

# Enteric Fermentation Flagship Project


## Rumen microbiome to predict methane emissions from cattle

Larissa Zetouni, M.K. Hess, R. Brauning, H. Henry, T. Van Stijn, A. McCulloch, J. Budel, J. C. McEwan, H. Flay, M. Camara and S.J. Rowe

# Acknowledgments

**GLOBAL RESEARCH ALLIANCE**  
ON AGRICULTURAL GREENHOUSE GASES

## AT A GLANCE



**57** member countries

**4** Research Groups

- Paddy Rice Research Group
- Livestock Research Group
- Croplands Research Group
- Integrative Research Group

**18** Science Networks

**17** partner organisations


Over **3000** scientists involved in activities of the GRA

**66** international collaborative projects supporting the GRA

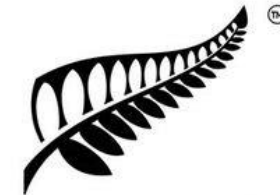
**87** fellowships awarded to recipients from **33** countries

**32** technical training workshops held

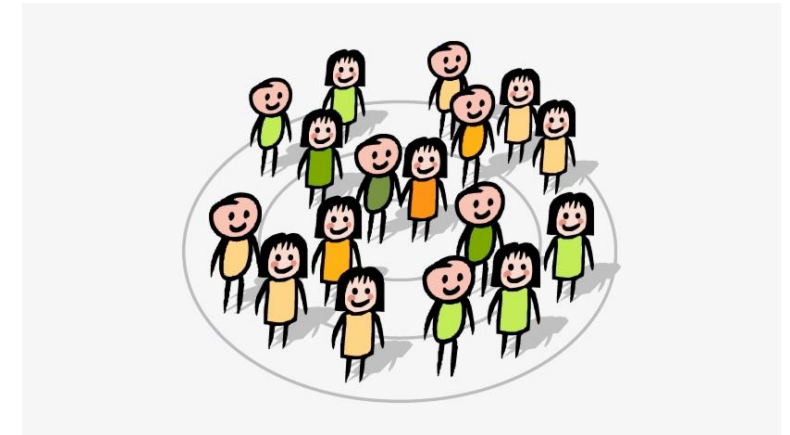
**21** technical guidelines, resource materials and databases produced



[globalresearchalliance.org](http://globalresearchalliance.org)      @GRA\_GHG      June 2019



**NEW ZEALAND**  
AGRICULTURAL GREENHOUSE GAS  
Research Centre

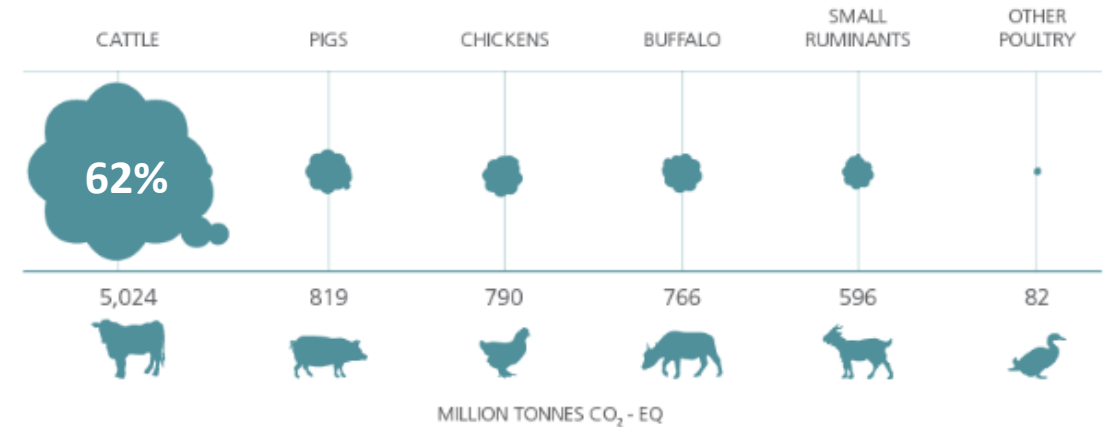
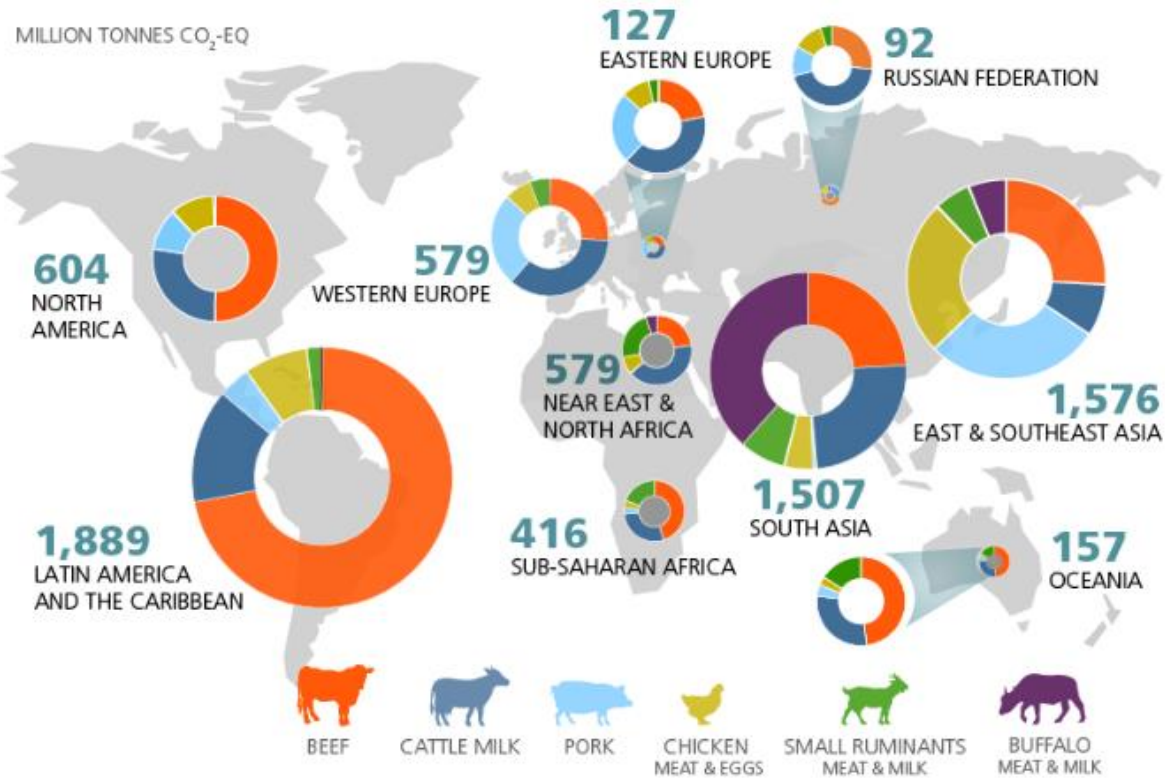


## The age of microbiomes



- Difford et al. (2018): methane production influenced by cow's individual genetic composition and rumen microbial composition
- Hess et al. (2019): microbial profile is heritable and phenotypically and genetically correlated to methane emissions in sheep

# Enteric Fermentation Flagship Project



**Global estimates of emissions by species.** It includes emissions attributed to edible products and to other goods and services, such as draught power and wool. Beef cattle produce meat and non-edible outputs. Dairy cattle produce milk and meat as well as non-edible outputs.

**Regional emissions.** Regional total emissions and their profile by commodity are shown. Results do not include emissions allocated to non-edible products and other services.

# Enteric Fermentation Flagship Project

## Collaborators

- Argentina
- Australia
- Belgium
- Ireland
- Mexico
- Philippines
- South Korea
- USA
- *Bos taurus*
- *Bos indicus*
- *Bubalus bubalis*
- Beef
- Dairy
- Dual purpose
- No data sharing
- Each collaborator gets their own results back
  - Sequencing
  - Predictors within dataset
- Developed countries
- Developing countries

# Rumen microbiome to predict methane emissions

**Restriction Enzyme Reduced  
Representation Sequencing  
(RE-RRS)**

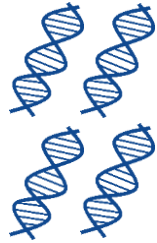
**Reduces genome  
complexity by digestion  
of DNA by restriction  
enzymes**



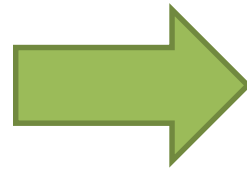
- High-throughput method that captures the diversity of the rumen microbiome
- Low cost – similar to 16S
- Microbial community profile shown to be heritable
- Useful methane predictor

# RE-RRS: Restriction Enzyme Reduced Representation Sequencing

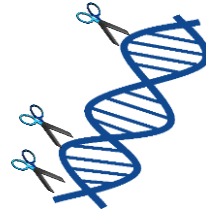
DNA extracted  
from sample



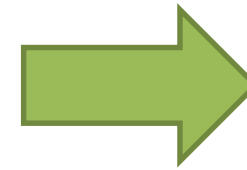
Global Rumen Census freeze-dried DNA extraction method



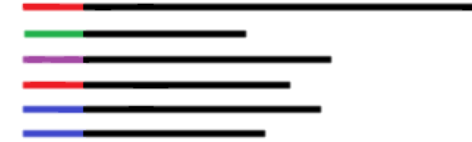
DNA digested by  
restriction enzyme(s)



ApeKI or PstI



Ligation with  
'barcode' oligos

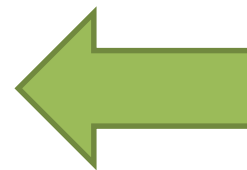


To link sequences  
with samples

Fragments  
sequenced



Illumina HiSeq2500  
236 samples across 2 lanes



Fragment  
size-selection



193-318bp



Hess et al. (2019)

## Reference-Based Bioinformatic Pipeline

### BLAST

against the Hungate 1000 Collection

Query ID	Database Hit	BitScore
A	Prevotella_ruminicola_D31d	104
A	Prevotella_sp_RM4	107
A	Prevotella_sp_KHP1	99



Assign at genus level  
using the MEGAN algorithm

Query ID	Genus
A	Prevotella



Obtain Metagenome Profile:  
table of counts for each genus

SampID	C1	C2	C3	C4	C5	C6
Samp1	1	0	0	0	4	0
Samp2	1	1	0	1	6	0
Samp3	2	0	0	0	9	1
Samp4	3	0	1	0	11	0
Samp5	4	0	0	1	6	0
Samp6	0	1	0	0	7	0
Samp7	2	0	0	0	9	0
Samp8	0	0	0	1	2	0
Samp9	0	0	1	0	3	1

## Reference-Free Bioinformatic Pipeline

### Identify common tags

Tag = first 65 bp of read

Common tag = present in  $\geq 25\%$  of samples

```

Name  Tag
Tag1  TGCAGAAGCCGAGAAAGCGATCTTTCTGGACCGGACGGCAGCATCAACCGCTATGTGGGCTTT
Tag2  TGCAGCAGAGGAGTTTCATGGGCAGGCTGAAGCTGATCGCCATTGAGACCCATGTAGCAGATGCTC
Tag3  TGCAGCGTTGGTGCCGACCTTGATATTCAGTTCGTGTGTGCCGAAAGTC TGTGCGGTGAGGATTA
Tag4  TGCAGGCCACGCTGGAATACCTGGACCGGAACCTGCCGGACAGGTGCCAGGCCGCGGAGATC
Tag5  TGCAGGTGGTAGAGCTGGAAGTCGAGTGCCGACGAGATGCGCGCTCCACCATGCTCATGGTCTC
Tag6  TGCAGTTGCTGACTGATCAGCAGGTGCGGCAGTTTCAGCAGCCTGTGCGGATGATGTGCCGCT
    
```



Obtain Metagenome Profile:  
table of counts for each tag

SampID	C1	C2	C3	C4	C5	C6
Samp1	1	0	0	0	4	0
Samp2	1	1	0	1	6	0
Samp3	2	0	0	0	9	1
Samp4	3	0	1	0	11	0
Samp5	4	0	0	1	6	0
Samp6	0	1	0	0	7	0
Samp7	2	0	0	0	9	0
Samp8	0	0	0	1	2	0
Samp9	0	0	1	0	3	1



# Comparison of Approaches

## Correspondence Analysis

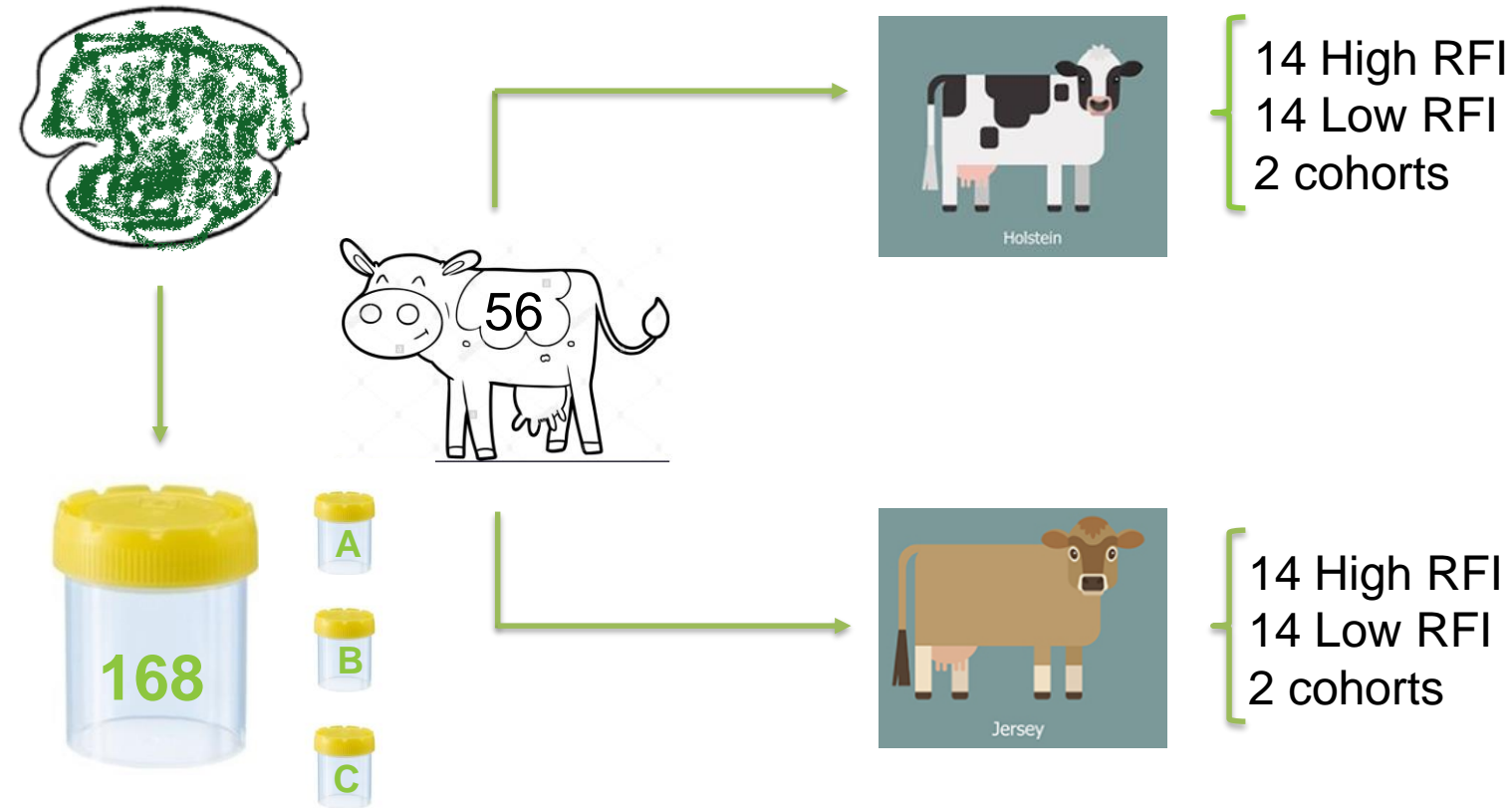
Pedigree-based estimates for the first dimension

Approach	$h^2$	Repeatability	Correlation CH <sub>4</sub> Yield
16S	0.26 (0.23)	0.45 (0.08)	0.63 (0.49)
ApeKI_RB	0.58 (0.32)	0.61 (0.06)	0.63 (0.31)
PstI_RB	NE	0.60 (0.06)	NE
ApeKI_RF	0.18 (0.25)	0.60 (0.06)	NE
PstI_RF	0.24 (0.27)	0.62 (0.06)	0.88 (0.25)

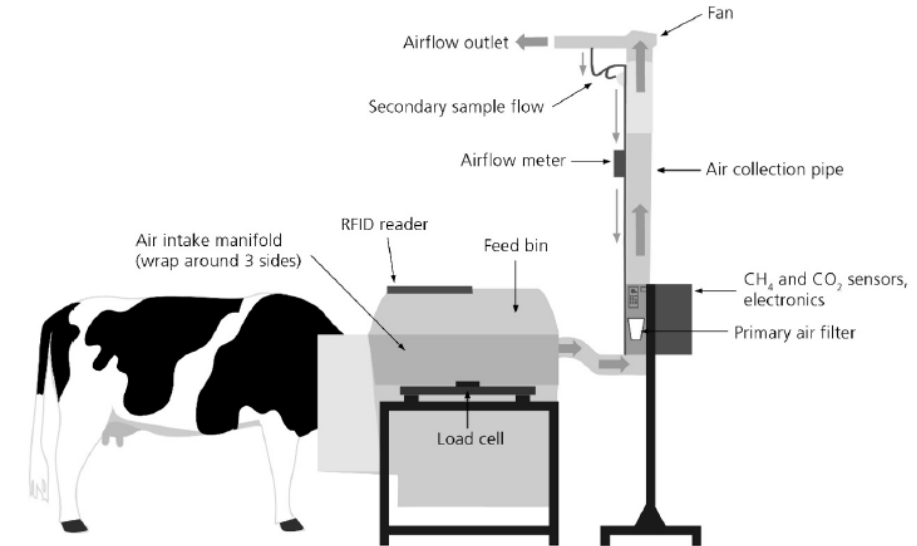
NE = Not Estimable

\* Permanent Environment constrained to zero

# Method validation with cattle rumen samples *DairyNZ*

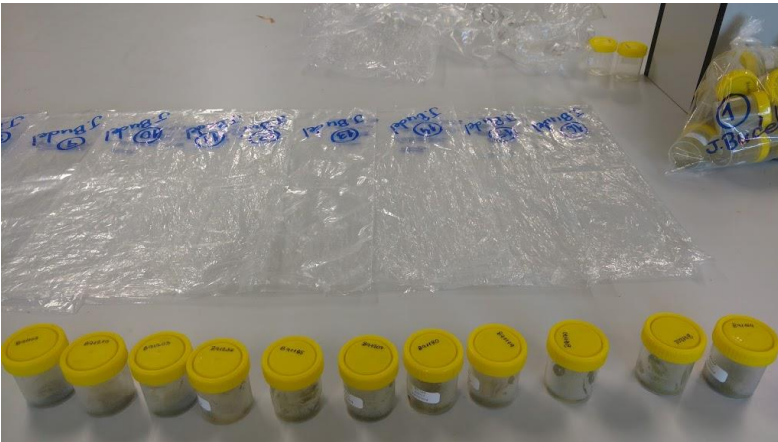


Sequences combined  
Low repeatabilities



- Adapted GreenFeed - DMI-CH<sub>4</sub> stations (Zimmerman et al. 2015)
- Animals sole diet can be fed from stations
- Same diet for all animals - lucerne hay cubes

# Workflow for processing rumen samples



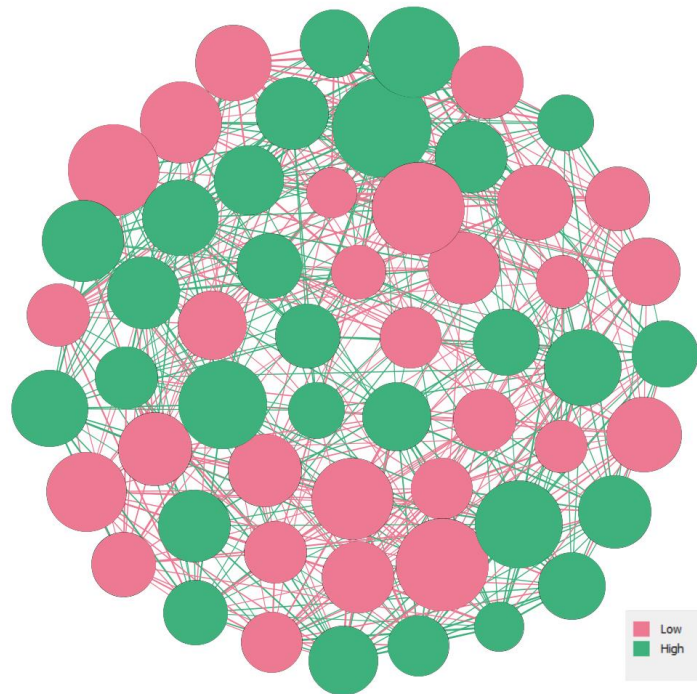
# Cattle sequencing results compared to sheep

Information	Parameter	Sheep1	Sheep2	Cattle
Sequencing	Number of Samples	236	654	186
	Samples per Lane	118	164	188
	Number of Reads/Sample	2.7M ± 680k	1.5M ± 586k	759k ± 147k
Reference-Based	Percent Assigned	6.8 ± 1.8	6.3 ± 1.2	9.3 ± 1.6
Reference-Free	Number of Tags*	503k	375k	423k
	Percent Assigned	38.1 ± 16.2	39.1 ± 3.2	64.3 ± 6.8

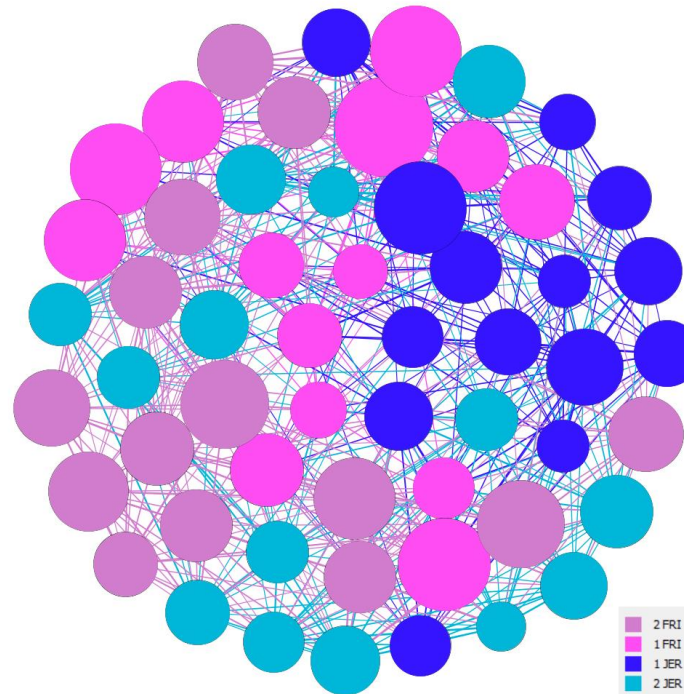
\* Tags are 65bp reads present in at least 25% of samples

# Network analysis – Reference based

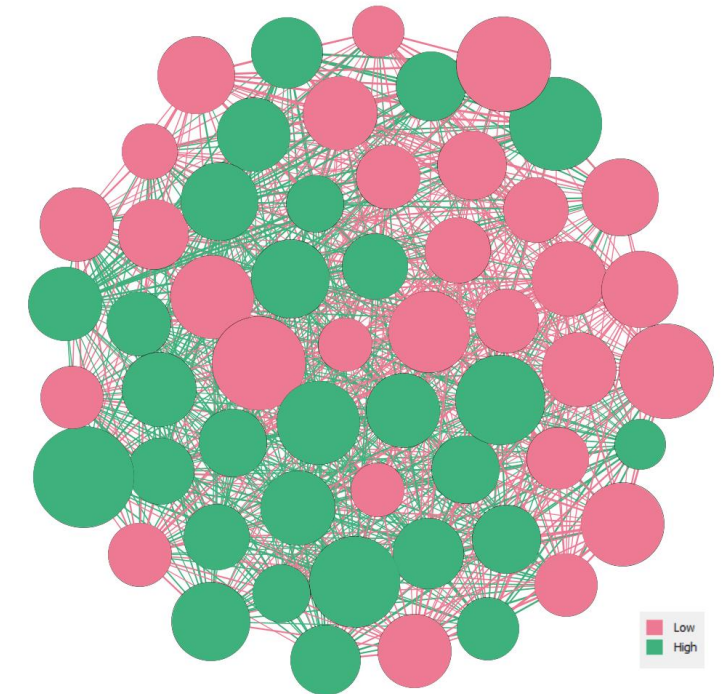
Unadjusted



Unadjusted

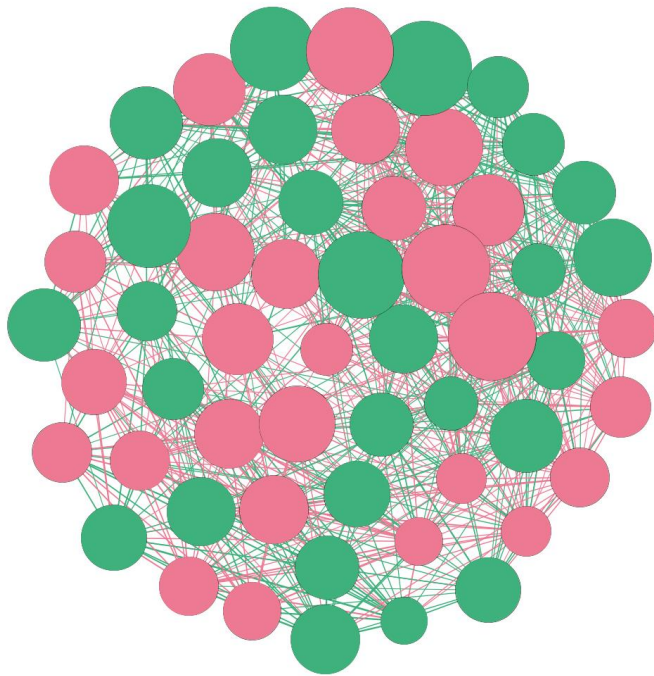


Adjusted by Cohort x Breed

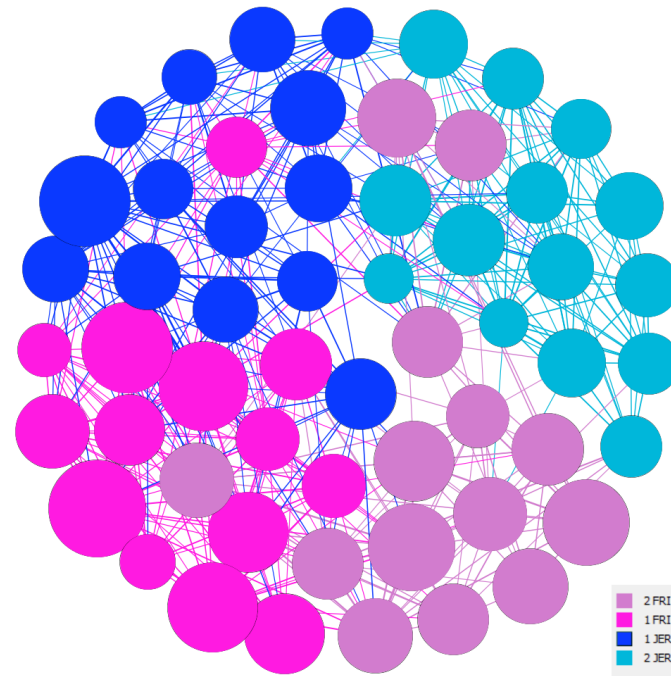


# Network analysis – Reference free

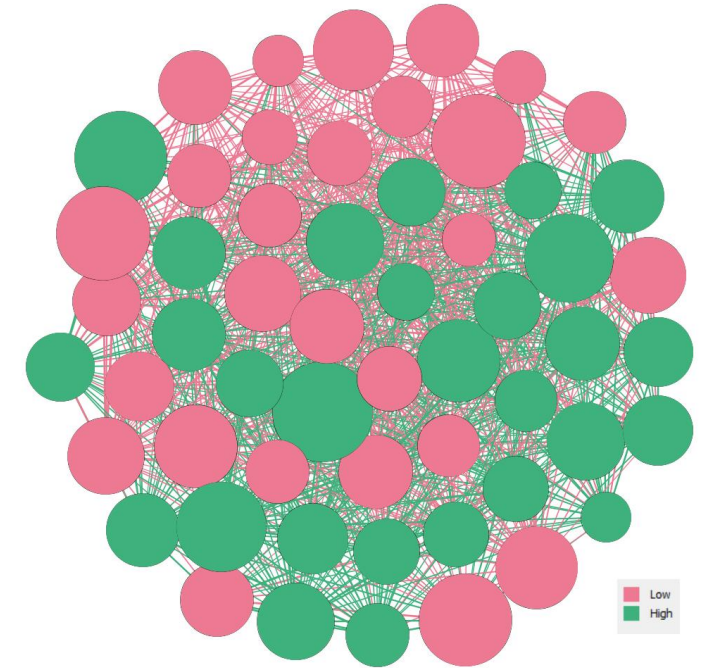
Unadjusted



Unadjusted



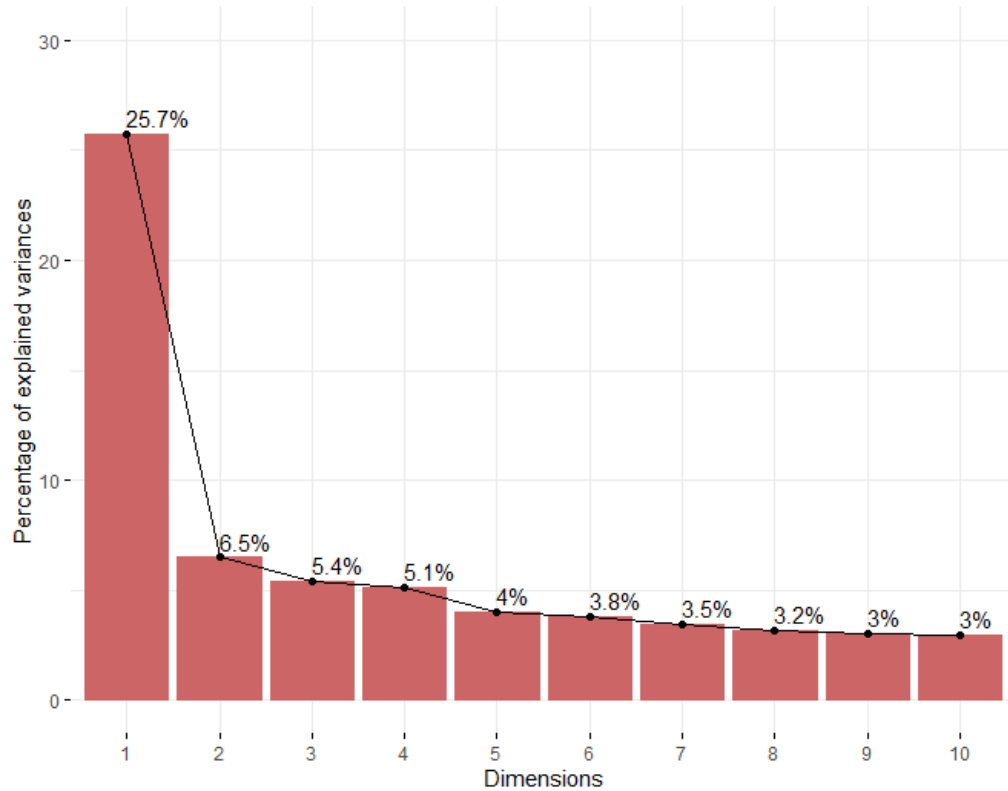
Adjusted by Cohort x Breed



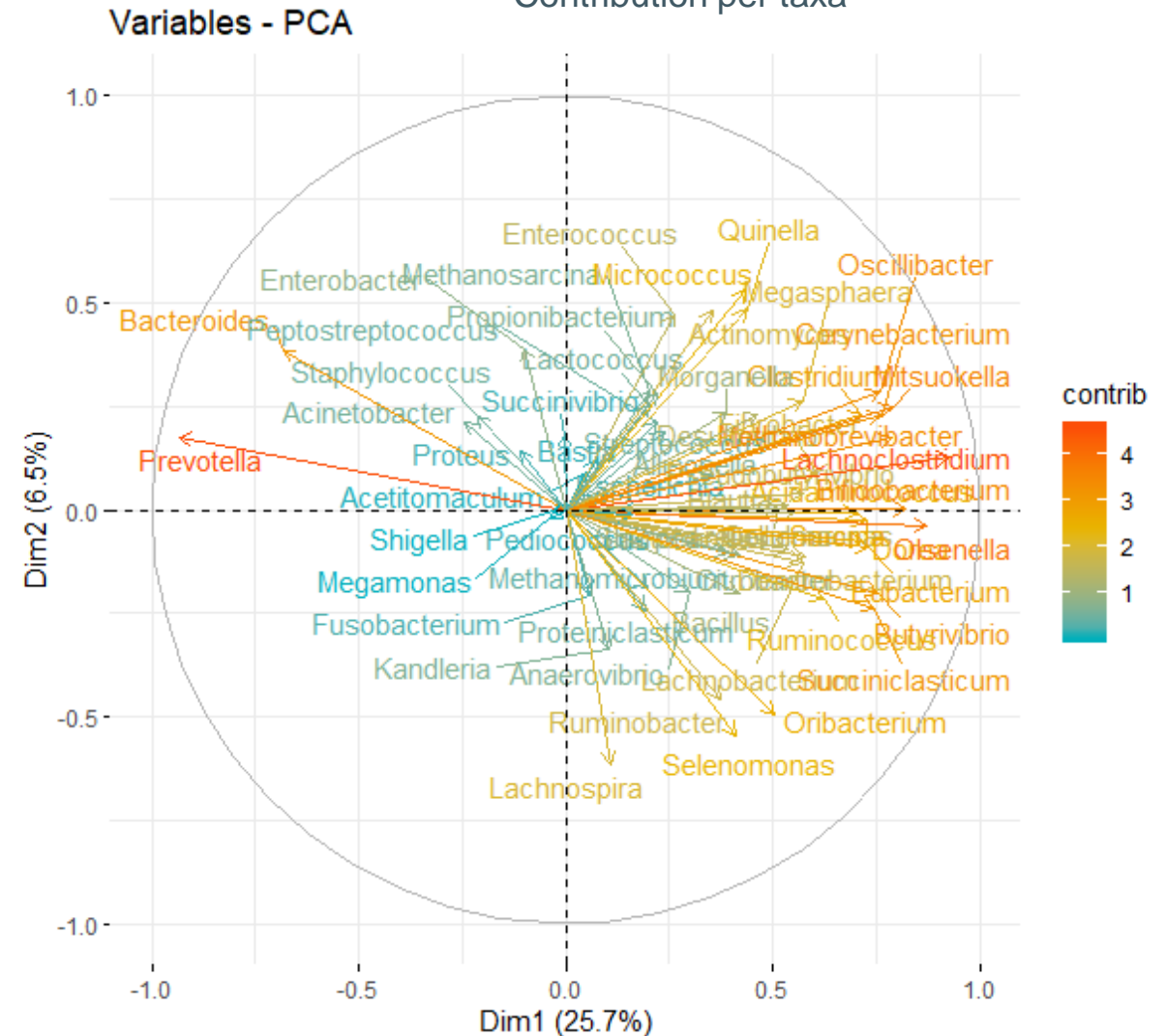
# Principal component analysis (Reference based)

Variance explained per component

Scree plot



Contribution per taxa



# Microbiability estimates (Reference based)

- ASReml 4.1 – univariate models
- Cohort x Breed effect
- Fitting PCs as a random effect



Number of PC fit	Variance explained by PC	Microbiability
1	26%	0.0003 (0.0020)
6	51%	0.0014 (0.0114)
15	77%	0.0685 (0.0935)
30	95%	0.3624 (0.1910)



# Enteric Fermentation Flagship Project – Next steps

- Make new sequencing methods available worldwide to partners
- Protocols and agreement to share samples and phenotypes
- 1,000 samples fully funded by GRA
- Allowance for the costs of genome sequencing for developing countries
- Greater understanding of the rumen microbiome and biological differences between high and low emitters
- Low cost methane predictor to compare and evaluate systems, feeds, individuals
- Potential for low cost global breeding solutions

**Developing microbial and genomic predictors for differences in methane emissions across the globe**

