancer growth [20]. However, there are no studies in which *PDZRN4* was indicated as a candidate gene related to the response to heat stress. Studies related to other livestock species showed that this gene could affect fat metabolism in pigs [21]. Furthermore, *PDDRN4* was found to be a significant gene associated with poor sperm motility in Holstein-Friesian bulls [22]. However, another gene identified in this study was *PRKG1* that encodes a protein called cGMP-dependent protein kinase 1 [23]. *PRKG1* was found as the gene associated with tick resistance in South African Nguni cattle [24]. Another study showed the importance of this gene in the local adaptation of indigenous Ugandan cattle to East Coast Fever [25]. However, the most interesting is that the gene *PRKG1* has already been found as a gene with a key role in body thermoregulation. In the study that focused on the adaptation to cold of indigenous Siberian populations, *PRKG1* has been shown to be the gene involved in cold acclimatisation [26]. Another study showed that this gene was the key to minimising heat loss by regulating blood vessel constriction in Yakutian horses [27]. There is also a study confirming the important role of *PRKG1* in temperature regulation in a cold environment in the Amur tiger [28]. Regarding the phenomenon of heat stress, it has been shown that *PRKG1* is associated with adaptation to heat stress in Egyptian sheep breeds [29].

Tolerance to heat stress is a complex trait that involves the interplay of multiple genetic and environmental factors [30]. SNPs located in *PDZRN4* and *PRKG1* provide valuable markers for selecting heat-stress-tolerant animals in breeding programmes. This may lead to the development of genomic selection programmes to improve heat stress resistance in cattle and improve animal welfare and productivity in hot climates.

It is important to note that our study has some limitations. First, the sample size may affect the statistical power to detect all SNPs associated with heat stress. However, having low power implies that the significant associations observed in our study may represent genes with an especially high impact on resistance to heat stress. Especially that *PRKG1* has already been confirmed as a heat stress-associated gene in other species, including humans. However, functional validation of SNPs located in both genes is warranted to further elucidate the underlying physiological mechanisms. Furthermore, more studies with larger sample sizes are needed to verify our findings and eventually identify additional SNPs and candidate genes with lower effects on heat stress.

Conclusion

In this study, we identified significant associations between SNPs located in the *PDZRN4* and *PRKG1* genes and heat stress in cattle, providing important information on the genetic basis of the tolerance to heat stress. Our findings contribute to the understanding of the physiological mechanisms underlying the response to heat stress and provide potential genetic markers for selecting animals tolerant to heat stress in breeding programmes. Due to the limited sample size, further research is needed to validate our findings and identify genes with low to moderate effects.

References

1. Idris M, Uddin J, Sullivan M, McNeill DM, Phillips CJC. Non-Invasive Physiological Indicators of Heat Stress in Cattle. Animals. 2021;11(1):71. doi:10.3390/ani11010071.

2. Summer A, Lora I, Formaggioni P, Gottardo F. Impact of heat stress on milk and meat production. Animal Frontiers. 2018;9(1):39–46. doi:10.1093/af/vfy026.

3. Liu J, Li L, Chen X, Lu Y, Wang D. Effects of heat stress on body temperature, milk production, and reproduction in dairy cows: a novel idea for monitoring and evaluation of heat stress — A review. Asian-Australasian Journal of Animal Sciences. 2019;32(9):1332–1339. doi:10.5713/ajas.18.0743.

4. Kitajima K, Oishi K, Miwa M, Anzai H, Setoguchi A, Yasunaka Y, et al. Effects of Heat Stress on Heart Rate Variability in Free-Moving Sheep and Goats Assessed With Correction for Physical Activity. Frontiers in Veterinary Science. 2021;8. doi:10.3389/fvets.2021.658763.

5. Parker AJ, Hamlin GP, Coleman CJ, Fitzpatrick LA. Dehydration in stressed ruminants may be the result of acortisol-induced diuresis1. Journal of Animal Science. 2003;81(2):512–519. doi:10.2527/2003.812512x.

6. Wankar AK, Rindhe SN, Doijad NS. Heat stress in dairy animals and current milk production trends, economics, and future perspectives: the global scenario. Tropical Animal Health and Production. 2021;53(1). doi:10.1007/s11250-020-02541-x.

7. Corvo MD, Lazzari B, Capra E, Zavarez L, Milanesi M, Utsunomiya YT, et al. Methylome Patterns of Cattle Adaptation to Heat Stress. Frontiers in Genetics. 2021;12. doi:10.3389/fgene.2021.633132.

8. Corazzin M, Saccà E, Lippe G, Romanzin A, Foletto V, Borso FD, et al. Effect of Heat Stress on Dairy Cow Performance and on Expression of Protein Metabolism Genes in Mammary Cells. Animals. 2020;10(11):2124. doi:10.3390/ani10112124.

9. Gallagher MD, Chen-Plotkin AS. The Post-GWAS Era: From Association to Function. The American Journal of Human Genetics. 2018;102(5):717–730. doi:10.1016/j.ajhg.2018.04.002.

10. Czech B, Wang Y, Wang K, Luo H, Hu L, Szyda J. Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition. Frontiers in Microbiology. 2022;13. doi:10.3389/fmicb.2022.998093.

11. Czech B, Szyda J, Wang K, Luo H, Wang Y. Fecal microbiota and their association with heat stress in Bos taurus. BMC Microbiology. 2022;22(1). doi:10.1186/s12866-022-02576-0.

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics. 2007;81(3):559–575. doi:10.1086/519795.

13. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics. 2012;44(7):821–824. doi:10.1038/ng.2310.

14. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016;17(1). doi:10.1186/s13059-016-0974-4.

15. Hu ZL, Park CA, Wu XL, Reecy JM. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Research. 2012;41(D1):D871–D879. doi:10.1093/nar/gks1150.

16. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Research. 2018;47(D1):D419–D426. doi:10.1093/nar/gky1038.

17. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Research. 2015;44(D1):D457–D462. doi:10.1093/nar/gkv1070.

18. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation. 2021;2(3):100141. doi:10.1016/j.xinn.2021.100141.

19. Hu T, Yang H, Han ZG. PDZRN4 acts as a suppressor of cell proliferation in human liver cancer cell lines. Cell Biochemistry and Function. 2015;33(7):443–449. doi:10.1002/cbf.3130.

20. Jin P, Wu L, Zhang G, Yang B, Zhu B. PDZRN4 suppresses tumorigenesis and androgen therapy-resistance in prostate cancer. Journal of Cancer. 2022;13(7):2293–2300. doi:10.7150/jca.69269.

21. Wang BB, Hou LM, Zhou WD, Liu H, Tao W, Wu WJ, et al. Genome-wide association study reveals a quantitative trait locus and two candidate genes on Sus scrofa chromosome 5 affecting intramuscular fat content in Suhuai pigs. Animal. 2021;15(9):100341. doi:10.1016/j.animal.2021.100341.

22. Hering DM, Olenski K, Kaminski S. Genome-wide association study for poor sperm motility in Holstein-Friesian bulls. Animal Reproduction Science. 2014;146(3-4):89–97. doi:10.1016/j.anireprosci.2014.01.012.

23. Coudert E, Gehant S, de Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics. 2022;39(1). doi:10.1093/bioinformatics/btac793.

24. Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks and Tick-borne Diseases. 2016;7(3):487–497. doi:10.1016/j.ttbdis.2016.02.005.

25. Vajana E, Barbato M, Colli L, Milanesi M, Rochat E, Fabrizi E, et al. Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever. Frontiers in Genetics. 2018;9. doi:10.3389/fgene.2018.00385.

26. Cardona A, Pagani L, Antao T, Lawson DJ, Eichstaedt CA, Yngvadottir B, et al. Genome-Wide Analysis of Cold Adaptation in Indigenous Siberian Populations. PLoS ONE. 2014;9(5):e98076. doi:10.1371/journal.pone.0098076.

27. Librado P, Sarkissian CD, Ermini L, Schubert M, Jónsson H, Albrechtsen A, et al. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. Proceedings of the National Academy of Sciences. 2015;112(50). doi:10.1073/pnas.1513696112.

28. Du H, Yu J, Li Q, Zhang M. New Evidence of Tiger Subspecies Differentiation and Environmental Adaptation: Comparison of the Whole Genomes of the Amur Tiger and the South China Tiger. Animals. 2022;12(14):1817. doi:10.3390/ani12141817.

29. Aboul-Naga AM, Alsamman AM, Allali AE, Elshafie MH, Abdelal ES, Abdelkhalek TM, et al. Genome-wide analysis identified candidate variants and genes associated with heat stress adaptation in Egyptian sheep breeds. Frontiers in Genetics. 2022;13. doi:10.3389/fgene.2022.898522.

30. Cheruiyot EK, Haile-Mariam M, Cocks BG, Pryce JE. Improving Genomic Selection for Heat Tolerance in Dairy Cattle: Current Opportunities and Future Directions. Frontiers in Genetics. 2022;13. doi:10.3389/fgene.2022.894067.

Chapter 4

Summary

The research presented in this dissertation revealed valuable relationships between heat stress in cattle and changes at the microbiome level, gene expression, and the entire genome. This chapter summarizes the research presented in scientific publications and outlines its limitations and directions for further research.

Microbiome and Heat Stress: Insights and Biomarkers

The first study underscored the profound impact of heat stress on the microbial composition of cattle digestive systems. The microbiome, often regarded as the "second genome", was shown to be responsive to physiological changes arising from heat stress. The identification of fecal microbiota linked to rectal temperature, drooling, and respiratory scores provided a novel perspective into the microbial response to heat stress. By applying advanced molecular techniques, we demonstrated that *Rhizobium* and *Pseudobutyrivibrio* emerged as pivotal genera associated with heat stress, while *Acidobacteria* and *Gemmatimonadetes* played key roles at the phylum level. These findings not only highlighted the microbial communities' sensitivity to heat stress but also raised the possibility of utilizing these microbial markers as heat stress indicators.

Transcriptomic Insights and Regulatory Networks

The second study used RNA sequencing techniques to identify genes whose expression was altered by heat stress. The interconnections between the transcriptome and microbiome were unveiled, illustrating the mutual influence between these molecular layers. The identification of differentially expressed genes within the context of various heat stress phenotypes offered insights into the intricate mechanisms underlying cattle responses. The integration of bioinformatics techniques revealed gene modules, linked to the cytokine-cytokine receptor interaction pathway, underpinning the vital role of the microbiome in gene expression regulation during heat stress. This interaction may be used to develop biomarkers of heat stress in cattle.

Genetic Markers for Heat Stress Tolerance

The third study focused on the genetic factors affecting the tolerance of Chinese Holstein cattle to heat stress. By conducting a genome-wide association study, we identified key single nucleotide polymorphisms (SNPs) distributed across chromosomes, signifying their associations with heat stress resilience. Notably, SNPs within the *PDZRN4* and *PRKG1* genes provided genetic markers with potential implications for breeding programs. The identification of these markers

provided a foundation for the selection aimed at enhancing heat stress tolerance within cattle populations.

Limitations and Future Directions

Collectively, these studies provide a holistic understanding of the multifaceted responses of cattle to heat stress. The integration of genomics, transcriptomics, and microbiomics unraveled intricate regulatory networks that govern biological reactions to heat stress. The identified biomarkers and genetic markers present promising avenues for future research, encompassing the development of strategies to mitigate heat stress effects on animal welfare and productivity. As the global climate continues to evolve, the insights garnered from this research are poised to guide evidence-based decisions in livestock management and breeding practices. Further studies are warranted to delve deeper into the mechanistic underpinnings of the identified biomarkers and genetic markers, paving the way for sustainable solutions that uphold animal welfare and maintain the economic viability of the cattle industry. Understanding the impact of heat stress on cattle, as these studies have shown, is the cornerstone of future efforts to address the challenges posed by changes in the environment.

It should be noted that the study has potential limitations. One of the main limitations of our research, as in many modern research initiatives, is the limited size of the sample. Due to practical and financial constraints associated with collecting and analyzing multi-omics data, the scope of the study population was limited. This limitation has a number of implications for the generality and statistical ability of our findings. The small sample size limits the possibility of extrapolating our results to larger populations or different environments. It is important to recognize that variability in our limited sample may not capture the entire spectrum of biological diversity, which may lead to biased results. This limitation emphasizes the importance of careful interpretation and calls for studies with larger and more diverse cohorts to validate our conclusions. Furthermore, the small sample size affects the statistical power of our analyses. With fewer data points, it becomes challenging to detect subtle or modest effects, increasing the risk of both type I and type II errors. Researchers should be aware that some associations or trends may remain undetected due to the limited sample, and this should be considered when interpreting the significance of the results. Furthermore, the small sample size affects the statistical power of our analyses. With fewer data points, it becomes challenging to detect subtle or modest effects, increasing the risk of both type I and type II errors. Researchers should be aware that some associations or trends may remain undetected due to the limited sample, and this should be considered when interpreting the significance of our results. Another limitation in our study relates to the issue of missing data across different molecular layers. The availability of complete multiomics data for all individuals is a considerable challenge in many multiomics investigations. In our study, not all animals had data for all omics, which impedes a comprehensive multiomics analysis involving all molecular layers simultaneously. However, the problem of small sample sizes and missing data is a common challenge in multiomics research, and researchers need to be careful when interpreting and generalizing the results. Future research should seek to overcome these limitations to advance our understanding of complex biological systems.

In conclusion, this research underscores the power of bioinformatics in unraveling the complexities of cattle responses to heat stress. The integration of

diverse omics data has facilitated a comprehensive view of how molecular, genetic, and ecological factors intertwine in the face of adversity. As our world grapples with climate change, this dissertation's findings have far-reaching implications for enhancing livestock management practices, breeding strategies, and sustainable agricultural practices. The connection between bioinformatics and life sciences holds the promise of deeper insights into the mechanisms governing adaptation, contributing to the welfare of livestock and the security of global food systems.

Bibliography

Abbas, Zaheer et al. (July 2020). "Glucose Metabolism and Dynamics of Facilitative Glucose Transporters (GLUTs) under the Influence of Heat Stress in Dairy Cattle". In: *Metabolites* 10.8, p. 312. DOI: 10.3390/metabo10080312.

Aboul-Naga, Adel M. et al. (Oct. 2022). "Genome-wide analysis identified candidate variants and genes associated with heat stress adaptation in Egyptian sheep breeds". In: *Frontiers in Genetics* 13. DOI: 10.3389/fgene.2022.898522.

Andretta, I. et al. (Sept. 2021). "Modeling the performance of broilers under heat stress". In: *Poultry Science* 100.9, p. 101338. DOI: 10.1016/j.psj.2021.101338.

Andrews, S. (2010). *FASTQC. A quality control tool for high throughput sequence data*.

Archana, PR et al. (Jan. 2017). "Role of Heat Shock Proteins in Livestock Adaptation to Heat Stress". In: *Journal of Dairy, Veterinary &amp Animal Research* 5.1. DOI: 10.15406/jdvar.2017.05.00127.

Ashburner, Michael et al. (May 2000). "Gene Ontology: tool for the unification of biology". In: *Nature Genetics* 25.1, pp. 25–29. DOI: 10.1038/75556.

Baek, Youl Chang et al. (Mar. 2020). "The impact of short-term acute heat stress on the rumen microbiome of Hanwoo steers". In: *Journal of Animal Science and Technology* 62.2, pp. 208–217. DOI: 10.5187/jast.2020.62.2.208.

Bai, Hanako et al. (Jan. 2020). "Effect of summer heat stress on gene expression in bovine uterine endometrial tissues". In: *Animal Science Journal* 91.1. DOI: 10.1111/asj.13474.

Bailey, Michael T. et al. (Mar. 2011). "Exposure to a social stressor alters the structure of the intestinal microbiota: Implications for stressor-induced immunomodulation". In: *Brain, Behavior, and Immunity* 25.3, pp. 397–407. DOI: 10.1016/j.bbi.2010.10.023.

Becker, C.A., R.J. Collier, and A.E. Stone (Aug. 2020). "Invited review: Physiological and behavioral effects of heat stress in dairy cows". In: *Journal of Dairy Science* 103.8, pp. 6751–6770. DOI: 10.3168/jds.2019-17929.

Benjamini, Yoav and Yosef Hochberg (Jan. 1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1, pp. 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.

Biffani, S. et al. (July 2016). "Short communication: Effect of heat stress on nonreturn rate of Italian Holstein cows". In: *Journal of Dairy Science* 99.7, pp. 5837–5843. DOI: 10.3168/jds.2015-10491.

Bohlouli, M. et al. (Apr. 2022). "Genome-wide associations for heat stress response suggest potential candidate genes underlying milk fatty acid composition in dairy cattle". In: *Journal of Dairy Science* 105.4, pp. 3323–3340. DOI: 10.3168/jds.2021-21152.

Bolyen, Evan et al. (2019a). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2". In: *Nature Biotechnology* 37.8, pp. 852–857. DOI: 10.1038/s41587-019-0209-9.

Bolyen, Evan et al. (July 2019b). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2". In: *Nature Biotechnology* 37.8, pp. 852–857. DOI: 10.1038/s41587-019-0209-9.

Brito, Luiz F. et al. (July 2020). "Large-Scale Phenotyping of Livestock Welfare in Commercial Production Systems: A New Frontier in Animal Breeding". In: *Frontiers in Genetics* 11. DOI: 10.3389/fgene.2020.00793.

Brivio, Francesca et al. (Feb. 2019). "Forecasting the response to global warming in a heat-sensitive species". In: *Scientific Reports* 9.1. DOI: 10.1038/s41598-019-39450-5.

Broad Institute (Accessed: 2022/07/15; version 2.27.4). *Picard Tools*. http://broadinstitute.github.io/picard/.

Bukin, Yu. S. et al. (Feb. 2019). "The effect of 16S rRNA region choice on bacterial community metabarcoding results". In: *Scientific Data* 6.1. DOI: 10.1038/sdata.2019.7.

Burt, David W (2009). "The cattle genome reveals its secrets". In: *Journal of Biology* 8.4, p. 36. DOI: 10.1186/jbiol137.

Callahan, Benjamin J et al. (May 2016). "DADA2: High-resolution sample inference from Illumina amplicon data". In: *Nature Methods* 13.7, pp. 581–583. DOI: 10.1038/nmeth.3869.

Cardona, Alexia et al. (May 2014). "Genome-Wide Analysis of Cold Adaptation in Indigenous Siberian Populations". In: *PLoS ONE* 9.5. Ed. by Manfred Kayser, e98076. DOI: 10.1371/journal.pone.0098076.

Carillier-Jacquin, C. et al. (2022). "Predicting pig digestibility coefficients with microbial and genomic data using machine learning prediction algorithms". In: *WCGALP*.

Cartwright, Shannon L. et al. (June 2023). "Impact of heat stress on dairy cattle and selection strategies for thermotolerance: a review". In: *Frontiers in Veterinary Science* 10. DOI: 10.3389/fvets.2023.1198697.

Chen, Qinghua and Dinghua Shi (Apr. 2004). "The modeling of scale-free networks". In: *Physica A: Statistical Mechanics and its Applications* 335.1-2, pp. 240–248. DOI: 10.1016/j.physa.2003.12.014.

Chen, Shifu et al. (Sept. 2018a). "fastp: an ultra-fast all-in-one FASTQ preprocessor". In: *Bioinformatics* 34.17, pp. i884–i890. DOI: 10.1093/bioinformatics/bty560.

Chen, Siyu et al. (Oct. 2018b). "Exposure to heat-stress environment affects the physiology, circulation levels of cytokines, and microbiome in dairy cows". In: *Scientific Reports* 8.1. DOI: 10.1038/s41598-018-32886-1.

Cheruiyot, Evans K. et al. (June 2022). "Improving Genomic Selection for Heat Tolerance in Dairy Cattle: Current Opportunities and Future Directions". In: *Frontiers in Genetics* 13. DOI: 10.3389/fgene.2022.894067.

Consortium, The Gene Ontology (Dec. 2020). "The Gene Ontology resource: enriching a GOld mine". In: *Nucleic Acids Research* 49.D1, pp. D325–D334. ISSN: 0305-1048. DOI: 10.1093/nar/gkaa1113. eprint: https://academic.oup.com/nar/article-pdf/49/D1/D325/35364517/gkaa1113.pdf.

Corazzin, Mirco et al. (Nov. 2020). "Effect of Heat Stress on Dairy Cow Performance and on Expression of Protein Metabolism Genes in Mammary Cells". In: *Animals* 10.11, p. 2124. DOI: 10.3390/ani10112124.

Corvo, Marcello Del et al. (May 2021). "Methylome Patterns of Cattle Adaptation to Heat Stress". In: *Frontiers in Genetics* 12. DOI: 10.3389/fgene.2021.633132.

Coudert, Elisabeth et al. (Dec. 2022). "Annotation of biologically relevant ligands in UniProtKB using ChEBI". In: *Bioinformatics* 39.1. Ed. by Peter Robinson. DOI: 10.1093/bioinformatics/btac793.

Cox, D. R. and N. Reid (Sept. 1987). "Parameter Orthogonality and Approximate Conditional Inference". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 49.1, pp. 1–18. DOI: 10.1111/j.2517-6161.1987.tb01422.x.

Czech, Bartosz et al. (2022a). "Fecal microbiota and their association with heat stress in Bos taurus". In: *BMC Microbiology* 22.1. DOI: 10.1186/s12866-022-02576-0.

Czech, Bartosz et al. (Nov. 2022b). "Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition". In: *Frontiers in Microbiology* 13. DOI: 10.3389/fmicb.2022.998093.

Dahl, Geoffrey E., Sha Tao, and Jimena Laporta (Mar. 2020). "Heat Stress Impacts Immune Status in Cows Across the Life Cycle". In: *Frontiers in Veterinary Science* 7. DOI: 10.3389/fvets.2020.00116.

Davies, J. (Mar. 2001). "In a Map for Human Life, Count the Microbes, Too". In: *Science* 291.5512, 2316b–2316. DOI: 10.1126/science.291.5512.2316b.

DeLuca, David S. et al. (Apr. 2012). "RNA-SeQC: RNA-seq metrics for quality control and process optimization". In: *Bioinformatics* 28.11, pp. 1530–1532. DOI: 10.1093/bioinformatics/bts196.

Diamanti, Andrea Picchianti et al. (Aug. 2016). "Microbiota and chronic inflammatory arthritis: an interwoven link". In: *Journal of Translational Medicine* 14.1. DOI: 10.1186/s12967-016-0989-3.

Diaz, Fabian A et al. (Oct. 2021). "Evaluation of Seasonal Heat Stress on Transcriptomic Profiles and Global DNA Methylation of Bovine Oocytes". In: *Frontiers in Genetics* 12. DOI: 10.3389/fgene.2021.699920.

Dobin, Alexander et al. (Oct. 2012). "STAR: ultrafast universal RNA-seq aligner". In: *Bioinformatics* 29.1, pp. 15–21. DOI: 10.1093/bioinformatics/bts635.

Du, Hairong et al. (July 2022). "New Evidence of Tiger Subspecies Differentiation and Environmental Adaptation: Comparison of the Whole Genomes of the Amur Tiger and the South China Tiger". In: *Animals* 12.14, p. 1817. DOI: 10.3390/ani12141817.

Freitas, P.H.F. et al. (2022). "Genome-wide association analyses for various thermotolerance indicators measured during lactation in maternal-line pigs". In: *WCGALP*.

Gallagher, Michael D. and Alice S. Chen-Plotkin (May 2018). "The Post-GWAS Era: From Association to Function". In: *The American Journal of Human Genetics* 102.5, pp. 717–730. DOI: 10.1016/j.ajhg.2018.04.002.

Gao, S. T. et al. (Aug. 2019). "Heat stress negatively affects the transcriptome related to overall metabolism and milk protein synthesis in mammary tissue of midlactating dairy cows". In: *Physiological Genomics* 51.8, pp. 400–409. DOI: 10.1152/physiolgenomics.00039.2019.

Garner, J. B. et al. (Nov. 2020). "Gene expression of the heat stress response in bovine peripheral white blood cells and milk somatic cells in vivo". In: *Scientific Reports* 10.1. DOI: 10.1038/s41598-020-75438-2.

Gonzalez-Rivas, Paula A. et al. (Apr. 2020). "Effects of heat stress on animal physiology, metabolism, and meat quality: A review". In: *Meat Science* 162, p. 108025. DOI: 10.1016/j.meatsci.2019.108025.

Guo, Zitai et al. (Mar. 2021). "Impacts of Heat Stress-Induced Oxidative Stress on the Milk Protein Biosynthesis of Dairy Cows". In: *Animals* 11.3, p. 726. DOI: 10.3390/ani11030726.

Hamady, M. and R. Knight (Apr. 2009). "Microbial community profiling for human microbiome projects: Tools, techniques, and challenges". In: *Genome Research* 19.7, pp. 1141–1152. DOI: 10.1101/gr.085464.108.

Hassan, Faiz ul et al. (Dec. 2019). "Prospects of HSP70 as a genetic marker for thermo-tolerance and immuno-modulation in animals under climate change scenario". In: *Animal Nutrition* 5.4, pp. 340–350. DOI: 10.1016/j.aninu.2019.06.005.

He, Jianwen et al. (Nov. 2019). "Heat stress affects fecal microbial and metabolic alterations of primiparous sows during late gestation". In: *Journal of Animal Science and Biotechnology* 10.1. DOI: 10.1186/s40104-019-0391-0.

Hering, D.M., K. Olenski, and S. Kaminski (May 2014). "Genome-wide association study for poor sperm motility in Holstein-Friesian bulls". In: *Animal Reproduction Science* 146.3-4, pp. 89–97. DOI: 10.1016/j.anireprosci.2014.01.012.

Hoque, M. Nazmul et al. (Sept. 2019). "Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis". In: *Scientific Reports* 9.1. DOI: 10.1038/s41598-019-49468-4.

Hu, Taotao, Hong Yang, and Ze-Guang Han (Oct. 2015). "PDZRN4 acts as a suppressor of cell proliferation in human liver cancer cell lines". In: *Cell Biochemistry and Function* 33.7, pp. 443–449. DOI: 10.1002/cbf.3130.

Hu, Zhi-Liang et al. (Nov. 2012). "Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era". In: *Nucleic Acids Research* 41.D1, pp. D871–D879. DOI: 10.1093/nar/gks1150.

Idris, Musadiq et al. (Jan. 2021). "Non-Invasive Physiological Indicators of Heat Stress in Cattle". In: *Animals* 11.1, p. 71. DOI: 10.3390/ani11010071.

Islam, Md. Ashraful et al. (Oct. 2021). "Automated Monitoring of Cattle Heat Stress and Its Mitigation". In: *Frontiers in Animal Science* 2. DOI: 10.3389/fanim.2021.737213.

Jairath, L. et al. (Feb. 1998). "Genetic Evaluation for Herd Life in Canada". In: *Journal of Dairy Science* 81.2, pp. 550–562. DOI: 10.3168/jds.s0022-0302(98)75607-3.

Jin, Peng et al. (2022). "PDZRN4 suppresses tumorigenesis and androgen therapy-resistance in prostate cancer". In: *Journal of Cancer* 13.7, pp. 2293–2300. DOI: 10.7150/jca.69269.

Kanehisa, M. (Jan. 2000). "KEGG: Kyoto Encyclopedia of Genes and Genomes". In: *Nucleic Acids Research* 28.1, pp. 27–30. DOI: 10.1093/nar/28.1.27.

Kanehisa, Minoru et al. (Oct. 2015). "KEGG as a reference resource for gene and protein annotation". In: *Nucleic Acids Research* 44.D1, pp. D457–D462. DOI: 10.1093/nar/gkv1070.

Karl, J. Philip et al. (Sept. 2018). "Effects of Psychological, Environmental and Physical Stressors on the Gut Microbiota". In: *Frontiers in Microbiology* 9. DOI: 10.3389/fmicb.2018.02013.

Khan, Imran et al. (July 2023). "Heat Stress as a Barrier to Successful Reproduction and Potential Alleviation Strategies in Cattle". In: *Animals* 13.14, p. 2359. DOI: 10.3390/ani13142359.

Kim, Dong-Hyeon et al. (July 2020). "Differential Dynamics of the Ruminal Microbiome of Jersey Cows in a Heat Stress Environment". In: *Animals* 10.7, p. 1127. DOI: 10.3390/ani10071127.

Kim, Seon Ho et al. (Feb. 2022). "Heat Stress: Effects on Rumen Microbes and Host Physiology, and Strategies to Alleviate the Negative Impacts on Lactating Dairy Cows". In: *Frontiers in Microbiology* 13. DOI: 10.3389/fmicb.2022.804562.

Kitajima, Kaho et al. (June 2021). "Effects of Heat Stress on Heart Rate Variability in Free-Moving Sheep and Goats Assessed With Correction for Physical Activity". In: *Frontiers in Veterinary Science* 8. DOI: 10.3389/fvets.2021.658763.

Koh, Hyunwook et al. (May 2019). "A Distance-Based Kernel Association Test Based on the Generalized Linear Mixed Model for Correlated Microbiome Studies". In: *Frontiers in Genetics* 10. DOI: 10.3389/fgene.2019.00458.

Krishnan, Govindan et al. (Sept. 2017). "Mitigation of the Heat Stress Impact in Livestock Reproduction". In: *Theriogenology*. InTech. DOI: 10.5772/intechopen.69091.

Krueger, F. (2015). *Trim Galore!*

Langfelder, Peter and Steve Horvath (Dec. 2008). "WGCNA: an R package for weighted correlation network analysis". In: *BMC Bioinformatics* 9.1. DOI: 10.1186/1471-2105-9-559.

Lees, Angela M. et al. (June 2019). "The Impact of Heat Load on Cattle". In: *Animals* 9.6, p. 322. DOI: 10.3390/ani9060322.

Li, Fuyong et al. (June 2019). "Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle". In: *Microbiome* 7.1. DOI: 10.1186/s40168-019-0699-1.

Li, Qiuling et al. (Dec. 2018). "Characterization of miRNA profiles in the mammary tissue of dairy cattle in response to heat stress". In: *BMC Genomics* 19.1. DOI: 10.1186/s12864-018-5298-1.

Librado, Pablo et al. (Nov. 2015). "Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments". In: *Proceedings of the National Academy of Sciences* 112.50. DOI: 10.1073/pnas.1513696112.

Liu, Jiangjing et al. (Sept. 2019). "Effects of heat stress on body temperature, milk production, and reproduction in dairy cows: a novel idea for monitoring and evaluation of heat stress — A review". In: *Asian-Australasian Journal of Animal Sciences* 32.9, pp. 1332–1339. DOI: 10.5713/ajas.18.0743.

Liu, Lily et al. (Aug. 2023). "Research progress on the regulation of production traits by gastrointestinal microbiota in dairy cows". In: *Frontiers in Veterinary Science* 10. DOI: 10.3389/fvets.2023.1206346.

Liu, Zhexi et al. (Apr. 2022). "Microbiome and ileum transcriptome revealed the boosting effects of selenium yeast on egg production in aged laying hens". In: *Animal Nutrition*. DOI: 10.1016/j.aninu.2022.04.001.

Love, Michael I, Wolfgang Huber, and Simon Anders (Dec. 2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome Biology* 15.12. DOI: 10.1186/s13059-014-0550-8.

Luengo, Raúl Miranda-Caso et al. (Jan. 2019). "Delayed differentiation of vaginal and uterine microbiomes in dairy cows developing postpartum endometritis". In: *PLOS ONE* 14.1. Ed. by Juan J. Loor, e0200974. DOI: 10.1371/journal.pone.0200974.

Luo, H. et al. (Apr. 2021). "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". In: *Journal of Dairy Science* 104.4, pp. 4390–4403. DOI: 10.3168/jds.2020-19192.

Luo, Hanpeng et al. (Aug. 2022). "Weighted single-step GWAS and RNA sequencing reveals key candidate genes associated with physiological indicators of heat stress in Holstein cattle". In: *Journal of Animal Science and Biotechnology* 13.1. DOI: 10.1186/s40104-022-00748-6.

Macciotta, N.P.P. et al. (June 2017). "Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle". In: *Journal of Dairy Science* 100.6, pp. 4683–4697. DOI: 10.3168/jds.2016-12249.

Manzoni, Claudia et al. (Nov. 2016). "Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences". In: *Briefings in Bioinformatics* 19.2, pp. 286–302. DOI: 10.1093/bib/bbw114.

Mapholi, N.O. et al. (Apr. 2016). "Genome-wide association study of tick resistance in South African Nguni cattle". In: *Ticks and Tick-borne Diseases* 7.3, pp. 487–497. DOI: 10.1016/j.ttbdis.2016.02.005.

Martínez-Álvaro, Marina et al. (Apr. 2022). "Bovine host genome acts on rumen microbiome function linked to methane emissions". In: *Communications Biology* 5.1. DOI: 10.1038/s42003-022-03293-0.

McInnes, Leland, John Healy, and James Melville (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. cite arxiv:1802.03426.

McLaren, William et al. (June 2016). "The Ensembl Variant Effect Predictor". In: *Genome Biology* 17.1. DOI: 10.1186/s13059-016-0974-4.

McMurdie, Paul J. and Susan Holmes (Apr. 2013). "phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data". In: *PLoS ONE* 8.4. Ed. by Michael Watson, e61217. DOI: 10.1371/journal.pone.0061217.

Mellado, Miguel et al. (Apr. 2023). "Effect of heat stress and body condition score on the occurrence of puerperal disorders in Holstein cows". In: *Spanish Journal of Agricultural Research* 21.2, e0501. DOI: 10.5424/sjar/2023212-19600.

Mi, Huaiyu et al. (Nov. 2018). "PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools". In: *Nucleic Acids Research* 47.D1, pp. D419–D426. DOI: 10.1093/nar/gky1038.

Miętkiewska, Klaudia, Pawel Kordowitzki, and Chandra S. Pareek (Dec. 2022). "Effects of Heat Stress on Bovine Oocytes and Early Embryonic Development—An Update". In: *Cells* 11.24, p. 4073. DOI: 10.3390/cells11244073.

Mondo, E. et al. (Jan. 2020). "Gut microbiome structure and adrenocortical activity in dogs with aggressive and phobic behavioral disorders". In: *Heliyon* 6.1, e03311. DOI: 10.1016/j.heliyon.2020.e03311.

Nardone, A. et al. (May 2010). "Effects of climate changes on animal production and sustainability of livestock systems". In: *Livestock Science* 130.1-3, pp. 57–69. DOI: 10.1016/j.livsci.2010.02.011.

Park, Da Som et al. (June 2021). "Dynamic changes in blood immune cell composition and function in Holstein and Jersey steers in response to heat stress". In: *Cell Stress and Chaperones* 26.4, pp. 705–720. DOI: 10.1007/s12192-021-01216-2.

Park, Tansol et al. (June 2022). "Heat stress impacts the multi-domain ruminal microbiota and some of the functional features independent of its effect on feed intake in lactating dairy cows". In: *Journal of Animal Science and Biotechnology* 13.1. DOI: 10.1186/s40104-022-00717-z.

Park, Youngjun, Dominik Heider, and Anne-Christin Hauschild (June 2021). "Integrative Analysis of Next-Generation Sequencing for Next-Generation Cancer Research toward Artificial Intelligence". In: *Cancers* 13.13, p. 3148. DOI: 10.3390/cancers13133148.

Parker, A. J. et al. (Feb. 2003). "Dehydration in stressed ruminants may be the result of acortisol-induced diuresis1". In: *Journal of Animal Science* 81.2, pp. 512–519. DOI: 10.2527/2003.812512x.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Petti, C. A., C. R. Polage, and P. Schreckenberger (Dec. 2005). "The Role of 16S rRNA Gene Sequencing in Identification of Microorganisms Misidentified by Conventional Methods". In: *Journal of Clinical Microbiology* 43.12, pp. 6123–6125. DOI: 10.1128/jcm.43.12.6123-6125.2005.

Polsky, Liam and Marina A.G. von Keyserlingk (Nov. 2017). "Invited review: Effects of heat stress on dairy cattle welfare". In: *Journal of Dairy Science* 100.11, pp. 8645–8657. DOI: 10.3168/jds.2017-12651.

Purcell, Shaun et al. (Sept. 2007). "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses". In: *The American Journal of Human Genetics* 81.3, pp. 559–575. DOI: 10.1086/519795.

Quast, Christian et al. (2012a). "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools". In: *Nucleic Acids Research* 41.D1, pp. D590–D596. DOI: 10.1093/nar/gks1219.

Quast, Christian et al. (Nov. 2012b). "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools". In: *Nucleic Acids Research* 41.D1, pp. D590–D596. DOI: 10.1093/nar/gks1219.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Ramayo-Caldas, Yuliaxis et al. (2021). "Leveraging host-genetics and gut microbiota to determine immunocompetence in pigs". In: *Animal Microbiome* 3.1. DOI: 10.1186/s42523-021-00138-9.

Ransom-Jones, Emma et al. (Jan. 2012). "The Fibrobacteres: an Important Phylum of Cellulose-Degrading Bacteria". In: *Microbial Ecology* 63.2, pp. 267–281. DOI: 10.1007/s00248-011-9998-1.

Rebez, Ebenezer Binuni et al. (Jan. 2023). "Heat Stress and Histopathological Changes of Vital Organs: A Novel Approach to Assess Climate Resilience in Farm Animals". In: *Sustainability* 15.2, p. 1242. DOI: 10.3390/su15021242.

Rio, Donald C. et al. (June 2010). "Purification of RNA Using TRIzol (TRI Reagent)". In: *Cold Spring Harbor Protocols* 2010.6, pdb.prot5439. DOI: 10.1101/pdb.prot5439.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth (Nov. 2009). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data". In: *Bioinformatics* 26.1, pp. 139–140. DOI: 10.1093/bioinformatics/btp616.

Robinson, Mark D and Alicia Oshlack (2010). "A scaling normalization method for differential expression analysis of RNA-seq data". In: *Genome Biology* 11.3, R25. DOI: 10.1186/gb-2010-11-3-r25.

Rodrigues, M.X. et al. (Dec. 2016). "The Lactococcus genus as a potential emerging mastitis pathogen group: A report on an outbreak investigation". In: *Journal of Dairy Science* 99.12, pp. 9864–9874. DOI: 10.3168/jds.2016-11143.

Roth, Zvi (Oct. 2020). "Reproductive physiology and endocrinology responses of cows exposed to environmental heat stress - Experiences from the past and lessons for the present". In: *Theriogenology* 155, pp. 150–156. DOI: 10.1016/j.theriogenology.2020.05.040.

Ruiz-González, Alexis, Daniel Enrique Rico, and Jorge Eduardo Rico (Feb. 2022). "Modulation of Fecal Metabolites by Heat Stress and Diet, and Their Association with Inflammation and Leaky Gut Markers in Dairy Cows". In: *Metabolites* 12.2, p. 142. DOI: 10.3390/metabo12020142.

Sales, Gustavo Felipe Correia et al. (Apr. 2021). "Heat stress influence the microbiota and organic acids concentration in beef cattle rumen". In: *Journal of Thermal Biology* 97, p. 102897. DOI: 10.1016/j.jtherbio.2021.102897.

Salonen, Anne et al. (Apr. 2014). "Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men". In: *The ISME Journal* 8.11, pp. 2218–2230. DOI: 10.1038/ismej.2014.63.

Scott, Karen P. et al. (Mar. 2013). "The influence of diet on the gut microbiota". In: *Pharmacological Research* 69.1, pp. 52–60. DOI: 10.1016/j.phrs.2012.10.020.

Shah, Tejas M. et al. (Sept. 2019). "Host transcriptome and microbiome interaction modulates physiology of full-sibs broilers with divergent feed conversion ratio". In: *npj Biofilms and Microbiomes* 5.1. DOI: 10.1038/s41522-019-0096-3.

Shi, Dayou et al. (June 2019). "Impact of gut microbiota structure in heat-stressed broilers". In: *Poultry Science* 98.6, pp. 2405–2413. DOI: 10.3382/ps/pez026.

Sigdel, Anil et al. (Oct. 2019). "Whole Genome Mapping Reveals Novel Genes and Pathways Involved in Milk Production Under Heat Stress in US Holstein Cows". In: *Frontiers in Genetics* 10. DOI: 10.3389/fgene.2019.00928.

Smid, Marcel et al. (June 2018). "Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons". In: *BMC Bioinformatics* 19.1. DOI: 10.1186/s12859-018-2246-7.

Sohail, Muhammad U. et al. (Mar. 2015). "Molecular analysis of the caecal and tracheal microbiome of heat-stressed broilers supplemented with prebiotic and probiotic". In: *Avian Pathology* 44.2, pp. 67–74. DOI: 10.1080/03079457.2015.1004622.

Solomon, Daniel L. (May 1975). "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". In: *The American Statistician* 29.2, pp. 101–102. DOI: 10.1080/00031305.1975.10477383.

Subramanian, Indhupriya et al. (Jan. 2020). "Multi-omics Data Integration, Interpretation, and Its Application". In: *Bioinformatics and Biology Insights* 14, p. 117793221989905. DOI: 10.1177/1177932219899051.

Summer, Andrea et al. (Oct. 2018). "Impact of heat stress on milk and meat production". In: *Animal Frontiers* 9.1, pp. 39–46. DOI: 10.1093/af/vfy026.

Triot, Alexa et al. (June 2014). "Inherited biallelic CSF3R mutations in severe congenital neutropenia". In: *Blood* 123.24, pp. 3811–3817. DOI: 10.1182/blood-2013-11-535419.

Vajana, Elia et al. (Oct. 2018). "Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever". In: *Frontiers in Genetics* 9. DOI: 10.3389/fgene.2018.00385.

Wang, B.B. et al. (Sept. 2021). "Genome-wide association study reveals a quantitative trait locus and two candidate genes on Sus scrofa chromosome 5 affecting intramuscular fat content in Suhuai pigs". In: *Animal* 15.9, p. 100341. DOI: 10.1016/j.animal.2021.100341.

Wang, Qing et al. (Jan. 2019). "Host and microbiome multi-omics integration: applications and methodologies". In: *Biophysical Reviews* 11.1, pp. 55–65. DOI: 10.1007/s12551-018-0491-7.

Wankar, Alok K., Sandeep N. Rindhe, and Nandkumar S. Doijad (Jan. 2021). "Heat stress in dairy animals and current milk production trends, economics, and future perspectives: the global scenario". In: *Tropical Animal Health and Production* 53.1. DOI: 10.1007/s11250-020-02541-x.

Wu, Tianzhi et al. (Aug. 2021). "clusterProfiler 4.0: A universal enrichment tool for interpreting omics data". In: *The Innovation* 2.3, p. 100141. DOI: 10.1016/j.xinn.2021.100141.

Yip, Andy M and Steve Horvath (Jan. 2007). "Gene network interconnectedness and the generalized topological overlap measure". In: *BMC Bioinformatics* 8.1. DOI: 10.1186/1471-2105-8-22.

Young, Matthew D et al. (2010). "Gene ontology analysis for RNA-seq: accounting for selection bias". In: *Genome Biology* 11.2, R14. DOI: 10.1186/gb-2010-11-2-r14.

Yu, Guangchuang et al. (May 2012). "clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters". In: *OMICS: A Journal of Integrative Biology* 16.5, pp. 284–287. DOI: 10.1089/omi.2011.0118.

Yu, Meng-Fei et al. (Oct. 2020). "Dihydropyridine Enhances the Antioxidant Capacities of Lactating Dairy Cows under Heat Stress Condition". In: *Animals* 10.10, p. 1812. DOI: 10.3390/ani10101812.

Yue, Shuangming et al. (June 2020). "Transcriptome Functional Analysis of Mammary Gland of Cows in Heat Stress and Thermoneutral Condition". In: *Animals* 10.6, p. 1015. DOI: 10.3390/ani10061015.

Zhang, Guoxing et al. (Aug. 2019). "The Association Between Inflammaging and Age-Related Changes in the Ruminal and Fecal Microbiota Among Lactating Holstein Cows". In: *Frontiers in Microbiology* 10. DOI: 10.3389/fmicb.2019.01803.

Zhang, Xiaohui et al. (Aug. 2022). "The Association between Gut Microbiome Diversity and Composition and Heat Tolerance in Cattle". In: *Microorganisms* 10.8, p. 1672. DOI: 10.3390/microorganisms10081672.

Zhao et al. (Nov. 2019). "Effect of Heat Stress on Bacterial Composition and Metabolism in the Rumen of Lactating Dairy Cows". In: *Animals* 9.11, p. 925. DOI: 10.3390/ani9110925.

Zhou, Xiang and Matthew Stephens (June 2012). "Genome-wide efficient mixed-model analysis for association studies". In: *Nature Genetics* 44.7, pp. 821–824. DOI: 10.1038/ng.2310.

Zhu, Baoli, Xin Wang, and Lanjuan Li (Aug. 2010). "Human gut microbiome: the second genome of human body". In: *Protein & Cell* 1.8, pp. 718–725. DOI: 10.1007/s13238-010-0093-z.

Scientific achievements

Education

- 2019-2023: **Doctor of Philosophy** – Biological Sciences – Wrocław University of Environmental and Life Sciences
- 2017-2019: **Master of Science** – Bioinformatics: biostatistics and bioinformatics programming – Wrocław University of Environmental and Life Sciences
- 2014-2017: **Bachelor of Science** – Bioinformatics – Wrocław University of Environmental and Life Sciences

Publications

- Prochowska, S., Napierkowska, S., **Czech, B.**, & Niżański, W. (2023). Feline sperm head morphometry in relation to male pedigree and fertility. In Theriogenology (Libk. 208, or. 119–125). Elsevier BV. https://doi.org/10.1016/j.theriogenology.2023.06.006
- Kwasnik, M., Socha, W., **Czech, B.**, Wasiak, M., Rola, J., & Rozek, W. (2023). Protein-Coding Region Derived Small RNA in Exosomes from Influenza A Virus–Infected Cells. In International Journal of Molecular Sciences (Libk. 24, Issue 1, or. 867). MDPI AG. https://doi.org/10.3390/ijms24010867
- **Czech, B.,** Wang, Y., Wang, K., Luo, H., Hu, L., & Szyda, J. (2022). Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition. In Frontiers in Microbiology (Libk. 13). Frontiers Media SA. https://doi.org/10.3389/fmicb.2022.998093
- **Czech, B.,** Szyda, J., Wang, K., Luo, H., & Wang, Y. (2022). Fecal microbiota and their association with heat stress in Bos taurus. In BMC Microbiology (Libk. 22, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1186/s12866-022-02576-0
- Suchocki, T., **Czech, B.**, Dunislawska, A., Slawinska, A., Derebecka, N., Wesoly, J., Siwek, M., & Szyda, J. (2021). SNP prioritization in targeted sequencing data associated with humoral immune responses in chicken. In Poultry Science (Libk. 100, Issue 11, or. 101433). Elsevier BV. https://doi.org/10.1016/j.psj.2021.101433
- Mielczarek, M., **Czech, B.,** Stańczyk, J., Szyda, J., & Guldbrandtsen, B. (2020). Extraordinary Command Line: Basic Data Editing Tools for Biologists Dealing

with Sequence Data. In The Open Bioinformatics Journal (Libk. 13, Issue 1, or. 137–145). Bentham Science Publishers Ltd. https://doi.org/10.2174/1875036202013010137

- Kotlarz, K., Mielczarek, M., Suchocki, T., **Czech, B.,** Guldbrandtsen, B., & Szyda, J. (2020). The application of deep learning for the classification of correct and incorrect SNP genotypes from whole-genome DNA sequencing pipelines. In Journal of Applied Genetics (Libk. 61, Issue 4, or. 607–616). Springer Science and Business Media LLC. https://doi.org/10.1007/s13353-020-00586-0

- **Czech, B.**, Guldbrandtsen, B., Szyda, J. (2020). Patterns of DNA variation between the autosomes, the X chromosome and the Y chromosome in Bos taurus genome. In Scientific Reports (Libk. 10, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/s41598-020-70380-9

- **Czech, B.**, Frąszczak, M., Mielczarek, M., & Szyda, J. (2018). Identification and annotation of breed-specific single nucleotide polymorphisms in Bos taurus genomes. In J. J. Loor (Arg.), PLOS ONE (Libk. 13, Issue 6, or. e0198419). Public Library of Science (PLoS). https://doi.org/10.1371/journal.pone.0198419

Conferences – presentations

- Jaśkowski, B. M., **Czech, B.**, & Niżański, W. (2023) Evaluation of the size of the recipient's corpus luteum after synchronization as an additional criterion for increasing the pregnancy rate after an embryo transfer in cattle. Talk presented during Japan-Poland Joint Seminar in Morioka, Japan

- **Czech, B.,** Szyda, J., & Wang, Y. (2022) Host transcriptome and microbiome data integration in Chinese Holstein cattle under heat stress. Talk presented during 73rd Annual Meeting of the European Federation of Animal Science in Porto, Portugal

- **Czech, B.,** Wang, K., Chen, S., Wang, Y., & Szyda, J. (2021) Faecal microbiota and their association with heat stress in Bos taurus. Talk presented during 72nd Annual Meeting of the European Federation of Animal Science in Davos, Switzerland

- Suchocki, T., **Czech, B.**, Siwek, M., & Szyda, J. (2019) Breed specific reference genomes in cattle. Talk presented during 70th Annual Meeting of the European Federation of Animal Science in Ghent, Belgium

- **Czech, B.,** Zając, M., Wasyl, D., Mielczarek, M., & Sztromwasser, P. (2019) Optimization of de novo assembly of genomes based on second and third generation sequencing data. Talk presented during XXIV International Conference of Student Scientific Circles at Wroclaw University of Environmental and Life Sciences, Wrocław, Poland

- **Czech, B.**, Mielczarek, M., Frąszczak, M., & Szyda, J. (2018) Breed specific reference genomes in cattle. Talk presented during 69th Annual Meeting of the European Federation of Animal Science in Dubrovnik, Croatia
- **Czech, B.** (2018) Construction of Bos taurus breed specific reference genomes. Talk presented during XXIII International Conference of Student Scientific Circles at Wrocław University of Environmental and Life Sciences, Wrocław, Poland
- **Czech, B.**, Dobkowski, E., Mielczarek, M., Frąszczak, M., & Szyda, J. (2017) Construction of Bos taurus breed specific reference genomes for Fleckvieh, Simmental, Guernsey and Brown Swiss. Talk presented during Vorttragstagung der DGfZ und GfT at University of Hohenheim, Stuttgart, Germany
- **Czech, B.,** Dobkowski, E., Barski, P., & Odziemczyk, M. (2017) Creating of the Brown Swiss breed specific reference genome. Talk presented during XXII International Conference of Student Scientific Circles at Wrocław University of Environmental and Life Sciences, Wrocław, Poland

Conferences – posters

- Stepien, R., Szyda, J., **Czech, B.**, & Mielczarek, M. (2023) The effect of transcriptomic annotations in breast cancer differential gene expression study. Poster presented during 17th International Symposium on Integrative Bioinformatics in Wrocław, Poland
- **Czech, B.**, Wang, K., Chen, S., Wang, Y., & Szyda, J. (2021) Challenges of 16S rRNA gene analysis in Chinese Holstein cows under heat stress condition. Poster presented during 72nd Annual Meeting of the European Federation of Animal Science in Davos, Switzerland
- Kotlarz, K., Mielczarek, M., Suchocki, T., **Czech, B.**, Guldbrandtsen, B., & Szyda, J. (2020) Don't play too much! Deep learning to classify true and false positive SNPs in whole genome sequence. Poster presented during 71st Annual Meeting of the European Federation of Animal Science in Porto, Portugal
- **Czech, B.**, Szyda, J., Wang, K., Chen, S., & Wang, Y. (2020) The effect of pipelines and databases on the analysis of the fecal microbiota of dairy cattle. Poster presented during 71st Annual Meeting of the European Federation of Animal Science in Porto, Portugal
- Szyda,J., **Czech, B.**, Wang, K., Chen, S., & Wang, Y. (2020) The application of mixed linear models for the analysis of microbiome influence on heat stress in Chinese Holstein cows. Poster presented during 71st Annual Meeting of the European Federation of Animal Science in Porto, Portugal
- **Czech, B.**, Szyda, J., & Guldbrandtsen, B. (2020) Patterns of DNA variation between the autosomes, the X chromosome and the Y chromosome in Bos taurus genome. Poster presented during Conference of Polish Bioinformatics Society in Warsaw, Poland

- Kotlarz, K., Mielczarek, M., Suchocki, T., **Czech, B.**, Guldbrandtsen, B., & Szyda, J. (2020) Deep learning algorithms for the imbalanced classification of correct and incorrect SNP genotypes from WGS pipeline. Poster presented during Conference of Polish Bioinformatics Society in Warsaw, Poland
- **Czech, B.**, Guldbrandtsen, B., & Szyda, J. (2019) Pattern of genetic variation between autosomes and sex chromosomes in Bos taurus genome. Poster presented during 70th Annual Meeting of the European Federation of Animal Science in Ghent, Belgium
- **Czech, B.**, Suchocki, T., & Szyda, J. (2019) SNP prioritisation in GWAS with dense marker sets. Talk presented during 70th Annual Meeting of the European Federation of Animal Science in Ghent, Belgium

Courses

- Oxford Statistical Genomics Summer School (2022). University of Oxford, Oxford, United Kingdom
- Machine Lerning Course by Krzysztof Mędrela (2020). Wrocław, Poland
- Python Course by Krzysztof Mędrela (2020). Wrocław, Poland
- RNA-seq data analysis (2018). Ideas4Biology, Poznań, Poland.
- Statistical Methods for Genome-Enabled Selection by Daniel Gianola (2018). Wrocław, Poland
- Writing and Presenting Scientific Papers (2018), EAAP, Dubrovnik, Croatia

Awards

- EAAP Scholarship Winner (2022). European Federation of Animal Science
- The Best Diploma (2019). Voivodeship Marshal of Lower Silesia
- Distinction for the best master's thesis in bioinformatics defended in 2019. Polish Bioinformatics Society
- First Prize on International Student Conference (2019). Wrocław University of Environmental and Life Sciences
- Second Prize on International Student Conference (2018). Wrocław University of Environmental and Life Sciences
- EAAP Scholarship Winner (2018). European Federation of Animal Science

Internships and professional experience

- July 2019-current: 7N, Bioinformatician and R/Shiny developer in big pharma. Warsaw, Poland
- July-August 2018: Center for Quantitative Genetics and Genomics, Aarhus University. Foulum, Denmark
- July-August 2016: Biostat sp. z o.o. – Department of Statistics. Rybnik, Poland

Załącznik nr 3 do Trybu postępowania w sprawie nadania stopnia naukowego doktora

Bartosz Czech

imię i nazwisko

Wrocław, 11.09.2023

miejscowość i data

Katedra Genetyki

afiliacja

OŚWIADCZENIE

Oświadczam, że w pracy Fecal microbiota and their association with heat stress in Bos taurus Czech Bartosz, Szyda Joanna, Wang Kai [et al.], BMC Microbiology, 2022, vol. 22, no. 1, pp.1-9, Article number:171. DOI:10.1186/s12866-022-02576-0 mój udział polegał na zaplanowaniu oraz wykonaniu bioinformatycznej oraz statystycznej analizy danych, wykonaniu wizualizacji oraz na napisaniu manuskryptu.

PODPIS ZAUFANY
BARTOSZ
CZECH
11.09.2023 13:52:51 [GMT+2]
Dokument podpisany elektronicznie podpisem zaufanym

……..

data i podpis

Potwierdzam treść oświadczenia.

PODPIS ZAUFANY
JOANNA ZOFIA
SZYDA
12.09.2023 14:13:04 [GMT+2]
Dokument podpisany elektronicznie podpisem zaufanym

……..……

data i podpis promotora

Załącznik nr 3 do Trybu postępowania w sprawie nadania stopnia naukowego doktora

Bartosz Czech

imię i nazwisko

Wrocław, 11.09.2023

miejscowość i data

Katedra Genetyki

afiliacja

OŚWIADCZENIE

Oświadczam, że w pracy Host transcriptome and microbiome interactions in Holstein cattle under heat stress condition Czech Bartosz, Wang Yachun, Wang Kai [et al.], Frontiers in Microbiology, 2022, vol. 13, pp.1-9, Article number:998093. DOI:10.3389/fmicb.2022.998093 mój udział polegał na zaplanowaniu oraz wykonaniu bioinformatycznej oraz statystycznej analizy danych, wykonaniu wizualizacji oraz na napisaniu manuskryptu.

PODPIS ZAUFANY
BARTOSZ
CZECH
11.09.2023 13:51:05 [GMT+2]
Dokument podpisany elektronicznie podpisem zaufanym

……..

data i podpis

Potwierdzam treść oświadczenia.

PODPIS ZAUFANY
JOANNA ZOFIA
SZYDA
12.09.2023 14:10:12 [GMT+2]
Dokument podpisany elektronicznie podpisem zaufanym

……..……

data i podpis promotora

Załącznik nr 3 do Trybu postępowania w sprawie nadania stopnia naukowego doktora

Bartosz Czech

imię i nazwisko

Wrocław, 11.09.2023

miejscowość i data

Katedra Genetyki

afiliacja

OŚWIADCZENIE

Oświadczam, że w pracy Genome-wide association study of heat stress response in Bos taurus Bartosz Czech, Yachun Wang, Joanna Szyda bioRxiv 2023.06.05.543663; doi: https://doi.org/10.1101/2023.06.05.543663 mój udział polegał na zaplanowaniu oraz wykonaniu bioinformatycznej oraz statystycznej analizy danych, wykonaniu wizualizacji oraz na napisaniu manuskryptu.

PODPIS ZAUFANY
BARTOSZ
CZECH
11.09.2023 13:56:08 [GMT+2]
Dokument podpisany elektronicznie
podpisem zaufanym

.....

data i podpis

Potwierdzam treść oświadczenia.

PODPIS ZAUFANY
JOANNA ZOFIA
SZYDA
12.09.2023 14:16:02 [GMT+2]
Dokument podpisany elektronicznie
podpisem zaufanym

.....

data i podpis promotora