e-ISSN 2966-0289

Eventos Técnicos & Científicos

3

December, 2024

Proceedings

São Paulo School of Advanced Science on Hologenomic Data Analysis for Agriculture

July 31st to August 9th
São Carlos , SP

Embrapa
Pecuária Sudeste

Brazilian Agricultural Research Corporation
Embrapa Southeastern Livestock
Ministry of Agriculture and Livestock

e-ISSN 2966-0289

Eventos Técnicos & Científicos 3

Proceedings

São Paulo School of Advanced Science on Hologenomic Data Analysis for Agriculture

July 31st to August 9th
São Carlos , SP

Embrapa Southeastern Livestock
São Carlos, SP
2024

Embrapa Pecuária Sudeste
Rod. Wasghinton Luiz, km 234
13560-970, São Carlos, SP, Brazil
Fone: 55 16 3411-5600
https://www.embrapa.br/pecuaria-sudeste
www.embrapa.br/fale-conosco/sac

Text Revision
Patricia Tholon

Bibliographic standartization
Vera Viana dos Santos Brandão (CRB 8/7283)

Book layout
Maria Cristina Campanelli Brito

Graphic design
Maria Cristina Campanelli Brito

Digital publication: PDF

Cataloging-in-Publication (CIP)
Embrapa Pecuária Sudeste

São Paulo school of advanced science on hologenomic data analysis for agriculture. – São Carlos, SP: Embrapa Pecuária Sudeste, 2024.

85 p. : il. -- (Eventos Técnicos & Científicos / Embrapa Pecuária Sudeste, e-ISSN 2966-0289 ; 3)

1. Congresso. 2. Ciência Avançada. 3.Análise de dados. 4. Agricultura. I. Regitano, Luciana Correia de Almeida Regitano. II. Cardoso, Daniel Rodrigues. III. Colnago, Luiz Alberto. IV. Título. V. Série.

CDD (21. ed.) 607

Vera Viana dos Santos Brandão (CRB 8/7283)

Organizing Committee

Coordinators

Luciana Correia de Almeida Regitano - Embrapa Pecuária Sudeste, São Carlos, SP
Daniel Rodrigues Cardoso - Universidade de São Paulo, São Carlos, SP
Luiz Alberto Colnago - Embrapa Instrumentação, São Carlos, SP

Members

Adriana Mércia Guaratini Ibelli - Embrapa Pecuária Sudeste, São Carlos, SP
Alessandro Pelegrine Minho - Embrapa Pecuária Sudeste, São Carlos, SP
Alda Juliana Castro - Universidade Federal do Pará
Aline Silva Mello Cesar - ESALQ/ Universidade de São Paulo, Piracicaba, SP
Aline Teixeira do Brasil Morais Trimidi - Bolsista Pos-Doc Embrapa Pecuária Sudeste, São Carlos, SP
Ana Rita de Araujo Nogueira- Embrapa Pecuária Sudeste, São Carlos, SP
Avelardo Urano - Embrapa Pecuária Sudeste, São Carlos, SP
Bárbara Cunha Padilha Antonio - Universidade Federal de São Carlos, São Carlos, SP
Bianca Baccili Zanotto Vigna- Embrapa Pecuária Sudeste, São Carlos, SP
Bruno Gabriel Nascimento Andrade - Munster technological University, Cork, Irlanda
Cristiane Vieira Peres Fragalle - Embrapa Pecuária Sudeste, São Carlos, SP
Gabriel A Colmenarez Pena - Bolsista TT Fapesp na Embrapa Pecuária Sudeste, São Carlos, SP
Gerson Barreto Mourão - ESALQ/ Universidade de São Paulo, Piracicaba, SP
Gisele Rosso - Embrapa Pecuária Sudeste, São Carlos, SP
Hélio Omote - Embrapa Pecuária Sudeste, São Carlos, SP
Jennifer Jessica Bruscadin - Universidade Federal de São Carlos, São Carlos, SP
José Alberto Bastos Portugal - Embrapa Pecuária Sudeste
Joyce Yumi Inoue Tosi – Embrapa Pecuária Sudeste
Juliana Afonso - Bolsista Pos-Doc Embrapa Pecuária Sudeste, São Carlos, SP
Juliana Priscila Sussai - Embrapa Pecuária Sudeste, São Carlos, SP
Lívia Ferreira Pinho - Universidade Federal do Pará
Liz Jiannine Fuentes Rojas - Universidade Federal de São Carlos, São Carlos, SP
Maria Cristina Campanelli Brito- Embrapa Pecuária Sudeste, São Carlos, SP
Mônica Ferreira Laurito - Embrapa Instrumentação, São Carlos, SP
Nizael Francislei Rosa - Embrapa Pecuária Sudeste, São Carlos, SP
Patrícia Tholon - Embrapa Pecuária Sudeste, São Carlos, SP
Priscila Silva Neubern de Oliveira - Bolsista Pos-Doc Embrapa Pecuária Sudeste, São Carlos, SP
Robson Rodrigues Santiago - Embrapa Pecuária Sudeste, São Carlos, SP
Sérgio Raposo de Medeiros - Embrapa Pecuária Sudeste, São Carlos, SP
Tainã Figueiredo Cardoso da Cruz - Bolsista Pos-Doc Embrapa Pecuária Sudeste, São Carlos, SP
Wanderson Anjos da Silva - Universidade Federal de São Carlos, São Carlos, SP
Wilson Malagó Junior- Embrapa Pecuária Sudeste, São Carlos, SP

Scientific Committee

Luciana Correia de Almeida Regitano (coordinator)
Daniel Rodrigues Cardoso (coordinator)
Luiz Alberto Colnago (coordinator)

Adriana Mércia Guaratini Ibelli
Aline Silva Mello Cesar
Ana Rita de Araujo Nogueira
Bianca Baccili Zanotto Vigna
Bruno Gabriel Nascimento Andrade
Gerson Barreto Mourão
Juliana Afonso
Patrícia Tholon
Priscila Silva Neubern de Oliveira
Tainã Figueiredo Cardoso da Cruz

Forewords

As the global population continues to grow, the demand for crop and livestock products is also expected to increase. At the same time, climate change and other environmental factors create new challenges for the agricultural industry. For example, changing weather patterns and extreme weather events can impact feed availability, disease incidence, and productivity.

To address these challenges, there is a growing interest in developing new approaches to improve agriculture's productivity and sustainability. Studying the microbial community associated with a host is a new research field that can help understand what is needed for the necessary improvements. Using high-throughput technologies such as next-generation sequencing and large-scale-based metabolomics, we can comprehensively analyze the microbiome and associate these data with the host's genomic, epigenomic, transcriptomic, proteomic, and metabolic profile. These approaches enable researchers to analyze the complex microbial communities associated with agricultural species at a systems-level, providing a more comprehensive understanding of their composition and function. By identifying key microbial players, it may be possible to develop targeted interventions that can help modulate the microbiome and improve host health and productivity.

In addition, developing and incorporating more sophisticated analytical methods, such as artificial intelligence and language models, further enhance our ability to analyze the vast amounts of data generated by these approaches. For example, machine learning algorithms can identify patterns in the data and predict which interventions are likely to be most effective, which could accelerate the development of new treatments and management strategies. Integrating omic and meta-omic approaches with advanced technologies can transform the food production industry, enabling more efficient, sustainable, and environmentally friendly production practices.

As it is imperative that graduate students have access to such knowledge, the "São Paulo School of Advanced Science on Hologenomic Data Analysis for Agriculture'' came to provide students and young researchers an interdisciplinary vision of hologenomics and state-of-the-art knowledge on how to analyze and integrate the analysis of different omics. Furthermore, this school also discussed how these tools could improve the development of new agricultural management strategies.

Finally, we would like to acknowledge the support of the organizing institutions Embrapa and IQSC-USP, São Carlos and also, the FAPESP and CNPq for funding the "São Paulo School of Advanced Science on Hologenomic Data Analysis for Agriculture'.

Summary

Abstracts

Expanded summary

Abstracts

Coding and non-coding uterine exosomal contents differ in early pregnancies produced by artificial insemination and in vitro fertilization in cows(1)

Amanda de Oliveira Furlan(1*), Liza Margareth Medeiros de Carvalho Sousa(1), Alessandra Bridi(2), Flavia Regina Florêncio de Athayde(1), Maria Fernanda da Silva Lopes(1), Mariana Cordeiro Almeida(1), Juliana de Souza Felix(1), Natália Francisco Scaramele(1), Paula de Carvalho Papa Keohane(4), Juliano Coelho da Silveira3, Roberto Sartori Filho(3), Danila Barreiro Campos(5), Flavia Lombardi Lopes(1)

(1) Departamento de Produção e Saúde Animal, Universidade do Estado de São Paulo (Unesp), Faculdade de Medicina Veterinária, Araçatuba, SP, Brasil.

(2) Departmento de Medicina Veterinária, Faculdade de Ciência Animal e Engenharia de Alimentos, Universidade de São Paulo (USP), Pirassununga, SP, Brasil.

(3) Departamento de Ciência Animal, "Luiz de Queiroz" Colégio de Agricultura (ESALQ), Universidade de São Paulo, Piracicaba, SP, Brasil.

(4) Departamento de Cirurgia, Faculdade de Medicina Veterinária e Ciência Animal, Universidade de São Paulo, SP, Brasil.

(5) Centro de Ciências Agrárias, Departamento de Ciências Veterinárias, Universidade Federal do Paraíba (UFPB), Areia, Brasil.

*Corresponding Author: flavia.lopes@unesp.br

The gestational process is described by several physiological changes to sustain and provide adequate development of the embryo. Before and during pregnancy, effective communication between embryo/fetus and mother is vital and extracellular vesicles can mediate cell-to-cell communication. Among the extracellular vesicles, small vesicles known as exosomes with approximately (30–150nm) can act in maternal-fetal communication and transport of RNAs (mRNAs, miRNAs and lncRNAs) to other cells. MiRNAs contain about 19-25 nucleotides and function through post-transcriptional regulation of mRNAs. LncRNAs contain over 200 nucleotides, and regulate gene transcription directly on gene promoters, on chromatin modifying enzymes on mRNAs an proteins, as well as acting as miRNA sponges by competing with mRNA for miRNA binding. In mammals, miRNAs and lncRNAs have been found at all stages of development. Considering the prevalence of in vitro fertilization on bovine production, despite its lower birth rate, we aimed to evaluate if there are differences in coding and non-coding uterine exosomal content between pregnancies produced by artificial insemination (AI) and in vitro fertilization (IVF). We collected uterine fluid from indicine cows at 18 and 32 days of pregnancy, produced by AI and IVF, and separated vesicles corresponding in size to exossomes. Exosomal total RNA was then extracted for transcriptome and microtranscriptome analysis by RNA-Seq and microarray. Gene set enrichment analysis was performed for all treatment groups. A total of 21 significant ($p \leq 0.005$) gene sets were enriched in IVF18 x AI18 and 24 gene sets in IVF32 x AI32. Within techniques, 3 gene sets were enriched in AI32 x AI18 and 25 gene sets in IVF32 x IVF18. Of particular interest, pathways cytokine-cytokine receptor interaction (IVF32 x IVF18) and cell adhesion molecules (IVF32 x AI32) were enriched. Differential expression of exosomal lncRNAs was also performed, 435 lncRNAs were up- and 5 down-regulated in IVF32 x AI32, and 30 were up- and 25 down-regulated in AI32 x AI18, whereas 1,208 were up- and 22 down-regulated in IVF32 x IVF18. Through microtranscriptome analysis, 9 miRNAs were up- and 8 down-regulated in IVF18 x AI18, 5 up- and 8 down-regulated in IVF32 x AI32, 6 up- and 5 down-regulated in AI32 x AI18 and, and 11 up- and 2 down-regulated in IVF32 x IVF18. Differential exosomal content of coding and non-coding was much more pronounced between days 32 and 18 in IVF pregnancies than in AI, suggesting that temporal changes in transcripts can be influenced by reproductive technique. In summary, exosomes derived from AI and IVF pregnancies have significantly different coding and non-coding transcript content, which could be associated to failures in maternal-fetal communication in IVF, as these differentially expressed transcripts are involved in several processes vital to early pregnancy.

Index Terms: Assisted reproduction, Epigenetic, Exosome, lncRNA, microRNA.

Funding Agency: AAmanda de Oliveira Furlan was supported by a scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.
This study was financed in part by the São Paulo Research Foundation (FAPESP, SP, Brazil, Project 2018/03798-8 and 14/00739-9) and BARD-US Israel Binational Agricultural Research and Development Fund grant IS-4799-15.

Genetic predictions and multi-breed GWAS for traits of economic impact in beef cattle

André Luis Romeiro de Lima[1*], Fernando Flores Cardoso[2], Gilberto Romeiro de Oliveira Menezes[3], Fabio Luiz Buranelo Toral[4], Paulo Sávio Lopes[1]

[1] Programa de Pós-Graduação em Zootecnia, Universidade Federal de Viçosa (UFV), Viçosa, MG.

[2] Embrapa Pecuária Sul, Bagé, RS.

[3] Embrapa Gado de Corte, Campo Grande, MS.

[4] Programa de Pós-Graduação em Zootecnia, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG.

*Corresponding author: andre.romeiro@ufv.br

The weight gain in the post-weaning phase evaluates the animals, considering their genetic potential in the absence of maternal effects. Tick counting is a distinctive trait that enables individual identification, offering an alternative to infestations that cause significant economic losses. In a multi-breed context, a specific breed's genomic information may contribute to the other breeds, having greater relevance in situations in which crossbred breeds are being proved together with their pure breeds. Allelic origin of crossed animals and analysis of the predictive capacity of the genetic values using joint information from metapopulations are relevant tools for the genetic evaluation of crossbreed animals. The aim of the study is to evaluate the traits of tick count (TC) and post-weaning weight gain (ADG) for the Nelore, Hereford and Braford breeds under single and multi-breed approaches. Genotype and phenotype data for the Hereford and Braford breeds were made available by the Brazilian Association of Hereford and Braford (ABHB), while those for the Nelore breed were made available by Mundo Novo Farm, Uberaba, MG. Firstly, analysis of (co)variance component estimates for TC and ADW for the Nelore, Hereford and Braford breeds will be performed separately for each breed. In the second step, the result of the single trait analysis of variance components and genetic values will be used to predict the effects of markers in Genome Wide Association Studies (GWAS). In post-GWAS analysis, candidate genes will be searched for in the SNP regions relevant to each trait. All analyses of estimates of (co)variance components, predictions of genetic values and GWAS will be carried out using BLUPF90 family software. After identifying the candidate genes associated with TC and ADW, the allelic origin in the Braford breed will be identified following the BOA methodology. A multibreed GWAS will be carried out, considering the Nelore, Hereford and Braford breeds as a single population, aiming to find overlapping genes in both breeds that have already been described in single-breed GWAS. The coincidence in candidate genes between the methodologies (single or multi-breed) will be evaluated, as well as whether the origin of the crossbreed alleles (Braford) comes from the Nelore or Hereford population. Finally, a multi-breed genomic genetic evaluation will be carried out for TC and ADG, with the results compared with those previously obtained for each breed. The goal is to verify gains in accuracy or bias in predictions for the three breeds under multi-breed genetic evaluation.

Index Terms: composite breed, candidate gene, genomic selection.

Funding Agency: Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPQ.

Telomere-to-Telomere genome assemblies of human-infecting *Encephalitozoon* species

Anne Caroline Mascarenhas Dos Santos[(1)], Alexander Thomas Julian[(1)], Pingdong Liang[(1)], Oscar Juárez[(1)], Jean-François Pombert[(1*)]

[(1)] Department of Biology, Illinois Institute of Technology, Chicago, IL, USA
*Corresponding author: jpombert@iit.edu

Microsporidia are diverse spore forming, fungal-related obligate intracellular pathogens infecting a wide range of hosts. This diversity is reflected at the genome level with sizes varying by an order of magnitude, ranging from less than 3 Mb in *Encephalitozoon* species (the smallest known in eukaryotes) to more than 50 Mb in Edhazardia spp. As a paradigm of genome reduction in eukaryotes, the small *Encephalitozoon* genomes have attracted much attention with investigations revealing gene dense, repeat- and intron-poor genomes characterized by a thorough pruning of molecular functions no longer relevant to their obligate intracellular lifestyle. However, because no *Encephalitozoon* genome has been sequenced from telomere-to-telomere and since no methylation data is available for these species, our understanding of their overall genetic and epigenetic architectures is incomplete. In this study, we sequenced the complete genomes from telomere-to-telomere of three human-infecting *Encephalitozoon* spp. *E. intestinalis* ATCC 50506, *E. hellem* ATCC 50604 and *E. cuniculi* ATCC 50602- using short and long read platforms and leveraged the data generated as part of the sequencing process to investigate the presence of epigenetic markers in these genomes. We also used a mixture of sequence- and structure-based computational approaches, including protein structure prediction, to help identify which *Encephalitozoon* proteins are involved in telomere maintenance, epigenetic regulation, and heterochromatin formation. The Encephalitozoon chromosomes were found capped by TTAGG 5-mer telomeric repeats followed by telomere associated repeat elements (TAREs) flanking hypermethylated ribosomal RNA (rRNA) gene loci featuring 5-methylcytosines (5mC) and 5-hemimethylcytosines (5hmC), themselves followed by lesser methylated subtelomeres and hypomethylated chromosome cores. Strong nucleotide biases were identified between the telomeres/subtelomeres and chromosome cores with significant changes in GC/AT, GT/AC and GA/CT contents. The presence of several genes coding for proteins essential to telomere maintenance, epigenetic regulation, and heterochromatin formation was further confirmed in the *Encephalitozoon* genomes. Altogether, our results strongly support the subtelomeres as sites of heterochromatin formation in *Encephalitozoon* genomes and further suggest that these species might shutdown their energy-consuming ribosomal machinery while dormant as spores by silencing of the rRNA genes using both 5mC/5hmC methylation and facultative heterochromatin formation at these loci.

Index Terms: Computational biology, *Encephalitozoon*, Heterochromatin, Methylation, Protein structure, Telomeres.

Funding Agency: This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health [grant number R15AI128627] to J.-F.P. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Use of systems biology tools to assess the effects of fetal programming in beef cattle

Guilherme Henrique Gebim Polizel[1,2*], Angela Cánovas[3], Wellison Jarles da Silva Diniz[4], Heidge Fukumasu[1] , Miguel Henrique de Almeida Santana[1]

[1] College of Animal Science and Food Engineering (FZEA/USP), Pirassununga, SP, Brazil.
[2] Programa de pós graduação em Biociência Animal.
[3] University of Guelph, Guelph, ON, Canada. 4Auburn University, AL, USA.
*Corresponding author: guilherme.polizel@usp.br

Systems biology is the field that studies the interactions between components of a biological system in a holistic way, however it is still little explored in science due to the difficulty of integrating data from different molecular levels. It is already known that the nutrition of pregnant mothers in livestock is an extrinsic factor that can interfere with characteristics of productive interest throughout the life of the offspring, however the mechanisms related to the effects of this highly complex process still require studies. This project seeks to elucidate part of the mechanisms for multiomic and tissue integration in male Nellore animals that have been identified with different prenatal supplementation strategies. This study will be based on 63 animals from 3 different nutritional prenatal treatments: NP - only mineral supplementation (common to all treatments), PP - protein-energy supplementation in the final third of pregnancy and PC - protein-energy supplementation throughout the gestation. This project is divided into 3 studies. In study 1, a multi-omic approach of liver from selected animals will be carried out. The differential expression and co-expression networks of the genes will be evaluated by transcriptomics. The metabolites present in that organ will be evaluated by the metabolomics. These data will integrate with genomic data of 35 thousand single nucleotide polymorphisms (SNPs) to relate to the performance of these bulls. In study 2, the systems biology approach using differential expression and gene co-expression networks in the liver and skeletal muscle of bulls will be carried out by transcriptomics, the profile of metabolites in plasma, meat, subcutaneous fat and liver will be evaluated by metabolomics and profile of ruminal microbial communities will be evaluated by metagenomics. In this study, fetal programming will be associated with multi-tissue transcriptomics and multi-tissue metabolomics and the integration of omics, tissues and phenotypes of bulls from birth to slaughter. This project may demonstrate the biological basis of the influence of fetal programming on the performance of Nellore animals.

Index Terms: Bioinformatics, Nutritional genomics, Omics sciences, Prenatal nutrition.

Funding Agency: São Paulo Research Foundation (FAPESP).

Genome-wide association study for reproductive traits in a maternal line of pigs.

Guilherme Oselame[1]*, Suelen Fernandes Padilha Aleixo[(2)], Maurício Egídio Cantão[(3)], Jane de Oliveira Peixoto[(3)], Luis Orlando Duitama Carreño[4], Marcelo Silva de Freitas[4], Jader Silva Lopes[4], Aline Zampar[(1)] Mônica Corrêa Ledur[(1,3)]

[(1)] Programa de Pós-Graduação em Zootecnia, Universidade do Estado de Santa Catarina (UDESC)-CEO, Chapecó, Santa Catarina.

[(2)] Programa de Pós-Graduação em Zootecnia, Universidade Federal do Paraná, Curitiba, Paraná.

[(3)] Embrapa Suínos e Aves, Concórdia, SC.

[(4)] BRF SA, Curitiba, PR.

*Corresponding author: gui.oselame89@gmail.com

Pork is the most widely consumed animal-derived protein worldwide. To meet the demand, animals are selected based on their performance. However, regarding reproductive and litter traits, studies indicate low heritability and high genetic complexity, which makes genetic improvement of these traits slower and more arduous. Therefore, the aim of this study is to identify single nucleotide polymorphisms (SNPs) associated with teat number (TN), age at first mating (AFM), total number born (TNB), and number of live piglets at day 5 (LP5) traits through a genome-wide association study (GWAS). The data for this study are from a maternal line of Large White pigs from a nucleus farm of the BRF Company, located in Santa Catarina State, Brazil. The phenotype and pedigree databases contain data collected since 2001, with 16,797 and 187,847 records, respectively. The genotype dataset was collected between 2011 and 2023 and contain information from 4,366 animals genotyped with the GGP Porcine 50K and Porcine SNP80 BeadChip panels from Illumina®. For the analyses, the BLUPF90 family programs and the SNPs common to both panels (47,680) will be used. Initially, the data will be renumbered using the RENUMF90 package to organize the data to be used by the subsequent programs and to recode the pedigree. Then, with PREGSF90, sample and SNP quality control will be performed, excluding those that do not meet the criteria for analysis, as well as animals with inconsistencies in the relationship matrix. Furthermore, the GIBBSF90+ package will be used to estimate variance components and genetic parameters based on the pedigree and genotypes of the animals. The PREGSF90 package will be used to construct the hybrid matrix H and POSTGSF90 for the GWAS analysis. Genomic windows of 1 MB will be used to estimate the SNP effects to identify relevant genes and SNPs. Genomic windows explaining more than 3% of the genetic variance will be considered significantly associated with the traits. Based on the initial and final positions of the significant windows, a search will be conducted in the Ensembl database to identify and explore positional candidate genes. The DAVID and REViGO tools will be employed to perform functional enrichment analysis. Moreover, for the gene interaction analysis, the STRING database will be accessed through the utilization of the NetworkAnalyst tool. These findings will make a substantial contribution towards a deeper understanding of the genetic mechanisms underlying the reproductive performance of the evaluated line, enhancing the accuracy of selection within this population.

Index Terms: genetic parameters, GWAS, candidate genes, BLUPF90, reprodutction.

Funding Agency: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001. G.O. is supported by a CAPES scholarship.

Comparative satellitome of distinct *Scleropages* species (Osteoglossiformes, Teleostei): Insights into repetitive DNA evolution in an ancient fish group.

Gustavo Akira Toma[1*], Jhon Alex Dziechciarz Vidal[1], Fernando Henrique Santos de Souza[1], Guilherme Mota Souza[1], Tariq Ezaz[2], Alongklod Tanomtong[3], Fabio Porto-Foresti[4], Ricardo Utsunomia[4], Marcelo de Bello Cioffi[1]

[1] Departament de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil.

[2] Institute for Aplied Ecology, University of Camberra, Camberra, Australia.

[3] Department of Biology, Faculty of Science, Khon Kaen University, Muang, Khon Kaen, Thailand.

[4] Faculdade de Ciências, UNESP, Bauru, São Paulo, Brazil.

*Corresponding author: gustavo_toma@hotmail.com

Satellite DNAs (satDNAs) are composed of tandem repeats of a DNA motif arranged from shorts to large sizes arrays. These repetitive DNAs are commonly found in centromeric/pericentromeric chromosome areas and in heterochromatin. Although initially considered as non-transcribed “Junk DNA”, new evidence gather have linked these sequences with regulatory functions, such as centromere epigenetics and chromosome recombination. The use of short reads sequencing data with specialized pipelines has enable the high throughput mining of satDNAs, resolving the limitations of previous isolation methods. Thus, it is now possible to assess and compare different species satDNAs catalogs, known as the “Satellitome”, shedding light into the evolutionary process and molecular dynamics of these DNA sequences. By choosing three species of arowanas, named Scleropages formosus, *S. jardinii* and *S. leichardti*, we query whether the recovered satDNAs would follow or not the evolutionary relationships of these species, and tested for long-term sharing of any satDNAs. These congeneric species, extensively analyzed by our research group, represents a good experimental model, as they have well resolved taxonomic status, represents an ancient group withing Teleostei and present a well-established phylogenetic tree. We sequenced each species in the BGISEQ-500 platform (BGI, Shenzhen Corporation, China) and characterized their satellitome using the TAREAN pipeline. We assessed the similarity and homology of each satDNA by combining different pipelines and python scripts, and group them into variants (>95% similarity), family (80%-95% similarity) and superfamily (50%-80% similarity). Altogether, we found 25, 22 and 23 satDNAs for *S. formosus*, *S. leichardti* and *S. jardinii* respectively, with great variation in monomers size (6bp-4000pb) and predominance of A+T% rich (>50%) DNA sequences. Our interspecific in silico homology search revealed 12 superfamily relationships and six variants among the sampled arowanas, with great sharing of satDNAs between *S. leichardti* and *S. jardinii*, but not *S. formosus*. Although being congeneric species, *S. leichardti* and *S. jardinii* diverged from each other approximately 9.3 Ma ago. These findings, demonstrate an unusual conservation and sharing of satDNA catalogs, despite such a long evolutionary divergence time. As Arowana genomes reveal ancestral patterns of evolution, by employing this multispecies data collection, we can assess how satDNA families have evolve over time. To better elucidate this scenario and prospect possible reasons for this sharing, future steps include the design of satDNA primers for PCR procedures and Fluorescence in situ Hybridization (FISH) experiments, aiming to characterize the chromosome location of these DNA sequences and understand their possible role in centromere regulation.
Funding Agency: FAPESP (2018/14677-6)

Index Terms: Cytogenomics; Satellitome; repetitive DNA; Arowana.

LPL gene, a potential regulator of tail fat deposition in sheep, revealed by allele-specific expression analysis

Hossein Mansourizadeh[1], Mohammad Reza Bakhtiarizadeh[1*], Luciana Correia de Almeida Regitano[2], Jennifer Jessica Bruscadin[2,3]

[1] Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Tehran, Iran;

[2] Embrapa Southeast Livestock, São Carlos, Brazil;

[3] Department of Genetics and Evolution, Federal University of São Carlos (UFSCar), São Carlos, Brazil

* Corresponding author: mrbakhtiari@ut.ac.ir

In various sheep populations, the accumulation of fat in the tails exhibit remarkable variability. However, the molecular mechanisms behind these phenotypic differences are not well understood. Allele-specific expression (ASE) is one of these molecular mechanisms that refers to the difference in expression levels between two alleles of a gene and is a common phenomenon in regulating the amount of expression in mammals. In addition, since this phenomenon is likely driven by cis-regulatory mutations, scanning them across the sheep genome is an important step to elucidate the underlying mechanisms of gene expression regulation. From this perspective, the analysis of ASE may help to elucidate the genetic structure of complex traits such as fat deposition in sheep. In the present study, 45 RNA-Seq datasets were used for variant calling to perform ASE analysis. These datasets were obtained from seven distinct studies that compared the fat-tail transcriptomes of sheep breeds with different tail fat contents. Using a rigorous computational pipeline, 115 genes exhibiting differential ASE were detected, which was narrowed down to four key genes (TCP1, LPL, LRPAP1, and SOD3) by only considering those were found in multiple studies. Functional analysis showed that ASE genes play important biological functions in the development and maintenance of fat tissue. Notably, LPL emerged as particularly significant due to its presence in five studies, previous recognition as an ASE gene, and well-established role in fat deposition. Our findings suggest that the complexity of traits like fat-tail development is more comprehensively explained through diverse genetic interactions, including ASE. These new insights suggest that biallelic expression may not be typical in sheep fat-tail development. Therefore, the allelic expression imbalance in genes associated with fat deposition presents an additional dimension for future genetic research aimed at enhancing sheep breeding practices' effectiveness.

Index Terms: Allelic Imbalance, Lipoprotein lipase, RNA-Seq, Sheep.

Funding Agency: Not applicable.

Utilisation of white and orange-fleshed sweetpotato in the growth and reproductive phases of rabbit (*oryctolagus cuniculus* lam)

Ibikunle Olaleru[1,2*], Ahmed Abu[2]

[1] Department of Animal Science, University of Ibadan, Oyo State, Nigeria

[2] Farming Systems Research Programme, National Root Crops Research Institute, Umudike,Abia State Nigeria

*Corresponding author: olaleru.ibikunle@gmail.com

The pressure on maize for livestock feeding has necessitated the use of non-conventional feedstuff such as sweetpotato as alternative feed resource. Sweetpotato root is limited by nutrient quality when compared to maize. However, the composite of root and vine could enhance the nutrient quality. Current information on formulation based on sweetpotato composite for rabbit feeding is insufficient. Therefore, effect of composites of sweetpotato on growth and reproductive performance of rabbit were investigated. Roots and vine of sweetpotato varieties (White-Fleshed-WF (TIS87/0087) and Orange-Fleshed-OF (CIP440293)) were analysed for total carotenoids (µg/mL), tannin (mg/100g), saponin (mg/100g), Crude Protein-CP, Ether Extract-EE and Crude Fibre-CF using standard procedures. Composites of vine and root (35:65) of each of WF and OF were used to replace maize in rabbit diets at: 0%-Control, 25%OF, 50%OF, 75%OF, 25%WF, 50%WF and 75%WF. Weaner rabbits (n=84), aged 7-8 weeks, weighing 567.6±13.7g were randomly allotted to the diets in triplicates of four rabbits/replicate and fed ad-libitum for 72 days. Feed intake-FI (g), survivability, Dressing Percentage-DP, CP and CF digestibility were determined using standard procedures. Does (n=25), aged 6-7 months, weighing 2169.8±150.01g were randomly allotted to five treatments: control, 25%OF, 50%OF, 25%WF and 50%WF with five replicates, in a completely randomised design. Body Weight Before Gestation-BWBG (g), Body Weight at Parturition-BWP (g), Post Gestation Feed Intake-PGFI (g), and Litter Size at Birth-LSB were recorded. Data were analysed using descriptive statistics and ANOVA at α0.05. Total carotenoids in OF (2,182.3±40.2) and WF roots (2,057.4±90.4) were significantly higher than 1,511.9±100.5 and 1,297.9±122.7 in the corresponding OF and WF vine. Tannins 0.03±0.2 of WF vine were higher than 0.02±0.1 (OF root), 0.02±0.1 (WF root) and 0.02±0.03 (OF vine). Saponins ranged from 0.3±0.04 (WF vine) to 0.41±0.1 (OF root) while CP of 11.5±0.3 (OF) and 11.4±0.6 (WF) vine were higher than 8.1±0.6 (OF) and 5.3±0.3 (WF) roots. Vine CF value of 6.9±0.3 (WF) was significantly higher than 6.6±0.1 (OF) vine, 6.4±0.3 (OF) root and 5.7±0.2 (WF) root. Diets 25%OF, 25%WF and control resulted in similar FI of 50.1±1.6, 48.0±2.1 and 59.2±4.9, while 50%OF, 75%OF, 50%WF and 75%WF had lower FI of 43.1± 3.4, 40.1±1.8, 37.3±2.2 and 33.8±3.1, respectively. Survivability of 91.7±1.4% in control, 25% (OF) and 25% (WF) were higher than 75.0±1.6% (50%OF), 66.7± 0.8% (50%WF), 58.3±0.6% (75%OF) and 58.3± 0.6% (75%WF). The DP ranged from 49.5±0.9% (75%OF) to 58.6±4.7% (50%WF). Crude protein digestibility of 82.3 ± 2.5% (50%OF) was higher than 77.1±3.6 (25%WF), 75.4±3.7 (75%WF), 75.3±0.5 (Control), 73.2±2.1 (50%WF), 73.1±2.4 (25%OF) and 72.1±1.6 (75%OF), while CF digestibility ranged from 25%WF (13.3±2.2%) to 75%OF (19.5±1.1%). The BWBG ranged from 2,041.8±256.2 (25%WF) to 2,408.8±321.8 (Control) while BWP ranged from 2,214.2±102.2 (WF 50%) to 2,707.0±122.5 (Control). The PGFI of 126.9±3.1 (25%WF) and 126.4±2.4 (25%OF) were higher than 124.2±3.1 (50%OF), 123.4±2.1 (50%WF) and 121.2±1.6 (control) while LSB ranged from 4.0±0.9 (50%OF) to 5.0±0.7 (25%WF;50%WF), showing the synergistic influence between composites of sweetpotato and LSB. Orange-fleshed sweetpotato vine and root composites successfully replaced maize in rabbit diets. The nutrient digestibility, overall growth and reproductive performance were enhanced.

Index Terms: Rabbit, Nutrient digestibility, Reproductive performance, Composite sweetpotato meal

Studies of candidate resistance genes (MHC) in horses and donkeys

Ingrid Pereira Pinto Oliveira[1*], Valdecy Aparecida Rocha da Cruz[1], Luís Fernando Batista Pinto[1], Raphael Bermal Costa[1], Gregório Miguel Ferreira de Camargo[1]

[1] Universidade Federal da Bahia (UFBA), Salvador, BA.

*Corresponding author: ingridppoliveira@gmail.com

Diseases in horses generate financial costs and concern for breeders, as these animals can be sensitive to various diseases and the symptoms that affect them. Knowing this sensitivity, several studies were carried out to identify which genes may be related to the resistance that some horse breeds present to these diseases. Genes belonging to the Major Histocompatibility Complex (MHC) emerge as the main candidates for identifying resistant animals. In equids, several studies have been observed that identify polymorphisms in MHC class II responsible for making some animals and/or specific breeds more resistant than others. Furthermore, it is observed that the environment in which the animal/breed is inserted directly interferes with its ability to survive. The Marwari breed, a horse adapted to inhospitable and desert regions of India, with polymorphisms being identified in the DRB3 gene as possibly responsible for resistance. A similar phenomenon occurs in the Brazilian Pantaneiro breed, which survives most of the year in flooded fields, and which commonly affects horses with hoof problems. In addition, the animals of this breed are resistant to Equine Infectious Anemia, as can be observed in the Lavradeiro breed. This fact can also be observed in donkeys that are naturally more resistant to clinical signs of illness, and may be related to the inhospitable environments in which these animals survive. Therefore, it is necessary to understand the genetic mechanisms that involve the resistance of some breeds as a result of diseases. To achieve this, we seek to identify polymorphisms in the DRB3 gene and other genes belonging to the MHC that may be related to the resistance of native Brazilian equine breeds to diseases such as equine infectious anemia. For this, biological material will be collected from 10 animals of each equine breed (Arabian, Lavradeiro, Marajoara, Nordestino, Pantaneiro, Puruca and Quarter Horse), and 10 animals of each donkey breed (Brazilian Donkey, Nordestino Donkey and Pêga Donkey), totaling 100 animal samples. Extractions of total genomic DNA, quantification of the extracted DNA, sample dilutions and PCRs will be carried out. The primers to be used refer to the eight exons of the DRB3 gene, and later other genes belonging to the MHC will be chosen and primers created. Next, the PCR products will be subjected to agarose gel electrophoresis, purified, quantified and sequenced for both primers. The sequences obtained will be edited, aligned and analyzed using the CodonCode Aligner 11.0 software. Therefore, the project proposes to carry out an analysis and characterization of the frequency of polymorphisms in the DRB3 gene and other genes belonging to the MHC in the aforementioned equine breeds, in order to identify possible polymorphisms related to the greater resistance of some breeds to diseases.

Index Terms: Immune system; DRB3; Equine infectious anemia; Locally adapted breeds; Equus.

Funding Agency:The authors wish to thank the “Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)” for the financial support and “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)” for the fellowship of the first author.

Potential target genes prospecting for RNAi-mediated control of *Mahanarva* spp. spittlebugs in forage and sugarcane

Isabela S. Begnami[1,2*], Marcos R. Gusmão[3], Bianca B. Z. Vigna[3], Anete P. de Souza[1]

[1] Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, Campinas, SP.
[2] Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Estadual de Campinas, SP.
[3] Embrapa Pecuária Sudeste, São Carlos, SP.
*Corresponding author: isabegnami@gmail.com

Spittlebugs of the genus Mahanarva emerge as one of the most limiting pests to grass production in Brazil, causing major production losses that mainly affect the cultivation of forage grasses (Mahanarva spectabilis) and sugarcane (M. fimbriolata). Currently, there are not many viable alternatives for containing these insects, with the use of chemical pesticides being ecologically discouraged and biological control by Metarhizium anisopliae still inefficient on a large scale. Interfering RNA technology has been gaining prominence in the fight against herbivores, making it possible for plants to be resistant to specific species of pests in a more ecological way. For development of this technique, it is necessary to understand the insect in a genomic and transcriptomic level to find the best target gene for silencing. Therefore, the objective of this project is to identify and test possible genes of interest for RNAi-mediated silencing in two species of spittlebugs, Mahanarva fimbriolata and M. spectabilis, representing an innovative approach to controlling these pests. Firstly, investigative genomic analysis will be carried out, with the assembly of genomes and transcriptomes, to prospect genes related to insect development to be silenced. Afterwards, the chosen molecules will be tested, validated in insects and then delivered in sugarcane. By using these methodologies, we hope to identify genes with the best silencing potential and carry out proof of concept for them, contributing to the development of alternative technologies for controlling two major pests in national agriculture. The project will be carried out in partnership with Embrapa Pecuária Sudeste and Instituto Agronômico de Campinas.

Index Terms: gene silencing, pest control, plant-sap-feeding insects

Funding Agency: CAPES, Embrapa.

Characterization of genes expressed in donkey milk: first step towards identifying biological potential

Jackeline Santos Alves[1*], Raphael Bermal Costa[1], Gregório Miguel Ferreira de Camargo[1]

[1] Escola de Medicina Veterinária e Zootecnia, Universidade Federal da Bahia (UFBA), Salvador, Bahia, Brazil.
*Corresponding author: jsalves.bio@gmail.com

The donkey population in Brazil is declining. Agricultural mechanization generates abandonment and consequently indiscriminate slaughter due to the interest of Asians in the cosmetic properties of the species' skin. It is necessary to propose alternatives for the production of donkeys that allow their conservation. Milk production is one of them. Donkey milk is an excellent substitute for human breast milk, since it is hypoallergenic. It is also used in the production of cosmetics. However, little is known about why this product has these properties (not even the skin, which is more intensely used, is known). The study of genes expressed in donkey milk and skin can be fundamental for better understanding the protein profile to better understand its properties. It valorizes products of donkey origin, promoting conservation of the species. Thus, organized production and conservation of the species can be promoted. Therefore, in this study, the objective is to investigate variations in genes of milk proteins of donkeys, using the Nordestino Donkey as a study model, an adapted genetic group that is undergoing a population reduction.

Index Terms: caseins, genetic diversity, local adapted breeds, milk allergy, neglected species.

Funding Agency: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brasil) apoio financeiro (Processo nº 423258/2021-4).

Effects of 3-nitrooxypropanol on the modulation of rumen microbiota and enteric methane emission of dairy calves during the first six months of life

Jaqueline Fernandes Bruno[1,2*], Teresa Cristina Alves[2], Ives Cláudio da Silva Bueno[1]

[1] Faculdade de Zootecnia e Engenharia de Alimentos da Universidade de São Paulo (FZEA-USP), Pirassununga, SP.

[2] Embrapa Pecuária Sudeste, São Carlos, SP.

Corresponding author: jaquelinebruno@usp.br.

Given the global emphasis on reducing enteric methane (CH_4) emissions from livestock, several mitigation strategies are being studied. In Brazil, most animal production is carried out extensively on pastures, which commonly present some level of degradation and a high intensity of CH_4 emissions, making it necessary to mitigate these emissions. Evidence suggests that changes in rumen microbial colonization before weaning of dairy calves may influence the ruminal microbiome and impact phenotypes later in the animal's life. Additionally, recent results show that the chemically synthesized CH_4 inhibitor 3-nitrooxypropanol (3-NOP) is one of the most effective approaches for reducing enteric CH_4 in ruminants. However, studies with dairy calves since birth under tropical and grazing conditions are still scarce. Thus, this project proposes to investigate how dietary manipulation since birth can affect performance, microbial protein synthesis, CH_4 production, serum and rumen metabolome, and ruminal microbial population of dairy calves, through modulation of the ruminal microbiome. The experiment will be carried out at the Milk Production System of Embrapa Southeast Livestock, located in São Carlos, SP, Brazil. A batch of dairy cows and heifers of the Kiwi breed from Embrapa herd underwent a fixed-time artificial insemination (FTAI) protocol to synchronize the animals' estrous cycles. Sixteen to twenty-three days later, these animals were inseminated during natural heat using sexed semen from the same bull to ensure calving occurs as closely as possible, avoiding heterogeneity among the animals. Fourteen newborn Kiwi female calves will be randomly distributed into two collective rearing groups: treatment and control. The treatment group will receive the food additive 3-NOP at a dose of 3 mg/kg BW/day orally from birth until 12 weeks of life, while the control group will receive a placebo at the same dose. Metagenomic analysis will be carried out at weaning and in the sixth month of each animal's life. The foundation of molecular techniques for studying rumen microbial ecology will be based on shotgun metagenomic sequencing analysis. For the metabolomic profile of serum samples, blood samples will be collected at birth, in the 1st, 2nd, 4th, and 6th months old, by venipuncture of the jugular vein before morning feeding. For the metabolomic profile of rumen samples, the contents of all calves will be collected in the 1st, 2nd, 4th and 6th months old. Metabolomic analysis of serum samples and rumen content will be carried out using Nuclear Magnetic Resonance (NMR) equipment at 600 MHz for hydrogen frequency. For the statistical analysis, each animal will be considered an experimental unit. The data will be analyzed as repeated measures over time.

Index Terms: 3-NOP; Greenhouse gases; Kiwi; Metabolomics; Metagenomics.

Funding: This research proposal has been submitted for funding to the São Paulo Research Foundation (FAPESP) under process #2024/02777-7, which is currently under review.

Identification of novel mRNA isoforms in abomasal lymph node tissue of Scottish Blackface lambs affected by Teladorsagia circumcincta gastrointestinal nematode

Krishani Sinhalage[1], Stephanie Lam[1], Victoria Asselstine[1], Angela Cánovas[1]

[1] University of Guelph, Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, 50 Stone Road East, N1G 2W1 Guelph, ON, Canada
Corresponding author: ksinhala@uoguelph.ca

Gastrointestinal nematode (**GIN**) infections are the most prevalent parasitic diseases in grazing sheep worldwide, resulting in loss of sheep productivity, high mortality rate, subsequent economic losses and emerging animal welfare problems. In temperate climates, including Canada, a wide variety of production-limiting nematodes, such as Teladorsagia circumcincta (**TC**), *Haemonchus contortus* (**HC**), and *Trichostrongylu*s spp., commonly infect grazing sheep. Control of GIN requires costly anthelmintic treatments and proper pasture management. However, anthelmintic residues in sheep and anthelmintic resistance in nematodes are major problems. Assessment of the underlying host genetics and selection for increased GIN resistance are alternative strategies to combat parasitic infections in sheep. Transcriptomics using RNA-Sequencing (**RNA-Seq**) can effectively assess the differences in transcriptome profile and its impact on GIN resistance in sheep by identifying differentially expressed (**DE**) mRNA isoforms (**DEIs**) generated by alternative splicing. We aimed to assess the underlying genetics of GIN resistance in sheep by identifying DE novel and annotated DEIs using publicly available RNA-Seq data (NCBI GEO accession id: GSE63547). Abomasal lymph node tissue samples from Scottish Blackface lambs (n=20), classified as highly resistant (**R**; n=10) or susceptible (**S**; n=10) based on fecal egg counts. Large Gap Read Mapping analysis were performed using CLC Genomics Workbench (|FC|>2; FDR<0.05). The analysis revealed 30 DEIs (10 upregulated and 20 downregulated in S), including 6 novel and 24 annotated DEIs between R and S sheep at 7 days post-infection (**psi**). At 14 days psi, 3 novel and 19 annotated (6 upregulated and 16 downregulated in S) DEIs were identified between R and S sheep. Among them, *WDFY4* involved in cell-associated antigens presentation (|FC|=9.48) and *DOCK9* involved in GTPase-mediated signal transduction (|FC|=2.05) were the most significant DEIs upregulated in resistant sheep in the 7 and 14 d psi groups, respectively. Functional genomics analysis of DEIs, including metabolic pathway analysis, gene ontology and gene networks, were performed using Reactome (FDR<0.05). A total of 9 and 4 metabolic pathways were significantly enriched (FDR<0.05) and are primarily responsible for cellular physiological events, including immune response-associated factors in S sheep at 7 and 14 d psi, respectively. This study will provide a better understanding of the etiology of the disease by identifying key regulatory novel mRNA isoforms to enhance selection strategies for higher GINresistant sheep with high productivity.

Index Terms: Gastrointestinal nematodes, sheep, resistance, transcriptomics.

Funding Agencies: Food from Thought Project, University of Guelph, and OMAFRA, ON, Canada Presenting.

Assessing the impact of selection on the Holstein cattle genome through selection signatures

Larissa Graciano Braga[1,2], Danísio Prado Munari[2], Tatiane Cristina Seleguim Chud[1,3], Marco Antônio Machado[4], João Cláudio do Carmo Panetto[4], Marcos Vinícius Gualberto Barbosa da Silva[4], Flávio Schramm Schenkel[1]

[1] Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada;

[2] Departamento de Ciências Exatas, Universidade Estadual Paulista, Jaboticabal, São Paulo, Brazil;

[3] PEAK, URUS Group LP, Madison, WI, USA;

[4] Embrapa Gado de Leite, Juiz de Fora, Minas Gerais, Brazil.

*Corresponding author: flavio.schenkel@gmail.com

Unique genetic patterns in the genome of populations that have undergone a selection process are called selection signatures. These signatures can point to genomic regions responsible for the expression of economic traits under selection. In dairy cattle, the Holstein breed has been intensively selected to maximize productivity, particularly in Canada and the United States. These countries are the leading exporters of Holstein cattle genetics worldwide. This study aims to assess the impact of selection on the Holstein breed genome by detecting selection signatures. Single nucleotide variants from 307 Holstein animals from Canada and the United States were obtained from the 1000 Bull genome project (Run 9). Tajima's D test was computed with 50 kb windows using the VCFtools software. Only the bottommost 1% of windows were considered potential selection signatures. Neighboring selection signatures were merged, resulting in 344 non-overlapping selection signatures. In the functional analysis, 1,556 QTLs and 445 genes were co-localized with 220 selection signatures, including PLAG1, CAV1, BAD, and AQP1 genes. PLAG1 is a transcription factor that regulates many genes and pathways. It is an important regulator of stature and has previously been associated with milk composition, growth traits, and heifer livability in Holstein cattle. The CAV1 gene was previously found downregulated in the mammary gland of Holstein cows with hemorrhagic mastitis, with minor evidence of CAV1 protein in the vascular cells of affected cows compared with healthy mammary glands. CAV1 was also previously observed to be upregulated in bovine ovaries exposed to elevated non-esterified fatty acids (NEFA) compared to basal levels. During the negative energy balance, there is increased lipolysis and high levels of NEFA in circulation, which may further affect fertility due to the excess accumulation of lipids in the oocytes. The BAD gene controls apoptosis, or programmed cell death, in early embryo development stages. While apoptosis is a normal occurrence in developing embryos, BAD overexpression can stimulate embryonic death by sensitizing cells to proapoptotic signals, which is associated with poor pregnancy outcomes. AQP1 plays an important role in cellular fluid dynamics and has been found to be expressed in the cumulus-oocyte complex, blastocyst embryo stage, and the bovine mammary gland, where it may be responsible for increasing permeability and, to a lesser extent, milk secretion. QTL related to milk production, stature, growth traits, and reproduction were enriched (FDR<0.01). In conclusion, extensive genetic selection to promote milk yield has led to a drastic increase in energetic demands and reduced fertility in Holstein cattle, which can also be detected at genome level. Selection signatures related to milk production, reproduction, growth, and health were found, identifying regions in the genome responsible for controlling those traits.

Index Terms: adaptation, dairy, selective sweep, taurine.

Funding Agency: CAPES (001 and 88887.802720/2023-00).

A new lncRNA may impact bovine meat quality through immune response pathways

Luan Gaspar Clemente[1], Aline Silva Mello Cesar[1], Luciana Correia Almeida Regitano[2], Luiz Lehmann Coutinho[1]

[1] Luiz de Queiroz College of Agriculture" – University of São Paulo, Piracicaba/SP.

[2] Embrapa Livestock, São Carlos/SP.

*Corresponding author: luan.clemente@usp.br

Long non-coding RNAs (lncRNAs) have emerged as regulators of various cellular processes, including meat quality-related ones. Understanding the molecular mechanisms underlying lncRNA-mediated regulation of these processes could provide valuable insights for improving meat quality through targeted interventions. To investigate the potential involvement of lncRNA expression in Nelore cattle meat quality, a cohort of 192 animals, previously phenotyped for meat quality traits and productivity, was selected for the study. To extract total RNA, a 100 mg sample of *Longissimus dorsi* muscle tissue was treated with Trizol reagent following the manufacturer's instructions. The integrity of the RNA was assessed using the Bioanalyzer 2100, with samples having RNA integrity numbers (RIN) above 7 being selected for further analysis. 2 µg of RNA from each sample was employed for cDNA library preparation, following the TruSeq RNA Sample Preparation kit v2 protocol. Subsequently, the libraries underwent sequencing on the Illumina HiSeq2500 system utilizing the TruSeq SBS kit v3-HS (200 cycles). The reads underwent quality control (TrimGalore), mapping (STAR), transcriptome reconstruction and quantification (StringTie), and detection of new lncRNA isoforms (FEELnc). The transcriptome reconstruction resulted in 141,614 transcripts distributed across 41,038 gene loci. Count and prevalence filters were applied, and gene-level count data were converted to log2CPM, resulting in 17,453 genes. Co-expression analysis using WGCNA identified 24 modules centered around lncRNAs (hub genes), which were further analyzed through partial correlation analysis to reveal potential lncRNA targets. Statistically significant correlations, with a coefficient greater than 0,30 (for phenotypes) and 0.70 (for mRNA expression), were selected for enrichment analysis. Our analysis revealed significant associations between lncRNA-mRNA-phenotypes, indicating potential associations to meat quality traits such as shear force, water retention, fat thickness, and pH. Notably, the novel lncRNA MSTRG.10975 (chr16: 67,424,465 - 67,424,709) demonstrated correlations with 33 mRNA transcripts and shear force. Functional enrichment analysis highlights its involvement in crucial biological pathways, including post-translational modifications (SUMO ligase complex), chromatin remodeling (Cohesin Loading onto Chromatin), and immune/adaptive systems (IL-1 signaling pathway). The interaction among SUMO ligase complex, Cohesin Loading into Chromatin, and IL-1 signaling pathway may be related through post-translational modifications and immune response modulation. SUMO ligase complex can affect IL-1 signaling proteins, while Cohesin Loading influences chromatin structure, possibly impacting IL-1- responsive genes. Moreover, interleukins may impact meat tenderness by modulating inflammation, muscle metabolism, and oxidative stress during pre-slaughter and meat maturation processes. Their role in enzyme activity and muscle protein turnover also can influence meat texture. In conclusion, our study highlights the potential role of lncRNAs in modulating meat quality in Nelore cattle. By revealing significant correlations between lncRNA-mRNA expression and meat quality traits, we offer valuable insights into the molecular mechanisms underlying these complex traits. Ongoing efforts are being made to further elucidate these correlations.

Index Terms: lncRNA, mRNA, gene co-expression, meat quality, immune response.

Fundings: This work was funded by Sao Paulo Research Foundation (FAPESP) – Thematic Project "Molecular bases of meat quality in Nelore cattle" 12/23638-8. Moreover, the student receives a fellowship from Coordination for the Improvement of Higher Education Personnel (CAPES) – 88887-705131/2022-00.

Meta-phenomics behind quinoa seed priming, dormancy, and longevity

Maryam Razzaq[1], Luke T. Dunning[2], Irfan Afzal[3]

[1] PhD scholar, Seed Science and Technology, Department of Plant breeding and Genetics, University of Agriculture, Faisalabad-38040, Pakistan.

[2] University of Sheffield, Sheffield, UK.

[3] Seed Physiology Lab, University of Agriculture, Faisalabad-38040, Pakistan.

Quinoa grain loss poses a significant worldwide productivity challenge, primarily due to storage deterioration. Abiotic factors like temperature, humidity, and seed moisture exacerbate seed decay. These continuously and rigorously occurring deteriorating events can be minimized via various enhancement techniques. Among them priming have a promising effect against reduced seed longevity conditions (temperature, and humidity). Mediating longevity-based losses against such aging conditions is a main element that spurs on researchers to unveil various plethora of seed longevity in relation to seed enhancements, specifically their relative concentrations and work board range including seed-water relations. Similarly, dormancy is another unprecedented life event in seeds, which interacts with seed longevity. Limited insights into the genomics and molecular mechanisms governing seed dormancy, and longevity in quinoa hinder effective seed storage practices. So, based on these two aspects, priming can be employed as a solution. Therefore, it was hypothesized that when quinoa seed is primed before storage at low Relative humidity aging then there must be a chance of less deterioration, prolonged seed longevity, and release of dormancy. In order to meet the study objectives, this study was designed in two parts viz. phenotyping and genotyping. In the first experiment, 130 quinoa accessions after harvesting were screened for dormancy associated seed longevity traits and then the screened germplasm was subjected to low relative humidity aging for a period of 5 months, in “Seed Physiology lab, University of Agriculture, Faisalabad”. Similarly, in the second experiment, pre-screened seeds were subjected to seed priming and then exposed to artificial aging treatments (38°C and 33% relative humidity). Phynotypic results of seed traits (germination, abnormal seedlings, and seed vigor index) have been found significant against seed aging, and dormancy. Though, for further clarification the study would be repeated. In the next phase of the study, genotyping data will be acquired for each experiment and analyzed in Luke Dunning’s lab, School of Biosciences, University of Sheffield, Sheffield, UK. Based on results, the conclusions would be drawn and marker trait association would be deducted through mixed linear model of GWAS. The markers of seed longevity, dormancy and priming would be identified and compared with each seed trait. Relying on phylogenic assessments of these traits, this study would also penlights the effect of harsh environmental shifts on some specific accessions and prediction of seed’s prolonged shelf lives would be more evident with respect to their seed dormancy patterns. The compelling protagonist of this research would be the use of data in favor of the systemic biology behind these two quantitative traits (seed dormancy and seed longevity) and to fill the remnants of research gaps.

Index Terms: GWAS, productivity, genotyping

Role of the microbial cycling genes in biocrusts from extreme environments

Matías Garcia[1,2,3*], Michel Abanto[1,3], Paola Duran[2]

[1] Scientific and Technological Bioresource Nucleus (BIOREN). Universidad de La Frontera, Temuco, Chile.

[2] Biocontrol research laboratory. Universidad de La Frontera, Temuco, Chile.

[3] Programa de Doctorado en Ciencias mención Biología Celular y Molecular Aplicada, Universidad de La Frontera, Temuco, Chile.

*Corresponding author: mdagarcia07@gmail.com

As climate change continues to impact the world, understanding the natural processes that regulate the carbon cycle has become increasingly crucial. Biological Soil Crusts (BSCs), composed of microorganisms such as cyanobacteria, algae, fungi, and lichens, have emerged as key players in this scenario. Predominant in arid and semi-arid ecosystems, BSCs play a vital role in biogeochemical cycles, including the fixation of nitrogen and carbon. These microbial communities not only contribute to soil fertility and the reduction of erosion but are also efficient in capturing CO_2, which is crucial for mitigating the effects of climate change. However, there are still gaps in our knowledge about the functional diversity and resilience of BSCs under stress conditions, such as increased CO_2 levels. BSCs have been found to be crucial for soil stability, water and nutrient cycles, and trace gas exchange. However, the global coverage of BSCs is projected to decrease due to human-induced climate change and land use intensification, negatively affecting these processes. We have conducted an analysis of these BSCs in both environments, the Atacama Desert and Antarctica, obtaining their chemical parameters (exchangeable cations, salinity, moisture, Olsen phosphorus, total carbon, nitrogen). Additionally, a complete shotgun metagenomics analysis of the samples was performed at zero-time, control, and treatment. This analysis shows differences in genes, as well as in the taxa present in both samples from the Atacama Desert and Antarctica. Besides observing differences in genes, we were also able to recover the genomes of different species through the Metagenome-Assembled-Genome approach.Addressing these gaps and challenges in BSC research is essential for developing effective and sustainable management strategies. This pioneering study on BSC in extreme environments provides a crucial basis for future research and sustainable management strategies to address the challenges of climate change. A deeper analysis is necessary to explore the functional diversity and potential of BSC for agronomic applications and soil treatment.

Index Terms: biocrusts, microbiology, metagenomics, ecology, carbon.

Funding: This research was funded by the Anillo project ATT223360.

Genomic association for marbling in different pig breeds

Mathias Sunye Netto[1*], Jane de Oliveira Peixoto[2], Maurício Egídio Cantão[2], Luis Orlando Duitama Carreño[3], Marcelo Silva de Freitas[3], Jader Silva Lopes[3], Mônica Corrêa Ledur[1,2]

[1] Programa de Pós-Graduação em Zootecnia, Universidade do Estado de Santa Catarina (UDESC), Chapecó, Santa Catarina.

[2] Embrapa Suínos e Aves, Concórdia, SC.

[3] BRF SA, Curitiba, PR.

*corresponding author: mathiassunyenetto@gmail.com

Pork is one of the most consumed animal proteins worldwide. Brazil ranks as the third-largest pork producer globally, with the Santa Catarina State accounting for 29% of the National production. Marbling, which represents the percentage of fat within the muscle fibers of the meat, has a direct impact on consumer purchasing preferences. The heritability estimates for marbling is approximately 39%, highlighting the significant genetic influence on this trait. However, the genes responsible for this variation remain unidentified. Hence, the project aims to identify genomic regions and candidate genes associated with marbling in four lines of pigs: Large White, Pietrain, Duroc and a synthetic line. Phenotypic data will be provided by BRF Company. Animals of each breed will be classified into two groups based on high and low marbling percentage, with tail tissue samples utilized for DNA extraction. Genotyping will be performed with the Illumina GGP Porcine 50K SNP chip. Quality control of samples and genotypes, along with genomic association analysis in a case and control design, will be executed using PLINK 2.0. Upon identifying the genomic regions associated with marbling, a search for candidate genes within these regions will be carried out using several public databases. It is expected through the analyses, to identify markers and candidate genes that can be used in assisted selection to enhance the marbling of different breeds of pigs.

Index Terms: genotyping, GWAS, swine.

Leaf senescence in sunflower: systemic approach to a complex event with impact on genetic improvement

Melanie Corzo[1*], Daniela Becheran[2], Edmundo Ploschuk[2], Sofia A. Bengoa Louni[1], Santiago Trukpin[3], Daniel Alvarez[4], Nicolas Heinz[4], Natalia Izquierdo[5], Carla Caputo[2], Sebastian Moschen[6], Paula Fernandez[1]

[1] IABIMO- UEDD INTA-CONICET

[2] Facultad de Agronomia, Universidad de Buenos Aires

[3] IF-INTA Castelar

[4] EEA-INTA Manfredi

[5] Facultad de Ciencias Agrarias. UNMdP Balcarce

[6] CONICET

* Corresponding author: corzo.melanie@inta.gob.ar

Leaf senescence represents the terminal phase of leaf development, characterized by a decrease in photosynthetic activity, active degradation of cellular structures and recycling of nutrients to growth-active regions of the plant, such as shoots, young leaves and reproductive structures. This process is important from an economic point of view, as it impacts crop yield by affecting the plant's ability to maintain an active photosynthetic system, especially during critical growth phases such as grain filling, which impacts total plant biomass and oil content. In the context of climate change, it is crucial to understand leaf senescence, as it could be influenced by various stresses or environmental conditions, thus limiting yield in agronomically important crops. Sunflower (*Helianthus annuus* L.) is a crucial economic crop for its protein and high quality oil content rich in polyunsaturated fatty acids. Leaf senescence, governed by genetic and environmental factors, affects crop performance by prolonging the duration of photosynthesis and allowing more efficient remobilization of nutrients to support plant reproduction. Advances in genomics and functional genomics have catalyzed the discovery and characterization of genes involved in these processes, highlighting the role of functional genomics in integrating transcriptomics, proteomics and metabolomics data for a robust analysis of gene function. In this work, the stability of two sunflower genotypes with widely described responses to nitrogen (N) conditions was evaluated in the breeding program of the National Institute of Agricultural Technology (INTA). Focusing on genotypes R453, which exhibits early Instituto de Agrobiotecnología y Biología Molecular (IABIMO)- IB Centro de Investigación en Cs. Veterinarias y Agronómicas (CICVyA)-CNIA INTA-CONICET Dr. N. Repetto y De Los Reseros s/n B1686IGC – Hurlingham Provincia de Buenos Aires, República Argentina Teléfonos: 54+011 3754-8400 “2024-Año de la defensa de la vida, la libertad y la propiedad” senescence, and B481-6, characterized by delayed senescence, under contrasting nitrogen availability but identical ontogenetic cycle length, this study seeks to elucidate their yield stability under control conditions and varying N supply. This research explores the impact of nitrogen fertilization on sunflower yield and grain quality, with emphasis on soil supply, genotypic traits and water availability throughout the crop cycle. In particular, nitrogen is critical in the processes of growth, development, leaf senescence and yield generation. Crop yield is directly influenced by the duration of leaf area between flowering and physiological maturity, with nitrogen use efficiency (NUE) being a critical parameter reflecting the crop's ability to assimilate and utilize nitrogen from edaphic sources. The results from this project will contribute to elucidate the pathways involved in the differential triggering and progression of the senescence process in sunflower and its stability under different N nutrition profiles, allowing the detection of biomarkers useful for genetic improvement of this edible oil crop of great importance for Argentine production. In addition, the study approaches the "stay green" genotype condition as a strategy to mitigate premature senescence, potentially extending the period of photosynthetically active radiation uptake, which affects several yield-related parameters. Instituto de Agrobiotecnología y Biología Molecular (IABIMO)- IB Centro de Investigación en Cs. Veterinarias y Agronómicas (CICVyA)-CNI.

Index Terms: ecophysiology, nitrogen fertilization, transcriptomics, metabolomics, photosynthetic activity.

Using Genome Foundation Models for alignment-free bacterial taxonomy classification

Mike Leske[1], Bruno Gabriel Nascimento Andrade[1], Haithem Afli[2]

[1] Department of Computer Sciences, Munster Technological University
[2] Adapt Centre, Munster Technological University, Cork, Ireland
Corresponding author: mike.leske@gmail.com

The genome is an advanced data storage system consisting of long sequences of nucleotide base pairs ('a' = Adenine, 'c' = Cytosine, 'g' = Guanine, 't' = Thymine) that completely encode the instructions to rebuild and regulate the cell and, by extension, the organism itself. The genome is also a language in itself, being the nucleotides letters, genes the words, and functional pathways the phrases, with its own syntactic rules defining how genes should interact with each other. Over the past few years, Transformer architectures have evolved to become the de facto foundational processing engines for learning linguistic patterns in massive text corpora (e.g. GPT-4, Gemini, Claude). These foundation models further provide a computational baseline supporting a plethora of specific tasks, like text generation, classification, translation, summarization, and many more. As the architecture of these models improved to process sequences of gradually increasing length, a new field of Genome Foundation Models has emerged in which Transformers and its potential successor technologies are trained on massive genome datasets. Typically, these models intend to classify segments of DNA sequences according to a certain classification task, like enhancer detection or splice site detection, or to generate new sequences. However, some problem statements can be better addressed by extracting the hidden representation of sequences after they have been processed by the Genome Foundation Model. These hidden representations are commonly known as "embeddings" and are stored as a numerical representation (vector) with 512 to 4096 dimensions depending on the Genome Foundation Model architecture. Calculating the inner product between two embeddings provides a similarity score (e.g. cosine similarity) to evaluate the likeness of two input sequences. Using sequence similarity scores can be an efficient and computationally lightweight solution where sequences must be classified into hundreds to thousands of classes as is required for the taxonomy assignment of sequences down to the family, genus, or species level. Here, we present a framework for an alignment-free bacterial taxonomy classification using Genome Foundation Model embeddings of bacterial 16S gene sequences.

Index Terms: bacteria, 16S, genome foundation models, dna, alignment, taxonomy, bioinformatics.

Profile of hiv-infected women of child-bearing and prevention of vertical transmissions of hiv/aids in sub-saharan africa: a machine learning approach

Olusola Afuwape[1], Oluwatosin Akinsola[1], Patricia Akintan[2], Ifeanyi Nwoke[1], Folasade Tolulope Ogunsola[2] and Lifang Hou[3]

[1] University of Lagos, Lagos, Nigeria

[2] College of Medicine, University of Lagos, Lagos, Nigeria

[3] Northwestern University, Chicago, USA

Corresponding author: olusolaxx@gmail.com

It has been documented that 90% of pediatric HIV transmission is from mother-to-child. Despite the milestone achievements of Prevention of Mother-to-child Transmission (PMTCT) of HIV program, Nigeria currently has the highest burden of mother-to-child transmission globally. This may be attributed to poor coverage and dearth of novel techniques to halt the spread of HIV/AIDS while mother to child transmission (MTCT) has remains a significant contributor to the pandemic disease. This study utilized an array of different top features machine learning techniques to investigate and predict HIV/AIDS transmission from mother-to-child in order to understand better the contributing factors to the pathogenesis of MTCT of the disease. An analytical retrospective cohort study of HIV-infected expectant mothers and their neonates following delivery and also receiving care in the Anti-retroviral clinic of Lagos University Teaching Hospital, Nigeria. The study was conducted on 600 HIV-infected women in July 2023. The variables for each patient were initially captured in the National PMTCT maternal cohort register for the mothers and National PMTCT child for the infants. Variables captured include date of enrollment, age, point of entry whether postpartum or gravid, viral load, family planning methods, infant testing using two polymerase chain reaction (PCR) tests either at birth and/or 6 weeks, and a second test 6 months later or a rapid test done at 9 months due to changing program protocol over the years, mother anti-retroviral treatment, infant anti-retroviral treatment. Four machine learning (ML) techniques were engaged using R statistical software. The four ML algorithms used were: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF) and Naïve Bayes (NB). The data for the machine learning model was split into training and testing dataset in ratio 80:20. The numerical features were also standardized while hyper-parameter tuning was carried out on the machine learning algorithms. The mean age of women was 37.4±6.1years, mean infant gestational age at birth was 38.1±0.5 weeks, mean infant weight at birth was 3.0±0.6kg. A larger percentage of the women have secondary education (61.2%) while close to half of them were married (47.5%). Infant anti-retroviral therapy ($\chi2=41.367$, $p<0.001$**) and type of delivery ($\chi2=5.304$, $p=0.021$**) were found to be significantly associated with HIV outcome of infants of the top 17 features (variables) considered for both mothers and infants in the analysis, The RF model gave the best evaluation metrics for the prediction of HIV/AIDS in mother-to-child transmission. with: accuracy (81.8%), precision (83.1%), recall (90.2%) and F1-score (74.9%), roc-auc (88.7%). ML techniques can help with early HIV testing in expectant women, individualize treatment regimens, and increase the precision of transmission risk evaluations. To lower transmission rates and improve mother and child health, healthcare providers can utilize machine learning innovations to assist them make well-informed decisions and allocate scare resources efficiently in Nigeria and other sub-Saharan African countries.

Index Terms: HIV-Infection, Vertical Transmission, Machine Learning, Random Forest, R programming.

Transcriptomics and metabolomics profiling in chickens exposed to concurrent parasitic infections

Oyekunle John Oladosu[1], Banny Silva Barbosa Correia[2], Hanne Christine Bertram[2], Henry Reyer[3], Rosemarie Weikard[3], Christa Kühn[4], Cornelia C. Metges[1], Gürbüz Daş[1]

[1] Research Institute for Farm Animal Biology (FBN), Institute of Nutritional Physiology 'Oskar Kellner', Dummerstorf, Germany

[2] Aarhus University, Department of Food Science, Aarhus, Denmark.

[3] Research Institute for Farm Animal Biology (FBN), Institute of Genome Biology, Dummerstorf Germany.

[4] Friedrich-Loeffler-Institute, Insel Riems, Greifswald, Germany.

Corresponding author: oladosu@fbn-dummerstorf.de

Gastro-intestinal ascarid infections with *Ascaridia galli* and *Heterakis gallinarum* and associated histomonosis infection pose a significant threat to poultry production by impairing feed intake, nutrient absorption, and weight gain (Stehr et al., 2019). While the liver plays a central role in metabolism, its functions may switch to immunerelated pathways when challenged by these infections. With the major advancement in multi-omics techniques, we can understand parasite interaction at the system biology level—an essential prerequisite for making progress towards sustainable agriculture. Here, using datasets from two -omics experiments, we provide the first mechanistic insight into the alteration of the chicken transcriptome and metabolome profiles due to nematode infections. The first dataset was obtained from 1H-NMR metabolomics experiment designed to explore the variation in the metabolite profiles of the liver and plasma in laying hens experimentally infected with a mix of *A. galli* and *H. gallinarum* and necropsied at 2 weeks post-infection (wpi 2). In the second dataset, RNA-Sequencing was carried out to investigate gene expression in the livers of three genotypes of infected broilers with extremely different growth rates.There were significant ($P<0.05$) alterations in the plasma metabolites. The concentration of all significantly different metabolites was upregulated in the infected chickens. However, no significant differences were found in the liver metabolites in both the infected and non-infected groups. Pathway analysis showed that infection induced the strongest significant changes in amino-acids metabolism pathways. Whereas in the RNA-Seq experiment, a total of 1927 differentially expressed genes (i.e. 18.1% of annotated genes) was identified due to the infection effect ($P < 0.05$). Gene ontology of upregulated genes revealed increased transcriptomic activity related to immune responses in two fast growing broilers. On the other hand, downregulated genes were involved in metabolic pathways, including amino acid and glucose metabolism.Our data collectively suggest that heightened metabolic responses represents the high metabolic cost of defence against parasitic infection while the transcriptome changes further revealed that the liver of infected chicken is more occupied with immune related activities than its primary metabolic functions indicating a switch in functions during concurrent parasitic infections.

Index Terms: Ascaridia galli, Heterakis gallinarum, Histomonas meleagridis, immune response, metabolic response, host-parasite interaction, omics.

Funding: This project was funded by EU-Horizon 2020, grant agreement No 955374. NMR data were generated at FOODHAY, AU.

A Metabological Approach Of Bovine Liver As A Search Of Biomarkers To Mitigate Methane Emission

Pâmela Thays da S. Baima[1*], Daniel R. Cardoso[1]

[1] Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP 13566-590 Brazil.
*Corresponding author: baima.pamela@usp.br.

Enteric methane emissions constitute a significant environmental concern, leading Brazil to commit at the 26th Conference of the Parties (COP26) in 2021 to reduce methane emissions by 30% by 2030. As the livestock sector is one of the main contributors to methane emissions, mainly through the process of enteric fermentation of ruminants, it is necessary to use strategies to achieve the goals defined by COP 26. Among the various methods, the manipulation of diet stands out. In this way, it is important to study how modulations in bovine diets cause physiological changes in the organisms of ruminants. Since the liver plays a key role in regulating energy balance and general metabolic physiology, it is essential to investigate the biomarkers associated with the metabolism of the bovine liver. In view of this, this study focuses on the evaluation of polar and apolar metabolites of the Nelore bovine liver to understand how the various nutritional interventions can contribute to the mitigation of methane. A total of 52 liver samples from castrated male bovine animals, Bos indicus (Nelore), subjected to various diets, including conventional and by-products-based diets, were analyzed. Extraction procedures have been employed for polar and apolar compounds, and analytical techniques such as nuclear magnetic resonance spectroscopy (1H NMR) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS) have been applied. The data from the polar compounds were collected by 1H NMR and processed using the Chenomx software, while the data from the apolar components were gathered by MALDI-MS and handled in R using the MALDIquanti package. The lipids were annotated using the CEU 3.0 mass mediator tool. Multivariate and univariate analyses were performed using the software MetaboAnalyst 5.0. The study so far has not found biomarkers, but the analyses carried out have allowed the identification and quantification of metabolites in samples of bovine liver and demonstrated significant differences between the food diets submitted, conventional and by-product.

Index Terms: bovine, methane, liver, diets.

Funding Agency: The São Paulo Research Foundation (FAPESP).

Phylogenetic Analysis of Balami and West African Dwarf Sheep Populations Based on mtDNA Sequences

Popoola M.A.[1,2], Oseni S.O.[3], Machuka E[4], Githae D[4], Tarekegn G.M.[5], Stomeo F[4]

[1] Federal College of Animal Health and Production Technology, Ibadan, Nigeria

[2] National Dairy Research Institute, Karnal, Haryana, India

[3] Obafemi Awolowo University, Ile-Ife, Osun state, Nigeria

[4] International Livestock Research Institute-Biosciences east and central Africa, Nairobi, Kenya

[5] Swedish University of Agricultural Sciences, Uppsala, Sweden

Understanding the geographical distribution of livestock species is crucial for their utilization, genetic enhancement, and conservation. It is therefore necessary to explore the genetic relationships and diversities between livestock breeds. In this study, two Nigerian indigenous sheep populations were evaluated based on their genetic diversities and phylogenetic structure. These populations are Balami and West African Dwarf (WAD) sheep, which are different in their geographical distribution and physical attributes. Sixty hair samples were collected from each of these two breeds for genetic analysis. The study revealed twenty-three haplotypes, with no shared haplotypes between Balami and WAD sheep, indicating a clear separation between the two populations. There was a high haplotype diversity index between the two populations. Balami had a higher haplotype diversity index (0.96±0.04) than WAD (0.91 ± 0.06). Similarly, the nucleotide diversity index between the two populations differed with WAD recording a higher value of nucleotide diversity (0.01115 ± 0.0023) than Balami (0.00788 ± 0.00015). The number of polymorphic sites varied between the populations, with WAD having higher polymorphic sites than Balami. Non-significant Raggedness index value was obtained for Balami, which implies population expansion for the breed, unlike WAD sheep with a significant Raggedness index, which implies that the breed showed more than one independent subpopulation with high homozygosity and low heterozygosity (population subdivision). The analysis of molecular variation (AMOVA) for the two populations revealed that 79.1% of total genetic variation was presented in Balami and WAD, which was explained by genetic differences between individuals within populations. Only 20.9% of the variation was found among populations compared to 79.1% of the variation obtained within populations. The results of this study contribute to the knowledge of the existing genetic status of Nigerian indigenous sheep populations. Most of the variations were found within the population rather than between populations. These results revealed a signal of demographic or spatial expansion in these indigenous Nigerian sheep breeds. These sheep breeds were genetically different due to their geographical distributions, with Balami located in the country's Northern regions while WAD is regarded as a Southern Nigerian sheep breed. The finding suggests a probable origin from the very first domestic sheep ancestors in Asia, which could have reached the country through various routes. Thus, successive migration and relative selection could explain the current mitochondrial DNA variability in the sheep populations. This possible indication of wide genetic variation between these two breeds could be exploited in genetic improvement programs for both breeds.

Index Terms: Breed, conservation, diversities, genetic, utilization.

Funding Agency**:** This project was supported by the BecA- ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. The ABCF program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA- CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID); and the Swedish International Development Cooperation Agency (Sida).

Isolation and characterisation of multi-trait Plant-promoting bacteria from rhizosphere of plants growing in Arid Regions of Rajasthan

Prashant Puri[(1)] , Shiv Swaroop[(1#)]

[(1)] Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Ajmer, 305817, India.

Arid regions have tough climatic conditions, including low water availability, which make it difficult for plants to grow and survive. Rhizospheric bacteria play critical roles in increasing plant tolerance to drought-related stresses via different mechanistic mechanisms. This research intended to isolate and characterise rhizospheric bacteria associated with plants that thrive under drought-stress conditions in arid environments and to dig out grossly under-explored factors responsible for drought tolerance. The bacteria from drought-stressed plants within Rajasthan's arid region were isolated using standard microbiological techniques. The isolates were investigated for plant growth-promoting traits, such as indole-3-acetic acid production, ACC deaminase activity, siderophore production, HCN, and phosphate solubilisation. The results indicate that diverse bacterial isolates belong to PGPBs, and these strains were recognised as isolates belonging to the genus Bacillus and Cytobacillus based on 16S rRNA sequence similarities. Furthermore, many isolates were found positive for multiple plant growth-promoting traits; some of them were found positive for ACC deaminase and nitrogen fixation activity, which help plants to utilise N2 from the environment and regulate amino acid formation and the formation of nitrogenous compounds, suggesting their potential role in enhancing plant resilience to drought stress. The presence of acdS and nifH genes were confirmed by PCR amplification. Further studies are envisioned to explore the diversity and functional characterisation of rhizospheric bacteria and for using them for the development of sustainable agricultural practices to improve crop productivity in arid regions.

Index Terms: ACC deaminase, nitrogen fixation, indole-3-acetic acid, acdS, nifH, PGPBs.

Acknowledgements

PP is thankful to the Central University of Rajasthan for providing research fellowship. SS is thankful to DST-SERB for providing funding to complete the project.

Exploring the Impact of Environmental Perturbations: A Multiomic Approach

Rafaela Martins[1*], Roberto Carvalheiro[2], Lucia Galvão de Albuquerque[2], Rodrigo de Almeida Teixeira[3], Laila Talarico Dias[3]

[1] Student in the Graduate Program in Animal Science, Federal University of Paraná, Curitiba, Paraná, Brazil, and CNPq fellowship recipient.

[2] São Paulo State University (UNESP), School of Agricultural and Veterinarian Sciences, Department of Animal Science, Jaboticabal, SP 14884-900, Brazil.

[3] Associate Professor at the Department of Animal Science, Graduate Program in Animal Science, UFPR, Curitiba, PR, Brazil.

*Corresponding author: rafaelamartins.zootecnia@gmail.com

Environmental disturbances are significant changes in the external environment that can affect animal productivity, as observed in dairy cattle during the lactation curve. However, some animals are resilient to these environmental disturbances, being able to quickly return to their normal level of perfomance or not having their performance affected when exposed to such disturbances. Accordingly, resilience is a key trait in farm animals, as it determines their ability to cope with and adapt to adverse environmental changes. Nevertheless, resilience is a complex trait that is difficult to measure. One way to detect resilient animals is through transcriptome profiling, which can generate a list of differentially expressed genes (DEGs) for resilient and non-resilient animals. Therefore, this study aims to characterize DEGs in blood samples by analyzing RNA-Seq data in resilient Holstein cows and non-resilient Holstein cows. For this purpose, a database with the phenotype of daily milk production of the AMS system (automatic milking system) will be used, where deviations in milk production will be used to detect resilient or non-resilient animals. After the identification of these animals, transcriptome analysis will be performed by RNA-sequencing (RNA-seq). For this, RNA extraction will be performed with Trizol (Invitrogen). Libraries will be prepared using the TruSeq RNA Sample Prep v2 kit (Illumina), and sequencing will be performed on the Illumina HiSeq 2500. Reads will undergo quality control (Trimmomatic) and mapping (STAR). Gene expression analysis will be performed using the edgeR package in R, and those with FDR ≤ 0.05 will be considered differentially expressed. Functional annotation will be carried out using the clusterProfiler package in R, and the genetic network will be built in STRING. In the literature, Liu et al. (2020) when investigating heat resistance in Holstein cattle, identified 200 DEGs in resistant and non-resistant animals. Therefore, the expected results include identifying which genes are being expressed at higher or lower levels in more resilient animals compared to less resilient animals under different environmental disturbances.

Index Terms: automatic milking system, dairy cattle, differentially expressed genes, robust genotype and resilience.

Funding Agency: This research did not receive any specific funding.

Structural variations identification in Nelore using whole-genome sequencing data and evaluation of their influence through cell culture

Renato Duarte de Araújo[(1*)], Luiz Lehmann Coutinho[(1)]

[(1)] Animal Science Department, University of São Paulo (USP)/Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, São Paulo, Brazil.

*Corresponding author: renato.d.araujo@usp.br

Brazil is the largest beef exporter in the world, with Nelore (*Bos indicus*) cattle representing the majority of the animals. While single nucleotide polymorphisms (SNPs) have revolutionized cattle genetic evaluation, structural variants (SVs) are gaining recognition for their association with production traits, particularly in Nelore cattle. SVs are defined as genomic sequence rearrangements exceeding 50 base pairs, including deletions, duplications, translocations, and inversions. Due to their size and location, SVs are hypothesized to exert substantial influence on gene expression, potentially exceeding the effects of SNPs. Despite their biological significance, SV identification remains challenging, primarily due to their complexity. Conventional SV detection methods rely on high density SNP panels or hand-crafted features derived from read alignment patterns, often resulting in high rates of false negatives and positives, thereby complicating downstream analyses. A promising alternative lies in the utilization of deep learning strategies, capable of discerning intricate patterns directly from data, offering a more accurate and adaptable SV detection pipeline. However, the efficacy of deep learning approaches depends on the training dataset, which is predominantly derived from human genomes. Regardless of the progress in the field, SV calling tools still display low Precision and recall. Therefore, persist the necessity for rigorous validation procedures to ascertain SV detection and the influence of SVs on gene expression and phenotypes. In this study, we propose the utilization of the Cue deep learning framework to identify SVs from short-read whole-genome sequencing data obtained from 86 Nelore animals, contrasted with results from traditional SV callers such as Manta and GRIDSS. This comparative analysis aims to elucidate the effectiveness of Cue within our experimental context. Furthermore, we will attempt to assess the putative impact of identified SVs on gene expression by correlating them with ATAC-seq peaks, expression quantitative trait loci (eQTLs), synteny breakpoints between Bos taurus and Bos indicus reference genomes, copy number variation-QTLs (CNV-QTLs) detected by genotyping panels, and other data available on the literature. This comprehensive approach seeks to create a catalog of SVs potentially influencing gene expression in Nelore. Subsequently, prime SV candidates will undergo further validation through PCR and cell culture experiments, aimed at corroborating our assumptions regarding SV's influence on gene expression. This integration, coupled with robust validation, promises to minimize the gap between scientific research and its practical integration into breeding programs, thereby enhancing efficacy and genetic improvement strategies, in addition to serving as a basis for future studies.

Index Terms: Validation of structural variations; Tissue culture; Gene editing.

Funding Agency: CNPq, FAPESP.

Genetic and metabolomics markers for fertility in dromedary camels

Seham Samir Soliman[1], Samir H. Haggag[2], Noha Said hussien[2], Abdel-Hamid Z. Abdel-Hamid[2], Ahmed Sabry Abdoon[1]

[1] Department of Animal Reproduction and Artificial Insemination, Veterinary Research Institute, National Research Centre (NRC), Dokki, Cairo 12622, Egypt.

[2] Therapeutic Chemistry Department, National Research Center, El-Tahrir St., Dokki, Cairo, Egypt .

Corresponding author: sehamsamir81@yahoo.com

The importance of camel breeding in Egypt is increasing day by day, especially in deserts and newly reclaimed areas. This is mainly due to its unique physiological properties, which help it survive and perform under these harsh conditions. However, it is known that one of the main obstacles in this industry is the very low reproductive efficiency of camels. Fertility in animals is affected by many factors, including hormonal, genetic, nutritional and environmental. To improve the reproductive efficiency of camels, focusing on aspects of hormonal measurement, biochemistry and genetics is useful for gaining a better understanding of camel reproduction and fertility. Therefore, the objective of this work is to improve the reproductive performance of the Arabian camel in Egypt using modern assisted reproductive technologies. We will use some of these techniques to deal with bull camels; they also deal with the reproductive efficiency of female camels. These studies include biology, genetics, blood biochemistry, hormones, histology, and metabolomics studies.

Index Terms: Dromedary camels, Reproduction, Gene expression, biochemical analysis, histological examination, metabolomics.

Funding: National Research Centre (NRC), project number 13050416.

Roots of Change: Assessing the Impact of Sorghum Breeding on the Root Microbiome

Sierra S. Raglin[1]

[1] University of Illinois at Urbana-Champaign

Sorghum bicolor is the fifth most cultivated crop, globally, and modern germplasm development has prioritized using S. bicolor as a second-generation bioenergy feedstock, which requires reduced ecosystem impacts when compared to fossil-fuel-based energy. Therefore, assessing the adverse ecosystem impacts of S. bicolor germplasm is a necessary step for its use as a bioenergy feedstock. The plant microbiome influences plant growth and health but also contributes to agroecosystem sustainability through biogeochemical cycling. Plants manipulate the microbiome associated with the roots, termed the rhizosphere, by altering carbon allocation patterns within the root system, via changes in root system architecture, morphology, and metabolomics, the latter of which drives hotspots of microbial activity within the rhizosphere. Priming of the microbiome by roots varies with species and genotype and can stimulate carbon sequestration or respiration. When compared to African and American landrace sorghum varieties, bioenergy sorghum varieties have drastically altered carbon allocation patterns due to breeding manipulations of their photoperiod sensitivity. Through these manipulations, bioenergy sorghum remains within the vegetative stage, accumulating biomass, and re-directing belowground carbon into vegetative biomass. Therefore, this research seeks to understand the impact of S. bicolor breeding history on the root metabolome and rhizosphere microbiome. This was done by comparing the root metabolome and microbiome of 26 Sorghum bicolor genotypes, including African and American landraces, and bioenergy hybrids. We assessed these plants associated microbial communities using 16S rRNA amplicon sequencing, and flash-frozen root tissue was analyzed with untargeted LC-MS metabolomics. Breeding history and genotype significantly influence microbiome composition (PERMANOVA, History x Genotype, R2 = 0.20, P = 0.001). Differential analyses identified major differences in copiotroph and oligotroph taxa through breeding. African landrace varieties were differentially abundant in slow-growing Acidobacteria, whereas fast-growing Bacteriodetes were enriched within the American and Bioenergy varieties, suggesting breeding-driven changes in carbon partitioning into the rhizosphere microbiome. Untargeted metabolomics data is currently being processed and analyzed. Ultimately, the aim is to understand how germplasm development can adversely influence the microbiome via changes in root chemistry, which is an important step for comparing the ecosystem impacts of Bioenergy feedstocks to classical fossil fuel-based energy.

Index Terms: bioenergy, rhizosphere, metabolome, microbiome, crop breeding.

Funding: Center for Advanced Bioenergy and Bioproduct Innovation (CABBI), Bioenergy Research Centers, United States Department of Energy.

Immunomodulatory Effects of Kefir Beverage on Intestinal Tissue Injury: Insights from a Murine Model

S. Dentice Maidana[1], J. Villena[1], H. Kitasawa[2]

[1] Laboratory of Immunobiotechnology, Reference Centre for Lactobacilli (CERELA-CONICET), San Miguel de Tucumán 4000, Argentina

[2] Food and Feed Immunology Group, Laboratory of Animal Food Function, Graduate School of Agricultural Science, Tohoku University, and Livestock Immunology Unit, International Education and Research Centre for Food and Agricultural Immunology (CFAI), Graduate School of Agricultural Science, Tohoku University, Sendai 980-8572, Japan

Corresponding author: stefi.dentice@gmail.com

Kefir, a fermented beverage containing probiotics, exhibits a symbiotic association with various microorganisms within kefir grains, offering a plethora of nutraceutical benefits. This study aimed to assess the immunomodulatory potential of immunobiotic strains within kefir microbiota in modulating the response of intraepithelial lymphocytes (IELs), using a murine model. Intraperitoneal administration of poly(I:C) induced inflammatory-mediated intestinal tissue damage, characterized by increased inflammatory cell infiltration and upregulation of pro-inflammatory mediators (TNF-α, IL-1β, IFN-γ, IL-15, IL-10, IL-6). Mice subjected to higher concentrations of kefir consumption exhibited drastically reduced cellular damages (LDH, GOT, GPT) post-poly(I:C) exposure. Pre-treatment with varying concentrations of kefir prior to TLR3 activation significantly attenuated levels of TNF-α, IL-15, and IL-6, while elevating serum and intestinal IL-10 levels. Notably, kefir-treated mice demonstrated lower levels of intestinal damage compared to controls, indicating the immunomodulatory capacities of lactobacilli in reducing intestinal tissue damage. In conclusion, this study elucidates the reduction of intestinal tissue injury through kefir beverage consumption, mediated by the modulation of intraepithelial lymphocytes response. These findings contribute to advancing our understanding of the cellular mechanisms underlying the immunomodulatory effects of immunobiotic strains, highlighting the therapeutic potential of kefir in mitigating intestinal inflammation.

Index Terms: gut microbiota, kefir microbiota, immunomodulation.

Funding: PICT-2018-03264 J.V.

Machine Learning Identification of Low Methane Emitting Dairy Cattle for Selective Breeding

Stephen R Ross[1, 2], Huiru J Zheng[1], Haiying Wang[1], Masoud Shirali[2, 3], Tianhai Yan[2]

[1] School of Computing, Ulster University, Belfast, Northern Ireland
[2] Agri Food and Biosciences Institute, Hillsborough, Northern Ireland
[3] School of Biological Sciences, Queens University, Belfast, Northern Ireland
Corresponding author: h.zheng@ulster.ac.uk

Methane (CH_4) emissions produced by dairy cattle (DC) are the result of an energy partition inefficiency within the animal, which individually lowers production levels, and collectively, contributes towards agricultural greenhouse gas (GHG) emissions. Up to 40% of CH_4 emission variation between individual DC can be explained by genes. Therefore, the selective breeding of low CH_4 emitting DC has been proposed as a novel mitigation strategy to reduce the GHG emissions produced on farms, as future generations inherit the low CH_4 emission characteristics of their more efficient ancestors, resulting in improved production efficiency. However, the direct recording of DC CH_4 emissions, to identify suitable candidates for selective breeding, requires complex and expensive equipment, preventing the strategy from being implemented on commercial farms. Fortunately, agriculture is currently experiencing a smart farming revolution, and there now exists a wealth of animal proxy information, which can be used to predict DC CH_4 emissions, instead of directly measuring them, replacing the need for exclusive recording technologies. Yet rather than determined by any individual trait, DC CH_4 emissions are the result of an intricate dance between multiple systems, including Biological, Environmental and Genetic (BEG) factors. Therefore, if the full potential of a DC CH_4 emission predictive modelling approach is to be realised, one must incorporate each BEG component during formation. Yet the addition of each new layer behind DC CH_4 emission production, brings with it an entirely new dominion that has its own complicated relationship with the subsequent layers. The sophistication of Machine Learning (ML), however, has the ability to appreciate the multidimensionality behind CH_4 production, by enabling the facilitation of the diverse BEG factors behind it, and identifying the underlying structures and complex relationships between them. The multitude of techniques ML models introduce to the challenge, whether as an alternative vehicle for direct prediction via diverse algorithms, a concierge for heterogeneous dependency patterns through hybridisation or stacking, or as a trampoline for data complexity through imputation, may perhaps hold the key to overcoming the limitations currently faced and surpassing the accuracies currently achieved, in the prediction of DC CH_4 emissions. It is within this research project that we will attempt to develop a hybrid ML approach, which can incorporate BEG data, in the prediction of DC CH_4 emissions. We hope to uncover new insights behind CH_4 production through the approach developed, as well as provide a deployable solution for commercial farms to assist in the identification of optimal candidates for selective breeding.

Index Terms: methane, dairy cattle, production efficiency, greenhouse gas emissions, selective breeding, prediction, multivariate, machine learning, hybridisation, stacking, imputation.

Funding Agency: This study is a collaboration between Ulster University and the Agri-Food and Biosciences Institute (AFBI) and is funded by the Northern Irish Executive Department of Agriculture, Environment and Rural Affairs (DAERA) through the Northern Ireland Farm Animal Biobank (NIFAB) (Project No. 21/5/01).

Integration of omics methodologies in the investigation of White Striping in broiler chickens

Suelen Fernandes Padilha[1*], Adriana Mércia Guaratini Ibelli[2], Jane de Oliveira Peixoto[3], Maurício Egídio Cantão[3], Laila Talarico Dias[1], Mônica Corrêa Ledur[3]

[1] Programa de Pós-graduação em Zootecnia, Universidade Federal do Paraná (UFPR), Curitiba, PR.

[2] Embrapa Pecuária Sudeste, São Carlos, SP.

[3] Embrapa Suínos e Aves, Concórdia, SC.

*Corresponding author: suelenfpadilha7@gmail.com.

White Striping (WS) is characterized by the appearance of white streaks composed of adipose and connective tissue parallel to the muscle fibers of the breast of broilers. This condition mainly affects birds with fast-growth and high breast yield, reducing meat quality and consumer acceptability, causing great economic losses. Although the occurrence of WS is high, its etiology remains poorly understood. Furthermore, it is known that epigenetic mechanisms do not alter the DNA sequence, but influence gene expression. Therefore, independent omics studies do not provide a comprehensive view of the etiology of WS. To date, there are no studies relating methylome with WS, nor the integration with other omics. Thus, the aim of this study is to analyze the pectoral muscle transcriptome integrated with its epigenome to better elucidate the etiology and mechanisms regulating the expression of WS. For that, 16 samples (8 controls and 8 affected) were collected from the pectoral major muscle of male chickens with 35 days of age, belonging to a paternal broiler line TT, developed by Embrapa Swine and Poultry (Concórdia, SC, Brazil). For the RNA-Seq analysis, RNA was extracted with Trizol (Invitrogen), libraries were prepared with the TruSeq RNA Sample Prep v2 kit (Illumina) and sequenced in the Illumina HiSeq 2500 (paired-end, 100 pb). Reads were submitted to quality control (Trimmomatic) and mapping (STAR). Differential expression analysis was performed using the edgeR package from R and genes with FDR ≤ 0.05 were considered differentially expressed (DE). Functional annotation will be performed with the clusterProfiler package from R, and the gene network will be constructed with STRING. We found 30 DE genes between normal and WS-affected broilers. Among them, 10 will be chosen for validation by qPCR. For the methylation analysis, DNA extraction will be performed using the Dneasy Blood and Tissue kit (Qiagen). Libraries will be prepared with the EZ DNA Methylation-Lightning (Zymo Research) and Nextera® (Illumina) kits and sequenced in the Illumina NovaSeq X (paired-end, 150 pb). Reads will be submitted to quality control (TrimGalore) and mapping (Bismark). The identification of differentially methylated regions (DMR) and differentially methylated genes (DMGs) will be carried out using the program MethylKit from R. The DE genes and DMGs will be integrated for functional annotation using the DAVID and Webgestalt, and the gene network will be carried out with Cytoscape software.

Index Terms: methylome, pectoral myopathy, RNAseq, transcriptome.

Funding Agency: This study was financed by National Council for Scientific and Technological Development (CNPq) [Grant no.: 423159/2021-6] and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001. S.F.P. is supported by a CAPES scholarship.

Assessing the effects of SNPs on mastitis resistance and milk production traits in selected genes through SNP-CHIP array in *Staphylococcus aureus* positive dairy cattle in Pakistan

Tahir Usman[(1*)], Mustafa Kamal[(1,2)], Shehryar Khan[(1,3)]

[(1)] College of Veterinary Sciences and Animal Husbandry, Abdul Wali Khan University Mardan, 23200 Mardan, Pakistan

[(2)] Department of Zoology, Abdul Wali Khan University Mardan, 23200 Mardan, Pakistan

[(3)] Department of Biotechnology, Abdul Wali Khan University Mardan, 23200 Mardan, Pakistan

*Corresponding author: tahirusman@awkum.edu.pk

Mastitis is an inflammation of the mammary gland of dairy animals. The present study was designed to investigate the effects of single nucleotide polymorphisms (SNPs) genotyped on Chinese Cow SNPs Chip-I (CCSC-I) in the *JAK2* and *TRAPPC9* gene on the milk production and mastitis related traits in dairy cattle in which milk samples were S. aureus positive. Blood and milk samples were collected from 707 lactating dairy cattle of five breeds, i.e. Holstein Friesian (HF), Achai (A), Red Sindhi (RS), Cholistani (C) and Sahiwal (S) maintained at wellestablished dairy farms across Pakistan. Out of 707 cattle, 202 milk samples were detected positive for *S. Aureus*. Three SNPs, i.e. SNP 1 (T>C, rs210148032), SNP 2 (T>C, rs135128681), and SNP 3 (T>C, rs377935625) in *JAK2* and 4 SNPs i.e. SNP 4 (T>C, rs109952887), SNP 5 (T>G, rs133738259), SNP 6 (G>A, rs111018678) and SNP 7 (G>A, rs55617160) in *TRAPPC9* were genotyped through CCSC-I in a population of 202 dairy cattle positive for S. aureus. The results were analyzed using the general linear model in SAS 9.4. Among the seven SNPs, SNP 2, SNP 3 and SNP 4 did not obey Hardy-Weinberg equilibrium ($P<0.01$). The association analysis of SNPs in the overall studied population showed significant association ($P<0.05$) of SNP 2 with Annual Milk Yield (AMY) and Somatic Cell Score (SCS). In SNP 3, the genotype CC showed significantly lower ($P<0.05$) Fat percentage (FP) than CT genotype. SNP 4 showed significant association ($P<0.05$) with AMY and Lactose percentage (LP). SNP 5 was significantly associated with the AMY, LP and Protein percentage (PP). SNP 6 was significantly associated with AMY, Somatic Cell Count (SCC) and SCS, Whereas SNP 7 was found significantly associated with AMY and SCS. The breed-wise association analysis depicted that these SNPs were significantly associated with mastitis related and production traits in all of the selected cattle breeds. Altogether, our findings suggested that the CCSC-I is a useful tool to assess SNPs association with phenotypic traits. Moreover, *JAK2* and *TRAPPC9* genes could be useful genetic markers in selection for milk production improvement and mastitis resistance phenotypic traits in dairy cattle positive for *S. aureus*.

Index Terms: chinese cow SNPs Chip-I, Single Nucleotide Polymorphisms, mastitis, JAK2, TRAPPC9, S. aureus.

Funding: This project was financially supported by Pakistan Science Foundation and Nature Science Foundation of China.

Maternal slow-release nitrogen diets during late gestation optimize the energy metabolism in calves' skeletal muscle

Thaís Correia Costa[1*], Diana Carolina Cediel Devia[2], Karolina Batista Nascimento[2], Marcio de Souza Duarte[1], Mateus Pies Gionbelli[2]

[1] Department of Animal Science, Universidade Federal de Viçosa, MG, 36570-900, Brazil

[2] Department of Animal Science, Universidade Federal de Lavras, MG, 37200-900, Brazil

*Corresponding author: thaiscosta.phd@gmail.com

The current study aimed to determine the enriched biological process, through proteomic and transcriptome data, associated with maternal slow-release nitrogen diets received during late gestation on the skeletal muscle of the offspring. At day 180 to day 268 of gestation a total of 16 pregnant Brahman cows, were assigned into Control treatment (CON; n = 7), where cows were fed ad libitum a low crude protein basal diet plus mineral mixture; or Slow-released Nitrogen (SRN, n = 9) based diet, where cows were fed a basal diet plus a slow-release N supplement. Muscle biopsy was performed at day 45 of age in calves and used to perform further RNA sequencing (RNA-seq) and proteomic (HPLC-MS/MS) analyses. The quality control of raw data from RNAseq was assessed through FASTQc software and the low-quality reads were trimmed by Trimmomatic. The reads were mapped against the Bos taurus reference genome using STAR. To determine the differentially expressed (DE) genes the limma package was used in the R environment. In addition, Cuffdiff tool was used to explore DE isoforms. The DE genes and isoforms were deemed significant and adjusted p-value ≤ 0.05. For the proteomic analysis, raw data was processed in MaxQuant using the Bos taurus reference proteome. Statistical analysis was performed by MSqRob package in R environment. The network analyses were performed to identify the biological processes of the proteins identified as differentially abundant proteins (DAPs) and were deemed significant when the adjusted p-value ≤ 0.05. Although the experimental treatment did not show effects on transcripts abundance, there were differences at protein level. Enriched (adjusted p-value ≤ 0.05) biological processes from the exclusive proteins identified in calves skeletal muscle from SRN group are related to Acetyl-CoA biosynthetic process from pyruvate , mitochondrial respiratory chain complex assembly, regulation of calcium ion transmembrane transporter, generation of precursor metabolites and energy, nitrogen compound transport , cellular respiration, protein folding, regulation of muscle system process tricarboxylic acid cycle, and isocitrate metabolic process. Protein-protein interaction network assessed in the DAPs revealed 4 main enriched biological processes, including ATP metabolic process, glucose metabolic, tricarboxylic acid cycle, and sarcomere organization. In general, the results of the current study highlight the positive impact of slow-release nitrogen enriched diets during late gestation on the energy metabolism in calves' skeletal muscle.

Index Terms: bovine, biological processes, energy metabolism, maternal nutrition, skeletal muscle, slow-release nitrogen nutrition.

Funding statement: This study was funded by Roullier Group received through the Rollier Innovation Awards 2020 in Animal Nutrition.

Deciphering microbial dynamics to boost biogas production efficiency

Thiago R. Bella[1*], Lucas M. de Carvalho[1], Marcelo F. Carazzolle[1]

[1] Departamento de Genética, Evolução, Microbiologia e Imunologia - Instituto de Biologia/UNICAMP.
*Corresponding author: t157414@dac.unicamp.br

The intensifying climate crisis, marked by a surge in greenhouse gas emissions and a global temperature rise surpassing 1.1 °C above pre-industrial levels, necessitates urgent mitigation efforts. Biogas production through anaerobic digestion emerges as a promising strategy to avert catastrophic outcomes such as heatwaves, droughts, and wildfires. The biogas industry, employing anaerobic digestion processes, holds the potential to play a crucial role in capturing organic wastes and reducing methane emissions, contributing to a targeted 10% reduction in global greenhouse gas emissions by 2030. However, optimizing anaerobic digestion processes faces challenges related to reactor pH control, nutrient imbalance, inhibitor presence, and the complex composition of microbial populations. Addressing these challenges, our study focuses on identifying key microorganisms and genes within microbial populations associated with efficient biogas production in an industrial-scale single-stage anaerobic digester. Metagenomic samples were collected from originally mesophilic full-scale anaerobic digesters after 80, 177, 218, 281, and 462 days of operation. Genomic data undergoes processing, including read-based analysis, assembly, binning, and refinement using state-of-the-art tools. Functional annotation of genes is performed to gain insights into their potential roles in the biogas production process. The study comprises two anaerobic digesters—one at mesophilic conditions and the other subjected to sub-mesophilic conditions, revealing the impact of temperature variation on microbial communities and genetic composition. The research culminates in a comprehensive database integrating microorganism abundance, annotated genes, reactor parameters, and biogas yield. Leveraging machine learning, specifically the CatBoost algorithm, we trained a model to predict biogas production based on these variables. The SHAP (SHapley Additive exPlanations) package is essential in the interpretation of correlations, elucidating the importance of each variable in predicting biogas production. Moreover, metabolic simulations of microbial communities will be carried out to identify the behavior of the community in relation to the metabolites shared by it. In conclusion, our findings deepen the understanding of microbial dynamics in anaerobic digestion and provide practical insights for optimizing biogas production processes. By identifying key microorganisms and genes correlated with efficient biogas production, this research sets the stage for targeted interventions, potentially enhancing the sustainability and effectiveness of anaerobic digestion in mitigating the impacts of the climate crisis.

Index Terms: metagenomic analysis, metagenome-assembled genomes, anaerobic digester, methane production, methanogenic archaea.

Funding Agency: Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP); Shell Petróleo Brasileiro LTDA; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Dynamics of changes in cecal and fecal microbial communities of swine fed a high-fat diet

Thiago Sugizaki dos Santos[1*], Aline Silva Mello Cesar[1, 2]

[1] Faculty of Animal Science and Food Engineering, University of Sao Paulo, Pirassununga, SP, Brazil.

[2] "Luiz de Queiroz" College of Agriculture, University of Sao Paulo, Piracicaba, SP, Brazil.

*Corresponding author: thiago.sugizaki.santos@usp.br

The pursuit of foods characterized by heightened nutritional density and reduced levels of saturated fatty acids necessitates strategic adaptations within the animal husbandry sector to align with evolving dietary guidelines. Such adaptations are geared towards facilitating the production of animal-derived products endowed with elevated biological value and enhanced health attributes. Of significance is the modulation of intestinal microbiota, a process influenced by the fatty acid composition of the diet, which plays a pivotal role in both disease prevention and therapeutic interventions. This study aims to evaluate the microbial community diversity within the cecum and feces of pigs subjected to diets enriched with canola, soy, and fish oils. Employing a randomized complete block design, each treatment group comprised six replicates, with four animals housed per enclosure. Treatments included a basal diet supplemented with 3% soybean oil, canola oil, fish oil, or no additional oil. The experimental phase spanned 98 days, during which animals were provided ad libitum access to feed and water. Slaughter occurred at 169 days of age, with an average final live weight of 133.9 ± 9.4 kg. Cecal (n = 61) and fecal (n = 71) specimens were promptly collected, flash-frozen in liquid nitrogen, and stored at -80 °C until further processing. Bacterial genomic DNA extraction ensued, followed by amplification and sequencing of the V3-V4 region of bacterial 16S ribosomal RNA (rRNA) genes using the Illumina NextSeq 1000 platform. The relative abundance of taxonomic groups is presently under evaluation utilizing the QIIME2 platform, with subsequent PERMANOVA tests earmarked to compare beta diversity across experimental groups. This research endeavors to elucidate alterations in meat fatty acid profiles and their consequential impacts on swine intestinal and physiological health, under the influence of various dietary energy sources. Additionally, this study is integrally linked to a project authorized by the São Paulo Research Foundation (FAPESP, process no. 2017/25180-2 - Young Researcher / JP), aimed at delineating diet-induced effects on immune response-associated biological processes through RNA sequencing (RNA-Seq) data analysis.

Index Terms: fatty acids, microbiome, nutritional modulation, swine.

Funding Agency: Process number 2021/01664-6, Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (Capes) - Finance Code 001.

Exploring allele-specific regulation of multi-tissue H3K27ac in chickens

Wenjing Yang[1], Jennifer Jessica Bruscadin[1,2], Claire P Prowse-Wilkins[1], Dailu Guan[1], Huaijun Zhou[1]

[1] Department of Animal Science, University of California, Davis, CA 95616, USA
[2] Center of Biological and Health Sciences, Federal University of São Carlos, São Carlos, SP, Brazil
Corresponding author: wenyang@ucdavis.edu

Genome-wide association studies (GWAS) have identified numerous agriculturally important trait-associated SNPs located in non-coding regions. These SNPs within regulatory regions like promoters and enhancers can potentially influence transcription factor binding sites, thus modulating gene expression. Due to the potential of analyzing regulatory mechanisms affecting each haplotype, characterizing allele specific (AS) epigenetic regulation can bridge the gap between genotypes, transcription level and phenotypic variation. H3K27ac is an epigenomic mark that exhibit a critical role in active enhancers, helping to understand the effect of candidate SNPs in functional regions. In this study, we profiled the AS H3K27ac binding in 19 distinct chicken tissues from two individuals, a F1 cross both derived from genetically distinct highly inbred lines, line 6 and line 7. We mapped 1.3 billion reads derived from whole genome sequences to the GRCg7b reference genome using BWA software and identified 5,967,787 SNVs by the GATK. Functional annotation revealed that the majority of the identified SNVs were located in non-coding regions, including intron (53.33%) and intergenic regions (19.85%). Exonic regions only contained 1.60% of the SNPs, including 61,873 synonymous and 33,109 nonsynonymous SNVs. By comparing to the chicken GTEx reference panel, we only found 37.90% of overlaps, indicating distinct genetic background of birds used in our study. We then concentrated on the set of 2,631,957 heterozygous SNPs (44.17% of the total) to identify AS imbalance events via the binomial test. This analysis has identified an average of 2,950 AS imbalance events across 19 different tissues (ranging from 40 in liver to 8,963 in cecum). The next steps are to extend this AS analysis to other histone marks (e.g. H3K4me1, H3K4me3, H3K27me3) and to explore the AS SNPs functional regulation, particularly concerning the enhancer activity, and the integration with genetic variants within regulatory regions associated with agriculturally important traits in poultry.

Index Terms: Allele-specific; epigenetics; chicken; gene regulation.

Comparative genomics of the Purple Land Crab (*Gecarcoidea lalandii*) sheds light on the adaptations to terrestrial environments

Ziwei Wu[(1)], Rongfeng Cui[*], and Ka Yan Ma[(1,*)]

[(1)] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), State Key Laboratory of Biocontrol, School of Ecology, Shenzhen Campus of Sun Yat-sen University, Shenzhen, 518000, China;

*Corresponding author: cuirf@mail.sysu.edu.cn (R.C.); majx26@mail.sysu.edu.cn (K.Y.M.)

The shift from marine to terrestrial habitats is a significant milestone in evolutionary history, fostering the diversification of species. However, research on the molecular mechanisms of terrestrial adaptation in terrestrial arthropods remains relatively scarce. Crustaceans transitioned to land or damp habitats many times, constituting therefore an ideal model system to understand the genomic basis of terrestrialization. This study uses *Gecarcoidea lalandii*, a crab that inhabits damp terrestrial burrows, to investigate the molecular mechanisms underlying terrestrial adaptation in crustaceans. The chromosome-level genome of *G. lalandii* (1.33 Gb, N50: 21.25 Mb), The chromosome-level genome of *Gecarcoidea lalandii*, was sequenced using short and long read sequencing, and Hi-C technologies, and assembled with high accuracy and completeness (N50: 21.25 Mb). By employing comparative genomics approaches, coupled with high-quality genome data from eight crustacean species and one insect species, identified expanded gene families, predominantly those involved in fatty acid synthesis. Additionally, the PAX gene family, which is crucial for the development of visual neurons, and the cadherin gene family, which relates to cytoskeletal organization, were also found to be expanded. The expanded gene families are enriched in KEGG pathways such as steroid hormone biosynthesis, thermogenesis, and fatty acid metabolism. Pathways crucial for maintaining body temperature and energy supply in terrestrial environments were found to under intensification of selection. To capture the dynamic changes in gene expression over a recent period of terrestrial adaptation, we compared the transcriptomic profiles of 28 crab species. Analysis of genes under intensification of positive selection in terrestrial species identified 63 genes. These genes are enriched in KEGG pathways involved in immune responses, growth regulation, and cell signaling, suggesting potential adaptative changes resisting solar radiation and combating pathogens. By comparing various habitat crabs, we have detected several instances of convergent molecular evolution within genes that play pivotal roles in critical pathways, including the NOD-like receptor, Wnt, and Hedgehog signaling cascades. These pathways regulate crucial biological processes such as immune responses, tissue repair, and gill organ development, which are integral during independent terrestrialization events. In summary, this study elucidates the molecular mechanisms underlying terrestrial adaptation in crustaceans through comparative genomics and transcriptomics. It enhances our understanding of the genetic drivers of terrestrial adaptive evolution in crustaceans.

Index Terms: Brachyura; terrestrial adaptation; genome; transcriptome; convergent evolution; comparative genomics; crustaceans.

Funding: This study was funded by the Natural Science Foundation of China (grant no. 32201376), and the Basic and Applied Basic Research Foundation of Guangdong Province (Grant no. 2021A1515110244).

Expanded summary

Microbiome vs fatty acids and production traits in pigs

Thiago Sugizaki dos Santos[2], Andrezza Maria Felício Ament[1], Camila Sabino de Oliveira[1], Aline Silva Mello Cesar[1,2,§]

[1] Luiz de Queiroz College of Agriculture, University of São Paulo, Avenue Pádua Dias 11, Piracicaba, São Paulo, 13418-900, Brazil

[2] Faculty of Animal Science and Food Engineering, University of São Paulo, Campus Fernando Costa, Avenue Duque de Caxias Norte 225, Pirassununga, São Paulo, 13635-900, Brazil

§ Corresponding author: alinecesar@usp.br

Abstract - The concept of food security has evolved over the years. Initially focused on food supply, it incorporated ethical and human rights dimensions in 1974, becoming a more comprehensive approach. Currently, food security refers to the state in which an individual has physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and preferences for an active and healthy life at all times of their life (FAO, 2016).

Changes in the population's lifestyle resulting from the effects of urbanization and globalization, as well as armed conflicts and health crises, increase inequalities and make access to food difficult (Bezerra et al., 2020). Among the various implications of food insecurity, the health consequences have been associated with the development of chronic non-communicable diseases such as heart disease, cancer, chronic respiratory disease, and diabetes (Murrell; Jones, 2020). In this context, the consumption of lipids has been associated with these diseases, however, the function of fatty acids in the body is varied and crucial for health, as they are part of several metabolic pathways related to disease susceptibility or prevention (Spector; Kim, 2015).

The consumption of saturated, non-essential fatty acids has been linked to the propensity for the development of cardiovascular diseases, as they are related to the increase in low-density lipoproteins (LDL) and pro-inflammatory factors (Gonçalinho et al., 2023). The most common saturated fatty acids include palmitic acid, found in cottonseed oil and palm oil, and stearic acid, found in ruminant animal fat and cocoa butter. MUFA, oleic acid (omega 9), can be synthesized by the body from enzymes that catalyze the removal of hydrogen and the introduction of saturations in SFA (saturated fatty acids) (Warude et al., 2006). It is important in increasing high-density lipoprotein (HDL) levels and is present in olive and canola oils. PUFAs, linoleic acid (omega 6) and α-linolenic acid (omega 3), are not synthesized by the body, therefore they must be acquired through the intake of oilseeds, such as soy and canola, and cold-water fish (Gogus; Smith, 2010).

Evidence indicates that high-fat diets can favor the occurrence of cardiometabolic diseases through changes in the composition of the intestinal microbiota. Studies with obese mice demonstrated a decrease in the abundance of bacteria from the Bacteroidetes phylum with a concomitant increase in those belonging to the Firmicutes phylum (Turnbaugh et al., 2008). Bacteria from the Firmicutes phylum exert pro-inflammatory and obesogenic effects, while those from the Bacteroidetes phylum promote host protection against such effects (Jayasinghe et al., 2016). Similar changes in the gut microbiota have also been observed in humans (Turnbaugh et al., 2008). It is important to highlight that the intestinal microbiota can be modulated not only by the amount of dietary fat, but also by the composition of fatty acids present in the diet. Metagenomic study revealed that diets rich in MUFA, from olive oil (rich in AO), can benefit the intestinal ecosystem of mice, since increases in the population of Bacteroidetes were detected, particularly species belonging to the Bacteroidaceae family (Patterson et al., 2014).

From a taxonomic point of view, efforts have been made to identify a unique microbiome of bacteria in the gastrointestinal tract, but in the distal part of the intestine, especially in fecal samples, there has been no success (Hillman et al., 2017). Evidence indicates that the composition of the rodent microbiome in the cecum and feces are not similar (Pang et al., 2012). Recently, Panasevich et al. (2015) suggested that the Western diet (rich in fructose, fats and cholesterol) promotes severe and distinct dysbiosis in the microbiome of the proximal (cecum) and distal (feces) portion of the intestine of Ossabaw pigs, a breed with a genetic predisposition to obesity. However, it is not yet fully understood whether the cecal or fecal microbiome is the most relevant for understanding their connection with the onset of obesity. Therefore, this is the first study that will evaluate the bacterial composition present in the cecal content, and feces of pigs fed diets rich in OA, to better understand the effect of OA on the bacterial profile of the microbiota of these animals.

In addition to the intestinal microbiota presenting distinct taxonomy and functionality in each portion of the gastrointestinal tract, progressive changes occur in the intestinal ecosystem of an individual throughout life depending on age, childhood transitions (gestational period of birth, type of birth and breastfeeding and end

of breastfeeding) and external factors, such as the use of antibiotics. In adulthood, although the intestinal microbiota is relatively stable, changes between individuals can be observed due to lifestyle, including, among other factors, eating habits and physical activity (Rinninella et al., 2019). Therefore, the great intra- and inter-individual variability combined with ethical concerns regarding research involving human beings make studies of the human intestinal microbiota limited. In vitro models have been proposed, including fecal material culture systems capable of mimicking the microbial ecosystem of the human intestine (Salminem et al., 1998). However, models in vitro do not completely simulate the biological effects provided by the human digestive and immune systems (Imaoka et al., 2004).

Animals, as part of the research process, have contributed to the progress of biomedical knowledge aimed at public health. Researchers are increasingly using experimental models of obesity induced by a high-fat diet, as these models more closely reflect the metabolic changes observed in human obesity (White et al., 2009). The animals commonly used for obesity models induced by high energy density diets are mice, due to a series of characteristics, such as small size, short reproductive cycle, numerous offspring, precocity, varied nutrition and adaptation to captivity (Hillman et al., 2017). However, there are great divergences between rodents and humans regarding the physiology and composition of the intestinal microbiota. Important members of the human intestinal microbiota, such as bifidobacteria, do not colonize the intestine of rodents. Furthermore, differences in the structure of the intestinal epithelium and nutritional requirements between these two species have been reported (Pang et al., 2007).

The pig is considered an attractive experimental model for the study of various aspects that contribute to the excessive accumulation of adiposity and its consequences, as the similarities between the anatomy, physiology and pathophysiology of pigs and those of humans are already widely established in the literature (Swindle et al., 2012; Hillman et al., 2017). The omnivorous eating habit and the period of development, especially during childhood, are similar between humans and pigs, a fact that increases the interest of this species as an experimental model (Garthoff et al., 2002). Additionally, pig metabolism is easily modulated due to rapid growth (Pang et al., 2007) and thus generates a model whose response is the most appropriate. Therefore, the use of pigs as an experimental model becomes more realistic and appropriate for the study of human obesity. Recently, genetic mutations associated with metabolic syndrome and insulin resistance have been successfully carried out in miniature pigs, such as the Ossabaw, to study obesity and human metabolic disorders (Yan et al., 2013; Almeida et al., 2021).

Differences in the fatty acid profile have been the subject of studies in nutrigenomic studies in animals (Almeida et al., 2021; Martins et al., 2015; Silva et al., 2023) by our group. So far, we observed that supplementing the pig diet with fatty acids the nutritional profile and the tenderness of the Longissimus dorsi muscle (Gomes et al., 2021). The difference in the fatty acid profile of canola, fish and soybean oils differs, mainly due to linoleic acid (18:2 n-6, AL), linolenic acid (18:3 n-3, AAL), oleic acid (18 :1 n-9, AO), EPA (C20:5 n-3) and DHA (C22:6 n-3). The animals increased the DHA content in liver, brain, and skeletal muscle (Silva et al., 2023).

In this study, our hypothesis was that the incorporation of different sources of fatty acids in higher concentrations into the pig diet will result in significant changes in the diversity of the microbial community of the cecum content and feces. Then, the study aimed to evaluate the diversity of microbial community in the contents of the cecum and feces of pigs fed different diets with canola, soybean, and fish oil. To perform this study, bacterial genomic DNA from the samples (feces n =71 and cecum n=61) was extracted using the ZymoBIOMICS®-96 MagBead DNA Kit (Zymo Research, Irvine, CA). The V3-V4 hypervariable region of the 16S rRNA gene will be amplified using the Quick16S™ NGS Library Prep Kit (Zymo Research, Irvine, CA). The 16S rRNA gene library was prepared using a process in which PCR reactions were performed on qPCR machines to control cycles and therefore limit the formation of PCR chimeras. The primers used here were 16S V3-V4 Primer Sequences 341f - CCTACGGGDGGCWGCAG, CCTAYGGGGYGCWGCAG, 17 bp and 806r - GACTACNVGGGTMTCTAATCC, 24 bp, where the forward primer 341f is a mixture of the two cited sequences. The final PCR products were quantified with qPCR fluorescence readings and grouped based on equal molarity. The final pooled library was purified with the Select-a-Size DNA Clean & Concentrator™ Kit (Zymo Research, Irvine, CA) then quantified with the TapeStation® (Agilent Technologies, Santa Clara, CA) and Qubit® (Thermo Fisher Scientific, Waltham, WA). A positive control, ZymoBIOMICS™ Microbial Community Standard (Zymo Research, Irvine, CA), and a negative control were included in the DNA extraction and library preparation steps. The final library was sequenced on the NextSeq1000 System (Zymo Research, Botucatu, SP) using a v3 reagent kit (600 cycles). Sequencing was performed with a spike-in 10% PhiX. Sequencing was carried out through the platform from PCR synthesis. The QIIME 2.0 platform (version 2023.5) was used for data analysis (Boylen et al., 2019) and data visualization

using QIIME2 View (https://view.qiime2.org/). Forward and reverse reads from each sample were merged (BBMerge, v.38.94), demultiplexed and denoising was performed with the DADA2 plug-in, where the first 20 base pairs were trimmed, and sequences were truncated to have a maximum length of 250 base pairs. The sequences were grouped into ASVs (amplicon sequence variants with at least 90% similarity with the sequences deposited in the databases (Tang et al., 2014) and the taxonomy was assigned using the reference database, SILVA v. 138.99 (Quast et al., 2012). For highlighting pivotal functional profiles, the microeco R package was used with metastat statistics [https://github.com/ChiLiubio/microeco].

In this study, no difference was observed on alpha diversity measured according to the Shannon diversity index ($p>0.05$). PCA plot based on the Jaccard distance metric was used to compare the composition of microbiota in the four diets. Upon PCA analysis, the first two principal components explained 20.5% of the variance among the diets. Significative differences on beta diversity were identified among the diets in different levels of the microbial diversity, such as Phylum, Family, Genus, and Species ($p<0.05$).

The results obtained from this study showed that the diet with different fatty acid profile affected the fatty acid profile of different tissues despite did not affect the alpha diversity of the microbiota of the cecum and feces of pigs fed with different sources of the oils. However, the beta diversity of the microbiota showed differences among the diets with different fatty acid profile. In conclusion, this study showed that the diets with different fatty acid composition can affect the cecum and feces microbiota of pigs fed with different sources and proportion of oil added in pig meal improving or not their metabolic processes, consequently the performance and health.

References

ALMEIDA, V. V.; SILVA, J. P. M.; SCHINCKEL, A. P.; MEIRA, A. N.; MOREIRA, G. C. M.; GOMES, J. D.; POLETI, M. D.; DARGELIO, M. D. B.;PATINHO, I.; CONTRERAS-CASTILLO, C. J.;COUTINHO, L. L.; MOURÃO, G. B.; REECY, J. M.;KOLTES, D.; SERÃO, N. V. L.; REGITANO, L. C. A.; FUKUMASU, H.; BRUSTOLINI, A. P. L.; ALENCAR, S. M.; LUCHIARI FILHO, A.; CESAR, A. S. M. Effects of increasing dietary oil inclusion from different sources on growth performance, carcass and meat quality traits, and fatty acid profile in genetically lean immunocastrated male pigs. **Livestock Science**, v. 248, 104515, 2021.

BEZERRA, M. S.; JACOB, M. C. M.; FERREIRA, M. A. F.; VALE, D.; MIRABAL, I. R. B.; LYRA, C. de O. Food and nutritional insecurity in Brazil and its correlation with vulnerability markers. **Ciência e Saude Coletiva**, v. 25, i. 10, pp. 3833-3846, 2020.

BOLYEN, E.; RIDEOUT, J. R.; DILLON, M. R.; BOKULICH, N. A.; ABNET, C. C.; AL-GHALITH, G. A.; ALEXANDER, H.; ALM, E. J.; ARUMUGAM, M.; ASNICAR, F.; BAI, Y.; BISANZ, J. E.; BITTINGER, K.; BREJNROD, A.; BRISLAWN, C. J.; BROWN, C. T.; CALLAHAN, B. J.; CARABALLO-RODRÍGUEZ, A. M.; CHASE, J.; COPE, E. K.; SILVA, R. da; DIENER, C.; DORRESTEIN, P. C.; DOUGLAS, G. M.; DURALL, D. M.; DUVALLET, C.; EDWARDSON, C. F.; ERNST, M.; ESTAKI, M.; FOUQUIER, J.; GAUGLITZ, J. M.; GIBBONS, S. M.; GIBSON, D. L.; GONZALEZ, A.; GORLICK, K.; GUO, J.; HILLMANN, B.; HOLMES, S.; HOLSTE, H.; HUTTENHOWER, C.; HUTTLEY, G. A.; JANSSEN, S.; JARMUSCH, A. K.; JIANG, L.; KAEHLER, B. D.; KANG, K. B.; KEEFE, C. R.; KEIM, P.; KELLEY, S. T.; KNIGHTS, D.; KOESTER, I.; KOSCIOLEK, T.; KREPS, J.; LANGILLE, M. G. I.; LEE, J.; LEY, R.; LIU, Y.-X.; LOFTFIELD, E.; LOZUPONE, C.; MAHER, M.; MAROTZ, C.; MARTIN, B. D.; MCDONALD, D.; MCIVER, L. J.; MELNIK, A. V.; METCALF, J. L.; MORGAN, S. C.; MORTON, J. T.; NAIMEY, A. T.; NAVAS-MOLINA, J. A.; NOTHIAS, L. F.; ORCHANIAN, S. B.; PEARSON, T.; PEOPLES, S. L.; PETRAS, D.; PREUSS, M. L.; PRUESSE, E.; RASMUSSEN, L. B.; RIVERS, A.; ROBESON II, M. S.; ROSENTHAL, P.; SEGATA, N.; SHAFFER, M.; SHIFFER, A.; SINHA, R.; SONG, S. J.; SPEAR, J. R.; SWAFFORD, A. D.; THOMPSON, L. R.; TORRES, P. J.; TRINH, P.; TRIPATHI, A.; TURNBAUGH, P. J.; UL-HASAN, S.; VAN DER HOOFT, J. J. J.; VARGAS, F.; VÁZQUEZ-BAEZA, Y.; VOGTMANN, E.; VON HIPPEL, M.; WALTERS, W.; WAN, Y.; WANG, M.; WARREN, J.; WEBER, K. C.; WILLIAMSON, C. H. D.; WILLIS, A. D.; XU, Z. Z.; ZANEVELD, J. R.; ZHANG, Y.; ZHU, Q.; KNIGHT, R.; CAPORASO, J. G. Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. **Nature Biotechnology**, v. 37, pp. 852-857, 2019.

FAO. **Methods for estimating comparable rates of food insecurity experienced by adults throughout the world**. Rome: FAO, 2016. (Voices of the Hungry technical report, n. 1).

GARTHOFF, L. H.; HENDERSON, G. R.; SAGER, A. O.; SOBOTKA, T. J.; O'DELL, R.; THORPE, C. W.; TROTTER, W. J.; BRUCE, V. R.; DALLAS, H. L.; POELMA, P. L.; SOLOMON, H. M.; BIER, J. W.; O'DONNELL JUNIOR, M. W.; CHI, R. K.; CHIRTEL, S. J.; BARTON, C. N.; BROWN, L. H.; FRATTALI, V. P.; KHAN, M. A. The Autosow raised miniature swine as a model for assessing the effects of dietary soy trypsin inhibitor, **Food and Chemical Toxicology**, v. 40, i. 4, pp. 487-500, 2002.

GOGUS, U.; SMITH, C. N-3 omega fatty acids: a review of current knowledge. **International Journal of Food Science and Technology**, v. 45, i. 3, p 417-436, 2010.

GOMES, J. D.; COSTA, K. A.; ALMEIDA, V. V. de; LUCHIARI FILHO, A.; CESAR, A. S. M.; FANALLI, S. L. Effects of dietary oil inclusion on meat quality of immunocastrated male pigs. **Revista Brasileira de Agrotecnologia**, v. 11, i. 2, pp. 382-385, 2021.

GONÇALINHO, G. H. F.; SAMPAIO, G. R.; FREITAS, R. A. M. S.; DAMASCENO, N. R. T. Stearic Acid, but not Palmitic Acid, is Associated with Inflammatory and Endothelial Dysfunction Biomarkers in Individuals at Cardiovascular Risk. **Arquivos Brasileiros de Cardiologia**, v. 120, i. 8, e20220598, 2023.

HILLMAN, E. T.; LU, H.; YAO, T.; NAKATSU, C. H. Microbial ecology along the gastrointestinal tract. Microbes and Environments, v. 32, i. 4, pp. 300-313, 2017.

IMAOKA, A.; SETOYAMA, H.; TAKAGI, A.; MATSUMOTO, S.; UMESAKI, Y. Improvement of human faecal flora-associated mouse model for evaluation of the functional foods. **Journal of Applied Microbiology**, v. 96, i. 4, pp. 656-663, 2004.

JAYASINGHE, T. N.; CHIAVAROLI, V.; HOLLAND, D. J.; CUTFIELD, W. S.; O'SULLIVAN, J. M. The New Era of Treatment for Obesity and Metabolic Disorders: Evidence and Expectations for Gut Microbiome Transplantation. **Frontiers in Cellular and Infection Microbiology**, v. 6, article 15, 2016.

MARTINS, J. M.; NEVES, J. A.; FREITAS, A.; TIRAPICOS, J. L. Rearing system and oleic acid supplementation effect on carcass and lipid characteristics of two muscles from an obese pig breed. **Animal**, v. 9, i. 10, pp. 1721-1730, 2015.

MURRELL, A.; JONES, R. Measuring food insecurity using the food abundance index: Implications for economic, health and social well-being. I**nternational Journal of Environmental Research and Public Health**, v. 17, i. 7, 2434, 2020.

PANASEVICH, M. R.; KERR, K. R.; DILGER, R. N.; FAHEY JUNIOR, G. C.; GUÉRIN-DEREMAUX, L.; LYNCH, G. L.; WILS, D.; SUCHODOLSKI, J. S.; STEER, J. M.; DOWD, S. E.; SWANSON, K. S. Modulation of the faecal microbiome of healthy adult dogs by inclusion of potato fibre in the diet. **The british journal of Nutrition**, v. 113, i. 1, pp. 125-133, 2015.

PANG W., et al. Selective inbreeding does not increase gut microbiota similarity in BALB/c mice. **Laboratory Animals**, v. 46, i. 4, pp. 335-337, 2012.

PANG, X.; HUA, X.; YANG, Q.; DING, D.; CHE, C.; CUI, L.; JIA, W.; BUCHELI, P.; ZHAO, L. Inter-species transplantation of gut microbiota from human to pigs. **The ISME Journal**, v. 1, i. 2, pp. 156-162, 2007.

PATTERSON, E.; CRYAN, J. F.; FITZGERALD, G. F.; ROSS, R. P.; DINAN, T. G.; STANTON, C. Gut microbiota, the pharmabiotics they produce and host health. **The Proccedings of the Nutrition Society**, v. 73, i. 4, pp. 477-489, 2014.

QUAST, C.; PRUESSE, E.; YILMAZ, P.; GERKEN, J.; SCHWEER, T.; YARZA, P.; PEPLIES, J.; GLÖCKNER, F. O. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. **Nucleic Acids Research**, v. 41, p. D590-D596, 2012.

RINNINELLA, E.; RAOUL, P.; CINTONI, M.; FRANCESCHI, F.; MIGGIANO, G. A. D.; GASBARRINI, A.; MELE, M. C. What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. **Microorganisms**, v. 7, i. 1, article 14, 2019.

SALMINEN, S.; BOULEY, C.; BOUTRON-RUAULT, M. C.; CUMMINGS, J. H.; FRANCK, A.; GIBSON, G. R.; ISOLAURI, E.; MOREAU, M. C.; ROBERFROID, M.; ROWLAND, I. Functional food science and gastrointestinal physiology and function. **The british journal of Nutrition**, v. 80, i. 1, suppl. 1, p. S147-S171, 1998.

SILVA, B. P. da; FANALLI, S. L.; GOMES, J. D.; ALMEIDA, V. V. de; FUKUMASU, H.; FREITAS, F. A. O.; MOREIRA, G. C. M.; SILVA-VIGNATO, B.; REECY, J. M.; KOLTES, J. E.; KOLTES, D.; BALIEIRO, J. C. de C.; ALENCAR, S. M. de; SILVA, J. P. M. da; COUTINHO, L. L.; AFONSO, J.; REGITANO, L. C. A.; MOURÃO, G. B.; LUCHIARI FILHO, A.; CESAR, A. S. M. Brain fatty acid and transcriptome profiles of pig fed diets with different levels of soybean oil. BMC **Genomics**, v. 24, article 91, 2023.

SPECTOR, A. A.; KIM, H. Y. Discovery of essential fatty acids. **Journal of Lipid Research**, v. 56, i. 1, pp. 11-21, 2015.

SWINDLE, M. M.; MAKIN, A.; HERRON, A. J.; CLUBB JUNIOR, F. J.; FRAZIER, K. S. Swine as Models in Biomedical Research and Toxicology Testing. **Veterinary Pathology**, v. 49, i. 2, pp. 344-356, 2012.

TANG, Y.; UNDERWOOD, A.; GIELBERT, A.; WOODWARD, M. J.; PETROVSKA, L. Metaproteomics Analysis Reveals the Adaptation Process for the Chicken Gut Microbiota. **Applied and Environmental Microbiology**, v. 80, i. 2, pp. 478-485, 2014.

TURNBAUGH, P. J.; BÄCKHED, F.; FULTON, L.; GORDON, J. I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. **Cell Host Microbe**, v. 3, i. 4, pp. 213-223, 2008.

WARUDE, D.; JOSHI, K.; HARSULKAR, A. Polyunsaturated fatty acids: biotechnology. **Critical Reviews in Biotechnology**, v. 26, i. 2, pp. 83-93, 2006.

WHITE, J.R., NAGARAJAN, N., POP, M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. **PLoS computational biology**, v. 5, i. 4. e1000352, 2009.

YAN, H.; POTU, R.; LU, H.; ALMEIDA, V. V.; STEWART, T.; RAGLAND, D.; ARMSTRONG, A.; ADEOLA, O.; NAKATSU, C. H.; AJUWON, K. M. Dietary fat content and fiber type modulate hind gut microbial community and metabolic markers in the pig. **PLoS One**, v. 8, i. 4, e59581, 2013.

Modeling microbiome in genomic selection

Daniela Lourenco[1]

[1] University of Georgia, Department of Animal and Dairy Science, Athens, GA 30602, USA
Presenting author, Corresponding author: danilino@uga.edu

Abstract - One of the significant developments in animal breeding and genetics in the last two decades is the use of genomic information for genomic selection (GS), which has been enabled by the availability of affordable single-nucleotide polymorphism (SNP) genotyping. This affordability allowed the genotyping of massive numbers of animals, including over 1.5 million for U.S. Angus and over 9 million for U.S. dairy cattle as of May 2024. The successful and extensive implementation of GS by the livestock industry (Wiggans et al., 2017, Lourenco et al., 2018, Misztal et al., 2020) has resulted in GS becoming the new standard in animal breeding and genetics. According to Guinan et al. (2023), using GS increased selection accuracy and decreased generation interval, which resulted in a twofold increase in the rate of genetic progress for traits that are economically important for the U.S. dairy industry. Similar rates of gain can be generalized to beef cattle, pigs, chicken, and fish. However, this twofold increase does not guarantee food security for 2030. An extra fivefold increase in the rate of genetic progress is needed to supply the demand for animal protein worldwide (Sciences, Engineering and Medicine, 2019). Therefore, how can the rate of genetic progress be increased further to secure animal-based food production? Several approaches can be taken, and one of them is having better predictive models. A better predictive model is one that provides higher selection accuracy, less bias, and minimal to no inflation. Selection accuracy is directly responsible for the rate of genetic gain. The increase in selection accuracy provided by GS compared with traditional selection based on pedigree and phenotypes varies from 10% to over 100% (Lourenco et al., 2015, VanRaden, 2017). This is because genomics is an extra source providing non-redundant information to that already contained in the phenotypes and pedigree.
Genomic selection is now considered a mature technology (Misztal et al., 2020), with single-step genomic best linear unbiased predictor (ssGBLUP) (Aguilar et al., 2010, Christensen; Lund, 2010) being the predominant methodology. This method combines genotypes, pedigree, and phenotypes into a single evaluation run, simplifying the evaluation system in the case there are genotyped and non-genotyped individuals in the dataset. When no individuals have genotypes, ssGBLUP becomes BLUP, and when all individuals are genotyped, ssGBLUP becomes GBLUP. Using a simple animal model as example,

$$y=X\beta+Z\_1\ u+e, \quad [1]$$

where **y,β, u**, e are vectors of phenotypes, fixed effects, breeding values, and residual; X and Z_1 are incidence matrices relating y with the effects in **β, u**. In BLUP, $var(\mathbf{u}) = \mathbf{A}\sigma_u^2;$ in GBLUP, $var(\mathbf{u}) = \mathbf{G}\sigma_u^2$; in ssGBLUP, $var(\mathbf{u}) = \mathbf{H}\sigma_u^2$. Therefore, A, G, and H are the pedigree, genomic, and realized relationship matrices, respec--tively; σ_u^2 is the additive genetic variance. The **G** matrix is often constructed as in VanRaden (2008), where the diagonals count the number of homozygous loci for each individual and off-diagonals count the number of alleles shared between individuals following the equation below:

$$\mathbf{G} = \frac{\mathbf{KK}'}{2\sum_{i=1}^{SNP} p_i(1-p_i)} \quad [2]$$

where **K** is the SNP content matrix centered by twice the allele frequence (p).
In the mixed model equations, the inverse (superscript -1) of **A**, **G**, or **H** is used, and the inverse of **H** is constructed as (Aguilar et al., 2010, Christensen; Lund, 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad [3]$$

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad [4]$$

with superscripts 1 and 2 referring to non-genotyped and genotyped individuals; $\mathbf{A}_{22}^{-1}$ is the inverse of the pedigree relationship matrix for genotyped individuals. When genotyped and non-genotyped are considered together in ssGBLUP, genotypes are implicitly imputed to non-genotyped individuals if there are relationships between these two classes through the pedigree.

Dealing with genomic information is a well-established phenomenon in animal breeding and genetics. Therefore, looking at new sources of variation to be included in predictive models is currently a hot topic. According to the genome to phenome USDA blueprint for animal genome research from 2018 – 2027 (Rexroad et al., 2019), GS could be improved by adding trait-relevant "omics" data like transcriptomic, proteomic, and metabolomic information. Additionally, using the microbiome as a source of variation for the host phenotypes could result in improvements in production, efficiency, resilience, and welfare. However, there is a question on how to include this information in GS models. Khanal et al. (2020) showed that when the microbiome information is available for all individuals, the following GBLUP-based model can be used:

A) Microbiome as a random, correlated effect:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z_1 u} + \mathbf{Z_2 o} + \mathbf{e}, \quad [5]$$

where o is the random microbiome effect and $\mathbf{Z_2}$ is a matrix relating **y** to **o**; $var(\mathbf{o}) = \mathbf{\Pi}\sigma_o^2$; σ_o^2 is the microbiome variance; Π is the microbial similarity matrix (resemblance of microbiome composition) constructed as $\mathbf{\Pi} = \frac{\mathbf{1}}{q}\mathbf{QQ}'$, with **q** the number of operational taxonomic units (OTU) and Π defined as:

$$\mathbf{\Pi}_{ij} = \frac{\mathbf{\log}(\mathbf{S}_{ij}) - \mathbf{\log}(\bar{\mathbf{S}}_{.j})}{\mathrm{sd}(\mathbf{\log S}_{.j})} \quad [6]$$

where S is a matrix of relative abundance of OTU, and $\mathbf{S}_{.j}$ is the vector of the *j*-th column of S. The OUT can be replaced by amplicon sequence variant (ASV). Based on this model, it's possible to calculate the percentage of phenotypic variation that is due to the microbiome information, the so-called microbiability.

B) Interaction between microbiome and animal effect:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z_1 u} + \mathbf{Z_2 o} + \mathbf{Z_3 uo} + \mathbf{e}, \quad [7]$$

where $\mathbf{Z_3}$ is a matrix relating **uo** to **y**; **uo** is the random interaction between the additive genetic and microbiome effects; $var(\mathbf{uo}) = \mathbf{G}\odot\mathbf{M}\sigma_{uo}^2$; $\odot$ is the Hadamard product; σ_{uo}^2 is the variance for the interaction uo.

The models in equations [6] and [7] have two main limitations. First, the microbiome variance can be inflated if the microbial similarity matrix is not constructed properly. Second, these models assume all individuals have genotypes and microbiome information, which is not true in animal populations. Usually, 5% to 20% of the individuals in the pedigree are genotyped. To overcome the second limitation, we can use a model proposed by Christensen et al. (2021) to include intermediate omics features in genomic evaluations. Although the model was developed for transcriptomics or metabolomics, it can fit metagenome data. According to Legarra and Christensen (2023), some "omics" data are intermediate between DNA and phenotypes; however, even when they are not, we can pretend they are and use this model. In a nutshell, the intermediate omics model is based on the fact that the genetic background of an individual codes for an intermediate omics phenotype (OP) that is part of the actual phenotype (P). As not all genetic variation is expressed through OP, the breeding value (BV) of an individual can be decomposed into omics-mediated BV (i.e., from the genes controlling omics that are transmitted to the offspring) and non-omics-mediated BV. Therefore, having OP is essential to estimate the omics-mediated BV better.

In the methodology described by Legarra and Christensen (2023), there are two hierarchical mixed models as follows:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z}_m\mathbf{M\alpha} + \mathbf{Z}_r\mathbf{u}_r + \boldsymbol{\epsilon}, \quad [8]$$

$$\mathbf{m}_i = \widetilde{\mathbf{X}}\widetilde{\boldsymbol{\beta}}_i + \widetilde{\mathbf{Z}}\mathbf{g}_i + \mathbf{e}_i, \quad [9]$$

Equation [8] models the relationship between phenotypes and omics, where $\mathbf{Z}_m$ is a matrix relating individuals with omics data to **y**; **M** is the omics matrix; α is a vector of r egression coefficients of omics on phenotypes, with $var(\boldsymbol{\alpha}) = \mathbf{I}\sigma^2_{\boldsymbol{\alpha}}$; $\mathbf{Z}_r$ is a matrix relating $\mathbf{u}_r$ to **y**; $\mathbf{u}_r$ is a vector of residual genetic effect, with ; $var(\mathbf{u}_r) = \mathbf{H}\sigma^2_{u_r}$, $\boldsymbol{\epsilon}$ is a vector of residuals. Equation [9] models the relationship between omics and genetic effects for individuals, where the subscript *i* defines the *i*-th omics level; mı is the omics expression levels; $\widetilde{\boldsymbol{\beta}}_i$ is a vector of fixed effects $\widetilde{\mathbf{X}}$ relates $\widetilde{\boldsymbol{\beta}}_i$ to $\mathbf{m}_i$; $\widetilde{\mathbf{Z}}$ relates individuals to omics levels; $\mathbf{g}_i$ is a vector of genetic effects, with $var(\mathbf{g}_i) = \mathbf{H}\sigma^2_{g_i}$; $\mathbf{e}_i$ the vector of residuals. Based on that, the overall estimated BV of an individual is the sum of the omics-mediated BV, u_m=g_i α, and the residual breeding value, $\mathbf{u}_r$; therefore, $u = u_m + u_r$.

Alternatively, the microbiome information can be combined into an index and treated as a proxy trait, meaning phenotypes can be derived from microbiome information. Hayes et al. (2017) showed that using derived phenotypes from near infrared or nuclear magnetic resonance (considered omics because of the high dimensionality) improved the accuracy of genomic predictions in plant breeding. Similarly, the microbiome-derived phenotype can be used in multi-trait models with regularly measured traits to help improve those that are hard or expensive to measure. If the microbiome-derived trait is highly correlated (> 0.85) with traits that are hard or expensive to measure, those traits can be replaced by the microbiome-derived trait.
Several methods are available to fit microbiome information into genomic evaluation models; however, collecting microbiome information at a national scale, which is needed for genomic selection, is still not a reality. Whenever large-scale data becomes available, these methods will be fine-tuned, and new ones will be developed. Still, adopting this technology will depend on the impact of this extra source of information on the accuracy of breeding values.

References

AGUILAR, I.; MISZTAL. I.; JOHNSON, D. L.; LEGARRA, A.; TSURUTA, S.; LAWLOR, T. J. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, v. 93, i. 2, pp. 743-752, 2010.

CHRISTENSEN, O. F.; BÖRNER, V.; VARONA, L.; LEGARRA, A. Genetic evaluation including intermediate omics features. **Genetics**, v. 219, i. 2, iyab130, 2021.

CHRISTENSEN, O. F.; LUND, M. S. Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution**, v. 42, article 2, 2010.

GUINAN, F. L.; WIGGANS, G. R.; NORMAN, H. D.; DÜRR, J. W.; COLE, J. B.; VAN TASSELL, C. P.; MISZTAL, I.; LOURENCO, D. Changes in genetic trends in US dairy cattle since the implementation of genomic selection. **Journal of Dairy Science**, v. 106, i. 2, pp. 1110-1129, 2023.

HAYES, B. J.; PANOZZO, J.; WALKER, C. K.; CHOY, A. L.; KANT, S.; WONG, D.; TIBBITS, J.; DAETWYLER, H. D.; ROCHFORT, S.; HAYDEN, M. J.; SPANGENBERG, G. C. Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. **Theoretical and Applied Genetics**, v. 130, i. 12, pp. 2505-2519, 2017.

KHANAL, P.; MALTECCA, C.; SCHWAB, C.; FIX, J.; BERGAMASCHI, M.; TIEZZI, F. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. **Genetics Selection Evolution**, v. 52, article 41, 2020.

LEGARRA, A.; CHRISTENSEN, O. Genomic evaluation methods to include intermediate correlated features such as highthroughput or omics phenotypes. **Journal of Dairy Science Communications**, v. 4, i. 1, pp. 55-60, 2023.

LOURENCO, D. A. L.; TSURUTA, S.; FRAGOMENI, B. O.; MASUDA, Y.; AGUILAR, I.; LEGARRA, A.; BERTRAND, J. K.; AMEN, T. S.; WANG, L.; MOSER, D. W.; MISZTAL, I. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. **Journal of Animal Science**, v. 93, i. 6, pp. 2653-2662, 2015.

LOURENCO, D. A. L.; TSURUTA, S.; FRAGOMENI, B. O.; MASUDA, Y.; AGUILAR, I.; LEGARRA, A.; MILLER, S.; MOSER, D.; MISZTAL, I. Single-step genomic BLUP for national beef cattle evaluation in US: From initial developments to final implementation. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 11., 2018, Auckland. Proceedings... [S.l.: s.n.], 2018. p. 495. WCGALP 2018.

MISZTAL, I.; LOURENCO, D.; LEGARRA, A. Current status of genomic evaluation. **Journal of Animal Science**, v. 98, i. 4, skaa101, 2020.

NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE. **Science Breakthroughs to Advance Food and Agricultural Research by 2030**. Washington DC: The National Academies Press, 2019.

REXROAD, C.; VALLET, J.; MATUKUMALLI, L. K.; REECY, J.; BICKHART, D.; BLACKBURN, H.; BOGGESS, M.; CHENG, H.; CLUTTER, A.; COCKETT, N.; ERNST, C.; FULTON, J. E.; LIU, J.; LUNNEY, J.; NEIBERGS, H.; PURCELL, C.; SMITH, T. P. L.; SONSTEGARD, T.; TAYLOR, J.; TELUGU, B.; VAN EENENNAAM, A.; VAN TASSELL, C. P.; WELLS, K. Genome to Phenome: improving animal health, production, and well-being – a new USDA blueprint for animal genome research 2018–2027. **Frontiers in Genetics**, v. 10, article 327, 2019.

VANRADEN, P. M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v. 91, i. 11, pp. 4414-4423, 2008.

VANRADEN, P. M. Genomic tools to improve progress and preserve variation for future generations. In: ANNUAL MEETING OF THE EUROPEAN FEDERATION OF ANIMAL SCIENCE, 68., 2017, Tallin, Estonia. **Abstracts...** Tallin, Estonia: Wageningen Academic, 2017. pp. 79. (EAAP Book of Abstracts, Volume: 23)

WIGGANS, G. R.; COLE, J. B.; HUBBARD, S. M.; SONSTEGARD, T. S. Genomic selection in dairy cattle: the USDA experience. **Annual review animal bioscience**, v. 8, i. 5, pp. 309-327, 2017.

Bifidobacteria in the mammalian tree of life

Francesca Bottacini[1,2]

[1] APC Microbiome Ireland, University College Cork, Ireland

[2] Biological Sciences & ADAPT Munster Technological University, Bishopstown, Cork

§ Presenting author, Corresponding author: francesca.bottacini@mtu.ie

Abstract - Bifidobacteria are nonmotile, nonsporulating, catalase-negative, anaerobic (or microaerophilic), Gram-positive bacteria, with a relatively high GC content genome (Goodfellow et al., 2012). The *Bifidobacterium* genus is currently composed of over 100 (sub)species, the majority found in mammalian hosts. In terms of ecological origin, bifidobacteria have been isolated prevalently from the mammalian gut (i.e. faecal samples), however some species are also found in the intestine of social insects and a small number of bird species (e.g., chicken, peacock and phasian) (Alessandri et al., 2021). A recent metagenomic profiling study based on amplicon sequencing of an Internal Transcribed Spacer (ITS) (Milani et al. 2014a) in fecal samples collected from 67 different mammalian species revealed that no strict host specialization is observed in this genus, and that members of this genus are broadly found across mammals. This approach also identified Bifidobacterium adolescentis, B*ifidobacterium bifidum*, *Bifidobacterium longum* subsp. *longum* and *Bifidobacterium pseudolongum* as the most abundant taxa in the gastrointestinal tract of mammalian species (Milani et al., 2017). Next-Generation Sequencing and the in silico analysis of bifidobacterial pan-genomes has allowed to define the set of core and accessory feature of bifidobacterial (sub)species across the whole genus. These investigations indicated that this genus possess all enzymes required for the fermentation of glucose and fructose to lactic acid and acetate which are core genome features of this genus (Milani et al. 2014b). A significant portion of the accessory genome in the Bifidobacterium pangenome contains genes encoding for glycosyl hydrolases (GHs), involved in the metabolism of complex carbohydrates abundant in the human and animal diet as well as host and diet glycans. The variety and combination of GHs is strain-specific, thus highlighting the high metabolic versatility of members of this genus. This represents a key feature allowing *Bifidobacterium* to colonise various niches and establish microbe-microbe and microbe-host interactions (Bottacini et al., 2017). In addition to glycosyl hydrolases, also accessory gene functions involved in bacterial adhesion, production of exopolysaccharides and defence systems against horizontal DNA transfer have been described for this genus, thus highlighting the genomic variability and versatility of members of this genus. Comparative genome analysis have highlighted how a significant portion of bifidobacterial pangenomes is constituted of Mobile genetic elements (MGEs) involved in the adaptation and evolution of these commensal species by promoting horizontal gene transfer (HGT). Understanding the dynamics of DNA thansfer mediated by MGEs in bifidobacteria is essential for comprehending evolution and functional capabilities of members of this genus. Recent advances in comparative and functional genomics have facilitated the development of molecular genetic tools, used to expand our knowledge of health promoting gene functions in a number of bifidobacterial species, thus expanding the range of resources available for the selection and genetic investigation of novel probiotic strains (Bottacini et al., 2017; Hoedt et al., 2021).

References

ALESSANDRI, G.; van SINDEREN, D.; VENTURA, M. The genus Bifidobacterium: from genomics to functionality of an important component of the mammalian gut microbiota. **Computational and Structural Biotechnology Journal**, v. 19, pp. 1472-1487, 2021.

BOTTACINI, F.; van SINDEREN, D.; VENTURA, M. Omics of bifidobacteria: research and insights into their health promoting activities. T**he Biochemical journal**, v. 474, i. 24, pp. 4137-4152, 2017.

GOODFELLOW, M.; KÄMPFER, P.; BUSSE, H.-J.; TRUJILLO, M. E.; SUZUKI, K.; LUDWIG, W.; WHITMAN, W. B. **Bergey's Manual of Systematic Bacteriology:** volume five - the Actinobacteria, part A. 2. ed. New York: Springer, 2012.

HOEDT, E. C.; BOTTACINI, F.; CASH, N.; BONGERS, R. S.; VAN LIMPT, K.; BEN AMOR, K.; KNOL, J.; MACSHARRY, J.; van SINDEREN, D. Broad Purpose Vector for Site-Directed Insertional Mutagenesis in Bifidobacterium breve. **Frontiers in Microbiology**, v. 12, 636822, 2021.

MILANI, C.; LUGLI, G. A.; DURANTI, S.; TURRONI, F.; BOTTACINI, F.; MANGIFESTA, M.; SANCHEZ, B.; VIAPPIANI, A.; MANCABELLI, L.; TAMINIAU, B.; DELCENSERIE, V.; BARRANGOU, R.; MARGOLLES, A.; VAN SINDEREN, D.; VENTURA, M. Genomic encyclopedia of type strains of the genus Bifidobacterium. **Applied and environmental Microbiology**, v. 80, i. 20, pp. 6290-6302, 2014b.

MILANI, C.; LUGLI, G. A.; TURRONI, F.; MANCABELLI, L.; DURANTI, S.; VIAPPIANI, A.; MANGIFESTA, M.; SEGATA, N.; van SINDEREN, D.; VENTURA, M. Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. **FEMS Microbiology Ecology**, v. 90, i. 2, pp. 493-503, 2014a.

MILANI, C.; MANGIFESTA, M.; MANCABELLI, L.; LUGLI, G. A.; JAMES, K.; DURANTI, S.; TURRONI, F.; FERRARIO, C.; OSSIPRANDI, M. C.; VAN SINDEREN, D.; VENTURA, M. Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life. **The ISME Journal**, v. 11, i. 12, pp. 2834-2847, 2017.

Decoding the Language of Life: Advances in Transformer-Based Language Models for Bioinformatics

Haithem Afli[(1)]

[(1)] ADAPT Centre, Munster Technological University, Cork, Ireland
Corresponding author: Haithem.afli@mtu.ie

Introduction

The intersection of artificial intelligence (AI) and bioinformatics is increasingly prominent as researchers seek innovative methods to decode the complexities of biological systems (Ventero et al., 2021; Andrade et al., 2022). In this context, transformer-based language models have emerged as powerful tools capable of interpreting vast amounts of biological data. These models, initially developed for natural language processing (NLP) tasks (Vaswani et al., 2017; Wolf et al., 2020), have been adapted for various applications within the biomedical field, showcasing their versatility and transformative potential (Kalyan et al., 2022). The shift from traditional bioinformatics methods to AI-driven approaches marks a significant evolution in the field, where deep learning models such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pretrained Transformer), and specialised models like DNABERT, ESM (Evolutionary Scale Modeling), Evo, and AlphaFold are leading the charge (Devlin et al., 2018; Ji et al., 2021; Rives et al., 2021).

Background on Transformer-Based Language Models

Transformer-based language models have revolutionised NLP by introducing a mechanism known as attention, which allows the model to focus on different parts of the input sequence when making predictions (Vaswani et al., 2017). This ability to capture contextual relationships within data is particularly advantageous in bioinformatics, where the complexity and variability of biological sequences present significant challenges. The architecture of transformers, unlike recurrent neural networks (RNNs), does not rely on sequential data processing. Instead, transformers process input data in parallel, significantly improving computational efficiency and performance, especially on large datasets (Vaswani et al., 2017). This efficiency is crucial in bioinformatics, where datasets can be enormous and computationally intensive to analyse.

Applications in Bioinformatics

1. Genomic Analysis

One of the most significant applications of transformer-based models in bioinformatics is genomic analysis. Genomic data, consisting of long sequences of nucleotides, can be likened to sentences in a language, where understanding the “syntax” and “grammar” is essential for interpreting genetic information (Ji et al., 2021). BERT models, for instance, have been adapted to perform tasks such as DNA sequence classification, mutation detection, and functional annotation of genes.
DNABERT is a notable example of such an adaptation. DNABERT extends the BERT architecture to handle DNA sequences by representing k-mers (subsequences of length k) as tokens, similar to how words are treated in NLP. This adaptation has proven highly effective for tasks like predicting transcription factor binding sites and identifying regulatory elements, where understanding the contextual relationships between nucleotides is crucial (Ji et al., 2021).

2. Protein Folding Predictions

Predicting the three-dimensional structure of proteins from their amino acid sequences is one of the longstanding challenges in bioinformatics. The success of models like AlphaFold, which leverages transformer architecture, has demonstrated the potential of AI in addressing this challenge (Jumper et al., 2021).

AlphaFold, developed by DeepMind, represents a significant breakthrough in this area. By applying a deep learning approach to predict protein structures, AlphaFold has achieved unprecedented accuracy, even surpassing traditional methods in several benchmarks (Senior et al., 2020).
This achievement has far-reaching implications for the understanding of molecular biology and the development of new drugs, as it enables researchers to model proteins that were previously too complex to predict with high accuracy.

3. Evolutionary Modeling with ESM

Another noteworthy application is Evolutionary Scale Modeling (ESM), which applies transformer models to the study of protein evolution. ESM models learn from large-scale datasets of protein sequences, capturing evolutionary patterns and relationships between proteins (Rives et al., 2021). This approach allows researchers to predict the function of unknown proteins and understand the evolutionary history of different protein families. ESM models have been particularly successful in tasks such as predicting the effects of mutations on protein function and identifying conserved regions within protein families that are critical for their biological function. This evolutionary perspective is invaluable for fields like evolutionary biology and molecular genetics, where understanding the changes in protein sequences over time can provide insights into the mechanisms of evolution and the development of new functions (Rives et al., 2021).

4. DNA Foundation Modeling with Evo

Evo is another advanced model that extends the capabilities of transformer-based models to DNA foundation modeling, spanning from the molecular to genome scale. Evo integrates a deep understanding of molecular interactions with large-scale genomic data, allowing for comprehensive modeling of DNA behavior across different scales. This model is particularly useful for understanding how molecular changes at the DNA level can lead to broader genomic alterations, influencing everything from gene expression to evolutionary adaptation. Evo's ability to bridge these scales makes it a powerful tool for studying complex biological phenomena, such as how specific mutations propagate through generations or how genetic networks evolve over time.

5. Single-Cell Data Interpretation

The advent of single-cell RNA sequencing (scRNA-seq) has provided unprecedented insights into cellular heterogeneity. However, the analysis of this data is complex due to the high dimensionality and noise associated with single-cell datasets. Transformer-based models have been employed to interpret scRNA-seq data, enabling the identification of cell types, states, and developmental trajectories (Stuart et al., 2019). By leveraging the attention mechanism, these models can focus on critical genes and pathways, offering a more nuanced understanding of cellular behaviour

Challenges and Considerations

Despite the successes of transformer-based models in bioinformatics, several challenges remain. One of the primary issues is the need for high-quality, annotated datasets. In NLP, large-scale corpora such as Wikipedia and Common Crawl provide abundant data for training models. However, in bioinformatics, obtaining similarly comprehensive datasets can be difficult. Moreover, biological data is often noisy and heterogeneous, requiring sophisticated preprocessing and normalisation techniques (Topol, 2019).
Another challenge is the computational resources required to train transformer models. The depth and complexity of these models mean they require significant processing power and memory, often necessitating the use of specialised hardware such as GPUs or TPUs (Vaswani et al., 2017). This requirement can be a barrier to entry for smaller research groups or institutions with limited resources.
Furthermore, there is an ongoing need to improve the interpretability of these models. In the biomedical field, where decisions can have significant consequences, it is vital that models not only make accurate predictions but also provide insights into the reasoning behind these predictions (Kalyan et al., 2022). Current research is focused on developing techniques that make the inner workings of transformer models more transparent, allowing researchers and clinicians to better understand how these models arrive at their conclusions.

Future Directions

The future of transformer-based language models in bioinformatics looks promising, with several areas ripe for exploration. One potential direction is the integration of multi-modal data, such as combining genomic, proteomic, and transcriptomic data, to create more holistic models of biological systems. This approach could lead to a more comprehensive understanding of how different layers of biological information interact, potentially unveiling new biological mechanisms and pathways (AlQuraishi, 2019). Another avenue is the development of models that can learn from smaller datasets, which is crucial for rare diseases where data is often scarce. Techniques such as transfer learning, where models pre-trained on large datasets are fine-tuned on smaller, task-specific datasets, are being explored to address this challenge (Li et al., 2024). Additionally, the integration of unsupervised learning methods could allow models to extract meaningful patterns from data without requiring extensive annotation. For instance, BiGAMi, a bi-objective genetic algorithm, has been designed to optimize feature selection in microbiome datasets, enhancing the discovery of meaningful patterns in complex biological data (Leske et al., 2022). As these models become more sophisticated, there is also the potential for real-time applications in personalised medicine. For instance, transformer-based models could be used to analyse a patient's genomic data and predict their response to different treatments, enabling more tailored and effective therapies. This could be particularly valuable in the context of diseases like cancer, where personalised treatment plans are often necessary (Topol, 2019). However, as these models deal with sensitive patient data, ensuring secure training frameworks is crucial. Techniques such as splitfed learning and multi-head split learning have been proposed to address information leakage and enhance data privacy in collaborative learning settings (Joshi et al., 2022).

Conclusion

Transformer-based language models represent a significant advancement in the field of bioinformatics, offering new tools for analysing complex biological data. From genomic analysis to protein folding, evolutionary modeling, DNA foundation modeling with Evo, and single-cell interpretation, these models have the potential to deepen our understanding of biological systems and drive new discoveries. However, challenges such as data quality, computational demands, and model interpretability must be addressed to fully realise this potential. As research continues, transformer models are likely to play an increasingly important role in the ongoing effort to decode the language of life.

References

ALQURAISHI, M. End-to-End Differentiable Learning of Protein Structure. **Cell Systems**, v. 8, i. 4, pp. 292-301, 2019. Available at: http://doi.org/10.1016/j.cels.2019.03.006. Access at: 10 jan. 2024.

ANDRADE, B. G. N.; DONATONI, F. A. B.; CUADRAT, R. R.; CARDOSO, T. F.; MALHEIROS, J. M.; OLIVEIRA, P. S. N. DE; PETRINI, J.; MOURÃO, G. B.; COUTINHO, L. L.; REECY, J. M.; KOLTES, J. E.; ZERLOTINI NETO, A.; MEDEIROS, S. R. de; BERNDT, A.; PALHARES, J. C. P.; AFLI, H.; REGITANO, L. C. de A. Stool and ruminal microbiome components associated with methane emission and feed efficiency in Nelore beef cattle. **Frontiers in Genetics**, v. 13, 812828, 2022.

DEVLIN, J.; CHANG, M.-W.; LEE, K.; TOUTANOVA, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Conference of the North American Chapter of the Association for Computational Linguistics: HumanLanguage Technologies, 2019, Stroudsburg, PA. **Proccedings...** Stroudsburg, PA: NAACL, 2019. p. 4171-4186. Available at: https://arxiv.org/abs/1810.04805. Access at: 10 out. 2023.

JI, Y.; ZHOU, Z.; LIU, H.; DAVULURI, R. DNABERT: Pre-trained Bidirectional Encoder Representations from Transformers Model for DNA-language in Genome. **Bioinformatics**, v. 37, i. 15, pp. 2112-2120, 2021. Available at: http://doi.org/10.1093/bioinformatics/btab117. Access at: 8 mar. 2024.

JOSHI, P.; THAPA, C.; CAMTEPE, S.; HASANUZZAMAN, M.; SCULLY, T.; AFLI, H. Performance and Information Leakage in Splitfed Learning and Multi-Head Split Learning in Healthcare Data and Beyond. **Methods and Protocols**, v. 5, i. 3, 60, 2022.

JUMPER, J.; EVANS, R.; PRITZEL, A.; GREEN, T.; FIGURNOV, M.; RONNEBERGER, O.; TUNYASUVUNAKOOL, K.; BATES, R.; ŽÍDEK, A.; POTAPENKO, A.; BRIDGLAND, A.; MEYER, C.; KOHL, S. A. A.; BALLARD, A. J.; COWIE, A.; ROMERA-PAREDES, B.; NIKOLOV, S.; JAIN, R.; ADLER, J.; BACK, T.; PETERSEN, S.; REIMAN, D.; CLANCY, E.; ZIELINSKI, M.; STEINEGGER, M.; PACHOLSKA, M.; BERGHAMMER, T.; BODENSTEIN, S.; SILVER, D.; VINYALS, O.; SENIOR, A. W.; KAVUKCUOGLU, K.; KOHLI, P.; HASSABIS, D. Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, i. 7873, pp. 583-589, 2021. Available at: http://doi.org/10.1038/s41586-021-03819-2. Access at: 11 abr. 2024.

KALYAN, K. S.; RAJASEKHARAN, A.; SANGEETHA, S. AMMU: a survey of transformer-based biomedical pretrained language models. **Journal of biomedical informatics**, v. 126, 103982, 2022. Available at: https://doi.org/10.1016/j.jbi.2021.103982. Access at: 10 nov. 2024.

LESKE, M.; BOTTACINI, F.; AFLI, H.; ANDRADE, B. G. N. BiGAMi: bi-objective genetic algorithm fitness function for feature selection on microbiome datasets. **Methods and Protocols**, v. 5, i. 3, 42, 2022.

LI, X.; CHEN, K.; YANG, J.; WANG, C.; YANG, T.; LUO, C.; LI, N.; LIU, Z. TLDA: a transfer learning based dual-augmentation strategy for traditional chinese medicine syndrome differentiation in rare disease. **Computers in Biology and Medicine**, v. 169, 107808, 2024.

RIVES, A.; MEIER, J.; SERCU, T.; GOYAL, S.; LIN, Z.; LIU, J.; GUO, D.; OTT, M.; ZITNICK, C. L.; MA, J.; FERGUS, R. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. **Proceedings of the National Academy of Sciences**, v. 118, i. 15, e2016239118, 2021. Available at: https://doi.org/10.1073/pnas.2016239118. Access at: 10 abr. 2024.

SENIOR, A. W.; EVANS, R.; JUMPER, J.; KIRKPATRICK, J.; SIFRE, L.; GREEN, T.; QIN, C.; ŽÍDEK, A.; NELSON, A. W. R.; BRIDGLAND, A.; PENEDONES, H.; PETERSEN, S.; SIMONYAN, K.; CROSSAN, S.; KOHLI, P.; JONES, D. T.; SILVER, D.; KAVUKCUOGLU, K.; HASSABIS, D. Improved protein structure prediction using potentials from deep learning. **Nature**, v. 577, i. 7792, pp. 706-710, 2020. Available at: https://doi.org/10.1038/s41586-019-1923-7. Access at: 23 abr. 2024.

STUART, T.; BUTLER, A.; HOFFMAN, P.; HAFEMEISTER, C.; PAPALEXI, E.; MAUCK, W. M. 3rd; HAO, Y.; STOECKIUS, M.; SMIBERT, P.; SATIJA, R. Comprehensive Integration of Single-Cell Data. **Cell**, v. 177, i. 7, pp. 1888-1902.e21, 2019. Available at: https://doi.org/10.1016/j.cell.2019.05.031. Access at: 11 fev. 2024.

TOPOL, E. J. High-performance medicine: the convergence of human and artificial intelligence. **Nature Medicine**, v. 25, i. 1, pp. 44-56, 2019. Available at: https://doi.org/10.1038/s41591-018-0300-7. Access at: 9 abr. 2024.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUHKIN, I. Attention Is All You Need. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS (NIPS), 31., 2017, Long Beach, CA. **Anais...** Long Beach, CA: NIPS, 2017. Available at: http://doi.org/10.48550/arXiv.1706.03762. Access at: 10 fev. 2024.

VENTERO, M. P.; CUADRAT, R. R. C.; VIDAL, I.; ANDRADE, B. G. N.; MOLINA-PARDINES, C.; HARI MORENO, J. M.; COUTINHO, F. H.; MERINO, E.; REGITANO, L. C. de A.; SILVEIRA, C. B.; AFLI, H.; LÓPEZ-PÉREZ, M.; RODRÍGUEZ, J. C. Nasopharyngeal microbial communities of patients infected with SARS-CoV-2 that developed COVID-19. **Frontiers in Microbiology**, v.12, 637430, 2021.

WOLF, T.; DEBUT, L.; SANH, V.; CHAUMOND, J.; DELANGUE, C.; MOI, A.; CISTAC, P.; RAULT, T.; LOUF, R.; FUNTOWICZ, M.; DAVISON, J.; SHLEIFER, S.; VON PLATEN, P.; MA, C.; JERNITE, Y.; PLU, J.; XU, C.; LE SCAO, T.; GUGGER, S.; DRAME, M.; LHOEST, Q.; RUSH, A. Transformers: State-of-the Art Natural Language Processing. In: Conference on Empirical Methods in Natural Language Processing, 2020, Stroudsburg, PA. **System Demonstrations:** proceedings. [S. l.]: Association for Computational Linguistics, 2020. p. 38-45. EMNLP 2020. Available at: https://aclanthology.org/2020.emnlp-demos.6/. DOI: 10.18653/v1/2020.emnlp-demos.6. Access at: 15 nov. 2024.

Integrating multi-omics data using multilayer networks for the analysis of the rumen microbiome

Mengyuan Wang[1], Haiying Wang[2], Huiru Zheng[2*]

[1] School of Health & Life Sciences, Teesside University, United Kingdom
[2] School of Computing, Ulster University, United Kingdom
*Corresponding author: h.zheng@ulster.ac.uk

Introduction

Rumen is a large fermentation chamber hosting diverse microbial communities, including bacteria, archaea, protozoa, and fungi. The rumen microbes enzymatically degrade these complex carbohydrates to produce volatile fatty acids (VFAs), which serve as a primary energy source for the host (Henderson et al., 2015). The rumen microbiome represents a complex and dynamic ecosystem. Central to this ecosystem's functionality is the metabolic interplay among diverse microbial species, which is linked to animal health, productivity, and environmental impacts (i.e. methane emissions).

The advent of omics technologies presents an opportunity to comprehend the mechanism behind microbial activities at a particular level. For example, genomics helps identify the various microbial species present in the rumen. Understanding the genomic composition of the rumen microbes can reveal their potential functions and interactions. Transcriptomics and proteomics can be used to examine the active biological processes within the rumen microbiome. These omics layers provide insights into which genes are being expressed and what proteins are being produced and provide information on the metabolic pathways that are active in the microbiome under different dietary conditions. Metabolomics focuses on the small-molecule metabolites found within the rumen ecosystem. Rumen metabolite provides a direct measurement of biochemical activities, offering insights into how microbial functions can be optimized or controlled. Each metabolite offers a snapshot of biochemical activities, providing a dynamic picture of how microbial populations respond to diet, health, and environmental conditions.

Multi-omics analysis, which integrates data from different biomolecular levels, has great potential to obtain a holistic view of the microbial ecosystem, including its functional capabilities, and its interaction with the host animal and environment (Kaur et al. 2023). This integration can help in constructing more accurate models of microbial activity and interactions, leading to better predictions of how changes in the microbiome affect the host animal. Combining metabolomic data with genomic, transcriptomic, and proteomic data can provide a more complete picture of the microbial activities in the rumen. Multi-omics integrative approach can help elucidate how changes at the molecular level influence overall microbial function and host interaction.

Multi-layer networks

Multi-layer networks are a model based on graph theory used to capture the complexity of interactions across different interconnected systems. In multi-layer networks, nodes represent the entities or actors, like those in single-layer networks. However, each node can participate in various types of relationships or interactions, which are segregated into multiple layers (Domenico et al., 2015). Within each layer, nodes are connected by edges that represent a specific type of relationship relevant to that layer. Additionally, nodes across layers (either the same node in different contexts or different nodes across contexts) can be connected through inter-layer edges. These inter-layer connections can represent the same entities across different contexts or different entities that are related through some cross-contextual interaction. Layers in a multi-layer network can be categorized as either multiplex networks, where the same set of nodes is connected by different types of relationships (e.g., different social interactions), or interconnected networks, where different layers contain different nodes, and the layers are connected by specific types of inter-layer edges.

Multi-layer networks permit the extension of traditional graph theoretical measures, such as centrality, clustering coefficients, and connectivity, to account for the added complexity of multiple layers (Wang et al., 2022). New metrics specific to multi-layer contexts, such as layer similarity or inter-layer centrality, are also developed. Modeling interactions between different types of biological entities (e.g., gene regulation, protein-

protein interactions, metabolic pathways) across different conditions or types of data.
Integrating omics data with a multilayer network-based approach represents a significant step forward to the methods already in use in which each type of data is treated independently (Wang et al. 2016). Due to the heterogenous and high dimensional nature of omics data, integrating microbial multi-omics data analysis requires special model design consideration.

Integrating microbial multi-omics data: Challenges and Opportunities

Microbial multi-omics refers to the integrated analysis of different omics datasets—such as genomics, transcriptomics, proteomics, and metabolomics—to understand the functioning, interaction, and dynamics of microbial communities (Lee et al., 2020). However, the composition of the rumen microbiome can vary significantly between different animals, diet types, and even over time within the same animal. This variability introduces additional complexity in data analysis and interpretation, as it can be challenging to discern patterns or draw general conclusions from multi-omics data. Constructing multi-layer networks for multi-omics data poses several unique challenges. Multi-omics involves the integration and analysis of datasets generated from various omics layers such as genomics, transcriptomics, proteomics, and metabolomics, each providing different but complementary information about biological processes. The multi-layer network approach aims to capture the complexity of interactions between these various biological data types (Barupal et al., 2018). Figure 1 shows some key challenges and considerations in constructing these networks.

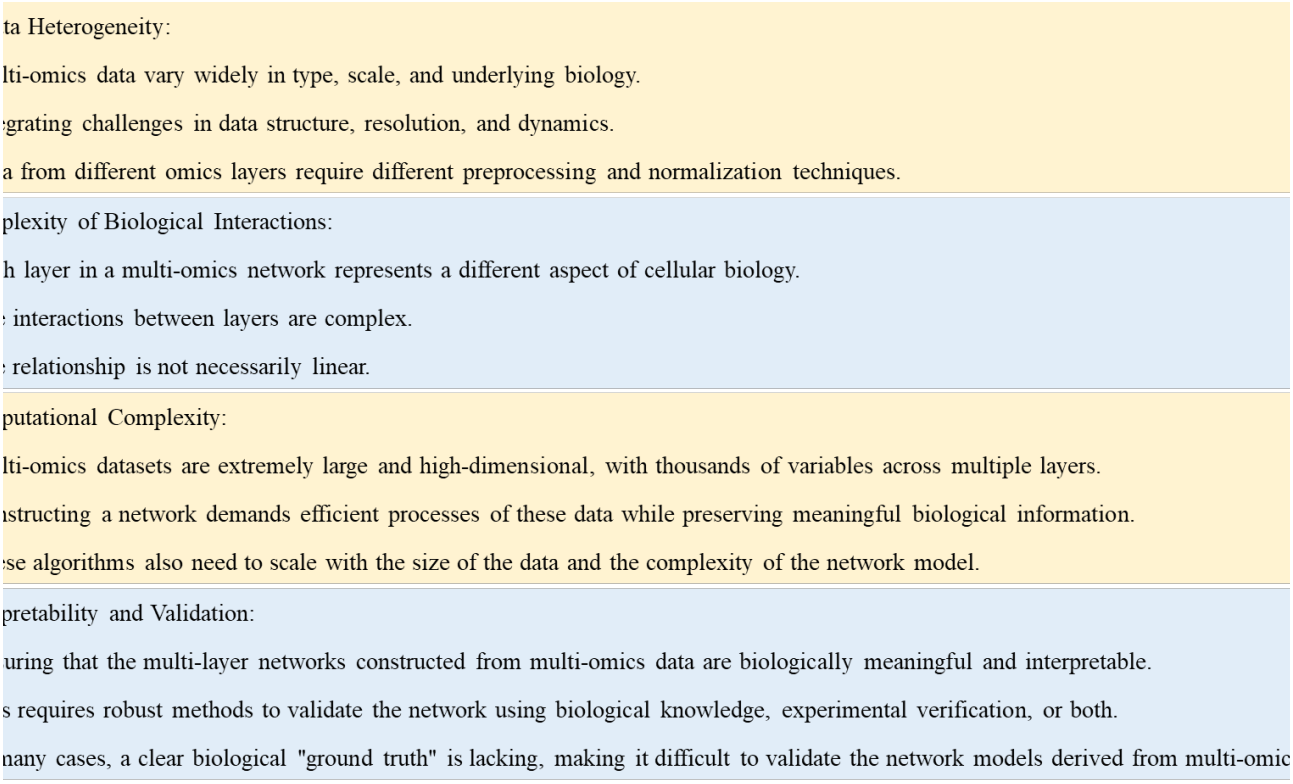

Figure 1. Key challenges in constructing a multilayer network.

To overcome the challenges outlined in Figure 1, several advanced methodologies could be considered. 1) Integration techniques and specific preprocessing methods are used to handle data heterogeneity, standardizing diverse data types and scales for effective analysis. 2) Complex network models, including nonlinear and graph-based models, capture intricate biological interactions across different data layers. 3) Ensuring the interpretability and validity of the network models are achieved through robust statistical methods and biological experiments, such as biological knowledge integration, to validate the computational predictions and ensure the biological relevance and reliability of the network analyses.

Integrating microbial multi-omics with multilayer networks offers significant opportunities to enhance our understanding of microbial ecosystems and the complex interplay between microbes, hosts, and their environment. For example, by using a multilayer network to model different omics layers, key regulatory nodes and interactions that drive community dynamics and responses to environmental changes could be identified (Wang et al., 2021). The study conducted by Zheng et al.(2020) developed a framework to analyze and interpret the co-occurrence networks of microbial communities in the bovine rumen for the methane emission phenotype. Two types of networks are established: co-presence networks, which reflect positive associations, and mutual-exclusion networks, which reflect negative associations. Significant modules within these networks are identified. Biological relevance is ascertained by mapping network nodes to KEGG functional pathways and identifying significant modules linked to methane emissions. Validation extends through cross-referencing findings with external datasets and prior studies to ensure biological significance.

Wang and her colleagues (Wang et al., 2021) explored a novel method for understanding the intricate dynamics of methane production in cattle through a multilayer network analysis with emphasis on the identification of key microbial genes and their interactions that contribute to methane emissions. Using data from 1,461 microbial genes and volatile fatty acids (VFAs) profiles obtained from rumen samples, the study constructed a multilayer heterogeneous network. This network integrates different layers of biological data, including microbial gene co-occurrence and gene-metabolite interactions (Figure 2). The heat diffusion algorithm was employed to navigate through this complex network structure, aiming to pinpoint functional clusters and biomarkers directly or indirectly associated with methane production. The approach revealed that the multilayer network could capture more comprehensive interactions compared to single-layer networks. The network analysis not only identified known biomarkers but also discovered 24 novel ones that show a significant correlation with methane emissions influenced by specific VFAs.

One of our studies (Wang et al., 2022) focuses on using a multilayer network to identify key hubs that play significant roles in microbial methane metabolism. The research constructed a three-layer heterogeneous network, comprising rumen metabolites, microbial genes, and microbial communities (Figure 2). The topological analysis identified 59 entities, including metabolites, microbial genes, and genera, as central nodes within this network, with 23 appearing across multiple topological rankings. Interestingly, the study identified specific microbial genes and microbial genera that consistently appeared in topological hubs, suggesting their significant involvement in methane metabolism. This research enhances understanding of the structural and functional connectivity within the rumen microbiome.

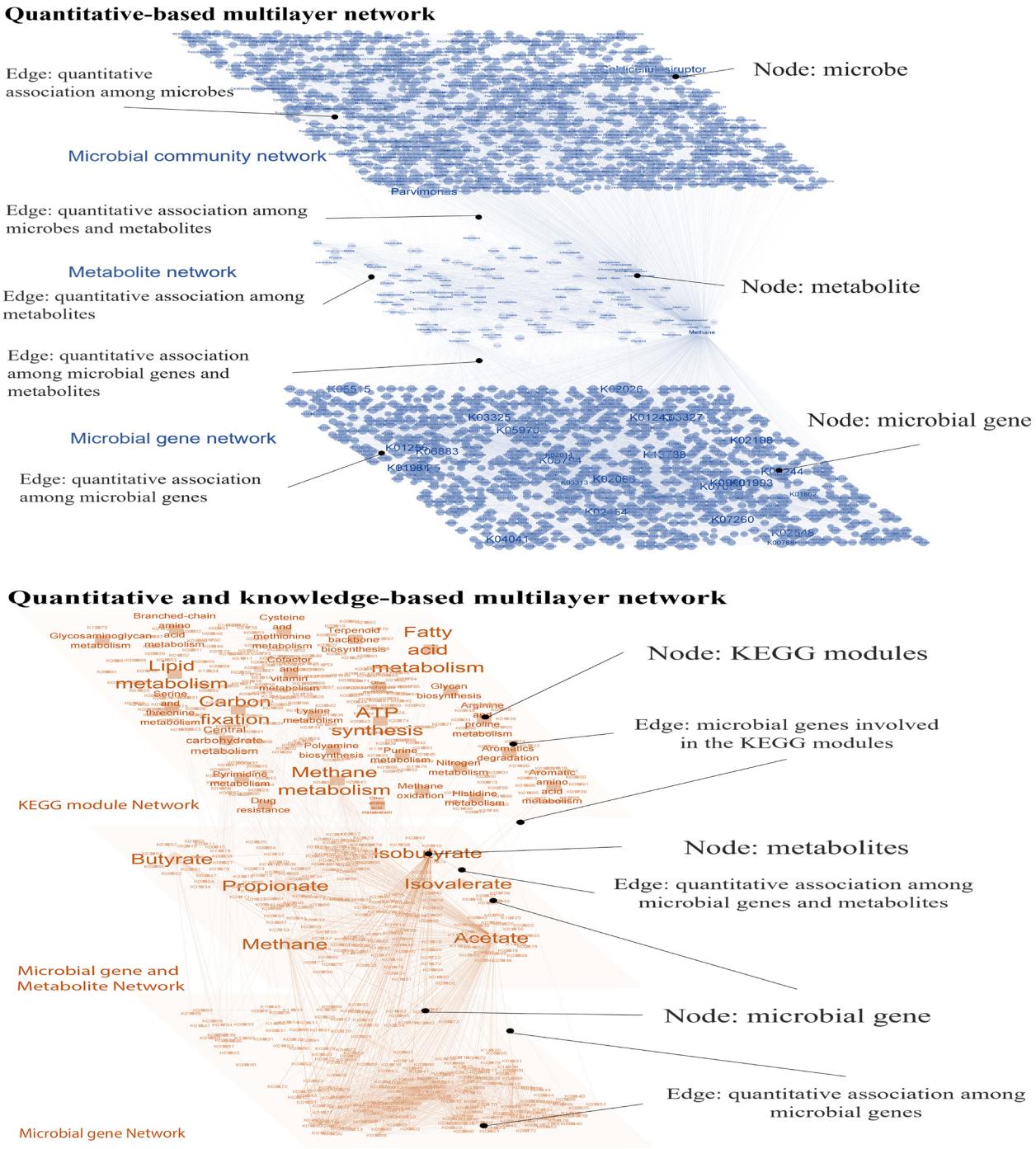

Figure 2. Example of multilayer networks.

Conclusions

Using a multi-layer network to represent the multi-omics of the rumen microbiome offers significant benefits for understanding and manipulating this complex system. This approach enables the integration of diverse datasets, such as genomic, transcriptomic, proteomic, and metabolomic data, providing a holistic view of the microbial interactions and functions within the rumen. By mapping these layers of biological information onto a network, key microbial species and their interactions can be identified to predict the microbiome's response to changes in diet or environment. This comprehensive model enhances the ability to optimize rumen function for better animal health and more efficient feed conversion, crucial for improving sustainable livestock production.

References

BARUPAL, D. K.; FAN, S.; FIEHN, O. Integrating bioinformatics approaches for a comprehensive interpretation of metabolomics datasets. **Current Opinion in Biotechnology**, v. 54, pp.1-9, 2018.

DOMENICO, M. de; SOLÉ-RIBALTA, A.; OMODEI, E.; GÓMEZ, S.; ARENAS, A. Ranking in interconnected multilayer networks reveals versatile nodes. **Nature Communications**, v. 6, article 6868, 2015.

HENDERSON, G.; COX, F.; GANESH, S.; JONKER, A.; YOUNG, W.; GLOBAL RUMEN CENSUS COLLABORATORS; JANSSEN, P. H. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. **Scientific Reports**, v. 5, article 14567, 2015.

KAUR, H.; KAUR, G.; GUPTA, T.; MITTAL, D.; ALI, S. A. Integrating Omics Technologies for a Comprehensive Understanding of the Microbiome and Its Impact on Cattle Production. **Biology (Basel)**, v. 12, i. 9, 1200, 2023.

LEE, B.; ZHANG, S.; POLEKSIC, A.; XIE, L. Heterogeneous Multi-Layered Network Model for Omics Data Integration and Analysis. **Frontiers in Genetics**, v. 10, article 1381, 2020.

WANG, H.; ZHENG, H.; WANG, J.; WANG, C.; WU, F.-X. Integrating omics data with a multiplex network-based approach for the identification of cancer subtypes. **IEEE Transactions on NanoBioscience**, v. 15, i. 4, pp.335-342, 2016.

WANG, M.; WANG, H.; ZHENG, H.; DEWHURST, R. J.; ROEHE, R. A heat diffusion multilayer network approach for the identification of functional biomarkers in rumen methane emissions. **Methods**, v. 192, pp.57-66, 2021.

WANG, M.; WANG, H.; ZHENG, H.; DEWHURST, R. J.; ROEHE, R. A multilayer co-occurrence network reveals the systemic difference of diet-based rumen microbiome associated with methane yield phenotype. In: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, 2020. **Proceedings...** Seoul: IEEE, 2020. pp. 2855-2860.

WANG, M.; WANG, H.; ZHENG, H.; DEWHURST, R. J.; ROEHE, R. Identifying Hub Nodes and Sub-networks from Cattle Rumen Microbiome Multilayer Networks In: INTERNATIONAL CONFERENCE, 11., 2021, Msida. **Advanced Computing:** proceedings. Cham: Springer International Publishing, 2022. pp.165-175. (Communications in Computer and Information Science, v. 1528). IACC 2021.

ZHENG, H.; WANG, H.; DEWHURST, R. J.; ROEHE, R. Improving the Inference of Co-Occurrence Networks in the Bovine Rumen Microbiome. **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, v. 17, i. 3, pp. 858-867, 2020.

Using Microbiome Information in the Context of Animal Sciences

Jeferson M. Lourenco([1&])

[(1)] Department of Animal and Dairy Science – University of Georgia, Athens, GA, USA
[&]Corresponding author: jefao@uga.edu

The animal microbiome, a complex ecosystem of microorganisms residing within and on an animal, plays a critical role in host health, performance, and disease susceptibility. For many years, culture-dependent methods were the only ones available to investigate microbiomes; however, recent advancements in sequencing technologies have made culture-independent techniques more accessible, and their development has generated large amounts of data, allowing researchers to understand the true nature and complexity of animal microbiomes.

In the context of applied Animal Science, research has shown that microbiome composition significantly impacts nutrient digestion, metabolism, and immune function of animals. For instance, recent studies found that important traits such as feed efficiency, the amount of fat in carcass, and even animal health can all be impacted by the microorganisms living in the animal's gastrointestinal (GIT) tract (Welch et al., 2020; Krause et al; 2020; Hu et al., 2023). However, since individual variations among animals play a significant role, there is a need for a more comprehensive understanding of the biological processes that occur in each individual animal. In addition, microbiome information should be evaluated in conjunction with other "omics" data such as metatranscriptomics, proteomics and metabolomics, as well as host phenotypic evaluations (Lourenco; Welch, 2022), since this can provide a more comprehensive view of microbial function and how it impacts the host animal.

Research performed in swine has identified the gut microbiome as a new source of variation to improve animal growth and efficiency (Lu et al., 2018; Khanal et al., 2019; Maltecca et al., 2019). Lu et al. (2018) combined gut microbiome information for over 1,200 pigs with phenotypes, pedigree, and genotypes that were already part of a genetic improvement program. They concluded that using the microbiome data under the form of relatedness among individuals based on similarity of microbial communities significantly improved their model for estimating breeding values for backfat and average daily gain (Lu et al., 2018). Similarly, Khanal et al. (2019) evaluated the impacts of gut microbiome on carcass composition and quality at different stages of the swine production cycle and found a significant contribution of the microbiome information to all traits analyzed. Moreover, adding the microbiome information to their model significantly improved the estimation of carcass composition traits, indicating that an important portion of the genomic variance was contributed by the gut microbiome. Results like these may lead to establishing a new approach for genetic evaluations through the addition of gut microbial information into the prediction models (Khanal et al., 2019).

In cattle, the GIT microbial population produces a huge array of enzymes that help animals digest their feed. Furthermore, there is evidence indicating that the microbial population of the GIT tract plays a key role in the way cattle perform (Hernandez-Sanabria et al., 2012; Myer et al., 2015; Shabat et al., 2016; Lourenco et al., 2019). In addition, research has shown that the host can exert some control over its own GIT microbiome through several biological mechanisms, such as by producing more or less compounds that serve as substrates for the microorganisms (e.g. production of mucus by the gut; Ouwerkerk et al., 2013). Since this kind of mechanism is inherent to the host animal and varies from individual to individual, certain animals can develop GIT microbiomes with unique characteristics. Moreover, in many instances, those individualities in microbial composition are associated with more desirable phenotypic traits such as better feed efficiency or greater carcass marbling scores (Welch et al., 2020; Krause et al; 2020).

Another application of animal microbiome analysis is providing evidence for the development of targeted probiotics and prebiotics for specific animal species and health needs. Probiotics are live microorganisms that confer health benefits to the host, while prebiotics are dietary fibers that selectively stimulate the growth of beneficial gut bacteria. By identifying fundamental microbial species associated with desired functionalities, Animal Scientists can develop targeted probiotic formulations to modulate the gut microbiota and promote animal health. Similarly, prebiotics can be designed to selectively nourish beneficial bacteria identified through microbiome profiling (Hill et al., 2014). Such knowledge can be used to formulate targeted diets that optimize gut microbial function and ultimately enhance animal performance.

Concerning animal health, a balanced gut microbiome can act as a barrier against pathogenic bacteria. Disruptions in this balance, known as dysbiosis, can predispose animals to infections. For example, in chicken, the onset of necrotic enteritis is associated with a shift in the gut microbiota (Birchler et al., 2001), and birds challenged with Clostridium perfringens showed signs of gut dysbiosis. Studies like that, exploring the link between specific gut bacteria and infection, can inform strategies to enhance disease resistance, and might allow us to identify bacterial strains that stimulate the host's immune system or even produce antimicrobial compounds. This information could then be incorporated into probiotic formulations to strengthen the animal's natural defenses against pathogens.

In conclusion, the exploration of the animal microbiome offers a transformative lens through which we can understand and improve animal health, performance, and welfare. Utilizing microbiome information in Animal Science holds significant promise for developing novel strategies in nutrition, disease prevention, and personalized animal care. As research progresses, integrating microbiome data into diverse aspects of animal management will undoubtedly translate into a more sustainable and responsible approach to animal production.

References

BIRCHLER, T.; SEIBL, R.; BÜCHNER, K.; LOELIGER, S.; SEGER, R.; HOSSLE, J. P.; AGUZZI, A.; LAUENER, R. P. Human Toll-like receptor 2 mediates induction of the antimicrobial peptide human beta-defensin 2 in response to bacterial lipoprotein. **European journal of immunology**, v. 31, i. 11, pp. 1521-1541, 2001.

HERNANDEZ-SANABRIA, E.; GOONEWARDENE, L. A.; WANG, Z.; DURUNNA, O. N.; MOORE, S. S.; GUAN, L. L. Impact of feed efficiency and diet on adaptive variations in the bacterial community in the rumen fluid of cattle. **Applied and environmental microbiology**, v. 78, i. 4, pp. 1203-1204, 2012.

HILL, C.; GUARNER, F.; REID, G.; GIBSON, G. R.; MERENSTEIN, D. J.; POT, B.; MORELLI, L.; CANANI, R. B.; FLINT, H. J.; SALMINEN, S.; CALDER, P. C.; SANDERS, M. E. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term prebiotic. **Nature review, gastroenterology and hepatology**, v. 11, i. 8, pp. 506-514, 2014.

HU, R.; LI, S.; DIAO, H.; HUANG, C.; YAN, J.; WEI, X.; ZHOU, M.; HE, P.; WANG, T.; FU, H.; ZHONG, C.; MAO, C.; WANG, Y.; KUANG, S.; TANG, W. The interaction between dietary fiber and gut microbiota, and its effect on pig intestinal health. **Frontiers in Immunology**, v. 14, 1095740, 2023.

KHANAL, P.; MALTECCA, C.; SCHWAB, C.; FIX, J.; TIEZZI, F. Microbiability of meat quality and carcass composition traits in swine. Journal of animal breending and genetics, v. 138, i. 2, pp. 223-236, 2019.

KRAUSE, T. R.; LOURENCO, J. M.; WELCH, C. B.; ROTHROCK, M. J.; CALLAWAY, T. R.; PRINGLE, T. D. The relationship between the rumen microbiome and carcass merit in Angus steers. **Journal of Animal Science**, v. 98, i. 9, skaa287, 2020.

LOURENCO, J. M.; CALLAWAY, T. R.; KIERAN, T. J.; GLENN, T. C.; MCCANN, J. C.; STEWART JUNIOR, R. L. Analysis of the Rumen Microbiota of Beef Calves Supplemented During the Suckling Phase. **Frontiers in Microbiology**, v. 10, 1131, 2019.

LOURENCO, J. M.; WELCH, C. B. Using microbiome information to understand and improve animal performance. **Italian Journal of Animal Science**, v. 21, i. 1, pp. 899-913, 2022.

LU, D.; TIEZZI, F.; MALTECCA, C. Gut Microbiome Provides A New Source of Variation to Improve Growth Efficiency in Crossbred Pigs. **Journal of Animal Science**, v.96, supp. 3, pp. 112-113, 2018.

MALTECCA, C.; LU, D.; SCHILLEBEECKX, C.; MCNULTY, N. P.; SCHWAB, C.; SHULL, C.; TIEZZI, F. Predicting growth and carcass traits in swine using microbiome data and machine learning algorithms. **Scientific Reports**, v. 9, i. 1, article 6574, 2019.

MYER, P. R.; SMITH, T. P. L.; WELLS, J. E.; KUEHN, L. A.; FREETLY, H. C. Rumen microbiome from steers differing in feed efficiency. **PloS One**, v. 10, i. 6, e0129174, 2015.

OUWERKERK, J. P.; VOS, W. M. de; BELZER, C. Glycobiome: bacteria and mucus at the epithelial interface. **Best practice & research, clinical gastroenterology**, v. 27, i. 7, pp. 25-38, 2013.

SHABAT, S. K.; SASSON, G.; DORON-FAIGENBOIM, A.; DURMAN, T.; YAACOBY, S.; BERG MILLER, M. E.; WHITE, B. A.; SHTERZER, N.; MIZRAHI, I. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. **The ISME journal**, v. 10, i. 12, pp. 2958-2972, 2016.

WELCH, C. B.; LOURENCO, J. M.; DAVIS, D. B.; KRAUSE, T. R.; CARMICHAEL, M. N.; ROTHROCK, M. J.; PRINGLE, T. D.; CALLAWAY, T. R. The impact of feed efficiency selection on the ruminal, cecal, and fecal microbiomes of Angus steers from a commercial feedlot. **Journal of Animal Science**, v. 98, i. 7, skaa230, 2020.

Can we breed cattle with more efficient rumen microbiome?

Leluo Guan[1]

[1] Faculty of Land and Food Systems, the University of British Columbia, Vancouver, Canada

Short title: genetic breeding of rumen microbiome

Production of adequate animal proteins for continuously increasing human population is an urgent global task. In many nations, ruminant protein is an integral component of the human diet and has contributed to the development and elimination of poverty in human society. Cattle industry, including both beef & dairy, is one of the most important agricultural commodities, which provides the animal proteins (milk and meat) to humans. Over the past decade, research performed globally suggests differences in rumen microbiota are associated with cattle production and health traits, such as feed efficiency (Guan et al., 2008), methane (CH_4) emission (Tapio et al. 2017), milk yield and composition (Xue et al., 2019), and so on, highlighting the important role of rumen microbial digestion and fermentation in regulating key production traits. Hence, rumen microbiota is a potential target for manipulation to improve ruminant productivity and health, as well as reduce CH_4 emissions. The rumen microbiome consists of hundreds of species that include prokaryotes (bacteria and methanogens), and eukaryotes (protozoa and fungi), making it one of the most sophisticated microbial ecosystems known. Bacteria are the most abundant ruminal microorganisms, despite the differences in ruminant species, animal breed, diet, environment, and geographic locations, the recent studies showed that there is a group of core bacteria in the rumen, with *Prevotella*, *Butyrivibrio*, and *Ruminococcus*, as well as unclassified Lachnospiraceae, *Ruminococcaceae*, *Bacteroidales*, and *Clostridiales* were the most abundant bacteria (Henderson et al., 2015). The second most abundant ruminal microorganisms are archaea with methanogens accounting for up to 99% (Moissl-Eichinger et al., 2018). In addition, the ciliated protozoa in the rumen have a larger size than bacteria and archaea, accounting for up to 50% of the rumen microbial biomass (Sylvester et al., 2004). Recent studies found that protozoal communities were more variable than bacteria and archaea (Henderson et al., 2015). The rumen anaerobic fungi all belong to phylum *Neocallimastigomycota*, which accounts for up to 20% of the microbial biomass (Edwards et al., 2017). The above microbial groups interact with each other, work synergically degrades most complex plant polysaccharides, producing volatile fatty acids (VFAs), microbial protein, and vitamins that provide nutrients to meet the host's requirement for maintenance, reproduction, growth, meat and milk production.

Rumen microbes are the key players of enteric CH_4 production, with CO_2 being reduced to CH_4 by methanogens through ruminal methanogenesis. This process contributes to the low partial pressure of hydrogen in the rumen needed for efficient fermentation of plant cell wall carbohydrates (Ungerfeld, 2020). In the meantime, enteric CH_4 emissions also represent a loss of approximately 2-12% of the gross energy intake of the host (Beauchemin; McGinn, 2008), negatively impacting production efficiency. Therefore, it has become an urgent global task to reduce CH_4 emissions and improve feed digestive efficiency from ruminants, while improving the efficiency of production so as to satisfy the growing demand for meat and milk.

To date, most studies to date only focused on bacteria and/or archaea, and only used correlation to reveal associations between rumen microbes and these traits. Proper methods are lacking to determine the specific causal contribution of rumen microbes to traits. Although an increasing number of studies are revealing factors affecting rumen microbiome composition and function, including diet, environment, and genetic background of host, these effects were nested and not readily separated, constraining the development and application of more effective rumen microbiome manipulation strategies. To date, it is unclear as to the relative impact of each of these factors on the rumen microbiome, and consequently, key production traits such as feed efficiency and CH_4 emissions. Although diet and environment can be key factors that change the rumen microbiome, recent research has highlighted that the rumen microbiome can vary between animals fed the same diet and reared in the same environment, suggesting the host animal could also play a role in determining which microbes inhabit the rumen. Recent research has highlighted the individual variation of rumen microbiota in both beef (Zhou et al., 2018) and dairy cattle (Jami; Mizrahi, 2012) when animals were fed the same diet and managed in the same environment, and recent evidence has revealed the heritability of rumen microbiomes. Difford et al. (2018) first reported the heritability of rumen bacteria and archaea communities in dairy cows, two recent landmark studies generated heritability estimates of rumen microbial features in both beef (Li et al., 2019) and dairy cattle (Wallace et al., 2019). These finding suggest that the host has an impact on the rumen microbiome.

Despite previous studies that reported heritable rumen microbes in beef and dairy cattle, a better understanding of the interactions, function and effect of heritable microbes on cattle phenotypes is needed (Zang et al. 2022). The latter work demonstrated that heritable bacteria make a stronger contribution to milk traits in dairy cattle (32 bacteria at species level; h^2>0.20; microbiability 0.16 - 0.33%), as compared to nonheritable bacteria (674 nonheritable bacteria at species level; h^2<0.20; microbiability 0.03 - 0.06%), confirming the importance of host genetics and rumen microbiota on performance traits (Zang et al. 2022). Studies have reported low positive genetic correlations between milk traits and CH_4 production traits in dairy cattle (Hossein-Zadeh, 2022; Pszczola et al. 2019). Heritable rumen microbiome features in beef cattle have also been studied, but their relationship to interactions between feed efficiency and CH_4 emission traits has yet to be investigated. Rumen microbiome features (diversity indices, relative abundance, total bacteria copy number) are moderately heritable (h^2>0.15) in beef cattle, while also being associated with host feed efficiency traits (Li et al., 2019). Host SNPs were also associated with specific heritable (h^2>0.15) rumen microbial taxa, and these same SNP markers overlapped with QTL for feed efficiency (Li et al., 2019). However, heritability of the rumen microbes was low to moderate as compared to other traits such as weaning weight. The dietary and farm environment effects on rumen microbiome are dominant when cattle go through dietary transition as well as move to the different production systems. In contrast, genetic impacts should be more important among cattle raised and managed within the same environment. Therefore, similar to other known traits, rumen microbes can be affected by genetics (G), environment (E) and interaction between them (G×E). To date, it is unclear how large the contribution is of each factor (genetics, diet, environment and interactions) on the rumen microbiome, and to what extent the differences in the presence/absence or abundance of rumen microbes and their fermentation metabolites can impact two key production traits (feed efficiency and methane). In summary, it suggests that host plays an important role in the rumen microbiota of cattle, and manipulating selected heritable microbiomes with genetic selection and breeding could be a possible strategy for improving feed efficiency of cattle and reducing methane emission from ruminants. The knowledge and information obtained through such aspects have paved the step stone for the research area to improve cattle production through genetic breeding of targeted rumen microbiome for improved feed efficiency and reduced methane emission.

References

BEAUCHEMIN, K.; McGINN, S. Reducing methane in dairy and beef cattle operations: what is feasible? **Prairie Soil Crop**, v. 1, pp. 17-21, 2008.

EDWARDS, J. E.; KINGSTON-SMITH, A. H.; JIMENEZ, H. R.; HUWS, S. A.; SKØT, K. P.; GRIFFITH, G. W.; MCEWAN, N. R.; THEODOROU, M. K. Dynamics of initial colonization of nonconserved perennial ryegrass by anaerobic fungi in the bovine rumen. **Fems Microbiology Ecology**, v. 66, i. 3, pp. 537-545, 2008.

GUAN, L. L.; NKRUMAH, J. D.; BASARAB, J. A.; MOORE, S. S. Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. **FEMS Microbiology Letters**, v. 288, i. 1, pp. 85-91, 2008.

HENDERSON, G.; COX, F.; GANESH, S.; JONKER, A.; YOUNG, W.; GLOBAL RUMEN CENSUS COLLABORATORS; JANSSEN, P. H. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. **Scientific Reports**, v. 5, article 14567, 2015.

HOSSEIN-ZADEH, G. N. Estimates of the genetic contribution to methane emission in dairy cows: a meta-analysis. **Scientific Reports**, v. 12, i. 1, article 12352, 2022.

JAMI, E.; MIZRAHI, I. Composition and similarity of bovine rumen microbiota across individual animals. **PLoS One**, v. 7, i. 3, e33306, 2012.

LI, F.; LI, C.; CHEN, Y.; LIU, J.; ZHANG, C.; IRVING, B.; FITZSIMMONS, C.; PLASTOW, G.; GUAN, L. L. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. **Microbiome**, v. 7, i. 1, article 92, 2019.

MOISSL-EICHINGER, C.; PAUSAN, M.; TAFFNER, J.; BERG, G.; BANG, C.; SCHMITZ, R. A. Archaea are interactive components of complex microbiomes. **Trends in Microbiology**, v. 26, i. 1, pp. 70-85, 2018.

PSZCZOLA, M.; CALUS, M. P. L.; STRABEL, T. Short communication: Genetic correlations between methane and milk production, conformation, and functional traits. **Jounal Dairy Science**, v. 102, i. 6, pp. 5342-5346, 2019.

SYLVESTER, J. T.; KARNATI, S. K. R.; YU, Z.; MORRISON, M.; FIRKINS, J. L. Development of an assay to quantify rumen ciliate protozoal biomass in cows using real-time PCR. **Journal of Nutrition**, v. 134, i. 12, pp. 3378-3384, 2004.

TAPIO, I.; SNELLING, T. J.; STROZZI, F.; WALLACE, R. J. The ruminal microbiome associated with methane emissions from ruminant livestock. **Journal of Animal Science and Biotechnology**, v. 8, article 7, 2017.

UNGERFELD, E. M. Metabolic Hydrogen Flows in Rumen Fermentation: Principles and Possibilities of Interventions. **Frontiers in Microbiology**, v. 11, article 589, 2020.

WALLACE, R. J.; SASSON, G.; GARNSWORTHY, P. C.; TAPIO, I.; GREGSON, E.; BANI, P.; HUHTANEN, P.; BAYAT, A. R.; STROZZI, F.; BISCARINI, F.; SNELLING, T. J.; SAUNDERS, N.; POTTERTON, S. L.; CRAIGON, J.; MINUTI, A.; TREVISI, E.; CALLEGARI, M. L.; CAPPELLI, F. P.; CABEZAS-GARCIA, E. H.; VILKKI, J.; PINARES-PATINO, C.; FLIEGEROVÁ, K. O.; MRÁZEK, J.; SECHOVCOVÁ, H.; KOPEČNÝ, J.; BONIN, A.; BOYER, F.; TABERLET, P.; KOKOU, F.; HALPERIN, E.; WILLIAMS, J. L.; SHINGFIELD, K. J.; MIZRAHI, I. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. **Science Advances**, v. 5, i. 7, eaav8391, 2019.

XUE, M. Y.; SUN, H. Z.; WU, X. H.; LIU, J. X.; GUAN, L. L. Multi-omics reveals that the rumen microbiome and its metabolome together with the host metabolome contribute to individualized dairy cow performance. **Microbiome**, v. 8, article 64, 2020.

ZANG, X.-W.; SUN, H.-Z.; XUE, M.-Y.; ZHANG, Z.; PLASTOW, G.; YANG, T.; GUAN, L. L.; LIU, J.-X. Heritable and Nonheritable Rumen Bacteria Are Associated with Different Characters of Lactation Performance of Dairy Cows. **mSystems**, v. 7, i. 5, e0042222, 2022.

ZHOU, M.; PENG, Y. J.; CHEN, Y.; KLINGER, C. M.; OBA, M.; LIU, J. X.; GUAN, L. L. Assessment of microbiome changes after rumen transfaunation: implications on improving feed efficiency in beef cattle. **Microbiome**, v. 6, article 62, 2018.

Capybara harbors unprecedented enzymatic systems to degrade plant polysaccharides

Lucélia Cabral[1,2§], Gabriela F. Persinoti[1], Douglas A. A. Paixão[1], Marcele P. Martins[1,3], Mariana A. B. Morais[1], Mariana Chinaglia[1,3], Mariane N. Domingues[1], Mauricio L. Sforca[4], Renan A. S. Pirolla[1], Wesley C. Generoso[1], Clelton A. Santos[1], Lucas F. Maciel[1], Nicolas Terrapon[5,6], Vincent Lombard[5,6], Bernard Henrissat[7,8], Mario T. Murakami[1]

[1] Brazilian Biorenewables National Laboratory, Brazilian Center for Research in Energy and Materials, Campinas, SP, Brazil.

[2] Biology Institute, University of Brasília (UNB), Brasília, DF, Brazil.

[3] Graduate Program in Functional and Molecular Biology, Institute of Biology, University of Campinas, Campinas, SP, Brazil.

[4] Brazilian Biosciences National Laboratory, Brazilian Center for Research in Energy and Materials, Campinas, SP, Brazil.

[5] The Institut National de la Recherche Agronomique, USC 1408 AFMB, 13288 Marseille, France.

[6] Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille Université, Marseille, France.

[7] Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, 2800 Kgs Lyngby, Denmark.

[8] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

§Corresponding author: luc.g.cabral@gmail.com

The digestive tract of herbivores has been an profuse source of diverse enzymatic mechanisms for lignocellulose depolymerization. The microbiota of foregut (rumen) fermenters has been studied as a model system along the last decades, which resulted in the discovery of sophisticated systems to degrade recalcitrant plant fibers (Krause, et al., 2003, Hess, et al., 2011). A less explored and equally effective class of herbivores is the hindgut fermenters. Similar to foregut fermenters, the digestion is accomplished by a symbiotic microbial community, but in a single fermentation chamber (Stevens; Hume, 1998).

The capybara (*Hydrochoerus hydrochaeris*) is the largest living rodent, typically found in the Pantanal wetlands and Amazon basin, and it is also known as “the master of the grasses” due to its diet based on gramineous and aquatic plants. In this animal, the fermentation takes place in the cecum that corresponds to almost three-quarters of the gastrointestinal tract, reaching a digestive efficiency comparable to that of ruminants (Kiani, et al., 2019). The wild capybara animals dwelling the Southeast region of Brazil have incorporated sugarcane in their diet for decades (Polo, et al., 2018), which makes their cecal microbiome particularly attractive for the discovery of enzymatic mechanisms for the depolymerization of this industrially relevant feedstock and related grasses.

In this way, to elucidate the enzymatic strategies employed by the Brazilian capybara microbiota for plant cell wall depolymerization, we comprehensively investigated this gut microbial community combining 16S rRNA gene targeting sequencing (16S), metagenomics (MG), metatranscriptomics (MT), and nuclear magnetic resonance (NMR) based metabolomics, with carbohydrate enzymology and X-ray crystallography, which ultimately led to the discovery of two families, according to the carbohydrate-active enzymes database (CAZy) The taxonomic structure of the capybara gut microbiota from fresh cecal and rectal samples is mainly constituted by Bacteria. 16S rRNA gene-based taxonomic analysis, corroborated by 16S rRNA reads recovered from metagenome (16S_MG), MG, and MT data sets, indicates that the most abundant bacteria found in this microbiota are members from the phyla Firmicutes and Bacteroidetes, Fusobacteria and Proteobacteria. A prevalence of Firmicutes, Bacteroidetes, and Proteobacteria was also observed in the gut microbiota of beaver, horse, rabbit, and koala (Pratama, et al., 2019, Armstrong, et al., 2018, Morrison, et al., 2018, Velasco-Galilea, et al., 2018, Barker, et al. 2013). In this regard, the reconstruction of metagenome-assembled genomes (MAGs) revealed several taxonomic novelties, representing either unknown species or genera from bacterial families that are recognized as plant fiber degraders such as Fibrobacteraceae and Bacteroidaceae. Among these taxonomic novelties, MAGs 41 to 44, assigned to the uncultured UBA932 family, are phylogenetically grouped and may represent an unprecedented Bacteroidetes genus with a genetic potential for lignocellulose degradation.

In order to understand the ability of capybara gut microbiota to convert plant polysaccharides into edible sugars, MG and MT data from cecal and rectal samples were investigated to determine the genomic potential associated with plant fiber depolymerization. A total of 7377 putative CAZymes genes encoding for 106 glycoside hydrolases (GH), 11 carbohydrate esterases (CE), and 10 polysaccharide lyases (PL) families were identified, of which 517 genes presented a modular architecture. The most abundant CAZymes identified are members of the families GH3, GH2, and GH1 (by decreasing abundance), which is in agreement with

that reported for other gut microbiomes such as human, swine, and cattle rumen (Li et al., 2020). In the CAZyme repertoire of this microbiota neither cellulases from families GH6, GH7, and GH48, nor cellulosomes, assessed by the presence of cohesin and dockerin domains associated with cellulases, could be identified. The CAZyome (CAZyme inventory) analysis of the capybara gut microbiota indicates Fibrobacteres as the main drivers for cellulose breakdown, whereas the numerous Bacteroidetes PULs and clusters of CAZymes confer to this community a myriad of enzymatic strategies to tackle with complex and diverse hemicelluloses and pectins commonly present in gramineous and aquatic plants, major components of capybara diet.

Metabolite profiling shows high performance on the conversion of dietary fibers into short-chain fatty acids. The major fermentation products measured in the capybara gut were short-chain fatty acids (SCFAs), among more than 40 metabolites detected by NMR spectroscopy-based metabolomics. The most abundant metabolites observed in cecal and rectal samples were acetate, propionate, and butyrate. Metabolic pathway reconstruction analysis shows that acetate can be produced by any of the bacterial MAGs recovered from capybara gut microbiome, which is in agreement with the high abundance of this metabolite in both cecal and rectal samples. On the other hand, the expression analysis of key genes involved in the butyrate pathway (atoA/D genes) indicates that *Firmicutes Ileibacterium* sp. MAG6 and Megasphaera sp. MAG33 are likely the major butyrate-producing bacteria in the capybara gut.

The joint MG and MT analysis of the capybara gut microbiome revealed several expressed genes annotated as hypothetical proteins. Some of these genes display remote similarity to CAZy members, with sequence identity ranging from 10 to 21%, suggesting a potential function in the processing of plant polysaccharides, but requiring further functional investigation. One of these hypothetical proteins (SEQ ID PBMDCECB_ 44807, named here as CapGH173), was recovered from *Bacteroidales bacterium* MAG42, a discovered genome that expands the uncultured UBA932 family. CapGH173 orthologues are found in Actinobacteria, Firmicutes, Verrucomicrobia and Bacteroidetes MAGs recovered from diverse sources such as rumen, feces, gut, and oral microbiotas, being the closest sequence from a rumen-derived MAG (UBA2817) from the uncultured RC9 group (Parks et al., 2017). In the *Bacteroidales bacterium* MAG42, CapGH173 is found in a predicted PUL that includes enzymes from families GH2 and GH78. In addition, as presented in the taxonomic analysis, *Bacteroidaceae bacterium* MAG57 encompasses a remarkable number of CAZyme-encoding genes including a gene cluster targeting arabinoxylan (CC102), an abundant hemicellulose in secondary cell walls of sugarcane and other grasses. This cluster encodes two exo-enzymes from families GH43 and GH97, and an unconventional GH10 member with an unknown 45 kDa N-terminal domain.

In this study, we investigated how the gut microbiota of the largest living rodent, capybara (*Hydrochoerus hydrochaeris*) known as “master of the grasses”, can efficiently depolymerize and utilize recalcitrant plant polysaccharides. An integrated meta-omics approach focused on the enzymatic and metabolic capabilities for plant fiber breakdown unveiled that cellulose degradation in this community is not accomplished by classical mechanisms involving cellobiohydrolases or cellulosomes. The complex and diverse composition of hemicellulosic and pectic polysaccharides present in gramineous and aquatic plants are tackled by a broad number of CAZymes organized in PULs found in the multiple recovered Bacteroidetes MAGs, which in part resembles to that from human gut Bacteroidetes species such as the PULs for mixed-linkage β-glucans (Tamura, et al., 2017) and xyloglucans (Larsbrink, et al., 2014, Hemsworth, et al. 2016).

In conclusion, this work sheds light on the enzymatic apparatus and metabolic pathways employed by the gut microbiota from the Amazon monogastric semi-aquatic herbivore, capybara, for the breakdown and utilization of recalcitrant dietary polysaccharides. The discovery of several taxonomic novelties associated with plant fiber degradation along with the founding of two CAZy families highlight this microbiota as an untapped source of CAZymes and enzymatic systems of biotechnological interest.

References

ARMSTRONG, Z.; MEWIS, K.; LIU, F.; MORGAN-LANG, C.; SCOFIELD, M.; DURNO, E.; CHEN, H. M.; MEHR, K.; WITHERS, S. G.; HALLAM, S. J. Metagenomics reveals functional synergy and novel polysaccharide utilization loci in the Castor canadensis fecal microbiome. I**SME Journal**, v. 12, i. 11, pp. 2757-2769, 2018.

BARKER, C. J.; GILLETT, A.; POLKINGHORNE, A.; TIMMS, P. Investigation of the koala (Phascolarctos cinereus) hindgut microbiome via 16S pyrosequencing. **Veterinay Microbiology**, v. 167, i. 3-4, pp. 554-564, 2013.

HEMSWORTH, G. R.; THOMPSON, A. J.; STEPPER, J.; SOBALA, Ł. F.; COYLE, T.; LARSBRINK, J.; SPADIUT, O.; GODDARD-BORGER, E. D.; STUBBS, K. A.; BRUMER, H.; DAVIES, G. J. Structural dissection of a complex Bacteroides ovatus gene locus conferring xyloglucan metabolism in the human gut. **Open Biology**, v. 6, i. 7, article 160142, 2016.

HESS, M.; SCZYRBA, A.; EGAN, R.; KIM, T.-W.; CHOKHAWALA, H.; SCHROTH, G.; LUO, S.; CLARK, D. S.; CHEN, F.; ZHANG, T.; MACKIE, R. I.; PENNACCHIO, L. A.; TRINGE, S. G.; VISEL, A.; WOYKE, T.; WANG, Z.; RUBIN, E. M. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. **Science**, v. 331, i. 6016, pp. 463-467, 2011.

KIANI, A.; CLAUSS, M.; ORTMANN, S.; VENDL, C.; CONGDON, E. R.; HERRERA, E. A.; KREUZER, M.; SCHWARM, A. Digestive physiology of captive capybara (Hydrochoerus hydrochaeris). **Zoo Biology**, v. 38, i. 2, pp. 167-179 2019.

KRAUSE, D. O.; DENMAN, S. E.; MACKIE, R. I.; MORRISON, M.; RAE, A. L.; ATTWOOD, G. T.; MCSWEENEY, C. S. Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. **FEMS Microbiology Reviews**, v. 27, i. 5, pp. 663-693, 2003.

LARSBRINK, J.; THOMPSON, A. J.; LUNDQVIST, M.; GARDNER, J. G.; DAVIES, G. J.; BRUMER, H. A complex gene locus enables xyloglucan utilization in the model saprophyte Cellvibrio japonicus. **Molecular Microbiology**, v. 94, i. 2, pp. 418-433, 2014.

LI, J.; ZHONG, H.; RAMAYO-CALDAS, Y.; TERRAPON, N.; LOMBARD, V.; POTOCKI-VERONESE, G.; ESTELLÉ, J.; POPOVA, M.; YANG, Z.; ZHANG, H.; LI, F.; TANG, S.; YANG, F.; CHEN, W.; CHEN, B.; LI, J.; GUO, J.; MARTIN, C.; MAGUIN, E.; XU, X.; YANG, H.; WANG, J.; MADSEN, L.; KRISTIANSEN, K.; HENRISSAT, B.; EHRLICH, S. D.; MORGAVI, D. P. A catalog of microbial genes from the bovine rumen unveils a specialized and diverse biomass-degrading environment. **GigaScience**, v. 9, i. 6, giaa057, 2020.

MORRISON, P. K.; NEWBOLD, C. J.; JONES, E.; WORGAN, H. J.; GROVE-WHITE, D. H.; DUGDALE, A. H.; BARFOOT, C.; HARRIS, P. A.; ARGO, C. M. The equine gastrointestinal microbiome: Impacts of age and obesity. **Frontiers in Microbiology**, v. 9, article 3017, 2018.

PARKS, D. H.; RINKE, C.; CHUVOCHINA, M.; CHAUMEIL, P.-A.; WOODCROFT, B. J.; EVANS, P. N.; HUGENHOLTZ, P.; TYSON, G. W. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. **Nature Microbiology**, v. 2, i. 11, pp. 1533-1542, 2017.

POLO, G.; MERA ACOSTA, C.; LABRUNA, M. B.; FERREIRA, F.; BROCKMANN, D. Hosts mobility and spatial spread of Rickettsia rickettsii. **PLoS Computational Biology**, v. 14, i. 12, e1006636, 2018.

PRATAMA, R.; SCHNEIDER, D.; BÖER, T.; DANIEL, R. First insights into bacterial gastrointestinal tract communities of the Eurasian beaver (Castor fiber). **Frontiers in Microbiology**, v. 10, article 1646, 2019.

STEVENS, C. E.; HUME, I. D. Contributions of microbes in vertebrate gastrointestinal tract to production and conservation of nutrients. **Physiological Reviews**, v. 78, i. 2, pp. 393-427, 1998.

TAMURA, K.; HEMSWORTH, G. R.; DÉJEAN, G.; ROGERS, T. E.; PUDLO, N. A.; URS, K.; JAIN, N.; DAVIES, G. J.; MARTENS, E. C.; BRUMER, H. Molecular mechanism by which prominent human gut Bacteroidetes utilize mixed-linkage beta-glucans, major health-promoting cereal polysaccharides. **Cell Reports**, v. 21, i. 2, pp. 417-430, 2017.

VELASCO-GALILEA, M.; PILES, M.; VIÑAS, M.; RAFEL, O.; GONZÁLEZ-RODRÍGUEZ, O.; GUIVERNAU, M.; SÁNCHEZ, J. P. Rabbit microbiota changes throughout the intestinal tract. **Frontiers in Microbiology**, v. 9, artice 2144, 2018.

Unraveling the Nelore (*Bos indicus*) hologenome

Luciana Correia de Almeida Regitano[1], Priscila Silva Neubern de Oliveira[2], Bruno Gabriel Nascimento Andrade[3]

[1] Embrapa Pecuária Sudeste, Caixa Postal 339, São Carlos, SP, Brasil, 13560-970, CNPq Fellowship,

[2] Programa de Pós Graduação em Genética e Evolução da Universidade Federal de São Carlos, São Carlos, SP, Brasil. Capes Scholarship.

[3] Department of Computer Sciences, Munster Technological University.

[1] Corresponding author: luciana.regitano@embrapa.br.

Ruminants have evolved to be able to digest complex molecules by segmenting their stomach to include a large fermentation chamber, the rumen, in which a diverse and abundant microbiota flourishes. Among other functions, microbiome-driven enzymatic fermentation allows complex saccharides such as cellulose to be broken down into short-chain fatty acids (SCFAs), which are the major source of energy for ruminant's metabolism.

Because of their role in the digestion and absorption of nutrients, microbiota constitution is expected to impact animal health and development, while there is growing evidence of a much broader contribution to the host, including neural and immune development (Yu et al., 2022, Sampson; Mazmanian, 2015). In turn, the host system also provides the necessary environment for microorganisms' colonization. Thus, a mutual benefit is established and the disequilibrium of this relationship frequently leads to disease.

Rumen microbiome composition has been associated with variation in feed efficiency, methane emission, milk, and fat yield (Zhang et al., 2021), as well as with susceptibility to diseases, e.g. mastitis (Hu et al., 2023). If the results with animal models apply to ruminants, a number of other traits including behavioral traits are expected to be influenced by the microbiome.

The term microbiability has been proposed to designate the fraction of phenotypic variation attributed to microbiome composition while the term holobiability was designated to refer to the host's genome x microbiome interaction, and recent research has demonstrated an inheritable microbiome composition. For instance, the microbiome proportion of variance explained by dairy cows' SNPs was found as highly variable across genera, with heritability for different bacterial genera ranging from 0 to 0.41 and up to 0.26 for specific archaeal genera (Zhang et al., 2020). A study with 16,234 samples collected from 585 wild baboons over a 14-year period revealed 97% of heritable microbiome taxa. In general low heritabilities were found, but most would increase with low diet diversity and with host age (Grieneisen et al.,2021). Notably, estimating microbiota heritability faces many challenges as it is a very complex and compositional system. For instance, the simple presence of one taxon may generate or consume a specific nutrient for other taxa, thus affecting its viability. Also, the proportion of one taxon may be significantly affected by the diversity and richness of others. In this scenario, methods need to be developed to account for these compositionality and complexity.

In fact, the equilibrium of this so-called "holobiont" system results from intense cross-talk between the elements of the different life kingdoms, probably driven by many molecules. Vesicles carrying metabolites, miRNAs, and proteins are good candidates for addressing these molecules to their targets, promising a biotechnological model for manipulating these systems.

As most of the literature deals with *Bos taurus* breeds in the northern hemisphere, in order to better understand the microbiome's contribution to Nelore's (*Bos indicus*) beef production and sustainability, we designed a multi-omics experiment using 52 animals split into two diets (conventional and by-product-based). To explore the taxonomic diversity and functionality of the rumen and fecal microbiomes of Nelore bulls, the rumen and stool microbial populations were accessed through sequencing of 16s rRNA gene's phylogenetic signals. The industrial by-product-based diet influenced the diversity of bacteria. Moreover, amplicon sequence variants (ASVs) were associated with residual methane emission (RME) and residual feed intake (RFI) in both ruminal and stool microbiomes (Andrade et al., 2022). While the rumen microbiome was expected to influence both traits, the relationship of the stool taxa Alistipes and Rikenellaceae with these traits was not reported before. As both are reported to produce anti-inflammatory compounds, this result may indicate a role of inflammation on RFI and RME variation, possibly implicating the microbiota in the inflammation effects on RFI described by Alexandre et al.(2015).

Shotgun metagenomics revealed higher microbial diversity in the rumen compared to the fecal microbiome. Differences in taxonomic profile and microbial functionality suggested different contributions of the microbiomes of these environments. Feed-efficient animals showed more functions related to

carbohydrate metabolism, while methanogenesis, hydrogen and formate functions were enriched from taxa associated with methane emission (Conteville et al., 2023). This data also allowed the recovery of 1,526 metagenome-assembled genomes (MAGs) from ruminal content and feces. These MAGs harbored ~2 million coding regions, which were annotated based on its functional potential, carbohydrate-active enzymes (CAZymes), antibiotic resistance genes, among others (Conteville et al., in preparation). Additionally, metatranscriptome data were mapped against the MAG's coding regions from rumen. About 960,240 coding regions were obtained from which 884,698 were active. Genes with functions related to the metabolism of carbohydrates, amino acids, lipids and energy were differentially expressed between diets. In addition to impacts on microbiome composition, these results suggest that diets can cause variations in rumen microbiome functionality through modulation of gene expression (Silva et al., 2023).

Regarding meta-metabolomics, thirty-eight ruminal fluid and 27 fecal metabolites were significantly different between the two diets. Ruminal fluid dimethylamine was significantly associated with dry matter intake (DMI), RFI, feed efficiency (FE), observed water intake (OWI) and predicted water intake (WI), while ruminal aspartate was associated with DMI, RFI, FE and WI. Fecal C22:1n9 was significantly associated with OWI and residual water intake (RWI), while C14:0 and hypoxanthine were significantly associated with DMI and RFI. The results demonstrated that nutritional interventions may alter ruminal and fecal metabolites and provided new insights into the relationship of these metabolites with feed efficiency and water intake traits in bovines. (Malheiros et al., 2023).

We also investigated the interplay among genome, transcriptome, and microbiome on feed efficiency and methane emission traits, providing new candidates for improving methane emission and feed efficiency phenotypes at the microbiota composition and transcriptional regulation levels. For instance, we identified epistatic SNP-SNP modules associated with RME and RFI traits. Genes in the SNP windows associated with RFI belong to several pathways related to the immune system and actin cytoskeleton organization, suggesting that immune response and cilium organization may play a significant role in feed efficiency, while methanogenesis may involve multiple general pathways. Moreover, many ruminal and fecal ASVs and metagenomes were correlated with these epistatic SNP-SNP modules (Cardoso et al., 2022).

Aiming to analyze fecal microbiota in a broader population of Nelore cattle, we collected 475 samples from four editions of the Young Sire Performance Test (PNAT) held by the Brazilian Zebu Breeders Association (ABCZ). Metabarcoding data revealed significant differences among the four years regarding bacterial diversity (Shannon index), but there were no significant differences regarding archaeal diversity. Association between bacterial ASVs and phenotypes identified five positive (FE, feed intake (FI), observed dry matter intake (DMIo), average daily gain (ADG), and average metabolic live weight (AMLW), and five negative (FE, FI, DMIo, ADG, and AMLW). In archaea, three positive associations were found between ASVs and phenotypes (average live weight (ALW), AMLW, and ADG) (Rojas et al., 2024). Also, seven upregulated fecal miRNAs were found in inefficient animals comparing RFI divergent bulls. The differentially expressed fecal miRNAs were highly correlated with taxa previously associated with RFI in Nelore cattle (Andrade et al., 2022), pointing to miRNAs as potential regulators of feed efficiency and possibly of the microbiome as well, contributing to the knowledge needed to develop future strategies to manipulate the microbiome (De Oliveira et al., in preparation). Genetic parameters estimates and modeling the microbiome effects on breeding equations for the PNAT bulls are ongoing and represent one of the biggest challenges in this research effort.

We also analyzed fecal microbiota from another ABCZ's project aiming to evaluate beef quality in Nelore cattle. Samples were collected at three different time points of the production system: grazing during the dry season, grazing during the rainy season, and at the end of the feedlot finishing. We observed greater bacterial diversity during the grazing period in the dry season, whereas the diversity of archaea was higher during the dry and rainy seasons than in the feedlot period (Antonio et al., 2023). These results are opposite to the literature reporting diversity increasing with age (Jami et al., 2013), and may reflect environmental effects. Association analyses between the fecal bacterial profiles and RFI showed eight ASVs from the dry season grazing profile, the earlier point after weaning, associated with reduced FE in the feedlot, and two with increased FE (Silva et al., 2023b).

From the above exposure, one can conclude that *Bos indicus* production and breeding may be better understood under the concept of hologenome, a.i., considering the complex relationships with their microbiome. Although rumen samples are frequently studied, feces are ready-accessible biological samples that may provide information on the microbiome and its effects on phenotypes. Using multi-omics, one might enlighten the underlying biological mechanisms and targets for manipulating the hologenome may be depicted.

References

ALEXANDRE, P. A.; KOGELMAN, L. J. A.; SANTANA, M. H. A.; PASSARELLI, D.; PULZ, L. H.; FANTINATO-NETO, P.; SILVA, P. L.; LEME, P. R.; STREFEZZI, R. F.; COUTINHO, L. L.; FERRAZ, J. B. S.; ELER, J. P.; KADARMIDEEN, H. N.; FUKUMASU, H. Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. **BMC Genomics**, v. 16, article 1073, 2015.

ANDRADE, B. G. N.; DONATONI, F. A. B.; CUADRAT, R. R.; CARDOSO, T. F.; MALHEIROS, J. M.; OLIVEIRA, P. S. N. DE; PETRINI, J.; MOURÃO, G. B.; COUTINHO, L. L.; REECY, J. M.; KOLTES, J. E.; ZERLOTINI NETO, A.; MEDEIROS, S. R. de; BERNDT, A.; PALHARES, J. C. P.; AFLI, H.; REGITANO, L. C. de A. Stool and Ruminal Microbiome Components Associated With Methane Emission and Feed Efficiency in Nelore Beef Cattle. **Frontiers in Genetics**, v. 13, article 812828, 2022.

ANTONIO, B. C. P.; CRUZ, T. F. C. da; CONTEVILLE, L. C.; SILVA, W. A. da; PEÑA, G. A. C.; ROJAS, L. J. F.; OLIVEIRA, P. S. N. de; MARCONDES, C. R.; MOURÃO, G. B.; REGITANO, L. C. de A. Analysis of fecal microbiome diversity in Nelore cattle under different environmental conditions. In: BRAZILIAN CONGRESS OF GENETICS, 68., 2023, Ouro Preto. **Paleogenomics:** sequencing ancient DNA. Ribeirão Preto: Sociedade Brasileira de Genética, 2023. p. 541.

CARDOSO, T. F.; BRUSCADIN, J. J.; DINIZ, W. J. da S.; BANERJEE, P.; AFONSO, J.; ANDRADE, B. G. N. B; MALHEIROS, J. M.; ZERLOTINI NETO, A.; VIEIRA, D. da S.; FERNANDES, A. C.; CESAR, A. S. M.; MOURÃO, G. B.; COUTINHO, L. L.; REGITANO, L. C. de A. Integration of different omics layers with SNP-SNP interaction networks associated with methane emission in nelore cattle. In: BRAZILIAN CONGRESS OF GENETICS, 67., 2022, Natal. **[Abstracts...]**. Ribeirão Preto: Sociedade Brasileira de Genética, 2022. p. 247.

CONTEVILLE, L. C.; SILVA, J. V. da; ANDRADE, B. G. N.; CARDOSO, T. F.; BRUSCADIN, J. J.; OLIVEIRA, P. S. N. de; MOURÃO, G. B.; COUTINHO, L. L.; PALHARES, J. C. P.; BERNDT, A.; MEDEIROS, S. R. de; REGITANO, L. C. de A. Rumen and fecal microbiomes are related to diet and production traits in Bos indicus beef cattle. **Frontiers in Microbiology**, v. 14, article 1282851, 2023.

GRIENEISEN, L.; DASARI, M.; GOULD, T. J.; BJÖRK, J. R.; GRENIER, J.-C.; YOTOVA, V.; JANSEN, D.; GOTTEL, N.; GORDON, J. B.; LEARN, N. H.; GESQUIERE, L. R.; WANGO, T. L.; MUTUTUA, R. S.; WARUTERE, J. K.; SIODI, L.; GILBERT, J. A.; BARREIRO, L. B.; ALBERTS, S. C.; TUNG, J.; ARCHIE, E. A.; BLEKHMAN, R. Gut microbiome heritability is nearly universal but environmentally contingent. **Science**, v. 373, i. 6551, pp. 181-186, 2021.

HU, R.; LI, S.; DIAO, H.; HUANG, C.; YAN, J.; WEI, X.; ZHOU, M.; HE, P.; WANG, T.; FU, H.; ZHONG, C.; MAO, C.; WANG, Y.; KUANG, S.; TANG, W. The interaction between dietary fiber and gut microbiota, and its effect on pig intestinal health. **Fronties in Immunology**, v. 14, article 1095740, 2023.

JAMI, E.; ISRAEL, A.; KOTSER, A.; MIZRAHI, I. Exploring the bovine rumen bacterial community from birth to adulthood. ISME journal, v. 7, i. 6, pp. 1069-1079, 2013.

MALHEIROS, J. M.; CORREIA, B. S. B.; CERIBELI, C.; BRUSCADIN, J. J.; DINIZ, W. J. S.; BANERJEE, P.; VIEIRA, D. S.; CARDOSO, T. F.; ANDRADE, B. G. N.; PETRINI, J.; CARDOSO, D. R.; COLNAGO, L. A.; BOGUSZ JUNIOR, S.; MOURÃO, G. B.; COUTINHO, L. L.; PALHARES, J. C. P.; MEDEIROS, S. R. de; BERNDT, A.; REGITANO, L. C. de A. Ruminal and feces metabolites associated with feed efficiency, water intake and methane emission in Nelore bulls. **Scientific Reports**, v. 13, i. 1, article 18001, 2023.

ROJAS, L. J. F.; CARDOSO, T. F.; COLMENAREZ PENA, G. A.; SILVA, W. A. da; ANTONIO, B. C. P.; CONTEVILLE, L.; BUSSIMAN, F. de O.; LOURENÇO, D. A. L.; ALMEIDA, L. F.; VENTURA, H. T.; JOSAHKIAN, L. A.; PASCHOAL, J. J.; MOURÃO, G. B.; REGITANO, L. C. de A.. Microbiota diversity and association with performance phenotypes in young Nelore bulls. In: CONGRESSO BRASILEIRO DE BIOINFORMÁTICA, 20., 2024, Salvador, BA. **X-Meeting 2024:** anais. Salvador, BA: Even3, 2024. Available at: https//www.even3.com.br/anais/xmeeting-2024/831261-MICROBIOTA-DIVERSITY-AND-ASSOCIATION-WITH-PERFORMANCE-PHENOTYPES-IN-YOUNG-NELORE-BULLS. Access at: 28 nov. 2024.

SAMPSON, T. R.; MAZMANIAN, S. K. Control of brain development, function, and behavior by the microbiome. **Cell host and Microbe**, v. 17, i. 5, pp. 565-576, 2015.

SILVA, J. V.; CONTEVILLE, L. C.; ANDRADE, B. G.; ZERLOTINI NETO, A.; MOURAO, G. B.; COUTINHO, L. L.; PALHARES, J. C. P.; REGITANO, L. C. A. Dietary Impacts On Rumen Metatranscriptome Of Nelore Cattle. In: ANNUAL INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY, 31.; ANNUAL EUROPEAN CONFERENCE ON COMPUTATIONAL BIOLOGY, 22., 2023, Lyon, Fr. **Proceedings...** Lyon: ISMB, ECCB, 2023. ISMB 2023. ECCB 2023.

SILVA, W. A. da; CONTEVILLE, L. C.; CARDOSO, T. F.; ANTONIO, B. P.; FUENTES ROJAS, L. J.; OLIVEIRA, P. S. N. de; PENÃ, G. A. C.; MARCONDES, C. R.; MOURÃO, G. B.; REGITANO, L. C. de A. Fecal bacterial microbiome of close to weaning Nelore associated with feed efficiency. In: BRAZILIAN CONGRESS OF GENETICS, 68., 2023, Ouro Preto. **Paleogenomics:** sequencing ancient DNA. Ribeirão Preto: Sociedade Brasileira de Genética, 2023b. p. 630.

YU, L. W.; AGIRMAN, G.; HSIAO, E. Y. The Gut Microbiome as a Regulator of the Neuroimmune Landscape. **Annual Review of Immunology**, v. 40, pp. 143-167, 2022.

ZHANG, Q.; DIFFORD, G.; SAHANA, G.; LØVENDAHL, P.; LASSEN, J.; LUND, M. S.; GULDBRANDTSEN, B.; JANSS, L. Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. **ISME Journal**, v. 14, i. 8, pp. 2019-2033, 2020.

ZHANG, X.; AMER, P. R.; STACHOWICZ, K.; QUINTON, C.; CROWLEY, J. Herd-level versus animal-level variation in methane emission prediction in grazing dairy cattle. **Animal:** an international journal of animal bioscience, v. 15, i. 9, article 100325, 2021.

NMR as a Powerful tools for metabolomics studies: Basic concepts and applications

Luiz Alberto Colnago[1]

[1] Embrapa Instrumentação, Rua XV de Novembro 1452, São Carlos, SP, Brazil, 13560-970, CNPq Fellowship.
email address: luiz.colnago@embrapa.br

Nuclear magnetic resonance spectroscopy (NMR) is a powerful technique that has been used in chemical analysis just after its introduction in 1946. The NMR information, used in metabolomics analyses, is based on the interaction of nuclear spin with electromagnetic radiation (radiofrequency) when the sample is in presence of a magnetic field. Figure shows a block diagram of a basic NMR spectrometer a photo of commercial 600 MHz instrument. he sample is placed in the probe head, inside the magnetic, and irradiate with a radiofrequency (transmitter) that depend on the strength of the magnetic field and isotopes constant property called magnetogyric ratio. The NMR signal is detected with a radiofrequency receiver, similar used in radio communication.

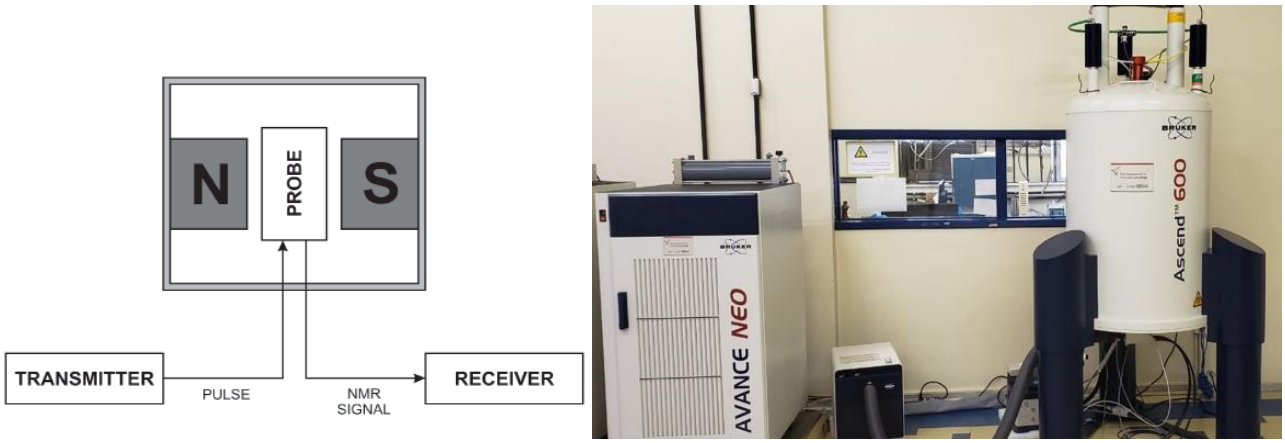

Figure 1. Block diagram of a simple NMR spectrometer (left) picture of a 400 MHz spectrometer

The sample must contain nuclei (isotopes) with odd number of proton and/or odd number of neutron that has nuclear spin, that is the property measured in NMR. For example, the three hydrogen isotopes 1H, 2H and 3H, known as hydrogen, deuterium and tritium, has odd number of proton and/or neutron and can be used in NMR analysis. Most NMR metabolomics studies is based on 1H isotope, because it is stable (not radioactive) has the high natural abundance (99.98%), and NMR high sensitivity and selectivity, allowing measurements in order of minutes.

One of the main parameters used in NMR metabolomics analysis is known as chemical shift that is related to chemical environment of the 1H in the molecules. Figure 2 shows a NMR spectrum of an ethanol sample from it is possible to obtain thrre information that can be used to characterize this molecule. Fist of them is the three peaks at approximately 1, 3 and 4 part per million (ppm) of the resonance frequency. The peaks at 1, 3 and 4 ppm are assigned to CH_3, CH_2 and OH group of the ethanol molecules, respectively. The peak splitting (separation) observed in the signals at 1 and 3 ppm is known as spin-spin coupling and is due to the number of hydrogens plus one (n +1 peaks) in the neighbor carbon. In this example the CH_3 signal, at 1 ppm, shows three peaks due to the two hydrons in CH_2 groups. On the other hand, the CH_2 signal at 3 ppm shows 4 peaks due to the three hydrogens in CH_3 group. The OH signal at 4 ppm has no splitting, in this spectrum, because it is a hydrogen that is weak bonded to the oxygen and it exchange with other ethanol molecules, and does not show the spin-spin coupling with the hydrogens bond to the carbons. The third information in this spectrum is the area of each signal is measured using the sigmoidal curve. The area (measured by the difference between the two plateaus) is proportional to the number of hydrogens in each group. The relative area of peaks at 1, 3 and 4 ppm are 3, 2 and 1, respectively, that is the number of hydrogens in each chemical environment or chemical shift.

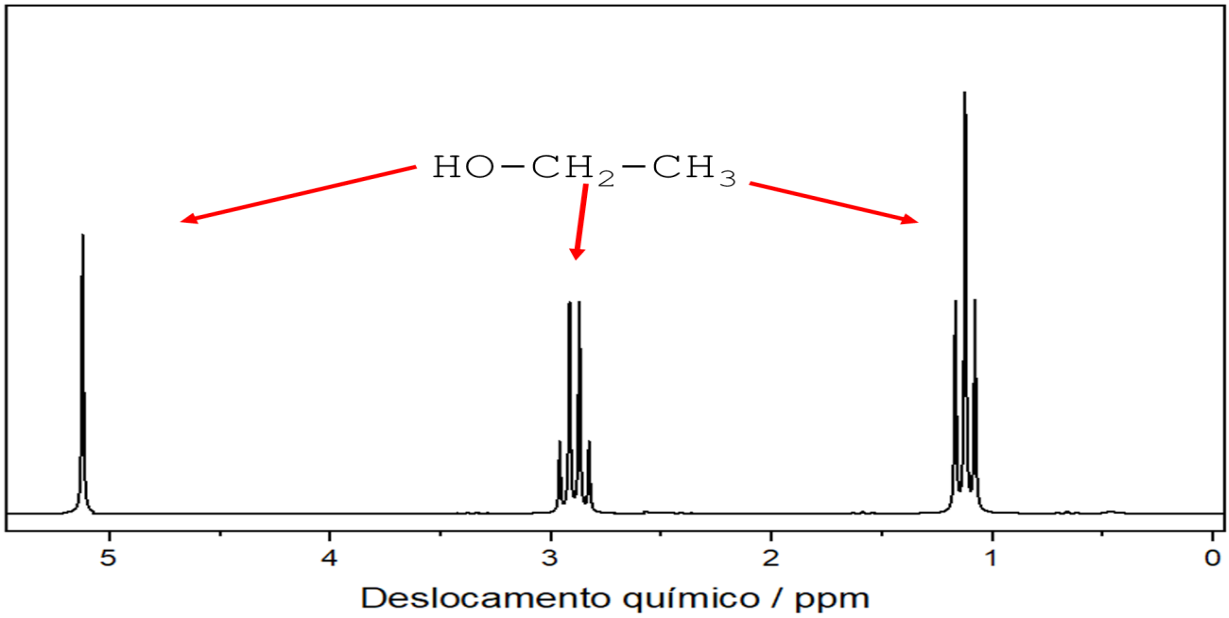

Figure 2. ^{1}H NMR spectrum of ethanol.

Therefore, using the information NMR spectroscopy has been a powerful tool to determine the chemical structure of small and large organic molecules (biopolymers).

Figure 3 shows a NMR spectrum of a blood serum sample from 0 to 6 ppm. The strong and broad line are assigned to lipoproteins, lipids and weak and sharp peaks to small molecules, like glucose, lactate, creatinine among several other metabolites. With this spectrum it is possible to quantify each metabolite, as the area of the metabolite signals are proportional metabolite concentration, the proteins and lipids. The protocol to performed the assignments is necessary to use NMR data base and the quantification can be doing by performed peaks integration or using multivariate analysis.

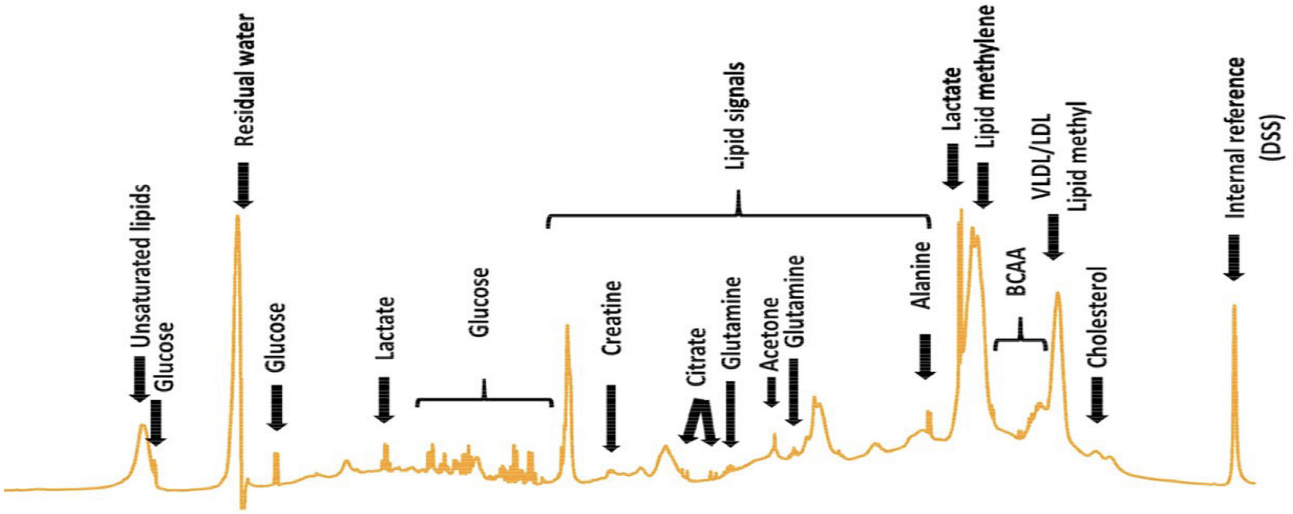

Figure 3. ^{1}H NMR spectrum of a blood serum sample obtain at 600 MHz.

Holobionts, a new view to improve the use of microorganisms in sustainable agriculture

Maria Carolina Quecine[1]

[1] Department of Genetics, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, São Paulo, Brazil
Corresponding author: mquecine@usp.br

The view of a single plant in a field looks like an individual organism from a single species, but if you are able to zoom a few million times, you will find that there are many other species living on and inside every plant. Together with all the organisms that interact with the host, a complex microbiota composed of bacteria, archaea, fungi, oomycetes, viruses, protozoa, algae, and nematodes can be found (Baedke et al., 2020). Collectively, the plant and its microbiome are holobionts, a term originally proposed by Adolf Meyer-Abich but most frequently associated with and popularized by the famous evolutionary theorist Lynn Margulis (Baedke et al., 2020). Initially, the term was applied to systems, such as fungi and algae, that combine in lichen-covered rocks and, more recently, has been popularized in plants. Certainly, the concept of holobionts in agricultural systems has broking paradigms and highlights the real role of microbes in plant health and productivity.

The idea of a holobiont has also inspired the concept of a hologenome, meaning the collective genetic character of all aspects of a holobiont. The hologenome theory of evolution is an intriguing and unresolved issue. More accept is the idea that the holobiont constitutes a unit of selection (Bordenstein & Theis, 2015). The authors strongly suggest that the hologenome concept is a holistic view of genetics in which animals and plants are polygenomic entities. For that purpose, some premises should be considered, such as the fact that holobionts and hologenomes are units of biological organization but are not organ systems, superorganisms, or metagenomes only. Moreover, hologenomes should be understood as a comprehensive gene system, and hologenomic variation integrates all the mechanisms of mutation. Thus, hologenomic evolution is most easily understood by equating a gene in the nuclear genome to a microbe in the microbiome, which fits squarely into genetics and accommodates multilevel selection theory (Bordenstein; Theis, 2015).

Current plant microbiome studies aim to investigate plant microbiome structure under a variety of conditions, for example, various host plant species, soil types and temporal and spatial differences (Sanchez-Canizares et al. 2017), which certainly contributes to answering the following question: “How is the microbiome assembled in plants?”. In general, a microbial community is the result of its succession over time. In plants, colonizing areas require fast growth, and the microbiome starts with autotrophic cyanobacteria, whose diversity increases sharply at the beginning and remains stable over time. The presence of cyanobacteria is linked to high organic carbon and nitrogen contents in soils, paving the way for the colonization of higher organisms (Liu et al. 2016). Disturbed environments support microorganisms with a greater response to nutrient inputs, and they evolve into more specialized but metabolically slower assemblages (Nemergut et al., 2016).

In agricultural systems, the input of exogenous microorganisms may affect the homeostasis of plant holobionts. This implies that we need to carefully consider adding or removing even single but important community members, as these modifications may have consequences for the whole ecosystem. For example, under field conditions, Ferrarezi et al. (2021) reported that the inoculation of beneficial bacteria did not affect the core microbiome of maize; however, differences in the richness of different plant niches were observed. Notably, the plant holobiont is a good example in nature, showing how good communication and interactive dialog between two partners lead to better performance (Gordon et al., 2013).

During agricultural practices, we should consider that the main origin of the plant microbiome is the soil, where microbial diversity is inversely correlated with the activity of its community. In addition to the microorganisms that potentially migrate from the soil to the leaves, the phyllosphere is also colonized by specialized airborne species. Another source in some plant species is the seed-borne community, which is inherited vertically. Soil health reflects the recruitment of “good” microbes to plant holobionts.

As plants strongly select their specific microbiome, their overall microbial diversity is reduced, paving the way for pathogen emergence (Niu et al., 2026). Crop monocultures are examples of continuous plant selection over vast areas and long time scales (Sanchez-Canizares et al. 2017). Thus, defining a model microbial community to increase plant productivity is of interest. The approach of assembling synthetic communities will need to incorporate community ecology concepts. These synthetic communities, which are usually based on biocontrol and biofertilizer species, should be supplemented with both accessory species to obtain robust

diversity and keystone species to maintain the functional stability of the plant holobiont. Ferrarezi et al. (2023) investigated whether the plant growth promotion efficiency of a key stone, Azospirillum brasilense (Ab-V5), is a result of niche occupation and persistence according to the holobiont conditions. The effects of Ab-V5 and the native microbiome were evaluated in microcosms via a combined approach of microbiomics (species-specific qPCR, 16S rRNA metataxonomics and metagenomics) and plant phenomics (conventional and high-throughput methods). The results revealed a weak ability of Ab-V5 inoculation to promote maize growth in the native microbiome, which contrasts with the positive effects of using a diluted community. Genes involved in bacterial metabolism of riboses and amino acids and cresol degradation were abundant in the diluted holobiont community, indicating that these pathways can contribute to plant growth promotion and might be a result of Ab-V5 performance as a microbial recruiter of beneficial functions to the plant. It is important to understand how beneficial bacteria and fungi, which are largely used as bioinputs and constitute a relatively small fraction of the microbiome, have had a disproportionately large influence on plant holobiont evolution.

Beneficial microbes can support plant holobiont nutrition through 1) biological nitrogen fixation (BNF), 2) solubilization of insoluble nutrients, and 3) increased root surface area. The synergies between members of the plant holobiont microbiome affect nutrient cycling in the rhizosphere and consequently shape plant nutrient status and crop yield in agricultural settings. For example, arbuscular mycorrhizal fungi create a suitable environment for the colonization of plant growth-promoting bacterial endophytes if inoculated together onto crop plants, and plant root and hyphal exudates provide a carbon source for bacterial endophytes. Compared with single-strain inoculants, coinoculation with both fungi and bacteria as a consortium can improve crop yields more (Nanjundappa et al., 2019).

Microbial communities present in the rhizosphere can control plant holobiont growth, development, and stress responses through the production and delivery of plant growth regulators, growth regulatory precursors, or their analogs (Singh et al., 2019). These plant–microbe interactions rely on a wide variety of long-distance chemical signaling compounds (Chargas et al., 2018), including plant hormones (indole-3-acetic acid (IAA), auxins, cytokinins, and gibberellins) and microbial-produced compounds that can mimic or induce plant hormone production (Lyu et al., 2021). For example, many bacteria produce auxin or manipulate host auxin signaling to, in the case of rhizosphere PGPB, promote plant root growth or, in the case of plant pathogens, interfere with plant development (Spaepen et al., 2010).

Multiomics technologies are leading to a better understanding of plant holobiont dynamics. Currently, the combination of automation with network analysis has improved the understanding of microbial ecology. Network analysis is used not only to identify microbial taxa that support diversity but also to identify specific taxa that show either a direct negative correlation with a known pathogen or an indirect correlation due to a positive interaction with a pathogen antagonist (Lyu et al., 2021, Jeffrey, 2018).

Finally, the concept of holobionts is a perfect example of how our understanding of biology continues to evolve over time. All of these plant breeding programs should consider those concepts during the development of new cultivars and varieties as well as during the development of plant holobiont engineering, considering the importance of the holobiont, where plants and microbes are interdependent. Moreover, improvements concerning the intrinsic associations of plants and microbes are key points that will contribute to the development of new management practices in fields contributing to more sustainable agriculture.

References

BAEDKE, J.; FÁBREGAS-TEJEDA, A.; DELGADO, A. N. The holobiont concept before Margulis. **Journal of Experimental Zoology Part B: Molecular and Developmental Evolution**, v. 334, i. 3, pp. 149-155, 2020.

BORDENSTEIN, S. R.; THEIS, K. R. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. PLoS Biology, v. 13, i. 8, e1002226, 2015.

CHAGAS, F.O.; PESSOTTI, R. D. C.; CARABALLO-RODRÍGUEZ, A. M.; PUPO, M. T. Chemical signaling involved in plant-microbe interactions. **Chemical Society Review Journal**, v. 47, i. 5, pp. 1652-1704, 2018.

FERRAREZI, J. A.; CARVALHO-ESTRADA, P. A.; BATISTA, B. D.; ANICETO, R. M.; TSCHOEKE, B. A. P.; ANDRADE P. A. M.; LOPES. B. M.; BONATELLI, M. L.; ODISI, E. J.; AZEVEDO, J. L.; QUECINE, M. C. Effects of inoculation with plant growth-promoting rhizobacteria from the Brazilian Amazon on the bacterial community associated with maize in field. **Applied Soil Ecology**, v. 170, article 104297, 2022.

FERRAREZI, J. A.; DEFANT, H.; SOUZA, L. F. de; AZEVEDO, J. L.; HUNGRIA, M.; QUECINE, M. C. Meta-omics integration approach reveals the effect of soil native microbiome diversity in the performance of inoculant Azospirillum brasilense. **Frontiers in Plant Science**, v. 14, article 1172839, 2023.

GORDON, J.; KNOWLTON, N.; RELMAN, D. A.; ROHWER, F.; YOULE, M. Superorganisms and holobionts. **Microbe Magazine**, v. 8, i. 4, pp. 152-153, 2013.

JEFFREY, M. J. What is the hologenome concept of evolution? **F1000 Faculty Review**, v. 7, article 1664, 2018.

LIU, J.; KONG, W.; ZHANG, G.; KHAN, A.; GUO, G.; ZHU, C.; WEI, X.; KANG, S.; MORGAN-KISS, R. M. Diversity and succession of autotrophic microbial community in high-elevation soils along deglaciation chronosequence. **FEMS Microbiology Ecology**, v. 92, i. 10, fiw160, 2016.

LYU, D.; ZAJONC, J.; PAGÉ, A.; TANNEY, C. A. S.; SHAH, A.; MONJEZI, N.; MSIMBIRA, L. A.; ANTAR, M.; NAZARI, M.; BACKER, R.; SMITH, D. L. Plant holobiont theory: the phytomicrobiome plays a central role in evolution and success. **Microorganisms**, v. 9, i. 4, article 675, 2021.

NANJUNDAPPA, A.; BAGYARAJ, D. J.; SAXENA, A. K.; KUMAR, M.; CHAKDAR, H. Interaction between arbuscular mycorrhizal fungi and Bacillus spp. in soil enhancing growth of crop plants. **Fungal Biology and Biotechnology**, v. 6, article 23, 2019.

NEMERGUT, D. R.; KNELMAN, J. E.; FERRENBERG, S.; BILINSKI, T.; MELBOURNE, B.; JIANG, L.; VIOLLE, C.; DARCY, J. L.; PREST, T.; SCHMIDT, S. K.; TOWNSEND, A. R.. Decreases in average bacterial community rRNA operon copy number during succession. **ISME journal**, v. 10, i. 5, pp. 1147-1156, 2016.

NIU, J.; RANG, Z.; ZHANG, C.; CHEN, W.; TIAN, F.; YIN, H.; DAI, L. The succession pattern of soil microbial communities and its relationship with tobacco bacterial wilt. **BMC Microbiology**, v. 16, i. 1, article 233, 2016.

SÁNCHEZ-CAÑIZARES, C.; JORRÍN, B.; POOLE, P. S.; TKACZ, A. Understanding the holobiont: the interdependence of plants and their microbiome. **Current Opinion in Microbiology**, v. 38, pp. 188-196, 2017.

SINGH, D.; RAINA, T. K.; KUMAR, A.; SINGH, J.; PRASAD, R. Plant microbiome: A reservoir of novel genes and metabolites. **Plant Gene**, v. 18, article 100177, 2019.

SPAEPEN, S.; VANDERLEYDEN, J. Auxin and plant-microbe interactions. **Cold Spring Harbor Perspective Biology**, v. 3, i. 4, a001438, 2010.

Support


Bronze Sponsorship

Silver Sponsorship

Gold Sponsorship

Funded by

CGPE 18881