Global Methane Genetics initiative

Led by WAGENINGEN UNIVERSITY & RESEARCH

In Partnership with BEZOS EARTH FUND, Global Methane Hub

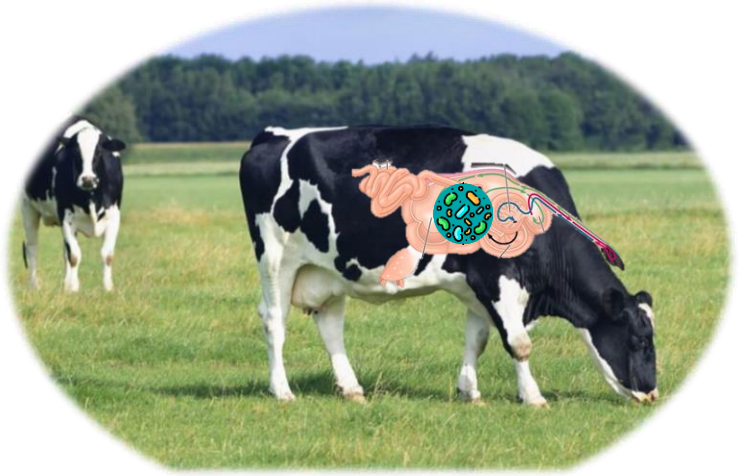
GMG Microbiome Working Group meeting

09-06-2026

WAGENINGEN UNIVERSITY & RESEARCH

Agenda

- Welcome
- Open discussion with all
- GMG microHub project
- Bioinformatics:
  - Boris Sepulveda (Australia)
  - Ben Perry (New Zealand)
  - Zexi Cai (Denmark)
  - Adrian López-Catalina (UK)
- Update GMG-initiative
- AOB



Open Discussion with all

- Microbiome/Methane projects beyond GMG? Running microbiome initiatives?
- Discussion rumen vs oral vs faecal microbiome & associations with methane
- Strategies to use microbiome in breeding
- ....

Updates of GMG-microHub project

Update GMG

NEW projects & proposals

- **Low-methane genetics expansion to small ruminants in Africa and Asia (ICARDA)**
- **Buffalo proposal** (Lead University of Adelaide, Australia, Asian countries & South America)
- **Asia cattle proposal** (Bangladesh, India, Indonesia, and Pakistan)

Share data,

and accelerate genetic progress to reduce methane emission

Global Methane Genetics Initiative

LOGIN

Please log in to continue to your account.

Email

Enter your email

Password

Enter your password

Remember me

Forgot password?

Log in

BEZOS EARTH FUND

WAGENINGEN UNIVERSITY & RESEARCH

Global Methane Hub

ICAR THE GLOBAL STANDARD FOR LIVESTOCK DATA

For registration, please contact us at methane-register@datagene.com.au

GMG project:

Framework adoption & incentivization system for genetic selection as methane mitigation tool

Methane breeding value validation protocol

Webinar
May 11, 2026 ✓

SWOT analysis trait definition Selection index Breeding objective

Webinar
April 13, 2026 ✓

Review of policy levers and market mechanisms

Webinar
May 7, 2026 ✓

Session at ICAR meeting in Verona: Adoption & Incentives to Breed for Methane Mitigation https://icar2026.it/meeting-programme/ & panel discussion

https://icar2026.it/wp-content/uploads/2026/05/Book-of-abstracts-2026.pdf

Webinars & WG meetings

News flash*

Towards guidelines to validate genetic evaluations for enteric methane emissions

Date and time
Monday, 11 May 2026
12:00-13:00h CET

Latin America working group meeting

Date and time
Tuesday, 12 May 2026
16:00-17:30h CET

Beef working group meeting

Date and time
Monday, 18 May 2026
13:00-14:30h CET

Microbiome working group meeting

Date and time
Tuesday, 9 June 2026
12:00-13:30h CET

Webinar planning GMG initiative 2026

Measuring methane emissions using Laser Methane Detection (LMD) in rural livestock systems: key considerations

Date and time
Wednesday 25 March 2026
9:00-10:30h CET

SWOT analysis of methane trait definitions and consequences for breeding programs

Date and time
Monday, 13 April 2026
12:00-13:00h CET

Dairy working group meeting

Date and time
Thursday, 23 April 2026
12:00-13:30h CET

Sniffer data aligning and editing pipelines

Date and time
Wednesday, 29 April 2026
14:00-15:30h CET

Sheep working group meeting

Date and time
Thursday, 30 April 2026
12:00-13:00h CET

Enabling adoption of methane-reducing breeding programs through GHG accounting and MRV frameworks

Date and time
Thursday, 7 May 2026
12:00-13:00h CET

Africa working group meeting

Date and time
Friday, 8 May 2026
13:00-14:30h CET

GMG newsletter:

WAGENINGEN
UNIVERSITY & RESEARCH

* https://www.wur.nl/en/news/methane-emissions-cows-and-sheep-can-be-reduced-25-using-breeding-programmes

In person meetings

World Congress on Genetics Applied to Livestock Production

Madison, Wisconsin
12-17 July 2026

Save the date

Courses:

Methane:
Data editing, Phenotypes Genetic analyses, Breeding program

Microbiome:
Bioinformatics, data analysis, trait definition, breeding program

Any other business & closure

Global Methane Genetics initiative

Led by WAGENINGEN UNIVERSITY & RESEARCH

In Partnership with BEZOS EARTH FUND, Global Methane Hub

Contact:
gmg@wur.nl

Micro HUB

June 9th

GMG presentation meeting

Micro HUB

Global Methane Genetics initiative

Led by WAGENINGEN UNIVERSITY & RESEARCH

In Partnership with BEZOS EARTH FUND, Global Methane Hub

THE UNIVERSITY of EDINBURGH The Royal (Dick) School of Veterinary Studies

WAGENINGEN UNIVERSITY & RESEARCH

AARHUS UNIVERSITY

AGRICULTURE VICTORIA

agresearch āta mātai, mātai whetū

Key objectives

- The microHub project will establish a **reference population** with metagenomic and genomic data and create a **genomic evaluation** system that can be used to select the parents of the next generation with microbiome profiles that reduce enteric methane emission
- Target different breeds and systems (GMG partners & external)
- Investigate the microbiome profile as a correlated trait for reducing methane emissions
- Lower environmental impact through metagenomics while maintaining rumen health

THE UNIVERSITY of EDINBURGH
The Royal (Dick) School of Veterinary Studies

- Oscar Gonzalez Recio
- Adrian López-Catalina
- Ridwan Ahmed
- Maddi Post

AGRICULTURE VICTORIA

- Jennie Pryce
- Boris Sepulveda
- Amanda Chamberlain
- Christy J Vander Jagt

agresearch
āta mātai, mātai whetū

- Suzanne Rowe
- Ben Perry
- Hannah Henry
- Timothy Bilton
- John McEwan
- Fern Booker

Micro HUB

WAGENINGEN UNIVERSITY & RESEARCH

- Birgit Gredler-Grandl
- Aniek C. Bouwman

AARHUS UNIVERSITY

- Goutam Sahana
- Zexi Cai
- Emre Karaman
- Maria Altaf Satti

Global Methane Hub

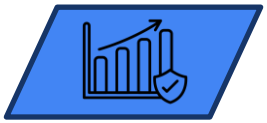
Global Methane Genetics initiative
Led by WAGENINGEN UNIVERSITY & RESEARCH
In Partnership with BEZOS EARTH FUND, Global Methane Hub

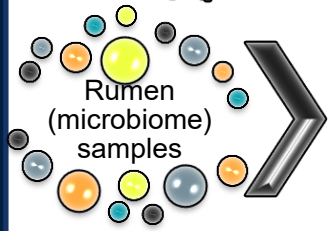
Micro HUB

SEQUENCING STRATEGY

- Sheep (4000)
- Dairy (7000)
- Beef (8000)
- African breeds/crosses (1200)

Bioinformatics working group:
-Annotate reads to taxonomy and functionality
-Assemble a pan-metagenome
-Determine a core metagenome associate to methane





Microbiome composition

Pedigree, genotypes and metadata

SNP coefficients

DNA extraction & Sequencing

Core microbiome (breed & system)

Genetic component

Genomic evaluations (breed & system)

Information to be provided to partners
(from their own samples)

Fastq files (reads)

Table with taxa and KO counts per animal/sample

SNP coefficients to estimate direct genomic breeding value for microbial features associated to methane

- Rumen metagenome as a genomic selection target to reduce enteric methane emissions
  - Boris Sepulveda – Agriculture Victoria
- Restriction Enzyme-Reduced Representation Sequencing (RE-RRS) of Rumen Microbial Communities: Methane Breeding at Scale in NZ
  - Ben Perry - AgResearch
- Assembly-based metagenomic profiling
  - Zexi Cai – Aarhus University
- Assigning functional annotation to microbiome reads
  - Adrián Lopez Catalina- The University of Edinburgh

Open discussion and collaborations

Feel free to reach out to:

Oscar Gonzalez Recio
oscar.gonzalezrecio@ed.ac.uk

Rumen metagenome as a genomic selection target to reduce enteric methane emissions

June 9th, 2026

Boris Sepulveda
Agriculture Victoria

Global Methane Genetics initiative

Led by WAGENINGEN UNIVERSITY & RESEARCH

In Partnership with BEZOS EARTH FUND, Global Methane Hub

TABLE OF CONTENTS

- Recent publications and Micro-HUB's aim
- Progress in Micro-HUB bioinformatics

Recent publications

communications biology

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › communications biology › articles › article

Article | Open access | Published: 13 April 2026

Reliable enteric methane prediction from the cattle (*Bos taurus*) rumen microbiome

Boris J. Sepulveda, Oscar González-Recio, Amanda J. Chamberlain, Ruidong Xiang, Benjamin G. Cocks, Jianghui Wang, Claire P. Prowse-Wilkins, Leah C. Marett, S. Richard O. Williams, Joe L. Jacobs, Aser García-Rodríguez, Jose A. Jiménez-Montero & Jennie E. Pryce

Communications Biology (2026) | Cite this article

J. Dairy Sci. 108:8619–8636
https://doi.org/10.3168/jds.2024-25436

Rumen metagenome as a genomic selection target to reduce enteric methane emissions

B. J. Sepulveda,[1,2*] O. González-Recio,[3] A. J. Chamberlain,[1,2] M. Khansefid,[1] B. G. Cocks,[2] J. Wang,[1] C. P. Prowse-Wilkins,[1,4] L. C. Marett,[5] S. R. O. Williams,[5] J. L. Jacobs,[2,5] A. García-Rodríguez,[6] J. A. Jiménez-Montero,[7] and J. E. Pryce[1,2]
[1]Agriculture Victoria Research, AgriBio, Centre for AgriBioscience, Bundoora, VIC 3083, Australia
[2]School of Applied Systems Biology, La Trobe University, Bundoora, VIC 3083, Australia
[3]Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC), Madrid 28040, Spain
[4]Department of Animal Science, University of California, Davis, CA 95616
[5]Agriculture Victoria Research, Ellinbank, VIC 3821, Australia
[6]Department of Animal Production, NEIKER-BRTA, Granja Modelo de Arkaute, 01080 Vitoria-Gasteiz, Spain
[7]Spanish Holstein Association (CONAFE), 28340 Madrid, Spain

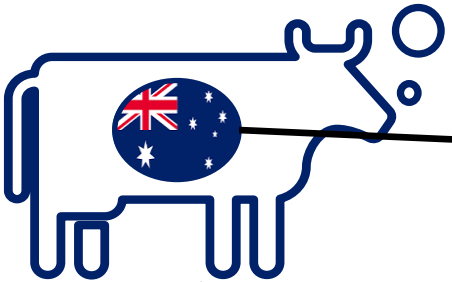
Dairy Australia · Dairy Bio · GARDINER FOUNDATION · DairyFeedbase · AGRICULTURE VICTORIA

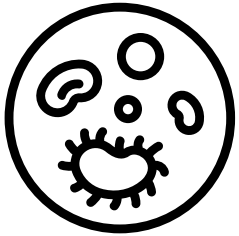
Prof Jennie Pryce
Agriculture Victoria

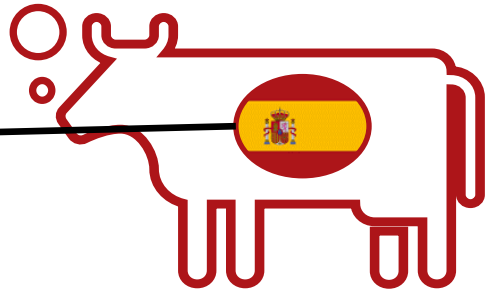
Dr Oscar González Recio
INIA - CSIC

MeP (g/d) with SF_6

MeC (ppm) with sniffers







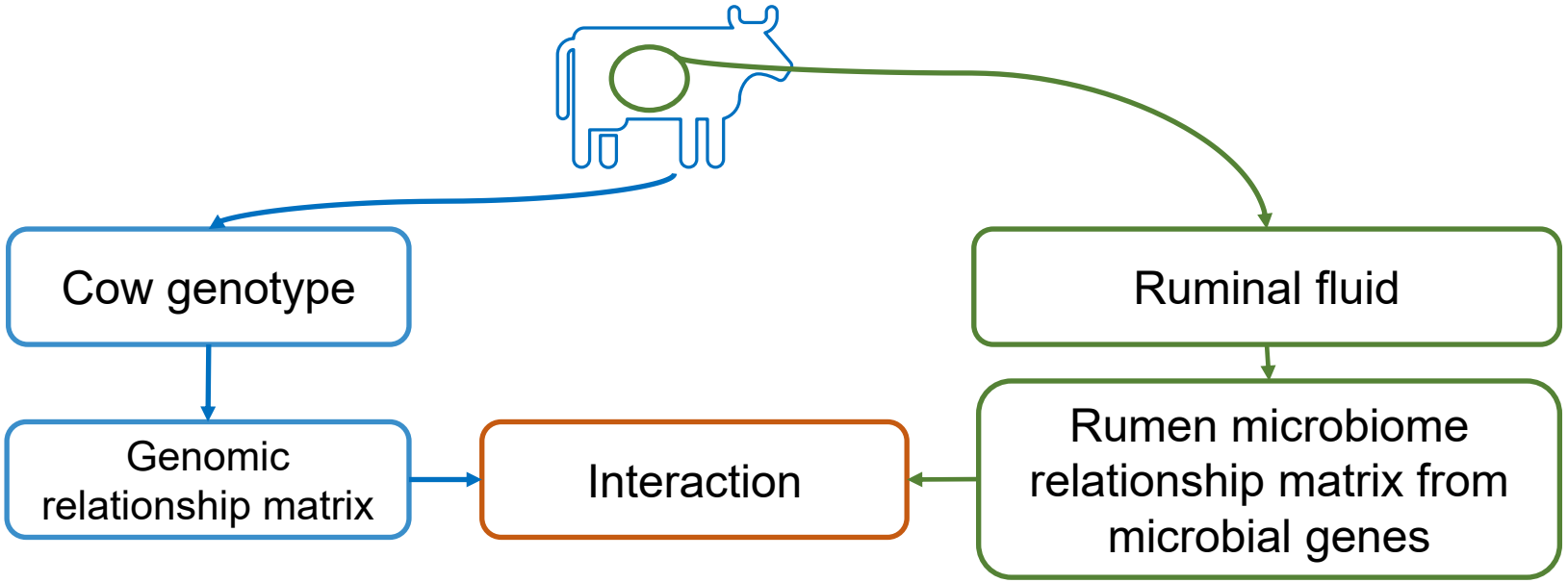
Rumen metagenome

>400 Holstein dairy cows

>400 Holstein dairy cows

Cow genome

$$\mathbf{Methane} = \mathbf{1}'\mu + \mathbf{Zg} + \mathbf{Um} + \mathbf{T(g \times m)} + \text{other effects} + \text{residual}$$

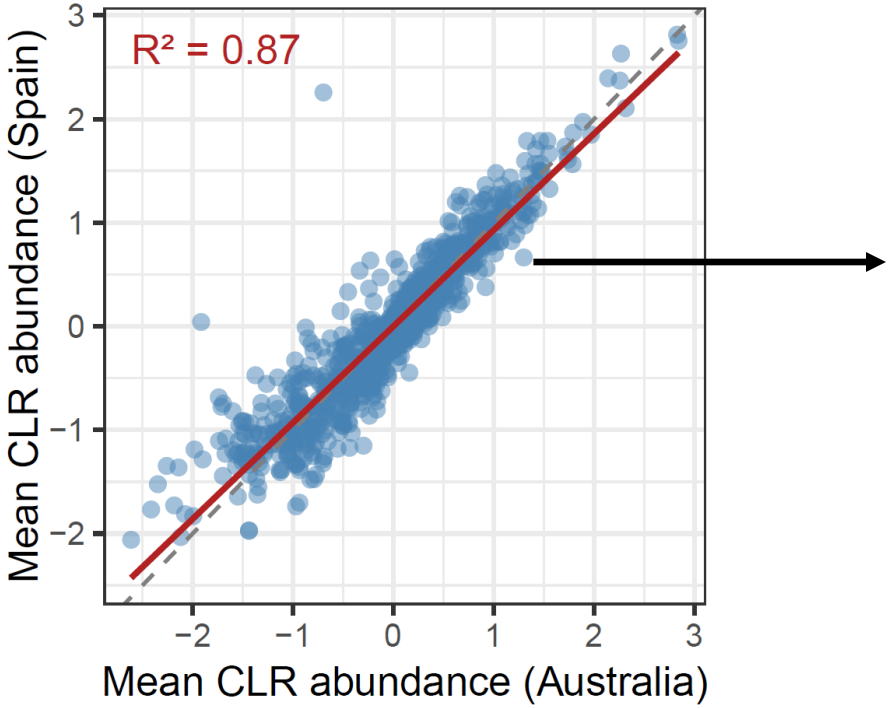


Total variance explained: up to ~ 0.60
Accuracy: up to ~0.40

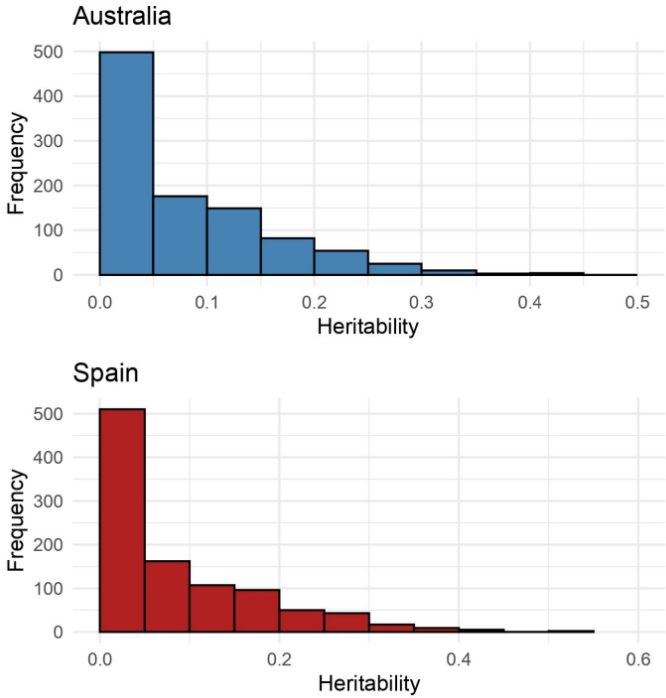
Sepulveda *et al. Commun Biol* (2026). https://doi.org/10.1038/s42003-026-10048-8

Similarities between Australia and Spain

Metagenome profile



Heritability (h^2) of metagenome features



Sepulveda *et al. Commun Biol* (2026). https://doi.org/10.1038/s42003-026-10048-8

Similarities between Australia and Spain

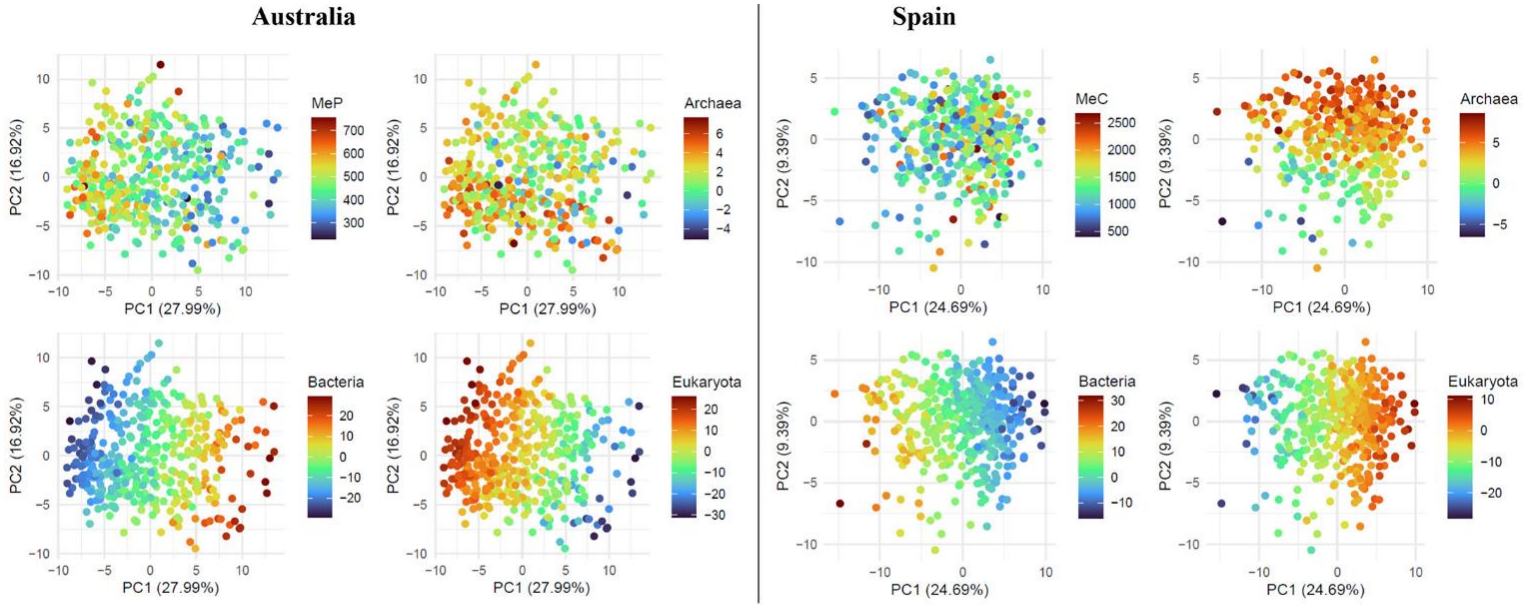
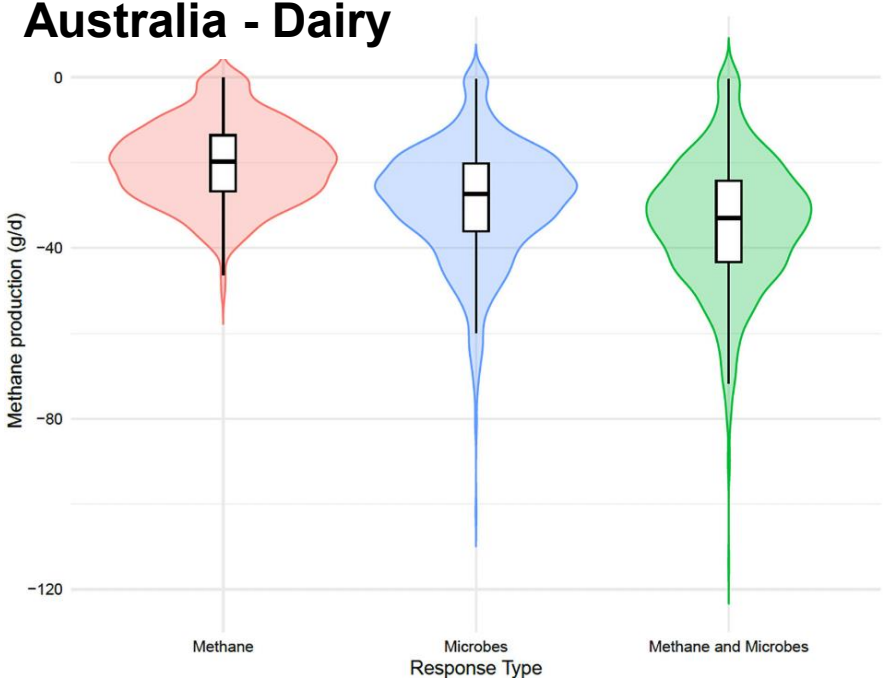


Figure 1. First 2 principal components (PC), depicting coordinates of the ruminal metagenome of dairy cattle located in Australia and Spain at genus taxonomic rank, colored by enteric methane emissions and content of *Archaea*, *Bacteria*, and *Eukaryota*. MeP = enteric methane production, recorded as grams per day (g/d) in Australia. MeC = enteric methane concentration, recorded as parts per million (ppm) in Spain.

Sepulveda *et al. J. Dairy. Sci. (2025) 08:8619–8636 https://doi.org/10.3168/jds.2024-25436*

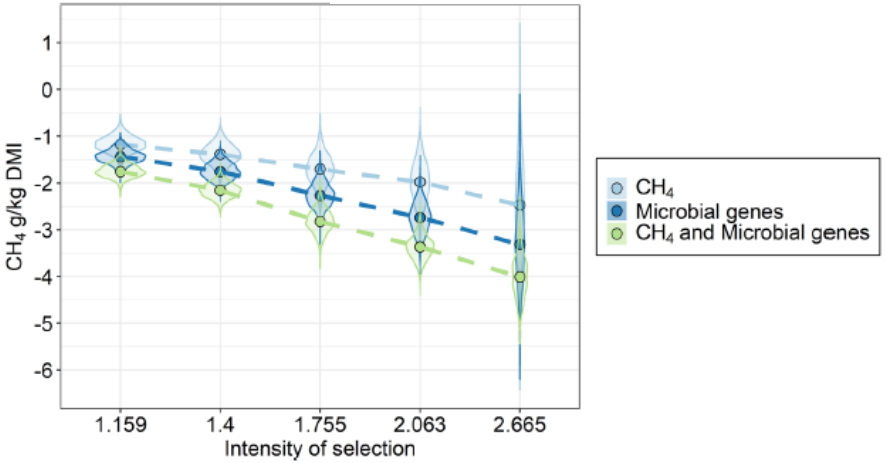
Estimated methane reduction per generation

Australia - Dairy



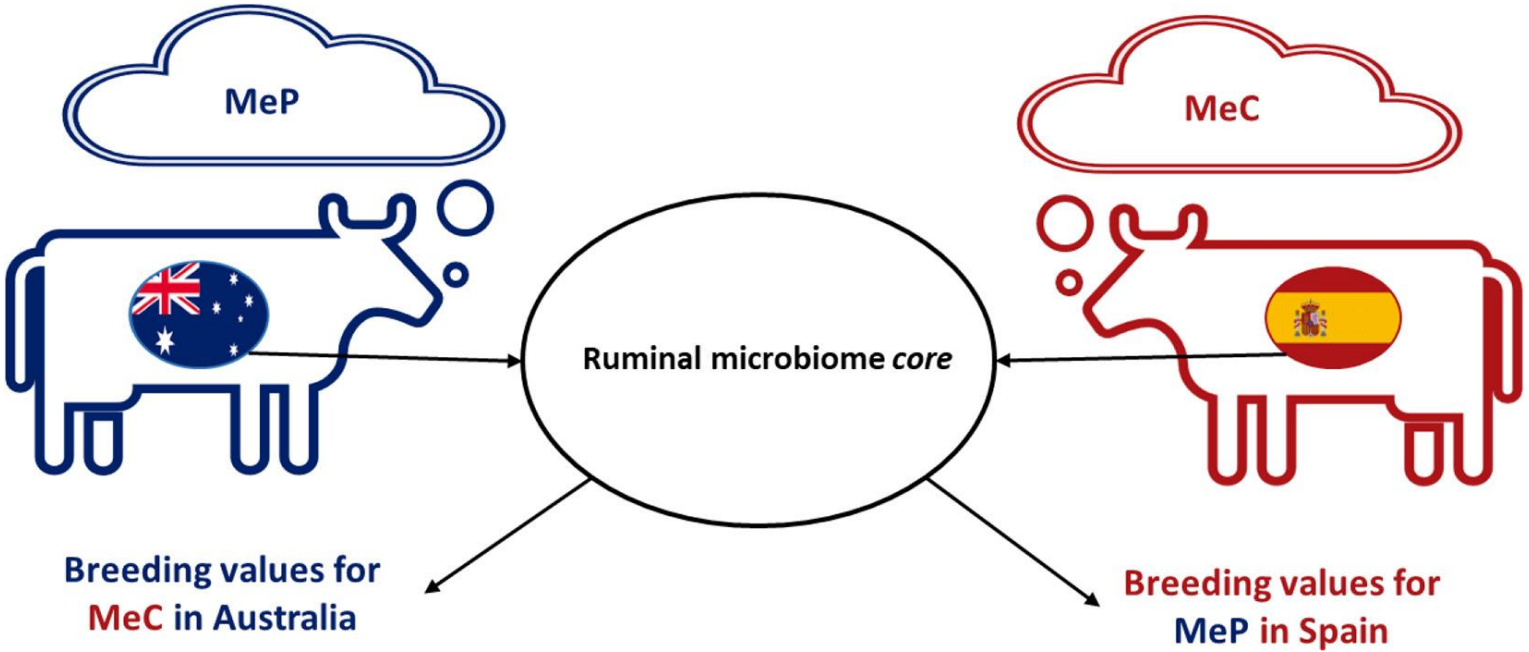
Sepulveda *et al. J. Dairy. Sci. (2025) 08:8619–8636*
https://doi.org/10.3168/jds.2024-25436

UK - Beef



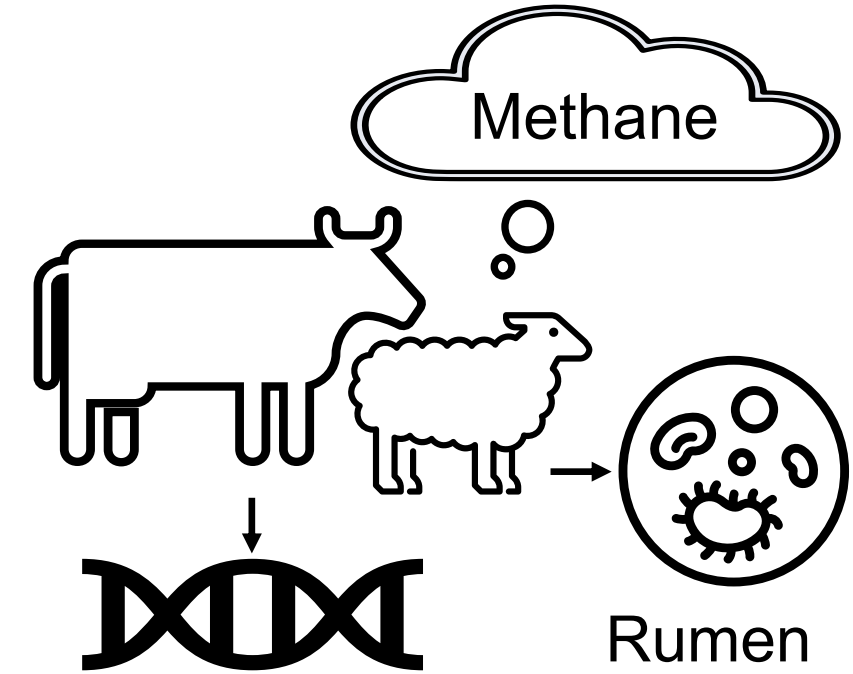
Martínez-Álvaro *et al. Commun. Bio.* 2022 Apr 12;5(1):350. doi: 10.1038/s42003-022-03293-0.

Merging reference populations using the metagenome



Sepulveda *et al. J. Dairy. Sci. (2025) 08:8619–8636*
https://doi.org/10.3168/jds.2024-25436

Micro-HUB



- 20K+ ruminants
- 50+ teams
- 25 countries
- Methane technologies: SF6, sniffers, GreenFeed, portable accumulation chambers, …

Micro-HUB Bioinformatics working group

THE UNIVERSITY of EDINBURGH
The Royal (Dick) School of Veterinary Studies

Oscar Gonzalez Recio

THE UNIVERSITY of EDINBURGH
The Royal (Dick) School of Veterinary Studies

Adrian Lopez

WAGENINGEN UNIVERSITY & RESEARCH

Aniek Bouwman

AARHUS UNIVERSITY

Zexi Cai

Bioeconomy Science Institute
AGRESEARCH GROUP

Ben Perry

AARHUS UNIVERSITY

Maria Satti

AGRICULTURE VICTORIA

Boris Sepulveda

Bioinformatics working group

- Aims
  - Unify sequencing quality control cutoffs
  - Improve annotation speed
    - Annotation: Assigning biological meaning to the metagenomic DNA sequences
  - Develop Standard Operating Procedures (SOPs) for bioinformatics
- Preliminary results
  - Unified sequence quality control cutoffs
  - Correlations >0.99 between functional annotations (“genes”) obtained by different teams
- Current tasks
  - Testing different approaches to improve annotation speed

THANK YOU

Questions?

Feel free to reach out to:

Oscar Gonzalez Recio
oscar.gonzalezrecio@ed.ac.uk

Bioeconomy Science Institute
Maiangi Taiao

Restriction Enzyme-Reduced Representation Sequencing (RE-RRS) of Rumen Microbial Communities: Methane Breeding at Scale in NZ

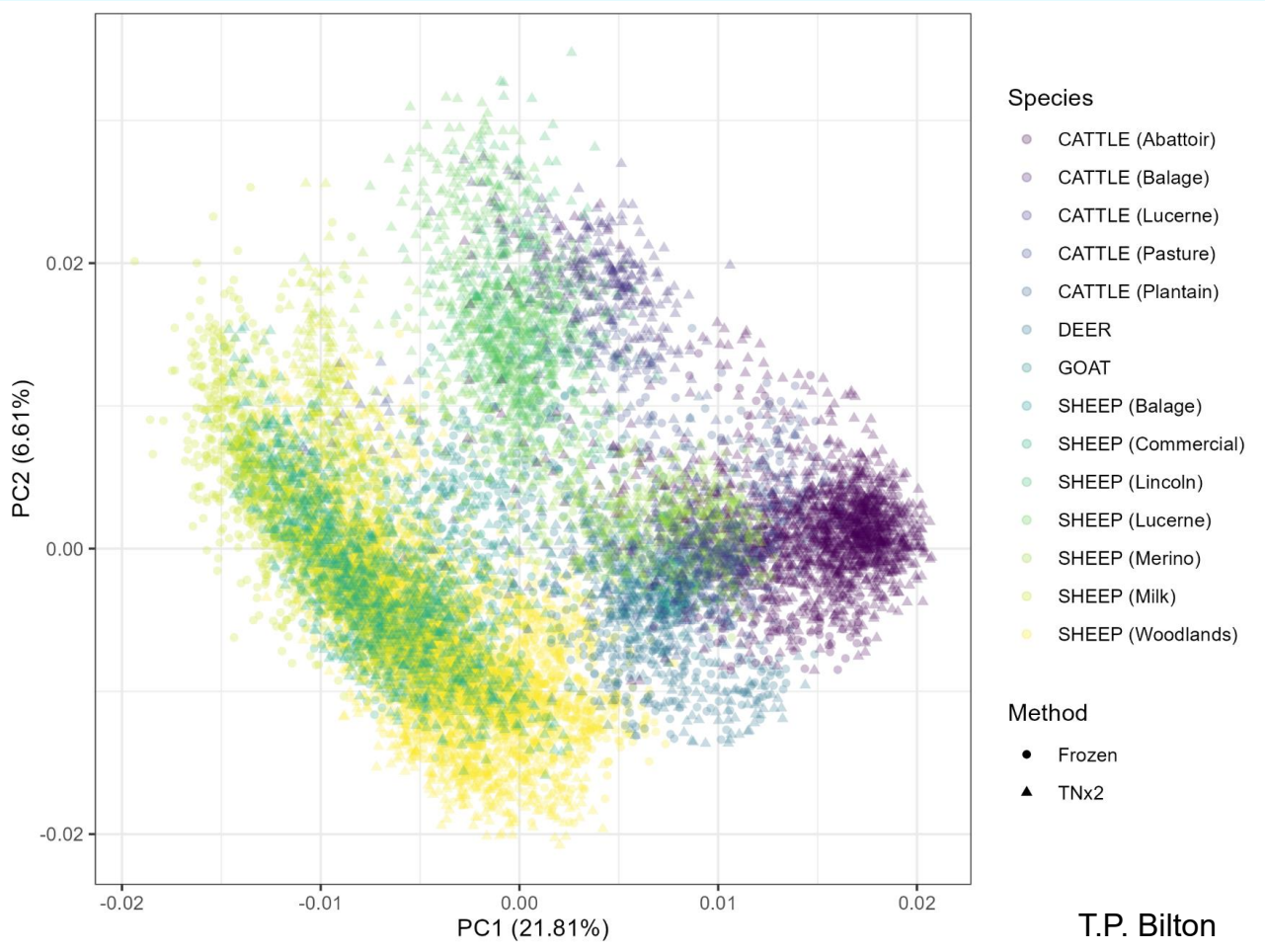
Ben Perry, Timothy Bilton, Hannah Henry, Suzanne Rowe

GMG Microbiome Working Group Meeting

09-06-2026

Rumen Sampling in New Zealand – Brief Context

- We've been sampling rumen content since about 2013
- Oldest samples ~2010
- **Sheep ~ 25,200**
- Cattle ~7600
- Deer ~1400
- Develop solutions to scaling:
  - Preservant – **TNx2**
  - Sequencing – RERRSeq in Microbiomes
  - Biofinformatics – Reference and MRM based
- Emphasis on cost, throughput, CH4-prediction
- **microHUB** – Handling Sequencing and Bioinformatics Analysis of Sheep Rumen Samples



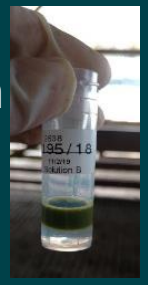
T.P. Bilton

Sample Processing and Data Generation

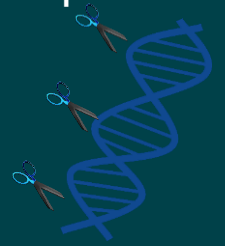
Rumen Sample:

Sample processing:
Preservation solution (TNx2)

(Budel *et al.* 2022)



Sequencing:
Restriction enzyme-reduced representation sequencing



PstI

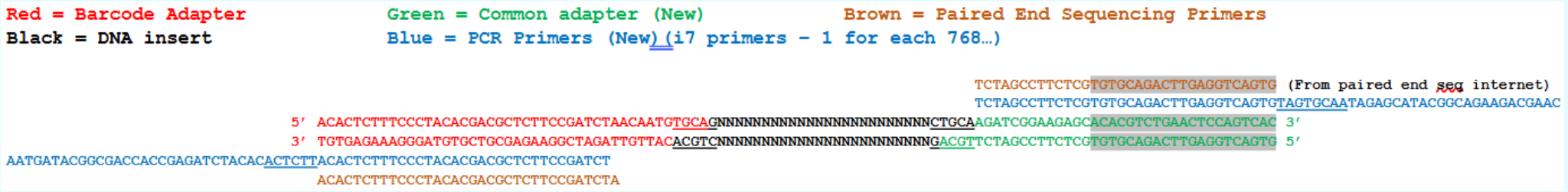
barcoding

- Low cost RT preservative – **TNx2** mixed 1:1 with rumen content
- In-house DNA extraction and library preparation
- Illumina sequencing – NovaSeq; moving to MGI T1+ and G99 DNBSeq
- Aim for 750,00 - 500,000 reads / sample from rumen material

genomNZ from agresearch

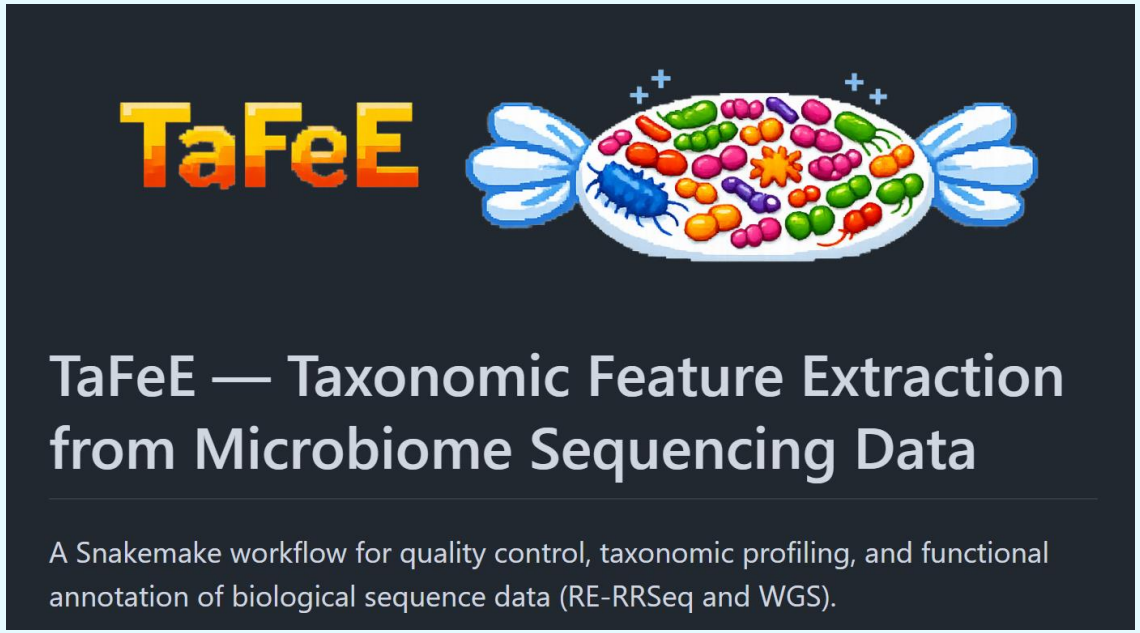
High-throughput Low Cost Sequencing Restriction Enzyme-Reduced Representation Sequencing RE-RRSeq

- DNA Extraction from unfiltered rumen content
- PstI – Digestion of DNA, ligation, and size selection
  - We did try other enzymes; gives us ~1% subsampling



- Size selection between 220-340 bp
- 100 bp single-end reads
- In-line barcodes for large numbers (384 samples / library); ~1000s / run
- Economies of scale, in-house oligos and reagents, hit a competitive price-point

Bioinformatic Processing: TaFeE Preprocessing and Reference-Based Profiling



- TaFeE is our in house bioinformatics pipeline for reference-based analysis of rumen RE-RRS and shotgun metagenomic or metatranscriptomic data

https://github.com/AgResearch/TaFeE

Workflow DAG

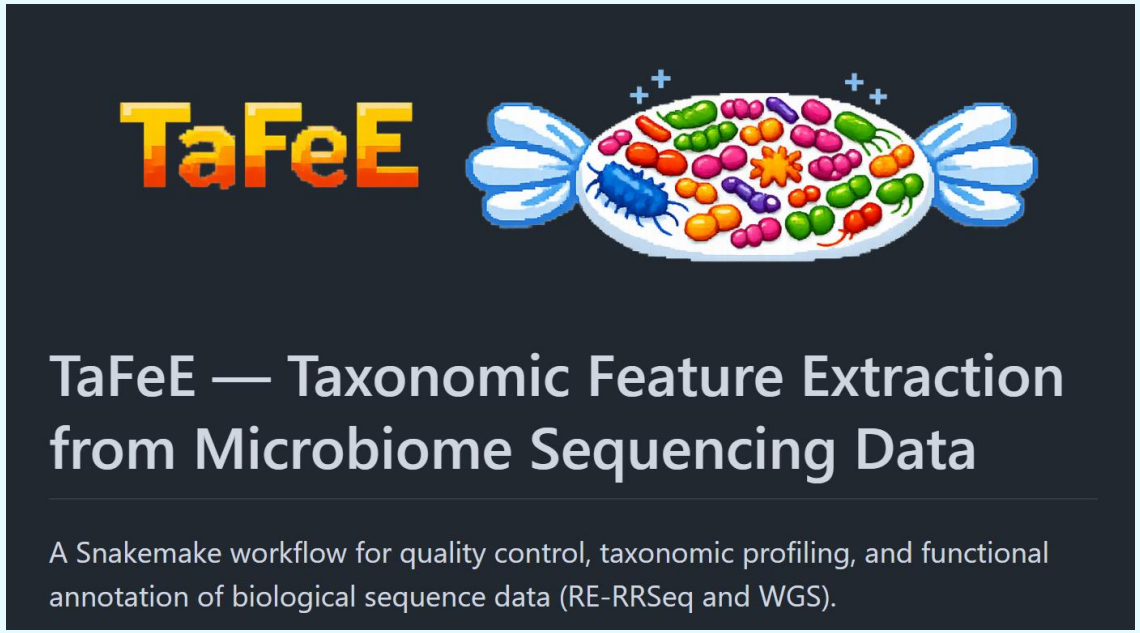
RE-RRSeq (RE-RRSeq-TaFeE)

```
Raw RE-RRSeq FASTQ
  → cutadapt (demultiplex)
    → BBDuk (entropy/quality filter)
      → PRINSEQ++ (low-complexity filter)
        → KneadData (Trimmomatic + TRF + SILVA rRNA removal)
          → Kraken 2 (taxonomic classification)
            → taxpasta (count matrices at domain → species)
```

- We build custom Kraken2 reference indices based on the GTDB, host genomes, protists, and a small selection of fungi
- **microHUB – Unified Reference Database**

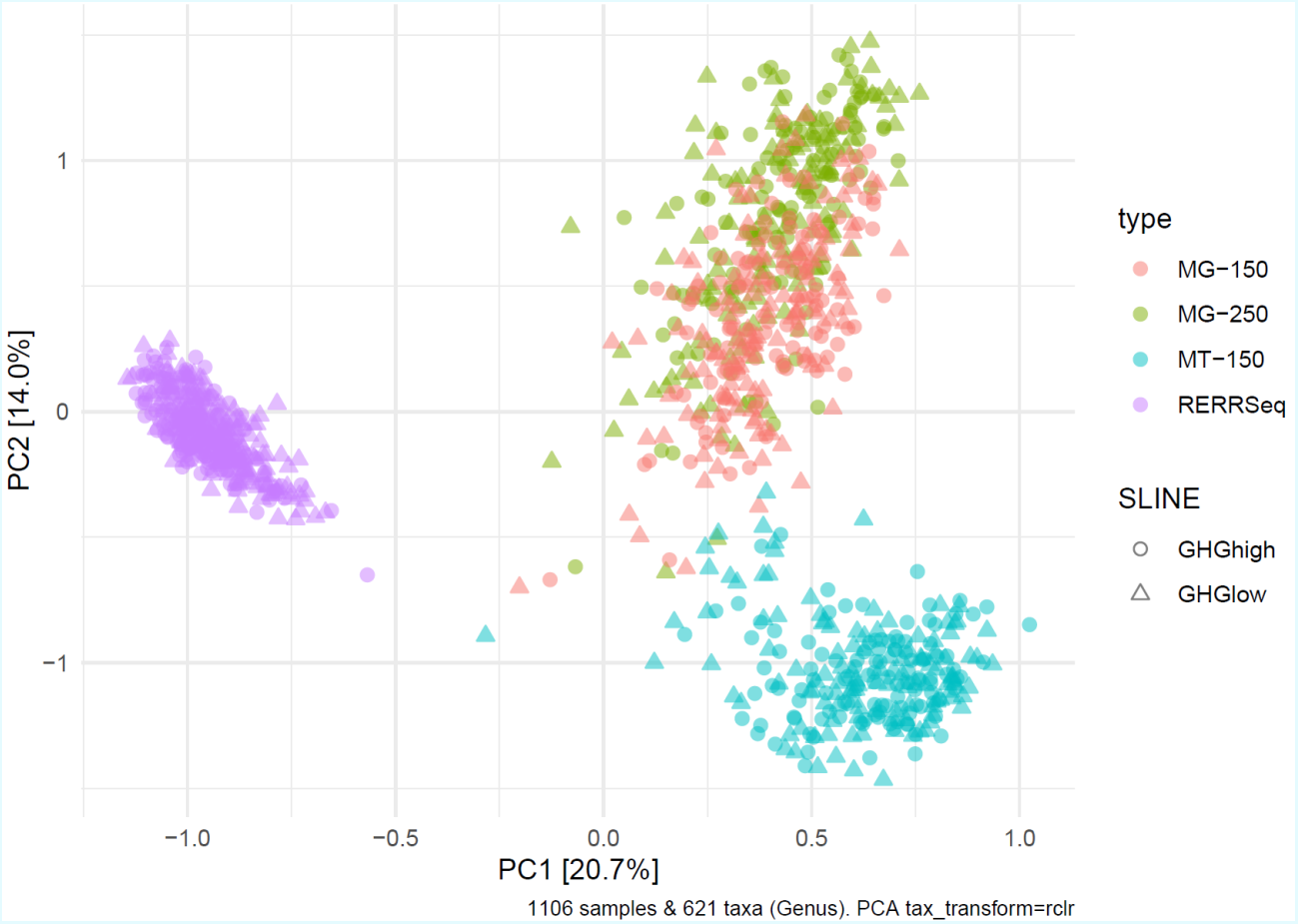
Perry et al. (2023) 31st PAG, San Diego, USA
Perry et al. (2024) 39th ISAG, Cape Town, SA.

Bioinformatic Processing: TaFeE Preprocessing and Reference-Based Profiling



https://github.com/AgResearch/TaFeE

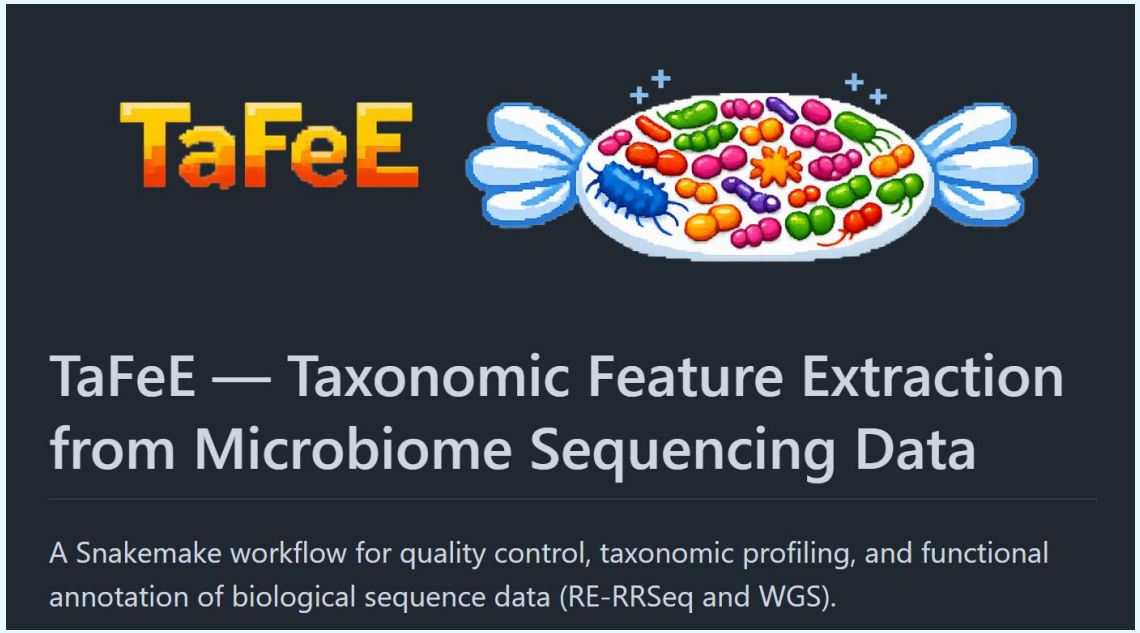
- Unified pipeline for RE-RRSeq and shotgun metagenomics and metatranscriptomics
- Kraken2 Indices are Easily Swapped



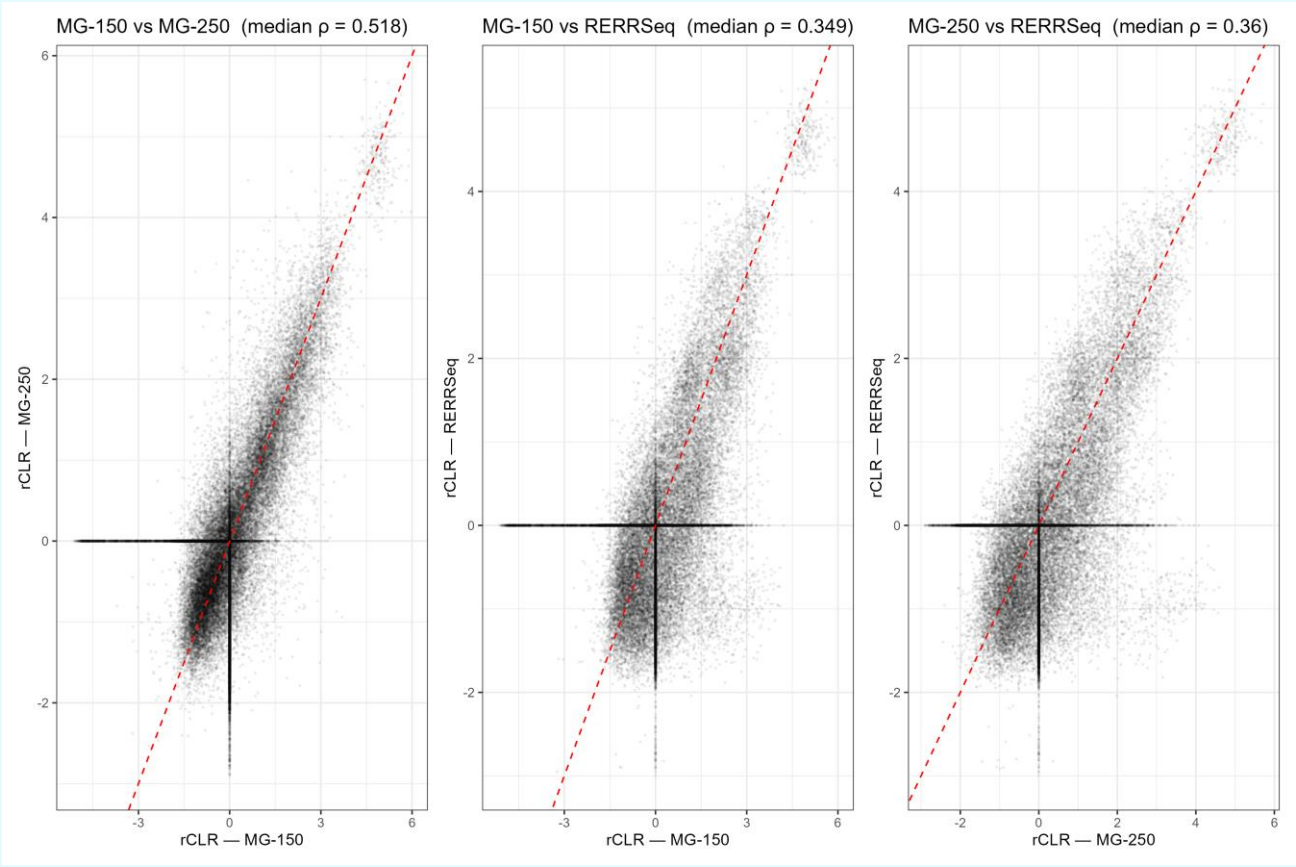
Perry et al. (2023) 31st PAG, San Diego, USA
Perry et al. (2024) 39th ISAG, Cape Town, SA.

snakemake
A framework for reproducible data analysis

Bioinformatic Processing: TaFeE Preprocessing and Reference-Based Profiling



https://github.com/AgResearch/TaFeE



- Unified pipeline for RE-RRSeq and shotgun metagenomics and metatranscriptomics
- Kraken2 Indices are Easily Swapped

Perry et al. (2023) 31st PAG, San Diego, USA
Perry et al. (2024) 39th ISAG, Cape Town, SA.

Bioinformatic Processing: Reference-Free – MRM & Methane Modelling / Prediction

- tags are trimmed 65 bp, less than 65 bp are removed
- tags is kept if count > 0 in 25% of training set
- Resultant ~200k tags in sheep for predictions (~OTUs or ASV)
- Cohort adjustment – per tag; mean zero; variance 1 (import for predicting across farms)
- Microbial relationship matrix (Hess *et al.* 2023)

Table 3 Heritability and microbiability estimates (standard errors)

Trait	Heritability	Microbiability
LW8	0.21 (0.023)	0.45 (0.060)
CH_4	0.16 (0.021)	0.54 (0.052)
CH_4Ratio	0.21 (0.021)	0.64 (0.054)
CO_2	0.21 (0.022)	0.18 (0.046)

nonredundant RERRSeq Tags

Microbial Relationship Matrix

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zm} + \mathbf{e}_m$$
$$\mathbf{m} \sim N\left(0, \sigma_m^2 \mathbf{M}\right),$$
$$\sigma_m^2/(\sigma_m^2 + \sigma_{e_m}^2).$$

Bilton et al. (2025) Genet Sel Evol **57**:25
Bilton *et al.* (2025) *AAABG* **26**:229-232

Budel *et al.* (2022) *Anim. Microbiome* **4**:39
Hess *et al.* (2020) *PLoS One* **15**:e0219882
Hess *et al.* (2023) BMC Genomics **24**:551

Bioeconomy Science Institute™
Maiangi Taiao

ben.perry@agresearch.co.nz

bioeconomyscience.co.nz

Thank you

Presentation disclaimer

Presentation for:
GMG Microbiome Working Group Meeting
2026-06-09

For more information contact:
Benjamin J Perry
DDI: +6434899003
Email: ben.perry@agresearch.co.nz

This report has been prepared by New Zealand Institute for Bioeconomy Science Limited (Bioeconomy Science Institute).

Assembly-based metagenomic profiling

Zexi Cai

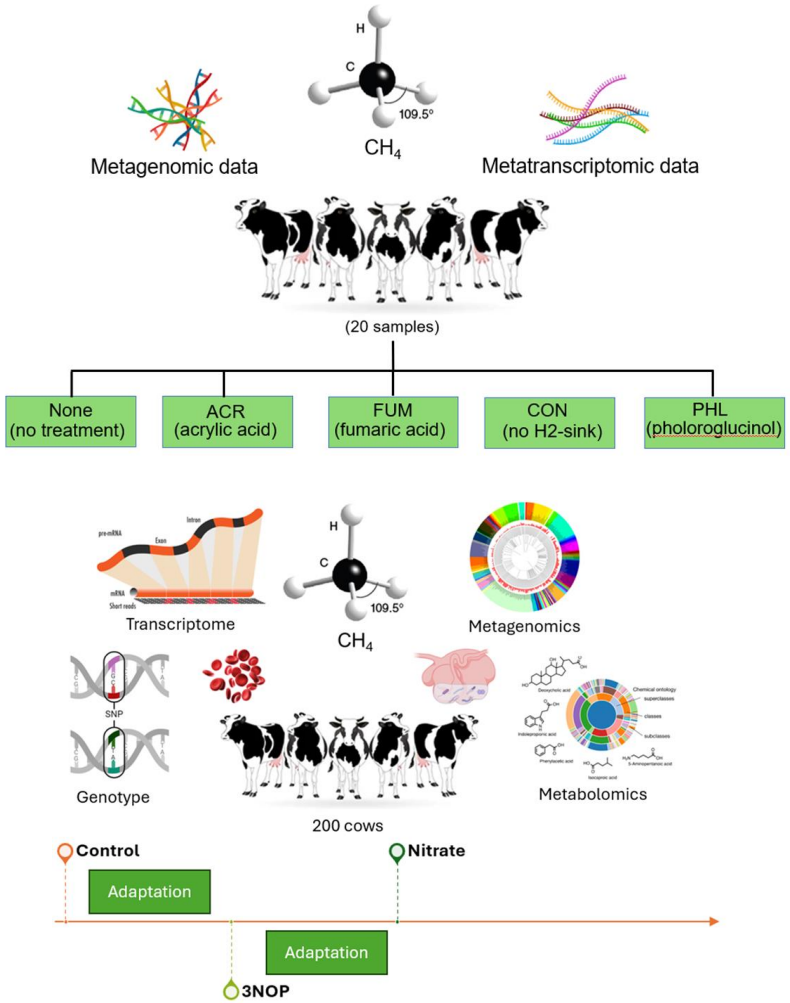
Aarhus University

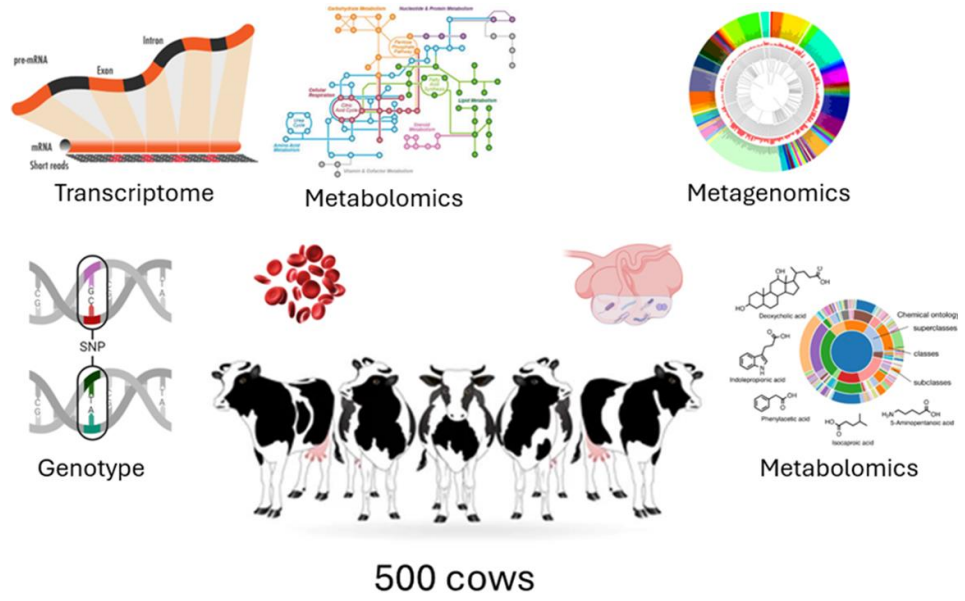
Method of choice

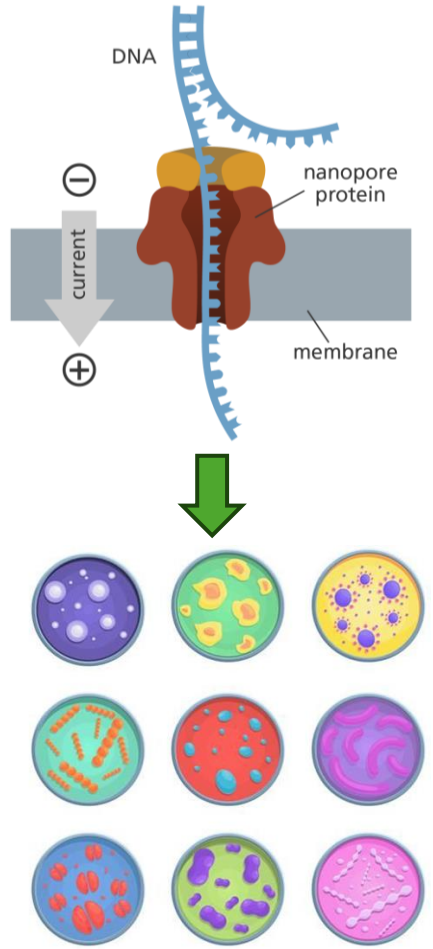
- Metagenomic genomes-based abundance profiling:
    - Possible to use per project information: every challenge for Marker-gene approach or protein search.
    - Possible for fucntional intepretation.

Metagenomic assembly -> custom database -> better taxonomy assignment

Current activity except Microhub







MAGs is the backbone

AARHUS UNIVERSITY

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

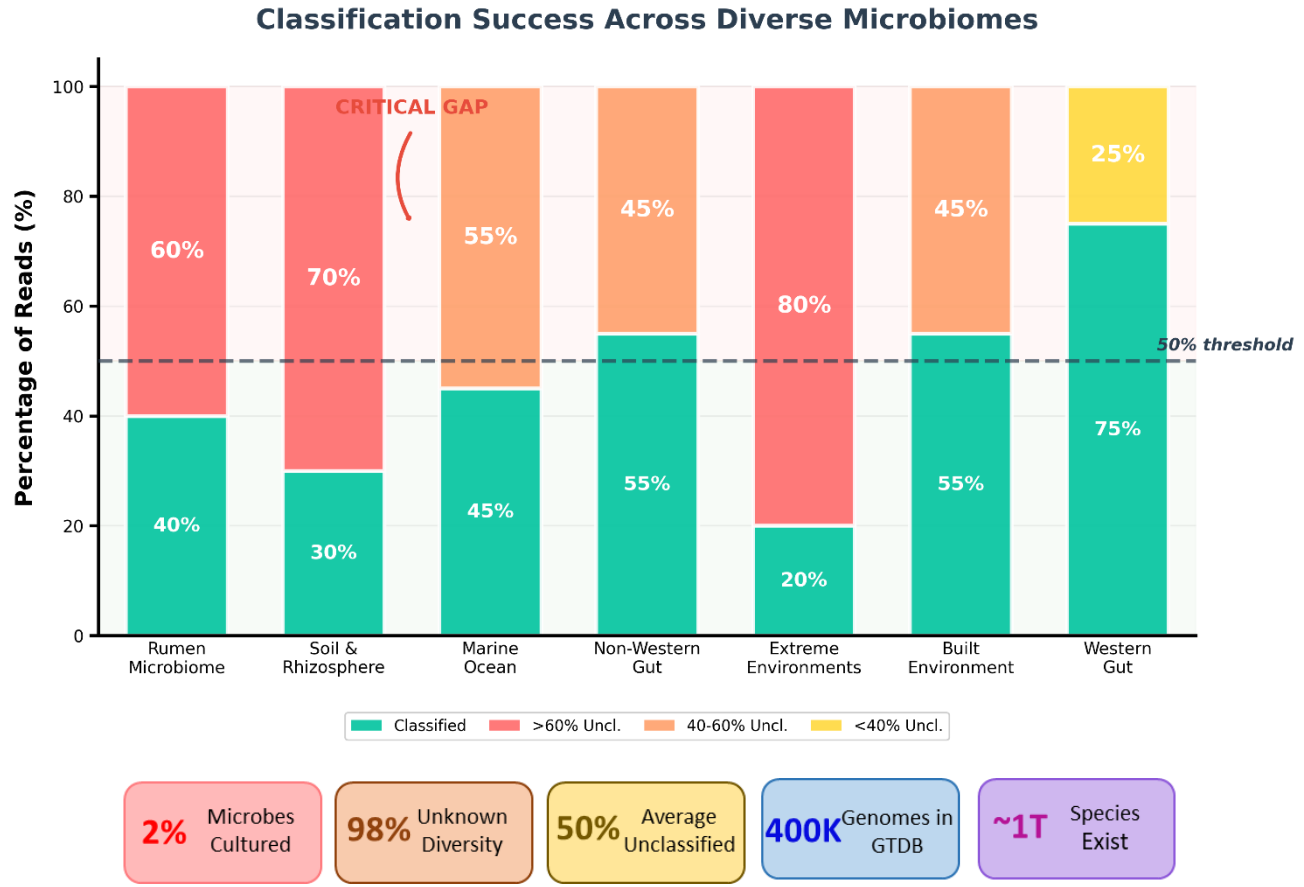
The Classification Gap in Metagenomics

In undersampled biomes, unclassified reads exceed 50% because references skew toward cultured genomes

Problem: Culture bias leaves majority of microbes unidentified

Impact: Missed taxa, poor ecological resolution

Solution: Recover MAGs → verify novelty → update the reference automatically.



AARHUS UNIVERSITY

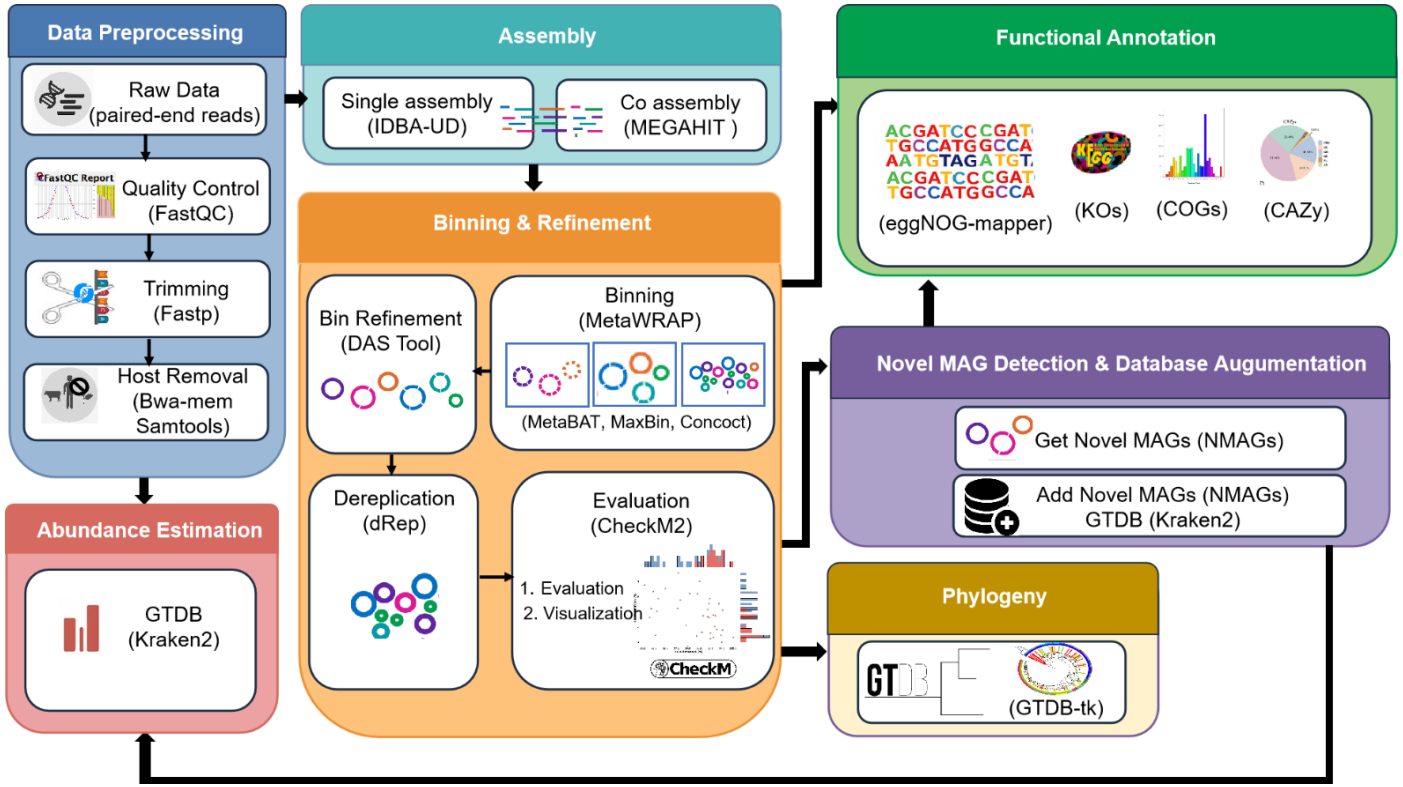
CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

MetaMAG Explorer

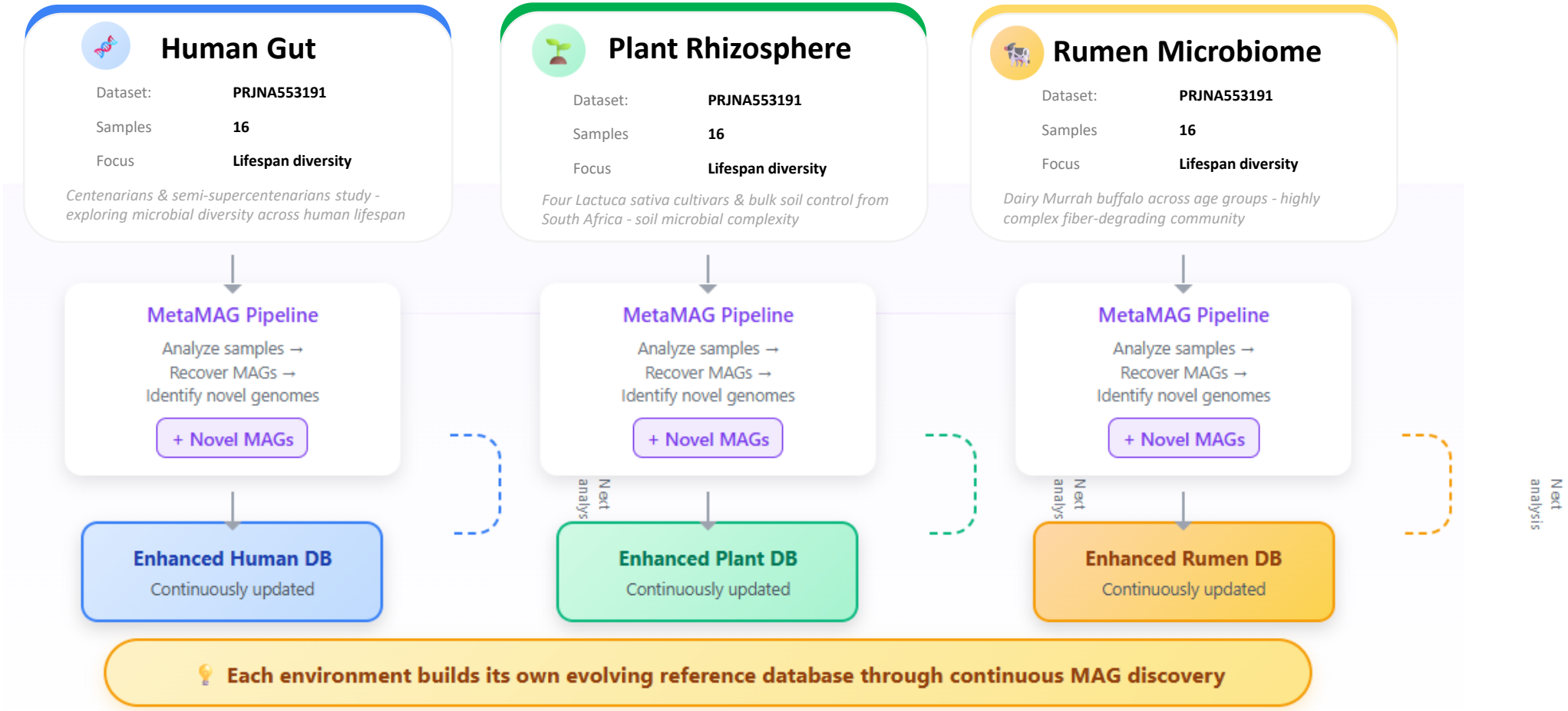
Maria Altaf Satti
Postdoc

A modular end to end pipeline that turns raw metagenomic reads into curated MAGs and updates the reference to improve classification in under-sampled biomes.

- Automatic Database Augmentation
- Novel MAG Detection & Validation
- Rumen-Specific MAG Processing



Pipeline Validation: Three Test Environments



AARHUS UNIVERSITY

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Database Augmentation Impact

+5.4M
Previously Unknown Reads Now Classified

100%
Samples Improved

+1.27%
Average Gain

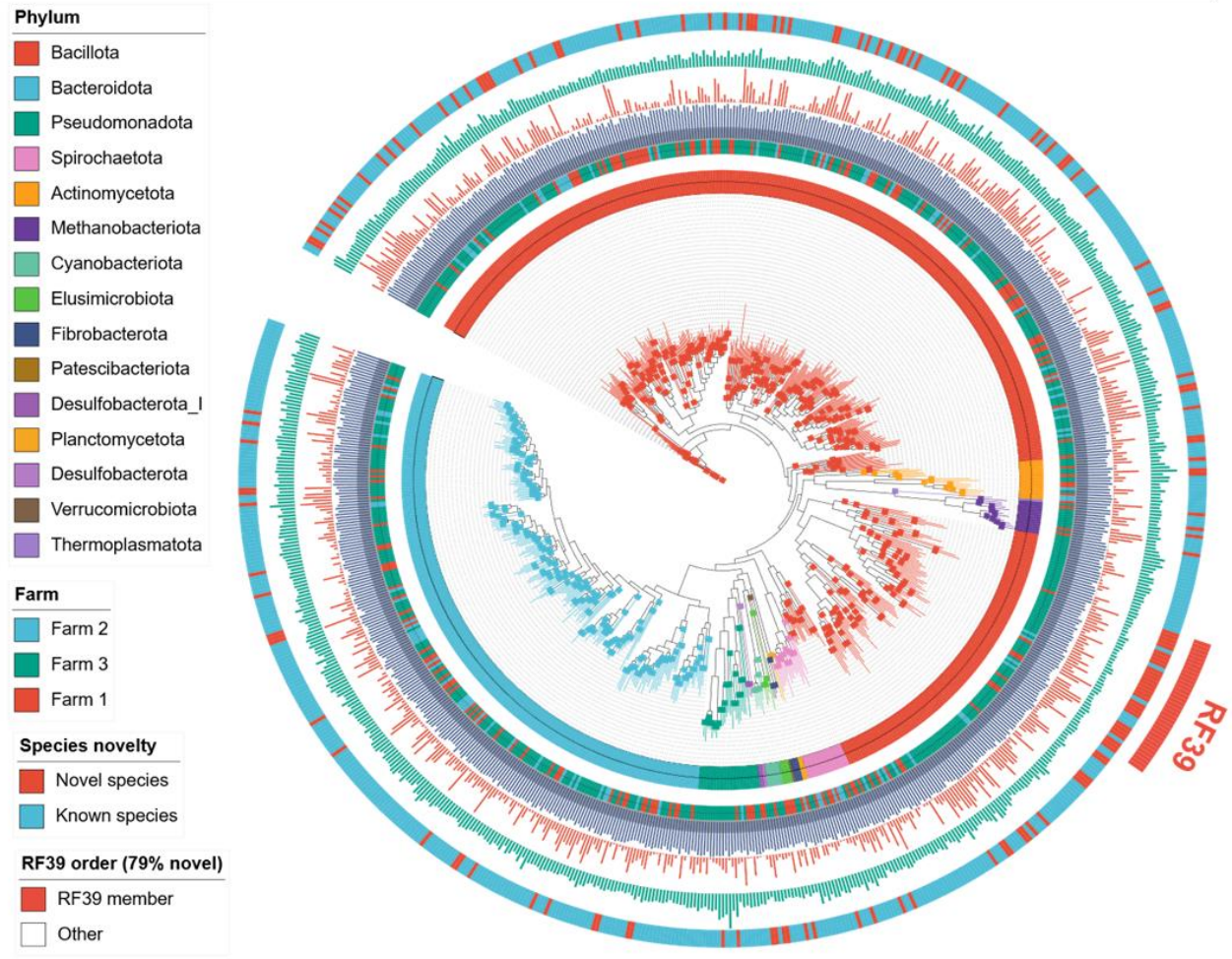
28
Novel MAGs

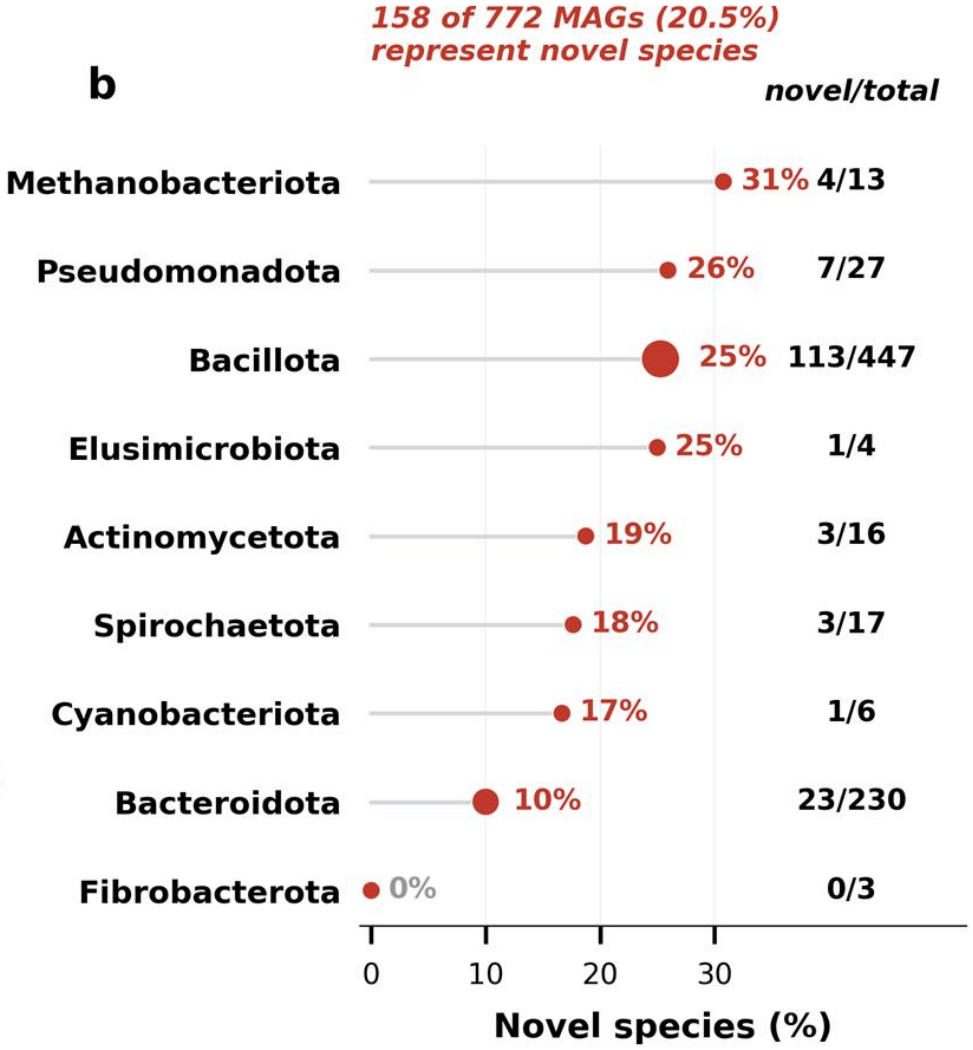
+1.78%
Best Result

Sample	Original	Enhanced	Change	New Reads
Sample_962	74.7%	76.5%	+1.78%	+522K
Sample_966	77.0%	78.6%	+1.56%	+491K
Sample_955	78.3%	79.8%	+1.54%	+419K
Sample_958	76.9%	78.3%	+1.45%	+431K
Sample_964	76.0%	77.4%	+1.39%	+410K

Novel MAGs found in **ALL 15 samples** – these are prevalent community members, not rare species

Novel species we found across various projects

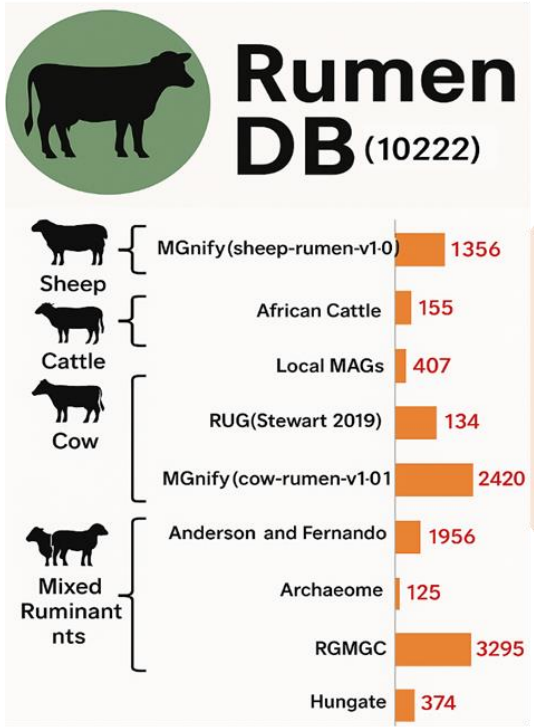


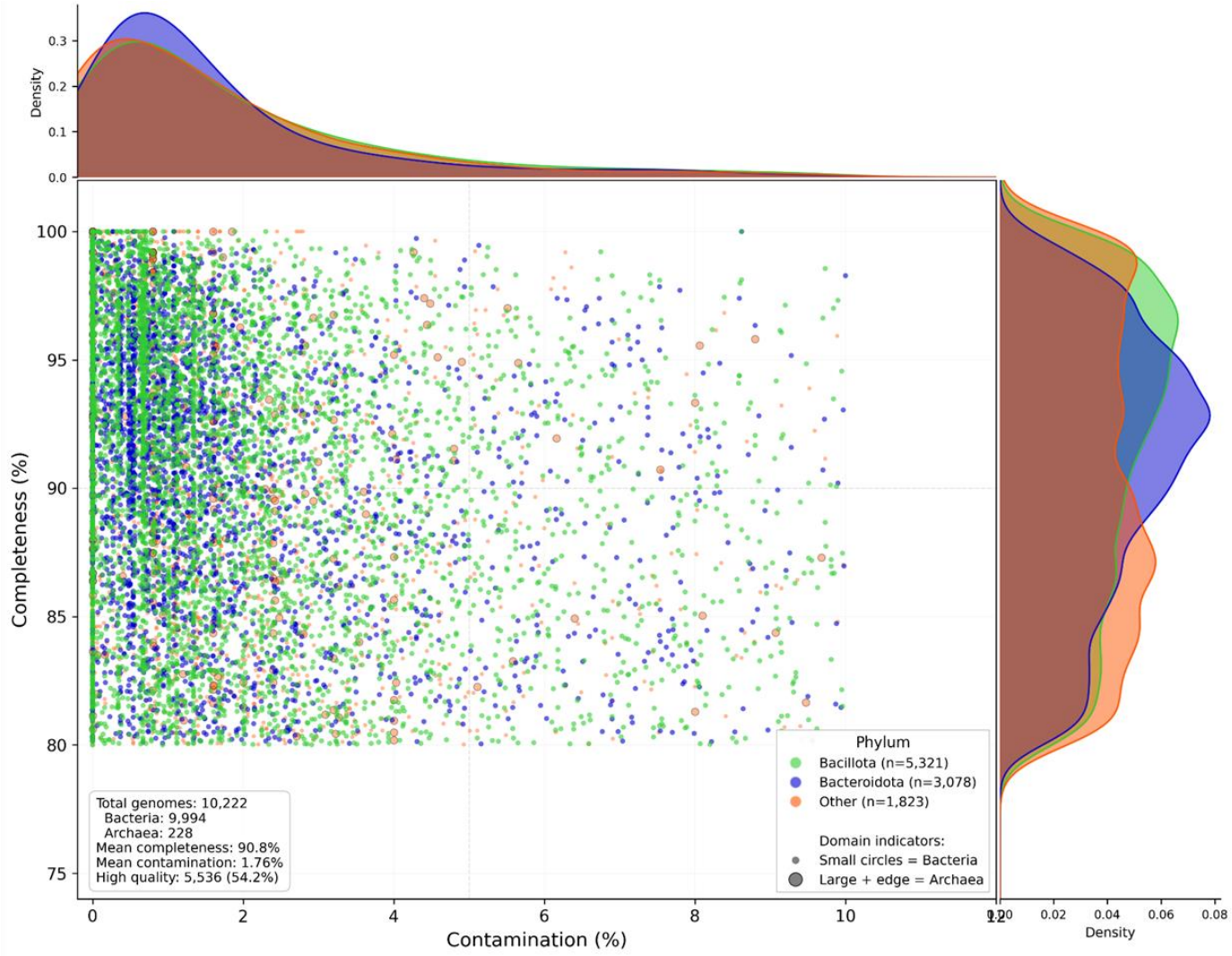


AARHUS UNIVERSITY

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Rumen database



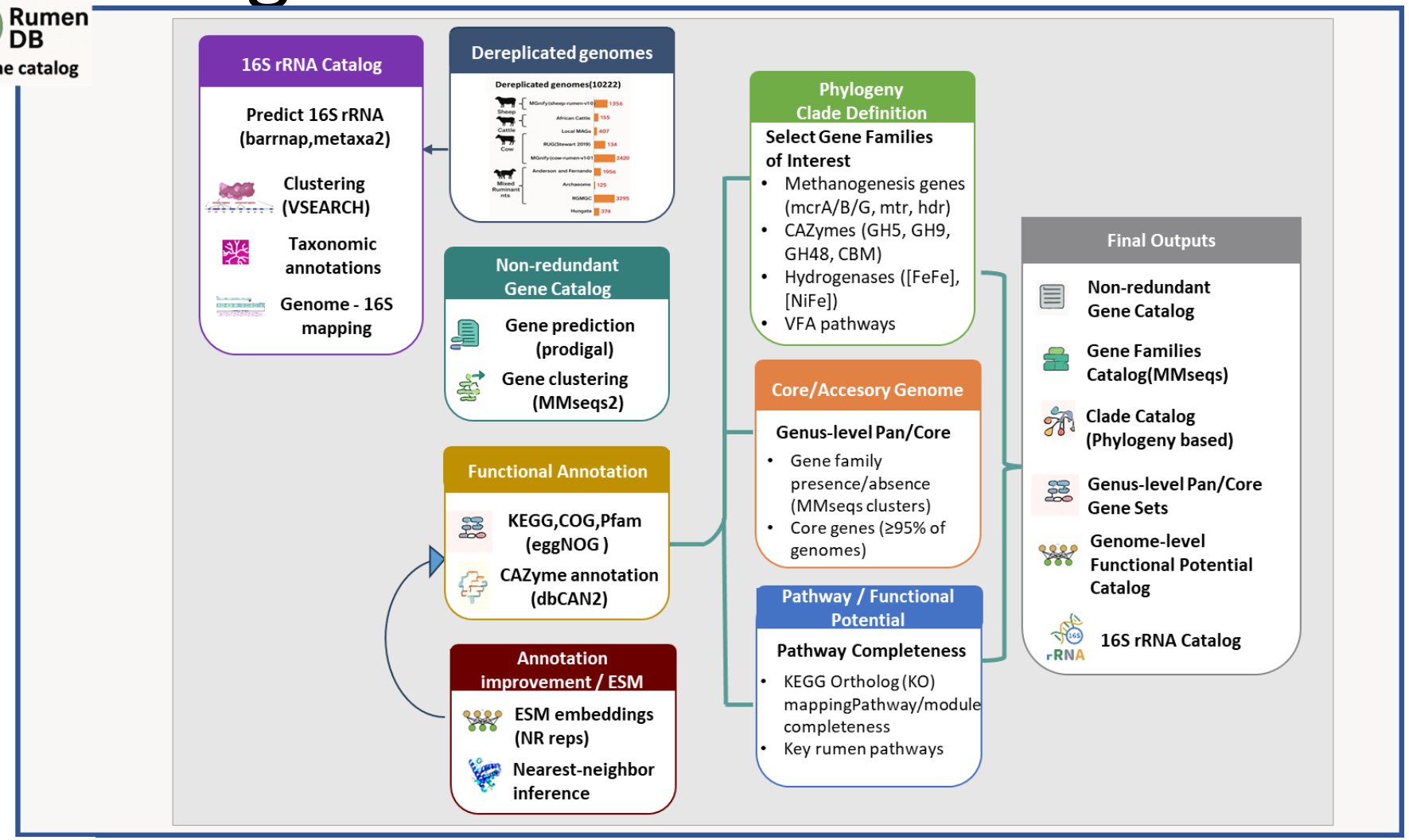


AARHUS UNIVERSITY

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Gene Catalog

Rumen DB
Gene catalog

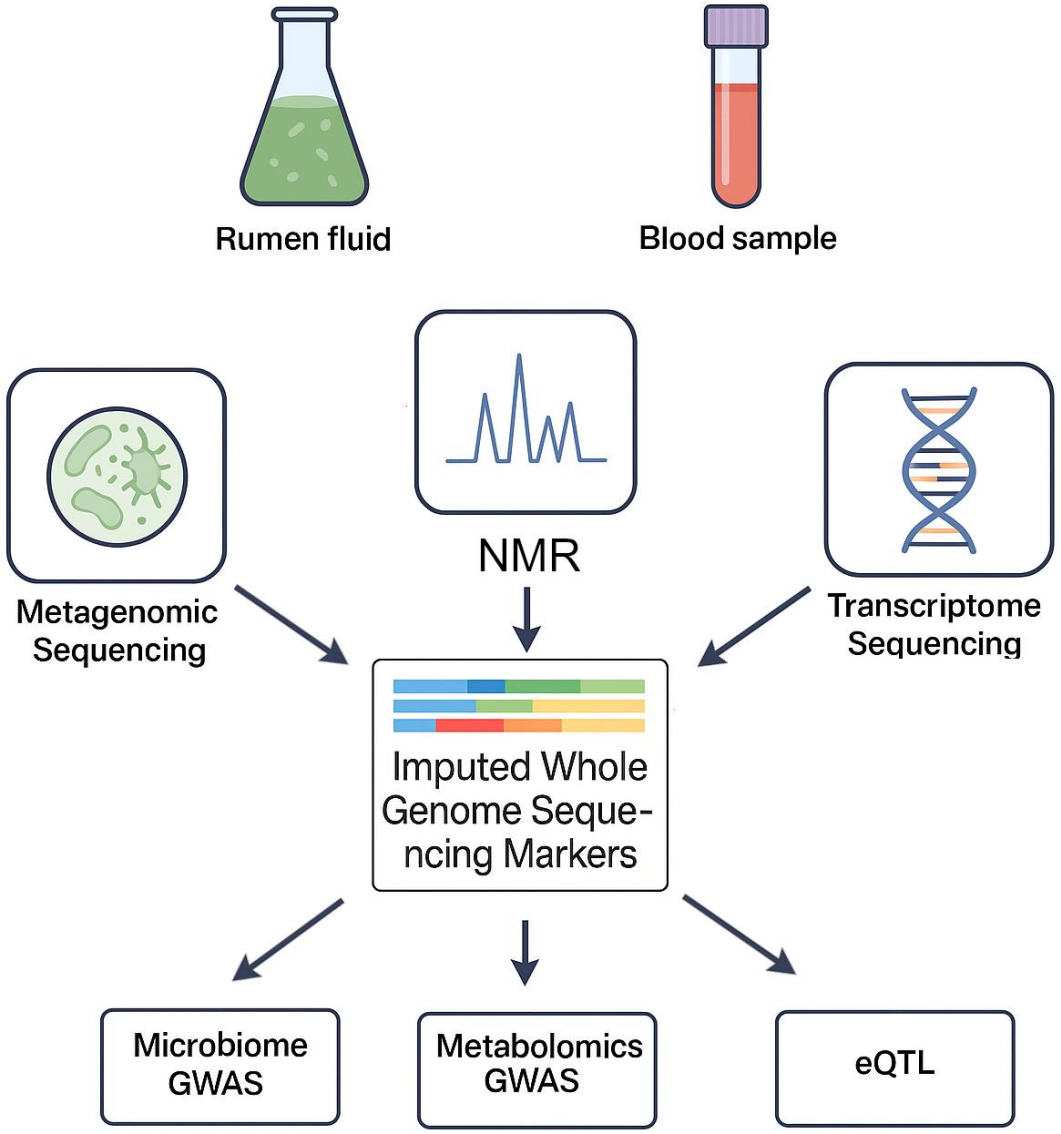


AARHUS UNIVERSITY

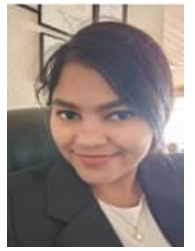
CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Holo-omics to distangle the host-microbiota interaction

Sha Zhang
Postdoc



Deep long read sequencing for enrichment media design



Israt Farhana
PhD student

4 × 4 EXPERIMENTAL DESIGN

Cow / Period	P1	P2	P3	P4
Cow 1	3-NOP + nitrate 71 Gbp	control 86 Gbp	nitrate 95 Gbp	3-NOP 97 Gbp
Cow 2	control 82 Gbp	3-NOP 118 Gbp	3-NOP + nitrate 82 Gbp	nitrate 117 Gbp
Cow 3	nitrate 113 Gbp	3-NOP + nitrate 99 Gbp	3-NOP 99 Gbp	control 81 Gbp
Cow 4	3-NOP 92 Gbp	nitrate 89 Gbp	control 112 Gbp	3-NOP + nitrate 91 Gbp

DATA VOLUME (Data)

Sample	Data Volume (Gbp)
1	71
2	86
3	95
4	97
5	82
6	118
7	82
8	117
9	113
10	99
11	99
12	81
13	92
14	89
15	112
16	91

DATA QUALITY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N50 (Higher is better)	7634	7867	7823	9132	7751	8653	7445	8635	9086	8669	7421	8683	8176	9955	7676	7873
Q score (Higher is better)	23.02	23.02	23.02	23.02	23.02	23.01	23.01	23.01	22.09	22.08	22.09	22.09	22.09	22.09	22.07	23.00

Treatments: control (gray) | nitrate (blue) | 3-NOP (yellow) | 3-NOP + nitrate (purple)

Synergy with Microhub

- With support from Microhub bioinformatics team:
  - We are contributing to the methodology for taxonomy and functional quantification.
    - We are leading the part of generation database for genome-based taxonomy quantification.
  - We are leading the genome assembly part.
- All activities in MicroHub are valuable to improve our current work.

Thanks!

ROSLIN

THE UNIVERSITY of EDINBURGH
The Royal (Dick) School of Veterinary Studies

Assigning Functional Annotation To Microbiome Reads

Optimising Pipelines for Large-Scale Ruminant Genetics

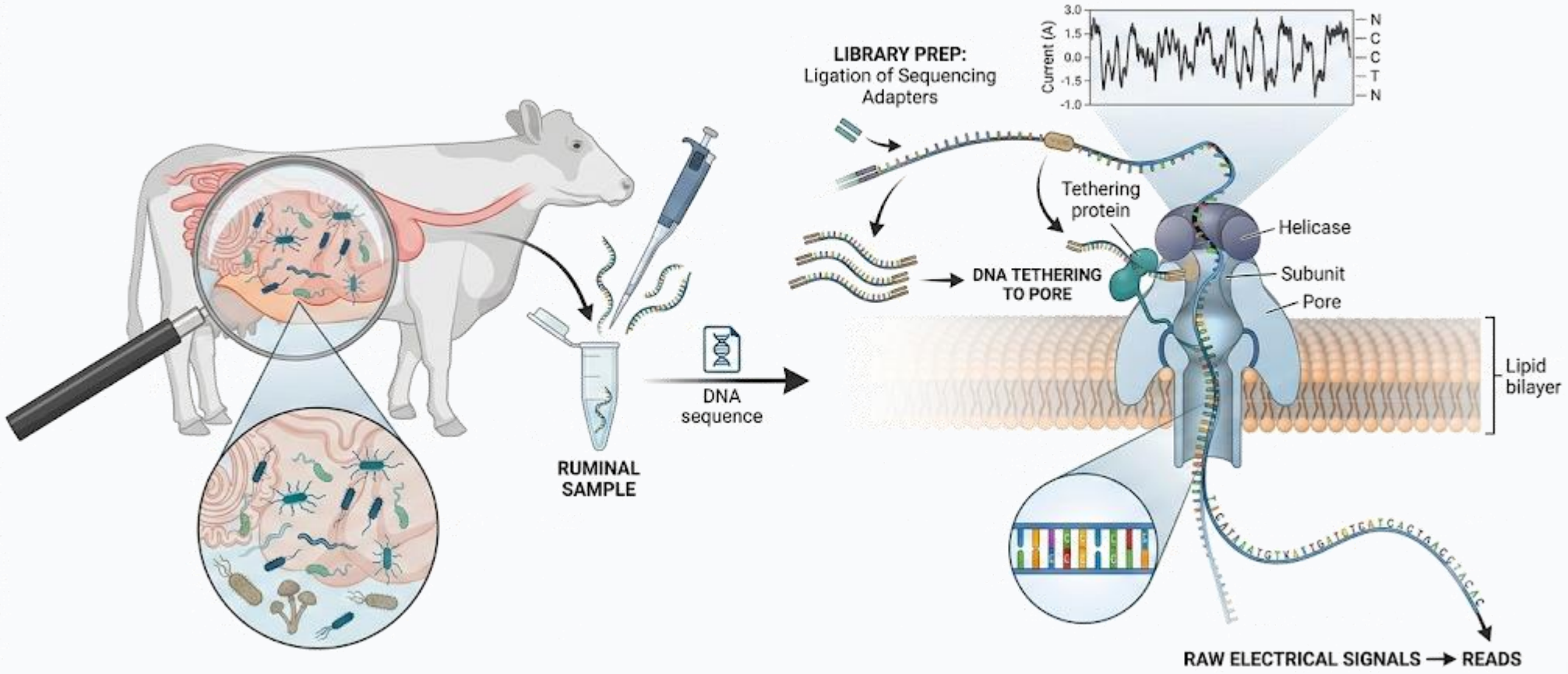
UKRI Biotechnology and Biological Sciences Research Council

Linking Microbiomes to Emissions

Our primary objective is mapping ruminal microbiomes to phenotypes of interest, specifically **methane emissions**.

By processing thousands of cattle ruminal samples through our streamlined functional pipeline, we isolate exact differential pathways (e.g. KEGG pathways) distinguishing high vs. low methane emitters.

Sequencing the ruminal microbiome



The Bioinformatic Fork

Pathway A: Taxonomic Annotation

WHO IS THERE?

Profiles the structural composition and species spectrum native to the rumen ecosystem

- **Reference Mapping:** Aligns raw reads directly against comprehensive genomic reference databases
- **Multi-Kingdom Detection:** Profiles Bacteria, Eukaryotes (protozoa/fungi), and Archaea simultaneously
- **Taxon Resolution:** Focuses on species-level classifications to identify exact names (e.g., Methanobrevibacter ruminantium)

Pathway B: Functional Annotation

WHAT ARE THEY DOING?

Deciphers the active metabolic potential and biochemical machinery within the microbiome

- **Enzymatic Pathways:** Maps genomic read signatures to chemical processes and pathways
- **KEGG Databases:** Aligns sequence reads to functional metabolic networks and structural biochemical pathways
- **COG Classifications:** Group proteins into broad, conserved evolutionary clusters to track core cellular mechanisms

Why Functional Profiling Matters

TAXONOMY ONLY

"Who is there?"

Provides a simple descriptive catalog of species (e.g., Escherichia coli 15%)

- **Uncertain Capabilities:** Highly heritable names but unclear physical pathways or functional outputs
- **High Redundancy:** Species taxonomy varies wildly across herds, while chemical capability is identical
- **Weak Associations:** Suboptimal correlation with complex animal phenotypes (methane, feed intake)

FUNCTIONAL PATHING

"What are they doing?"

Provides mechanistic insight by directly measuring biological pathways and enzymes

- **Conserved Roles:** Focuses on active pathways (butyrate production, methane metabolism) despite shifts in individual taxons
- **Direct Selectability:** Measures biochemical targets directly affected by the host genome
- **Biological Model:** Uncovers antibiotic resistance, cellulose degradation, and core metabolism

Our pipeline

SqueezeMeta: a fully automated metagenomics pipeline, from reads to bins

NEW: Now you can inspect the results of SqueezeMeta using our interactive viewer Watermelon.

sqm_longreads.pl

Taxonomy and functional annotation on the raw reads
Diamond Blastx

sqm_annot.pl

Only functional annotation
Prodigal

The Computational Bottleneck

Traditional Pipeline

The standard script **sqm_longreads.pl** executes Diamond Blastx on the raw reads to acquire taxonomy & function simultaneously.

50–60h Per Sample

40 Years Total CPU Time

The Streamlined Strategy

Splitting functional profiling and taxonomic profiling into separate dedicated lightweight steps via the custom script **sqm_annot.pl**.

2h Func. Annot

+2h Taxonomy (Bracken)

Faster Pipelines

Over 90% Processing Time Saved

By utilizing decoupled pipelines, we avoid massive Diamond Blastx sequence space runs, which completely eliminates unnecessary processing overhead.

Scale calculation for 7,000 samples:

- Standard Pipeline: **385,000 Core Hours**
- Decoupled Pipeline: **28,000 Core Hours**

Accelerated Pipeline Validation

ACCURACY PROOF

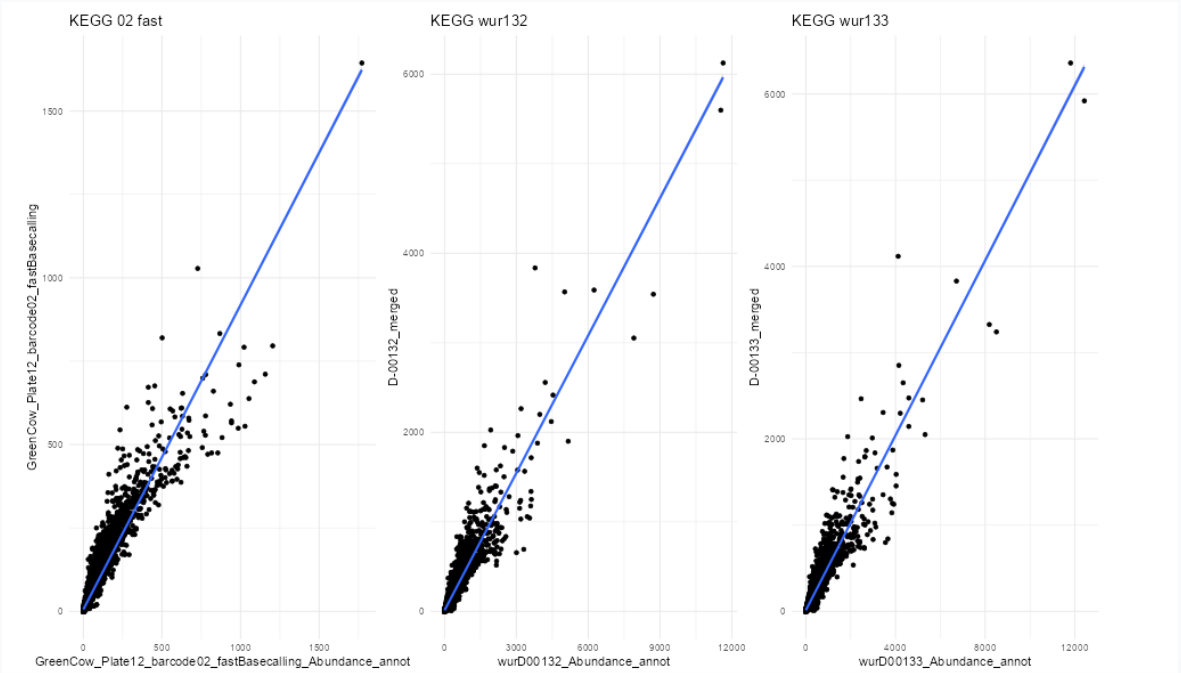
Identical Biological Insight

Plotting relative abundance profiles derived from the full pipeline against the fast functional annotator results shows remarkable consistency

Functional Correlation = 0.94

Highly stable alignment was verified across COG and KEGG databases on multiple samples

The fast pipeline delivers the exact same selection values with negligible variance



Mapping the Methane Phenotype

DIFFERENTIAL ABUNDANCE

Isolating Methane Biomarkers

To decipher pathways driving emissions, cows were ranked into quartile groups (Low, L-Mid, H-Mid, High) using Breath Infrared detectors

- **Linear Regression (Limma):** Identified 279 Differentially Abundant KEGGs (KEGGs-DA) with adjusted P < 0.05
- **Emission Clusters:** 182 KEGGs linked to High Methane, and 97 KEGGs linked to Low Methane emitters
- **Functional PC1:** Aggregating KEGGs-DA explained 53.0% of functional metagenome variation

LIMMA LINEAR REGRESSION AND BENJAMINI-HOCHBERG CORRECTION

ELSEVIER

Journal of Dairy Science

Volume 104, Issue 7, July 2021, Pages 8135-8151

Research

A dimensional reduction approach to modulate the core ruminal microbiome associated with methane emissions via selective breeding

Alejandro Saborío-Montero [1 2], Adrían López-García [1], Mónica Gutiérrez-Rivas [1], Raquel Atxaerandio [3], Idoia Goiri [3], Aser García-Rodriguez [3], José A. Jiménez-Montero [4], Carmen González [1], Javier Tamames [5], Fernando Puente-Sánchez [5], Luis Varona [6], Magdalena Serrano [1], Cristina Ovilo [1], Oscar González-Recio [1 7]

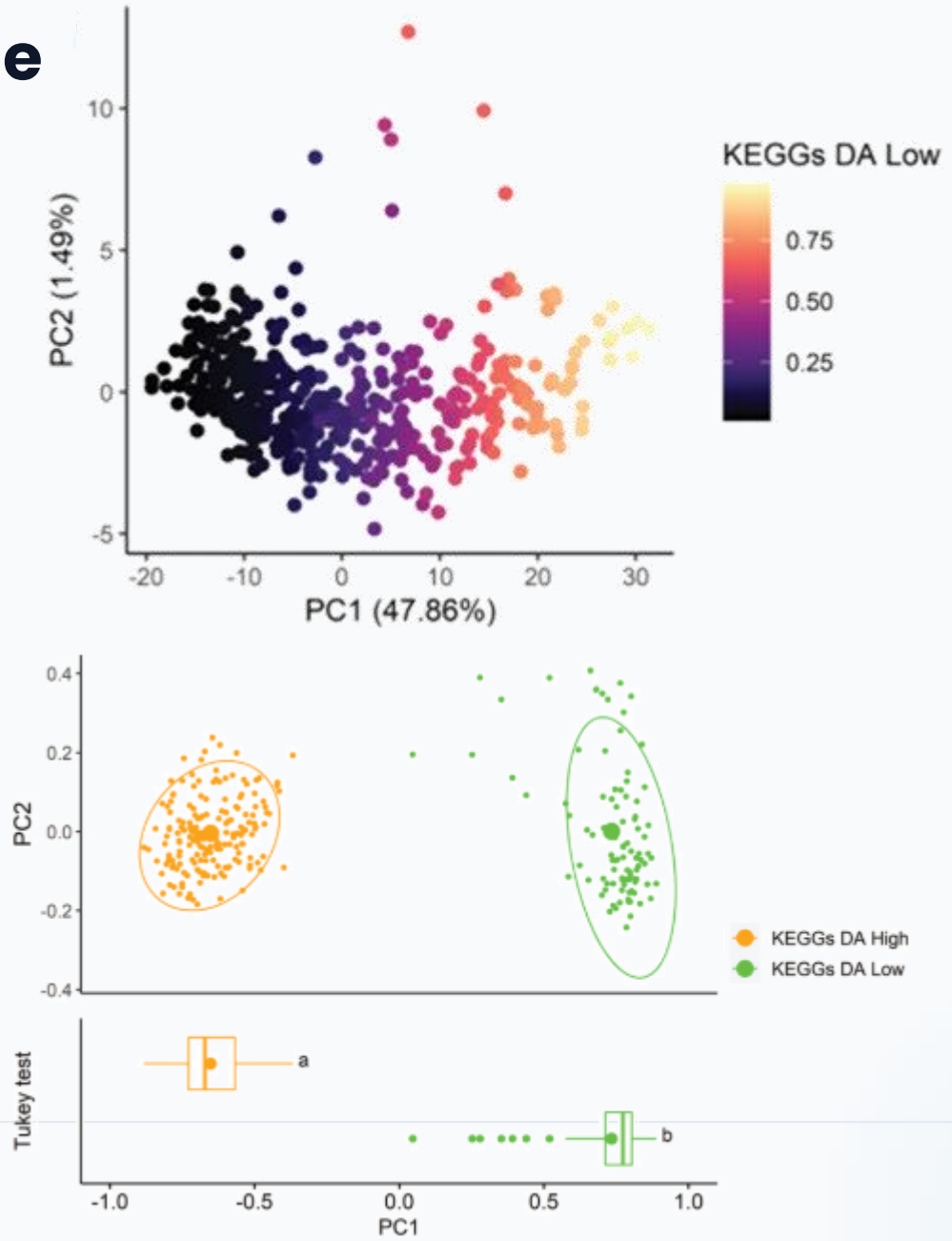
Host Genomic Control of Microbiome

HOLOGENE HERITABILITY

The Core Rumen is Heritable

By running bivariate animal models with genomic relationship matrices, we proved the host genotype regulates the PC1 aggregated metagenome.

- **PC1 Heritabilities:** Consistent ~0.30 across all taxonomic levels.
- **Functional Heritability:** KEGGs-DA heritability reached
- **Direct Selectability:** Rumen properties are more heritable than physical methane measurements (~0.15)

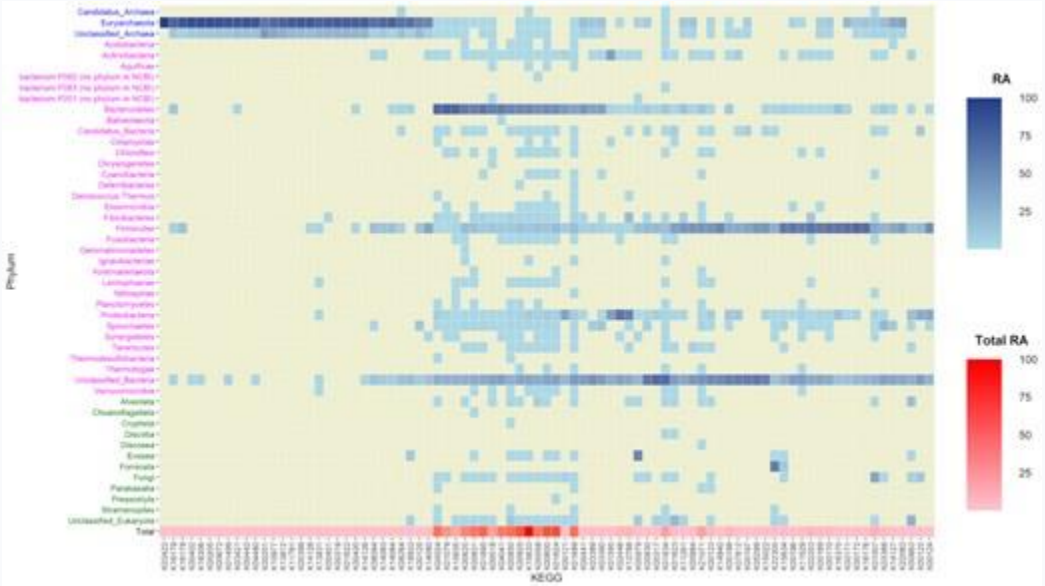


Host Genomic Control of Microbiome

We do not lose information

We can still know:

- What genes are more active in different taxa
- Keep track of what genera/species are associated with functions of interest
- Important to understand the micro-ecology



Final ideas

Dimension Reduction

Condensing thousands of differential KEGG pathway variables into structured Principal Components (PCs) for cleaner analysis.

Heritability Mapping

Variations in these microbiome PCs are highly heritable and controlled by the host genome.

Selective Breeding

Incorporating these host-microbiome genetic traits into national breeding indexes to select for low-emission cattle.

ROSLIN

Thank you