Australian Government
Department of Agriculture

CSIRO

mla
MEAT & LIVESTOCK AUSTRALIA

final report

Project code: B.CCH.6620

Prepared by: Stuart Denman

Commonwealth Scientific and Industrial Research Organisation Agriculture Flagship

Date published: August 2015

ISBN: 9781741919530

PUBLISHED BY
Meat & Livestock Australia Limited
Locked Bag 991
NORTH SYDNEY NSW 2059

Comparative analyses of rumen microbiomes to mitigate ruminant methane and improve feed utilisation

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Acknowledgements

This work was funded by grants from Meat and Livestock Australia and the Australian government through the Department of Agriculture.

Executive summary

The rumen microbiota of cattle and sheep are responsible for the bioconversion of feedstuffs through hydrolysis; then pathways involving both hydrogen-producing and hydrogen-consuming steps into various fermentation end products such as short chain fatty acids (SCFA). Methanogenic archaea (methanogens) consume metabolic [H] and produce methane, which accounts for a 2-12% loss of the metabolic energy from feed. Production gains to the livestock industry will be achieved through redirection of this metabolic energy into energy yielding products with a concurrent reduction in methane emissions.

Findings from this research have identified the variation in rumen microbial populations which helps explain the difference between low and high residual feed intake (RFI) as it pertains to lower methane production. The microbiota of low RFI animals were found to harbor similar dominant populations to those that were found in the gut of the low methane producing Tammar Wallabies. These populations are responsible for efficient turnover of soluble sugars without the release of H_2 into the gut environment.

Monitoring of gene transcripts in low methane fermenters found an acetogen species responsible for redirection of H_2 from methanogenesis to acetate production. Strategies that promote these species within cattle are likely to result in lower methane emissions and will revolve around increased starch and soluble sugar content in diets.

Table of Contents

1. Background

Recent projects have provided the first genomic information of the structure and function of rumen microbial communities in “real world” samples. The objective of this project is to further develop and use (meta)genomics approaches to understand the greater rumen microbiota populations in livestock, using the datasets produced in Australia and abroad.

Australia’s livestock industries are being confronted with public concerns and challenges to change with respect to resource use, environmental impact, and public health. The biology underpinning many of these challenges is microbiological in cause-and-effect, and perhaps the most pressing challenge presented to our livestock industries by society right now is to reduce methane emissions.

Some of the transformational advances required by our livestock industries will need to be microbiological in nature, and this project seeks to improve “carbon capture” by the animal itself, first by reductions in methane emissions per kilogram of meat produced and second by extracting more useable nutrients from recalcitrant feedstuffs. This should help protect against increasing costs of production, whether in the form of greenhouse gas taxes or via emissions trading, as well as the transport and use of expensive feed supplements. Accordingly, the research is directly relevant to Australia’s National Research Priority: “An environmentally sustainable Australia” via the Priority Goal: “Transforming Existing Industries”, while the key feature of the Emissions Reduction Fund (ERF) is to encourage land managers to earn carbon credits by reducing greenhouse gas emissions. Red meat production still remains one of the principal targets for further greenhouse gas reductions despite data that show that the red meat industries have reduced gross methane emissions, and emissions per tonne of red meat production, since 1990 (see http://www.mla.com.au). In other words, the “clean and green” image for our livestock industries continues to be eroded away, and efforts need to be made that successfully recover that image, by making a step-change reduction in greenhouse gas emissions from livestock production systems.

The intrinsic strength of this study design lies in the statistical interrogation of data relating to species composition, functional metabolic diversity and gene expression. It is possible with these methods, to determine how the differences in the rumen microbiota contribute to an animal’s “phenotype” with respect to methane production, and identify those characteristics unique to Australian livestock. The ability to identify these key microbes and the functional role that contributes to the low methane phenotype should thereby enable development of novel approaches for reducing methane emissions from ruminant sources. This is likely to be based on an integrated analysis of the host animal, its rumen microbiome and options to manipulate its components along with environmental and management factors associated with our livestock production systems.

The main objective of this project was to use the latest metagenomic and related computational and statistical approaches, some developed by the CSIRO group, to understand the critical control points of developing a rumen microbiota with a “low methane” phenotype in Australian livestock production systems. A key feature of this project is the international partnerships enabling our access to a much greater breadth of data to achieve our project goals, rather than initiate an independent study of our own. In this way, it was possible to achieve the objectives of the proposed project while at the same time providing additional insights into other key studies in Australia and abroad. By doing so, we have helped refine and strengthen candidate methane abatement strategies, whether they be feeding trials, or laboratory based studies seeking to identify strategies redirect rumen fermentation away from the formation of methane, using model consortia.

The outcomes of this project include the knowledge required to develop practical ways that will result in a “low methane phenotype” in animals; either by animal selection for the desired populations and (or) augmenting the metabolic capacity of the greater microbial community (as a complement to methanogen inhibition).

2. Methodology

Total DNA extraction:

DNA extractions were carried out on rumen samples collected from animals. Mixed rumen fermentations and pure microbial cultures using a modified version of the RBB+C method (Yu & Morrison, 2004). A 1ml sample of material is transferred to a 2ml screw-capped Eppendorf tube containing 200 mg of silica/zirconia beads (0.1 mm/1 mm) and centrifuges at 15k g for 5 minutes. The supernatant was removed and the pellet containing microbial cell material was resuspended in RBB+C lysis buffer (NaCl 500 mM, tris-HCl pH 8.0 50 mM, EDTA 50mM, SDS 4%). Samples were homogenised for 2 X 90 seconds in a fast prep machine (MP Biomedicals) at 6.5 m/s. Samples were incubated at 70 ^{o}C for 15 minutes, then centrifuged at 15k g for 5 minutes. The supernatant wash transferred to a new tube and 200 µl of QIAGEN AL buffer was added to the supernatant, followed by an equal volume of ethanol (100%). The resulting mixture was transferred to a QIAamp DNA extraction kit and purified following the manufactures instruction.

Total RNA extraction for RNAseq:

Total RNA was extracted from a 1ml sample by first adding 200 µl of phenol/ethanol (95:5 vol:vol) to preserve RNA transcripts. The samples were centrifuged for 2 min at 14K in a 2ml screw cap eppendorf tube and the supernatant removed. 200 mg of glass beads and 800 µl of cell lysis buffer (300µl 100mM TE, pH 6, 400µl phenol:chloroform (pH 4.3), 100 µl 10% SDS) are add to each cell pellet. The sample is placed in a bead beater for 3X 40 sec on speed 5 with the Fastprep bead beater and then centrifuged at 14K, 2 min, 25°C. Remove the supernatant to a DEPC treated tube and add 500 µl RLT (QIAGEN RNeasy kit). Mix well with pipette and add 500µl 96% ethanol. The sample is applied to an RNeasy column in 2 lots (750µl at a time) and the RNA is washed and eluted from the column following the manufacturer's instructions (QIAGEN RNeasy).Quality of extracted RNA was visualized on a bioanalyzer using the Agilent RNA 6000 Nano Kit (Agilent), before preparation of material for RNA sequencing.

RNAseq data was generated at MACROGEN using bar-coded RNAseq library prep kits. Firstly ribosomal RNA was depleted by using the Ribozero kit (Epricentre) with 'universal' rRNA hybridization probes before individual samples were barcoded and 250 bp paired end libraries prepared. Libraries were pooled and then loaded onto two lanes of an Illumina Hiseq2000. RNAseq data was extracted from the image data in fastq format after quality filtering.

Meta-genomic sequencing:

A metagenomic assessment of the rumen microbiome from seven animals using Hiseq Illumina was undertaken at DOE-JGI. For each sample a 250bp insert paired end shotgun library was created, each were individually bar-coded and pooled prior to loading onto two lanes of a Hiseq Illumina machine. Data was then passed through our processing pipeline for gene calling and annotation of reads. The sequence data was first filtered based on quality settings and passed through three assembly algorithms (Velvet, Meta velvet and idba-ud) using an assortment of kmer settings to find the optimum assembly parameters.

Assembly, gene calling and annotation:

The Illumina datasets were subjected to quality clipping for the removal of adapters and low quality regions using Trimmomatic (Bolger *et al.*, 2014), by specifying all Illumina adaptor sequences (allowing 2 seed mismatches, using a 40bp palindrome clip threshold, and a 15 simple clip threshold), as well as non-default parameters of LEADING=3, TRAILING=3, SLIDINGWINDOW of width 10 and minimum quality 20. Furthermore, reads shorter than 50bp were removed from the dataset.

Metagenomic assemblies were constructed on each dataset individually, using three different methods velvet, meta-velvet and idba-ud to identify the optimal assembly method and parameters.

Open-reading frames were called using FragGeneScan (Rho *et al.*, 2010), which identifies full and partial genes within the assembly. Called open-reading frames were scanned for PFAM-A domains using

HMMER3 (with a maximum e-value of 1). These results were then filtered for matches to PFAM-A domains commonly found in proteins that participate in metal uptake sytems (eg. Siderophores).

The number of domains across each sample was tabulated, and a principle components analysis conducted to identify which domains explained the largest proportion of variance across samples.

Analysis of microbial diversity from metagenomic data: Microbial composition analysis of the generated short read sequences were performed using the clade-specifc marker gene database incorporated into MetaPhlAn (Segata *et al.*, 2012). MetaPhlAn calculates the relative abundance of taxa by mapping reads against a set of clade-specific marker sequences that can identify specific microbial clades at the species level and cover all of the main functional categories. The results from each animal were merged into a single results file that could be used to generate hierarchical clustering and heat map visualisation.

Classification and binning of assembled contigs

16S rRNA classification

Scaffolds from all best idba-ud assemblies were scanned using the algorithm of Huang and colleagues (Huang *et al.*, 2009) to identify 16S rRNA fragments. These fragments were then processed using the QIIME package. A reference-based clustering was performed using the Greengenes database version Feb4.2011 (DeSantis *et al.*, 2006) with 97% identity cut-off, 16S rRNA fragments that did not match any reference were clustered to form de novo 97% identity OTU clusters. The OTU labels for de novo (unclassified) OTUs are given the 'None' prefix by QIIME. The scaffolds from which the 16S rRNA fragments originated were assigned the OTU classification of the 16S rRNA fragment contained therein. 3,295 scaffolds were assigned using this approach.

All read-pairs from each sample were mapped using bwa (mem algorithm) against the 16S rRNA-containing scaffolds, and retained if one read in the pair matched (95% identity for 50% of the read length, primary alignment) but the other did not. These unmatched mates were then mapped, again using bwa, against the scaffolds from all assemblies (filtered using the same thresholds as above), to identify potential adjacent scaffolds. A relationship between a 16S rRNA scaffold and a potentially associated scaffold was retained if four or more read pairs mapped uniquely between them. Scaffolds that mapped uniquely to a single 16S rRNA fragment were assigned to the taxonomic classification and OTU assigned to the fragment. This classified 4,645 scaffolds.

For scaffolds that mapped to multiple 16S rRNA sequences (192 scaffolds), the scaffold was assigned to a unique OTU id (with the prefix 'consensus'), and classified to the finest rank for which all of its associated 16S rRNA sequences agreed. Only 19 of these 192 could be classified at best to Phylum.

Tetramer windows

Non-overlapping tetramers (with one tally per reverse complement pair) were counted over 2Kb windows within the scaffolds. Scaffolds of length less than 4Kb were omitted from classification.

Read Coverage

A random subset of 2.5 million reads was extracted from each sample dataset, and mapped using bowtie2 (--very-sensitive setting) against the scaffolds of the best idba-ud assembly for all the samples. The mean depth of coverage of each scaffold was calculated using GATK-2.6-4. The coverage of individual 2Kb windows was not used as it would likely be influenced by PCR or repeated elements.

PCA

A PCA (using prcomp in R – no centring or scaling) was conducted on each of the coverage and tetramer matricies separately, as they were on different scales.

The first 5 principal components for each PCA were extracted, explaining 90.56% and 98.17% of the tetramer count and coverage variance respectively.

These ten principal components formed the features for each scaffold, to be used for SVM classification.

Training data for supervised vector machine

The above 16S rRNA-classified data was used as training data for the SVM. The SVM was fit using the kernlab package in R, with the de novo OTU cluster as the classification class, and the principal components as main effects. The type of SVM fit was 'c-svc', cost of constraints, C, set to 60, with k=10 cross-fold validation, due to the small size of the training data. (7.2% CVE). Based on the classification of all the training data using the SVM, only one 2Kb-window was misclassified out of 1,290. It was this:

OTU prediction was also conducted for the unknown scaffolds. Scaffolds were called to their majority OTU classification if 75% or more of the tetramer windows were classified to that OTU, or classified as 'unknown' otherwise. 67% of scaffolds were confidently assigned using this approach.

Analysis of microbial diversity from 16S rDNA:

Using high throughput sequencing platforms and barcode “pyrotagging”, phylogenetic based methods targeting the 16S rDNA gene were used to deeply characterise the microbial populations present in the rumen and faecal samples of cattle. 16S rDNA gene pyrotagging was performed using two methods (across the life of the project) either using modified universal bacterial primers 27f and 515r or 515F and 806R. Specific sequences matching the Roche 454 sequencing adaptor B were added to the 27f primer, while adaptor A was added to the 515r. In addition between the adaptor A sequence and the 16S 515r sequence a 10 bp barcode was inserted. Each individual DNA sample was amplified using the 27f primer and a uniquely bar-coded 515r primer. After amplification products were visualised by performing gel electrophoresis. Product quantities were calculated and an equal molar amount of each product was pooled. The pooled products were run in a 1.5 % agarose gel and the product gel extracted and purified prior to submission for 454 pyrosequencing. For 515F and 806R specific sequences matching the Illumina Miseq sequencing adaptor P5 were added to the 515f primer, while the P7 adaptor was added to the 806r. In addition a dual index primer approach was used with index primers attached to both the forward and reverse primers. Allowing for a multiplexing of 384 samples from 24 unique forward indexes and 36 reverse indexes. Each individual DNA sample was amplified using a unique index combination. After amplification products were visualised by performing gel electrophoresis. Product quantities were calculated and an equal molar amount of each product was pooled. The pooled products were run in a 1.5 % agarose gel and the product gel extracted and purified prior to submission for Illumina Miseq pyrosequencing.

Short read sequence data generated using The Miseq platform was analysed using the QIIME: Quantitative Insights Into Microbial Ecology software package (Caporaso *et al.*, 2010), for generation of operational taxanomic units (OTU) clusters, alpha and beta diversity measures and distance calculations and further analysis using R with the Ade4, phyloseq and DESeq2 packages (Chessel *et al.*, 2004, McMurdie & Holmes, 2013, Love *et al.*, 2014).

Inspection of the relationship between samples across experimental factors suggested that gender, cohort within year, and group within cohort had a significant bearing on species variation. To investigate the relationship between individual species and methane yield, while accounting for such factors as gender, cohort within year and group within cohort, a base model was fitted to the data as follows

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{(j)k}$$

where Y_{ijk} = response to animal gender i (i = M or F), from cohort within year j (j = 1 to 8, categorical) and group within cohort k (k = 1 to 4, categorical); μ = overall mean; α_i = fixed effect of gender, β_j = fixed effect of cohort within year, and $\gamma_{(j)k}$ = fixed effect of group within cohort nested within cohort within year. For each OTU, *l*, within the OTU table, the base linear model was updated to include the read count for OTU *l*,

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_{(j)k} + \tau_l$$

where Y_{ijk} = response to animal gender i (i = M or F), from cohort within year j (j = 1 to 8, categorical) and group within cohort k (k = 1 to 4, categorical); μ = overall mean; α_i = fixed effect of gender, β_j = fixed effect of cohort within year, and $\gamma_{(j)k}$ = fixed effect of group within cohort nested within cohort within year. For each OTU, *l*, within the OTU table, the base linear model was updated to include the read count for OTU *l*

To appropriately handle the library size and diversity differences between samples, the DESeq2 modelling package was used. The DESeq2 approach aims to identify OTUs that are differentially abundant across two groups, based on count data derived from next-generation sequencing. As the count-data is discrete and over dispersed relative to the Poisson model, DESeq2 models the counts using the negative binomial distribution. The mean is taken as a quantity, proportional to the concentration of sequences present in the sample, scaled by a normalisation factor which adjusts for library size differences across samples. The dispersion is estimated for each OTU, but information is pooled from OTUs with similar abundance levels to allow for accurate dispersion estimates when the number of samples is low.

The samples were divided into groups based on their methane yield, with yields less than 30 assigned to 'low', yields greater or equal to 30 but less than 35 assigned to 'medium', and greater than 35 assigned to 'high'. The negative binomial was fit in DESeq2, including an additive factor for cohort within year only.

Microbial culturing methods:

Anaerobic culturing media was prepared in culture tubes and batch fermentation vessels for the isolation and purification of bacterial species using the anaerobic techniques of Hungate (Hungate, 1969) as modified by Bryant (Bryant, 1972). Basal media M8C was used for most strains with the addition of ground lucerne hay for purification of fibrolytic strains. Media for the construction of synthetic microbiomes was based on M8C with addition of lucerne hay and soluble sugars (Table 1). Media was comprised per litre of 150 ml mineral solution 4 (3g/l K_2HPO_4) and 150 ml mineral solution 5 (3g/l KH_2PO_4, 6g/l $(NH_4)_2SO_4$, 6g/l NaCl, 0.6g/l $MgSO_4$, 0.6g/l $CaCl_2$) 300ml of clarified rumen fluid, peptone, 10g; yeast extract, 2.5g; cellobiose, 5g; $NaHCO_3$, 6g; Haemin solution(0.5% w/vol), 1ml; and resazurin (0.1%w/vol), 1ml. The media was sparged with CO_2 for 1 hour, and dispensed into appropriate sized anaerobic culturing vessels (Balch tubes, serum bottles, Schott bottles with modified seals) and autoclaved at 120^0C for 15mins. Prior to inoculation, 1ml of vitamin mix solution (biotin, 200mg; folic acid, 200mg; pyridoxine monohydrochloride, 1g; riboflavin, 500mg; thiamine, 500mg; nicotinic acid, 500mg; pantothenic acid, 500mg; 4-aminobenzoic acid per litre) was added to each 100ml of media. Initial fermentations focused on select species that are critical for rumen fermentation and have been observed at significant numbers in both low and high methane animals (table 1). The three main fibre degrading species and *Butyvibrio fibrosolvens* along with a representative of the most dominant genus, *Prevotella ruminocola* and with the addition of a potent methanogen *Methaobrevibacter smithii* were assigned as the primary cultures. Low methane cultures included these species and the addition of two acetogenic bacteria (*Blautia Hydrogenotrophica* and *Lachnospiraceae sp.*A4)

For the detection of SCFA, a 1ml-sample of culture fluid media is used for analysis. Samples were centrifuged (12 000g, 10 min), 600 µl of sample was combined with 150 µl of meta-phosphoric/internal standard solution (20% meta-phosphoric acid/0.24% 4-methyl valeric acid). Organic acids analysis was performed using a Shimadzu GC-17A gas chromatograph on a packed glass column (2 m length; 0.6 mm OD, 0.2 mm ID) containing 10% FFAP/1% H_3PO_4 on Chromasorb WAW 100/120 mesh. The C2–C5 acids were separated over 16 min using nitrogen as a carrier at 12 ml/min. Peaks were detected by a flame ionization detector (Playne, 1985). Lactate concentrations were determined from the supernatants of the centrifuged samples with the D-lactate and L-lactate assay kit (K-DLATE, Megazyme International Ireland Ltd., Wicklow, Ireland) according to the manufacturer's instructions.

Gas samples collected in gas tight syringes and CO_2, H_2 and CH_4 are analysed by gas chromatography using a Shimadzu GC-17A gas chromatograph with a thermal conductivity detector (GC-TCD).

Digestibility assays were performed by accurately weighing plant material into culture bottles prior to inoculation. After culturing plant biomass was harvested by centrifugation at 10,000xg for 15 mins and the supernatant removed. The centrifugation step was repeated to maximize the removal of liquid. The pellet

was transferred to pre-weighted 50ml falcon tubes and dried to completion at 80^0C. Weights of plant digestion were calculated by subtracting from the original plant dry weight. Statistical analysis of the data was performed by ANOVA, with differences determined by the method of least significant differences at the 5% level ($P < 0.05$). All statistical analyses were run with Statistica 6 (StatSoft Inc., Tulsa, OK).

Data analysis

Identification of rRNA

Fastq files were separated into original samples based on the barcodes used for each library. Each sample was aligned to the reference multiFASTA using Burrows-Wheeler Aligner bwa ('mem' algorithm) using the multithreaded option (-t 6). The resulting alignment files were sorted by read identifier using samtools, and the alignment results were tabled using featureCounts, focusing on only rRNA genomic features (parameters -t rRNA -p -g ID).

Assignment to Kegg

In order to map all reference genomes to Kyoto Encyclopaedia of Genes and Genome (KEGG) pathways, the KEGG orthology database (2011 version) was downloaded, and the genes.pep file was filtered for protein sequences that had at least one kegg pathway assignment and prepared as a blast database. RNAseq transcript data was extracted from the multiFASTA genome, translated using the bacterial codon table, and BLASTed (Protein-Protein BLAST 2.2.25+) against the reference database ('blastp -num_alignments 5 -evalue 1e-5 ').

The blast results were filtered to select the first hit that had sequence identity exceeding the minimum identity threshold. The threshold was determined by $p^S(n) = n + 420.L^{-0.335.(1+e-L/2000)}$, with L being the alignment length, n being the distance in percentage points from the curve on minimum identity – we let n = 0 (Rost, 1999).

Differential expression analysis

The binary alignment map (bam) and GFF files for the mapped transcripts to the reference database were loaded into R (version 3.1.0 (2014-04-10)) using the rtracklayer and GenomicRanges libraries. Read counts for each sample were generated using the function summarizeOverlaps (parameters singleStrand=FALSE, ignoreStrand=TRUE). The resulting SummarizedExperiment was then analysed using the DESeq2 package (DESeq2_1.4.4).

Technical replicates were pooled using the collapseReplicates function in DeSeq2, as recommended by the authors. Rather than fitting the negative binomial model with two explanatory variables and their interaction (community + time_point + community:timepoint), we instead made a dummy variable consisting of four levels with each combination of community and timepoint (phenotype =Low or High, time = 24 or 30).

Initial analysis considered all species simultaneously. The size factors were estimated using the option 'shorth', as there were large differences in usable library size due to varying rRNA depletion efficacy. The gene-wise dispersions were estimated with a maximum of 1000 iterations. The negative binomial wald test was used to identify significantly different gene expression between the phenotypes and/or timepoints.

This analysis was improved by subsetting the count data and analysing each species separately. This allowed the size factor estimation to account for both library and species abundance differences simultaneously. Kegg pathway maps were generated using the R package pathview.

Gene alignment analysis and qPCR primer design

For each gene of interest a separate database was generated including all gene fragments identified from the metagenomic sequencing of low and high animals as well as the gene of interest from representative microbial genomes. Alignments were performed using clustal X and importing into the ARB software package (Ludwig *et al.*, 2004). Regions within the alignments that were homologous across species were

identified as sequence motifs by using the MEME software toolkit (Bailey *et al.*, 2009). Quantitative PCR Primers were designed to target these motifs for specific functional genes. Primers were then compared with sequences available at the National Centre for Biotechnology Information (NCBI) via a Basic Local Alignment Search Tool (BLAST) search to ascertain primer specificity (Altschul *et al.*, 1990).

3. Results

***Microbial ecology of low and high methane animals*:**

The initial stage of the project focused on the interrogation of the microbiome of animals identified by researchers as part of the Beef CRC as producing "Higher" than or "Lower" than predicted methane outputs. Five animals were classified as producing higher than predicted methane levels while six animals fell into the lower category. In addition, it was possible to use DNA samples from animals whose measured methane output was close to that predicted from their feed intake as "controls". Replicate rumen samples across time were used for microbial profiling of the rumen microbiome. Microbial populations within the rumen separated based on these groupings with the largest variance observed for separation of the low and high methane from the control animals, while the second level of variance explained the separation between the high and low animals (Figure 1).

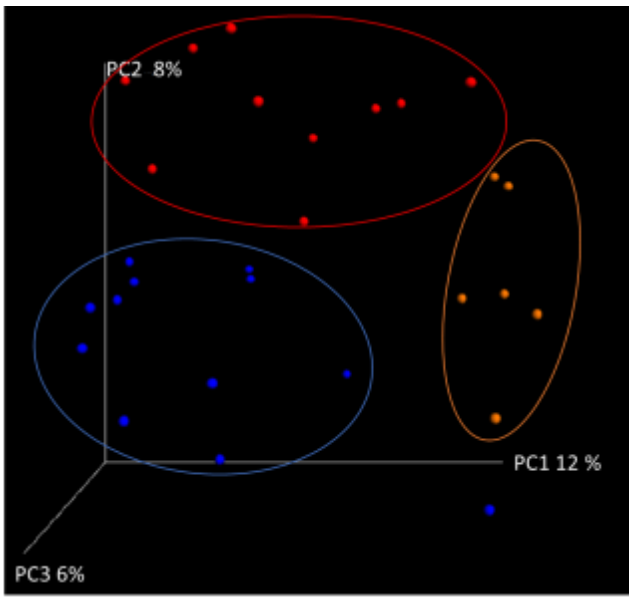


Figure 1. Principal coordinate of analysis of rumen microbiomes for individual animals with top three variances explained. Control animals (orange), "low" (blue) and "high" (red).

Overall changes to the microbial populations at taxonomically levels did reveal subtle changes between the low and high methane animals with most notably a decrease in Bacteroidetes and a concurrent increase in Firmicutes, Proteobacteria and Cyanobacteria phyla (Figure 2 and Appendix 1). Further classification of the Taxonomic assignment indicated that the major group within the Bacteroidetes phylum to decrease was those that belonged to the order S24-7 (21% and 10% for high and low respectively), thought to be a gut butyrate producer. However, increases in the Bacteroidetes phylum associated with the Porphyromondaceae family were observed as were increase to specific *Prevotella* species. While the increase in Firmicutes was assigned to bacteria that could only be accurately assigned to the Lachnospiraceae family. These OTUs were the most abundant group in the low animals and were approximately 5 times higher compared to the high animals. Another Lachnospiraceae member was more deeply classified to the genus Roseburia and was 30 times more prevalent in low methane animals. Species assigned to the cellulolytic genus Ruminococcus were 5 times lower in the low animals while Fibrobacter species increased 3 times. The Veillonellaceae family produced the greatest decrease in the Firmicute phylum for the low animals with species assigned to the genera Selenomonas and

Succiniclasticum reduced. However, there was an increase for Veillonellaceae family members in the Megasphera genus. Interestingly a large increase in numbers for taxa associated with the Cyanobacteria phylum close to the rumen group YS2 was observed in the low animals with a 6 times increase. Within the Proteobacteria phylum the major increase in the low animals was attributed to taxa assigned to the family level of Succinivibrionaceae. Also increases in the Synergistetes phylum for the low animals were evident over the high animals.

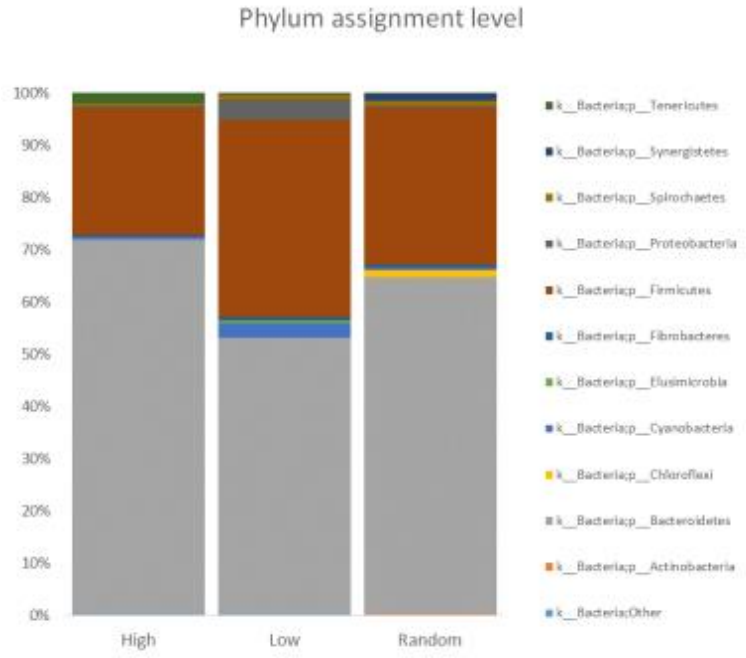


Figure 2. Phylum assignment taxonomy level analysis of rumen microbiomes for animal groups.

Archaeal populations did vary in their composition between the high and low phenotypes. Animal 212 presented as an outlier, which could be attributed to differences in the abundance of the methanogen populations and not that it harboured a divergent population of methanogen species compared to the other animals (Figure 3). Alpha diversity measures performed on each data set produced a chao1 estimate of 85.3 ± 3.9 and 83.2 ± 10.3 for the high and low animals respectively, indicating similar levels of diversity within the samples.

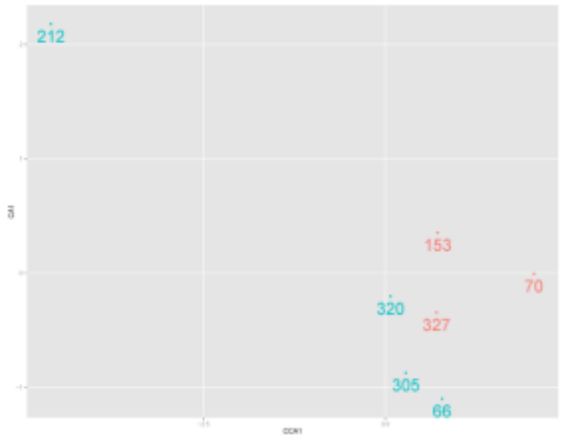


Figure 3. Canonical correspondence analysis of archaeal rumen microbiome data for low (blue) and high (red) animals (numbers) using primer set for V2-V4.

Taxonomic assignment of the OTUs showed that all animals were dominated by *Methanobrevibacter* sp., contributing in excess of 97% of the data (Figure 4). Analysis of variance (ANOVA) of archaeal species did not detect any species that were correlated with a methane phenotype. With no significant change to the methanogen populations it is likely that the low methane phenotype is from either reduced total numbers of methanogens or bacterially driven and could suggest that less efficient methane production in the "low" animals is due to the changes in bacterial populations that produce/consume hydrogen. Consumption of methylated compounds that can be utilised by methanogen species from the *Methanosphaera* and *Methanoplasmatales* genus would seem to contribute very little to methane production based on the observed numbers for these species and more likely reflects the diet composition.

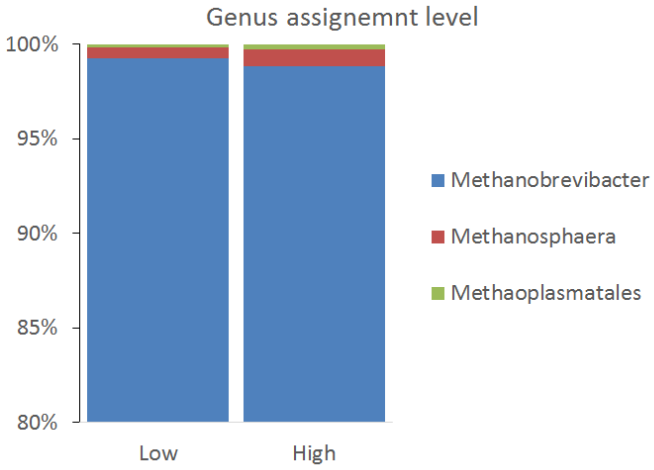


Figure 4. Taxonomic distribution analysis at the genus level of archaeal rumen microbiome data for low and high methane phenotypes. Percentage of total sequence data assigned to each genus is plotted on the Y axis.

Access to samples from a larger dataset of rumen and faecal samples collected from Black Angus cattle that were included in the genetics project (technologies to reduce methane emissions from Australian beef cattle, Robert Herd NSW DPI: 01200.044; B.CCH.6310) were subjected to microbial ecology studies to ascertain microbiome shifts related to methane emissions. DNA was successfully extracted from 320 paired rumen and faecal samples and used as a template for the amplification of the 16S V4 (variable region 4) using dual indexed Illumina primer sets. Utilizing the Illumina Miseq platform, just over 33.5 million paired sequences were obtained for the rumen and faecal samples. Quality filtering of sequences resulted in an average of 35,700 sequences per animal for the faecal samples and 45,900 for the rumen samples. Clustering of the sequences into operational taxonomic units (OTUs) based on sequence identity of 97% to define the equivalent of a species level resulted in a mean of 3,816 OTUs for faecal samples and 6,628 OTUs for the rumen. The rumen and faecal microbiomes were significantly divergent and were therefore analysed separately. Separation of rumen or faecal microbiome samples could not be strongly correlated with methane emissions (Figure 5A&C), with the main variance being contributed by the animal cohorts within the year. Similar patterns were also observed for taxonomic summary of the data (Appendix 2). Due to the nature of the animal trial with respect to gender, age, weight, diet, location and time of year a linear model was devised to incorporate these factors, most of which were captured by the animal cohort within year value (eq 1).

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_{(j)k} + \tau_l \qquad (1)$$

Where Y_{ijk} = response to animal gender i (i = M or F), from cohort within year j (j= 1 to 8) and group within cohort k (k = 1 to 4); μ = overall mean; α_i = fixed effect of gender, β_j = fixed effect of cohort within year, and $\gamma_{(j)k}$ = fixed effect of group within cohort nested within cohort within year and read count for OTU *l*.

This identified 27 significant OTUs in the rumen fluid dataset (Figure 6 shows 3 examples and Appendix 3 all 27) while no significant OTUs were found in the faecal sample data. The OTUs associated with low methane emission in the rumen were dominated by *Prevotella* species or OTUs of the Bacteroidales order, while the high emission animals were dominated by OTUs from the Clostridiales order including *Ruminococcus* species.

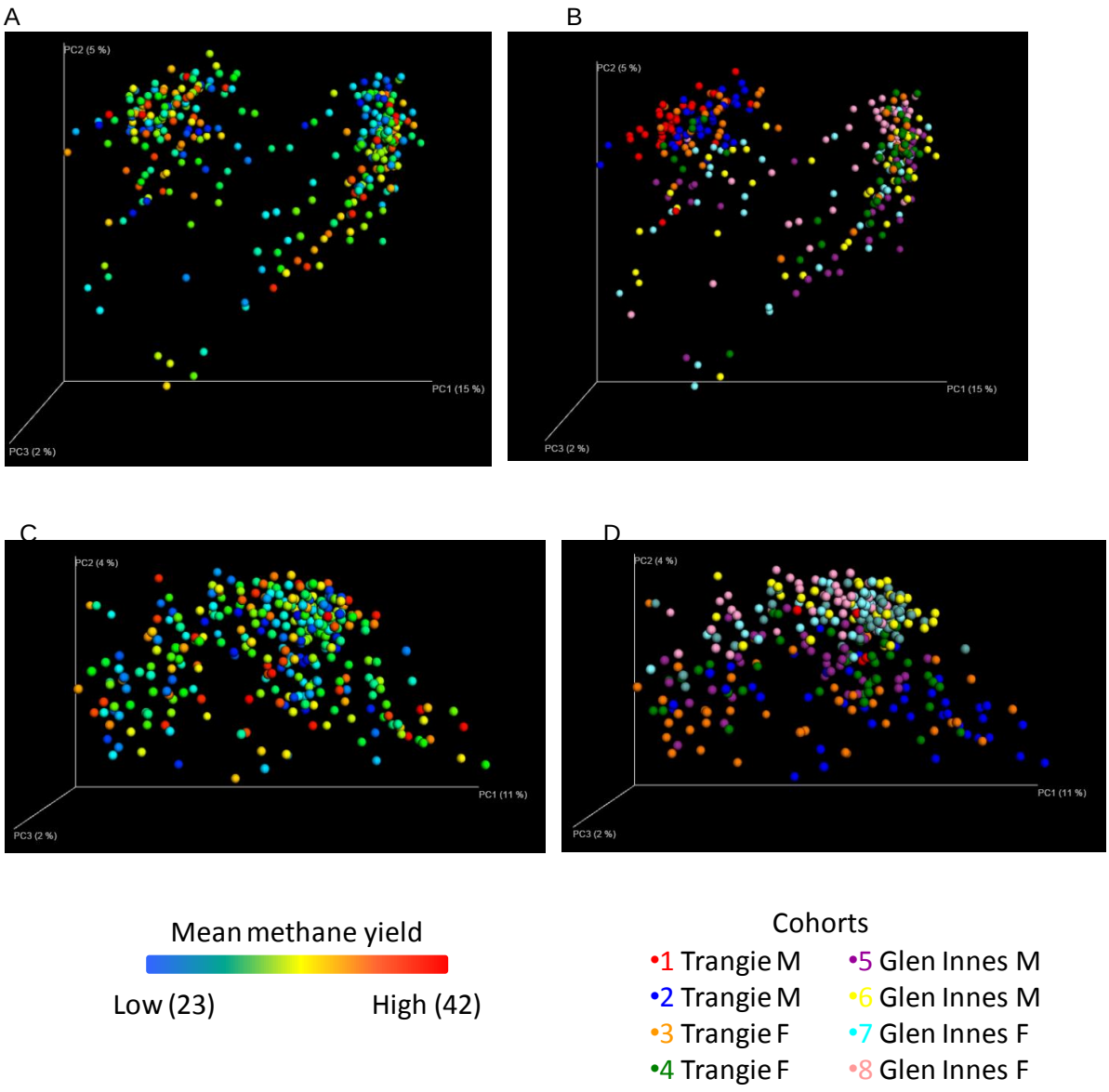


Figure 5. Principal coordinate of analysis of rumen (A&B) and faecal (C&D) microbiomes for individual animals with top three variances explained. **(**A&C) Colouring of samples related to measured methane emissions from low to high (blue to red gradient). (B&D) Colouring of samples related to cohort. Animal cohorts were comprised of groups of animals by sex M = male, F = female and from two locations Trangie or Glen Innes.

The linear model approach, whilst conceptually ideal with continuous changes in methane explained by continuous changes in species abundance, does not account for the range of species diversity and sample coverage found across the samples, with between 1,147-8,322 OTUs, and 6,270 -76,215 reads (Figure 7), with diversity noticeably lower for the first two rumen cohorts. The substantial difference between sample diversity and abundance metrics likely reduces the sensitivity of the linear model.

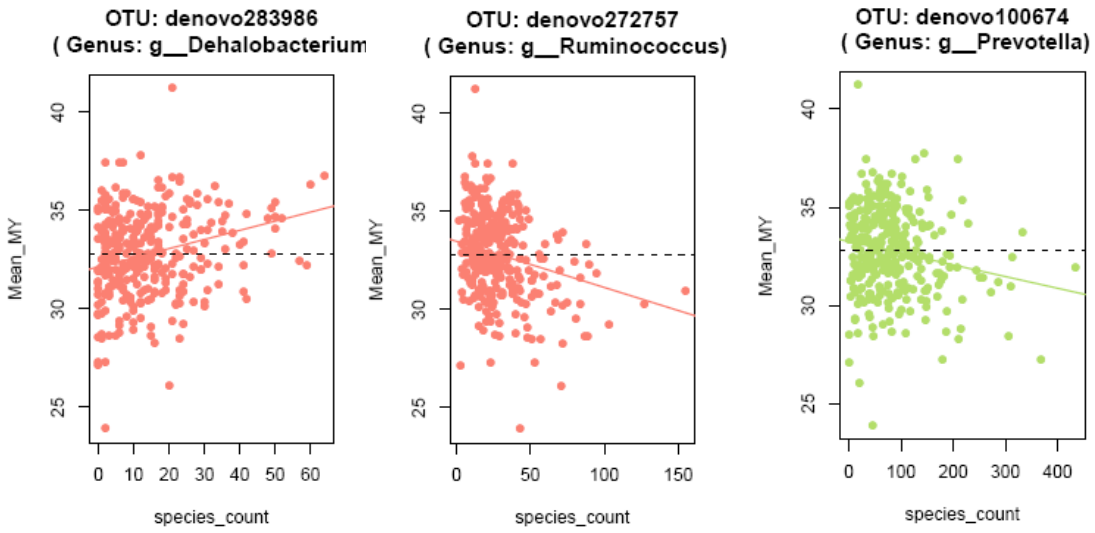


Figure 6. Examples of microbial OTUs significantly associated with low or high methane emissions. Points represent the number of times the OTU was observed in individual animals plotted (X axis) against their measured methane yield (Y axis).

To appropriately handle the library size and diversity differences, the DESeq2 modelling package was used. The DESeq2 approach aims to identify OTUs that are differentially abundant across two groups, based on observations. The mean OTU value is taken as a quantity proportional to the total sequencing effort present in the sample, scaled by a normalisation factor which adjusts for library size differences across samples. The samples were divided into groups based on their methane yield, with yields less than 30 assigned to 'low', yields greater or equal to 30 but less than 35 assigned to 'medium', and greater than 35 assigned to 'high'. The negative binomial was fitted in DESeq2, including an additive factor for cohort within year only. There were 299 differentially abundant OTUs between the high and low treatments (FDR corrected, $q < 0.05$) for rumen samples and 16 for faecal samples (Figure 8). Again the low methane OTUs in the rumen were dominated by *Prevotella* species, but also identified increases in the fibrolytic *Fibrobacter* species. Interestingly this analysis identified a methanogenic OTU from the novel methanogen cluster Methanoplasmatales as being increased in low methane animals. Very few OTU's were associated with either the low or high methane measures from the faecal microbiome and these were dominated by members of the phylum Firmicutes (Figure 8B).

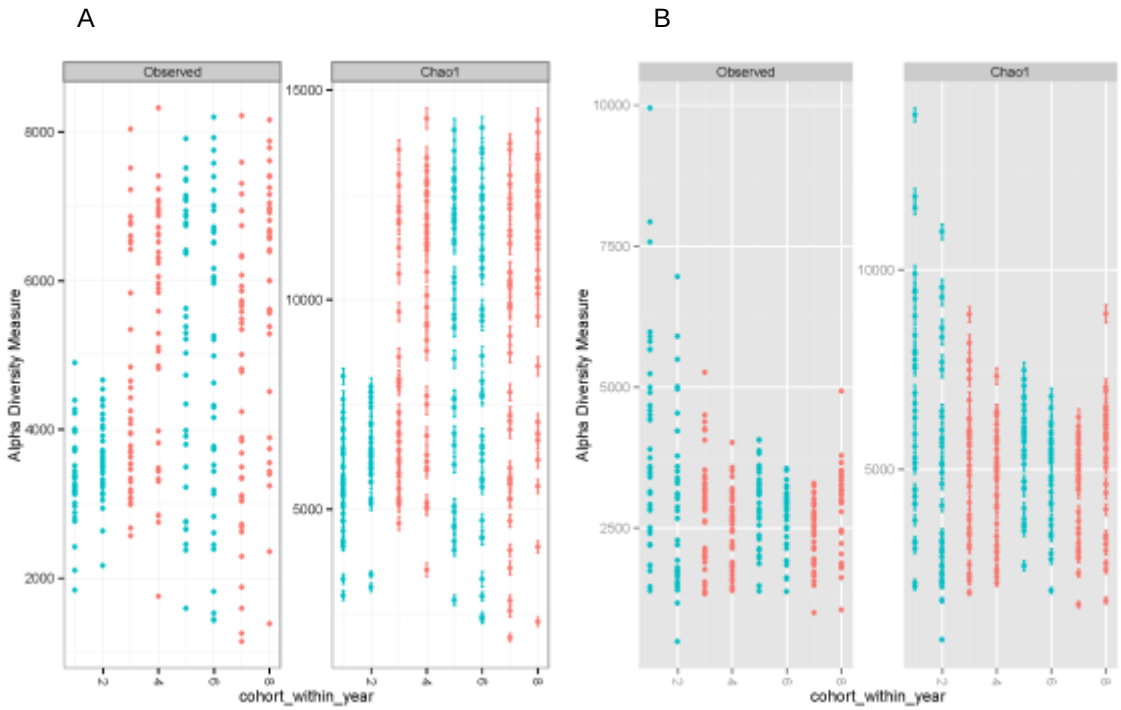


Figure 7. Alpha diversity measures for the number of observed OTUs and Chao1 diversity estimates for (A) rumen and (B) Faecal samples. Samples are separated on the x axis by the animal cohort within year value.

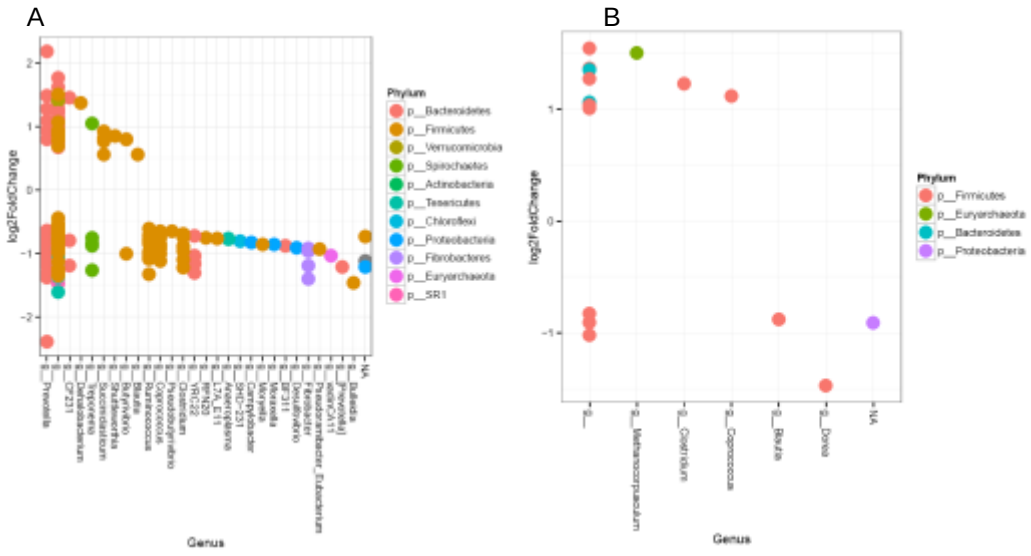


Figure 8. Microbial OTUs significantly associated with low or high methane emissions. Points represent the log2 fold change from the base line, with negative values indicating an increase in low methane animals and positive values indicating an increase in high methane animals. Points are coloured by the phylum taxonomic level and separated on the x axis for genus assignment for A) rumen and B) faecal samples.

Functional gene analysis (metagenomics) of low and high methane rumen microbiota

Total microbial DNA was isolated from the individual rumen samples of three “high” and four “low” animals and prepared for metagenome sequencing. Approximately 90 million and 108 million reads were collected from “low” and “high” methane animals respectively. Identification of sequence reads containing 16S rDNA data accounted for approximately 3% of the sequence data. Sequences could be identified with varying accuracy to the different phylogenetic ranks and provided evidence of changes to microbial abundance. Low methane animals were found to have increases in bacterial populations assigned at the family level to Fibrobacteraceae, Lachnospiraceae, Lactobacillaceae, Porphyromonadaceae, Prevotellaceae, Ruminococcaceae and Succinivibrionaceae (Appendix 4). Generally this was in agreement with the 16S amplicon data. Further phylogenetic assessment of all metagenomic sequencing data for the microbiomes from “high” and “low” methane animals established to accurately assign to a species level using MetaPhlAn (Table 1). Based on the number of genes contained within the reference database it is estimated that this equates to an average of only 3% identification to the species level for the diversity within the rumen microbiome. Although the reference database contains signature gene sets that uniquely describe 1,221 genomes, the database is underrepresented by rumen relevant genomes and therefore explains the low coverage based on this method.

Table 1. Sequence reads produced form cattle rumen samples assigned to phylogeny

Animal	Methane group	% of predicted methane	Total number of reads	MetaPhlAn Hits	% assigned	% estimated diversity covered
A066	LOW	62	44017458	523689	1.19	3.11
A070	HIGH	151	23270456	210699	0.91	2.37
A153	HIGH	151	22244890	264607	1.19	3.11
A212	LOW	84	21038372	358807	1.71	4.46
A305	LOW	49	22132462	234777	1.06	2.77
A320	LOW	75	20677480	300075	1.45	3.79
A327	HIGH	128	44789942	479403	1.07	2.80

Hierarchical clustering of the microbiomes based on the detected phylogenies for the most abundant 50 species did not clearly separate all animals based on their percentage of predicted methane (Figure 9). This clustering pattern was maintained across all taxonomic levels from phylum to species (data not shown). Two low animals (212 and 320) produced the most divergent microbiomes from the other animals, but as this is based on only a proportion of the identified phylogeny, strong conclusions should not be drawn. However, some methanogen populations were observed at varying levels in the “low” and “high” methane animals. *Methanobrevibacter ruminantium,* a rumen methanogen, contributed 1.28% in the high animals and was reduced to 0.63% in the low, *M smithii* was 1.48% in the high and 0.83% in the low. Other populations like *Methanosphaera stadtmanae* did not alter significantly between the groups with 0.19% and 0.17% for the high and low animals, respectively, while species of *Methanococcus* did alter with 0.35% for the high animals and 0.69% for the low animals.

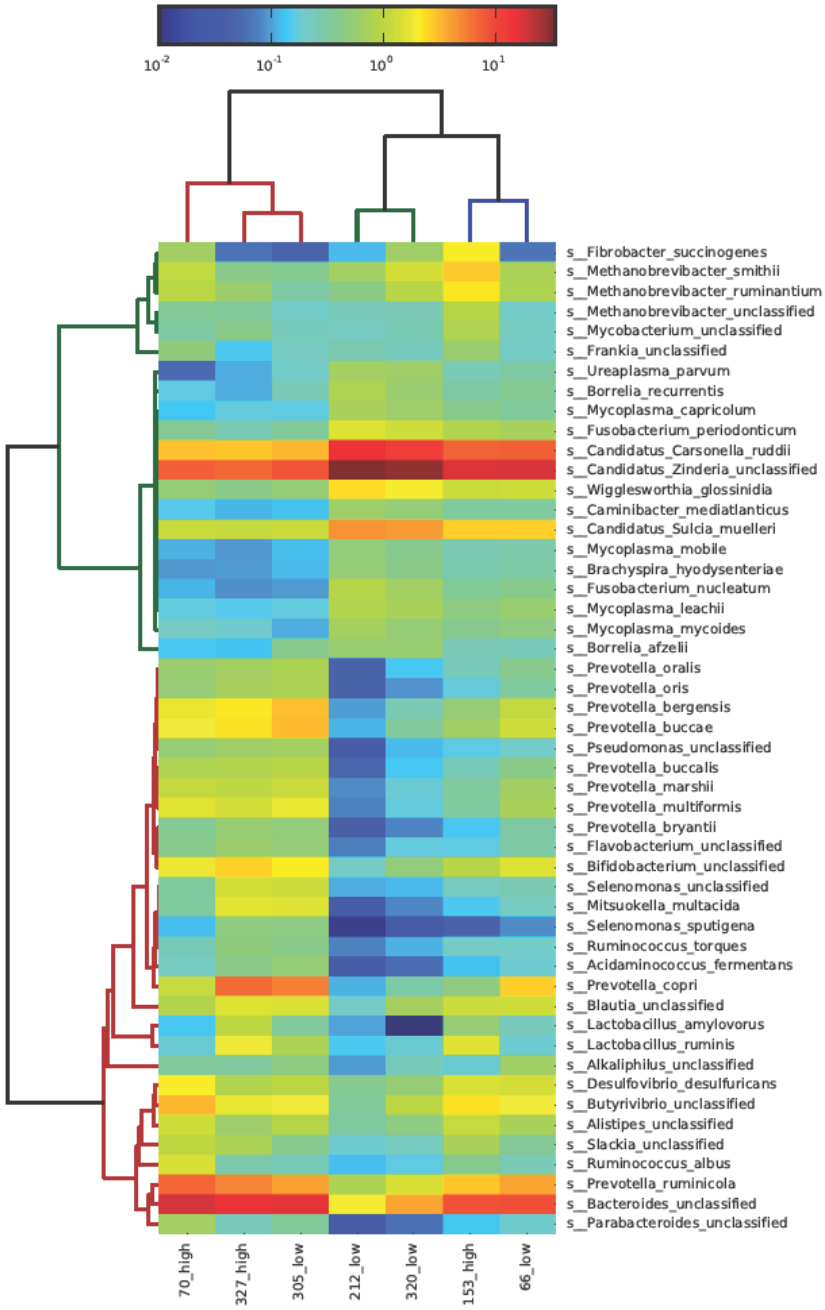


Figure 9. Hierarchal clustering of the top 50 abundant species identified from metagenomic data. Heatmap log scale indicates normalised abundance data for the observed species.

Metagenomic assemblies were constructed on datasets individually, using three different methods velvet, meta-velvet and idba-ud to identify the optimal assembly method and parameters. While the N50 metric favoured meta-velvet as the optimum assembler, this software does not attempt to assemble all the data, resulting in an artificially high N50. Instead, the algorithm that produced the largest contiguous sequence assembly (contig) was used to identify the best assembly conditions (Appendix 5). IDBA-UD consistently built the largest contigs at kmer settings of 100 and 120 while using the majority of the metagenomic reads for the assembly (Table 2).

The optimised assemblies resulted in the production of between 105,000 and 425,000 contigs for the different animals with the largest contigs ranging from 123,022 to 580,844 base pairs. The N50 values

are not particularly high and indicate a large number of small assemblies, which is typical of metagenomic assemblies from environments rich in genome diversity. The increased number of reads obtained from the deeper sequencing of samples 66-low and 327-high did allow for a larger N50 1,432 and 2,046 bp respectively compared to the lower sequenced samples. On average 55% of the sequence reads were used by the assembler, except for animal 66-low which used only 37%.

Table 2. Assembly metrics produced form cattle rumen samples using IDBA-UD

Animal	Kmer	Number of contigs	n50	Max contig (bp)	Total bases	Number of reads
66-Low	120	105,847	1,432	351,459	110,958,202	16,205,604
70-High	100	270,758	1,059	229,996	224,727,088	14,401,549
153-High	100	312,474	782	580,844	225,208,450	12,052,604
212-Low	100	425,281	706	340,757	256,622,310	10,630,343
305-Low	100	328,922	1,081	187,270	234,537,228	13,568,673
320-Low	100	416,865	720	123,022	262,948,044	9,266,709
327-High	120	199,792	2,046	208,088	219,662,849	25,853,618

Clustering of assembled read data based on sequence compositional attributes allowed for the “binning” of genome fragments that are likely to have originated from the same organism. Identification of 16S gene fragments in these bins then allowed for taxonomic labelling of the bins to various phylogenetic ranks. Approximately 95% of the data could ranked to the order level but were not always classified to a known isolate (Figure 10). The changes in the proportion of data assigned to the order levels was similar to that of the 16S amplicon data detailing decreases in methanobacteriales and increase in Fibrobacteriales in the low methane animals.

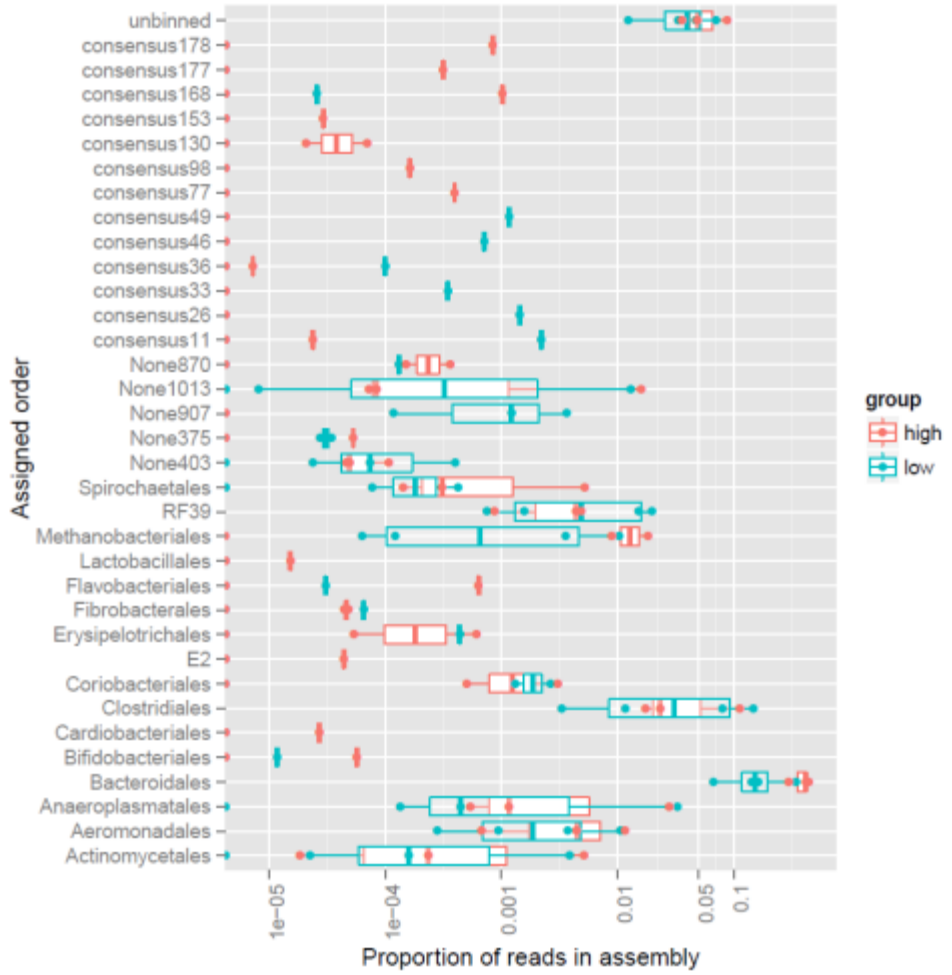


Figure 10. Proportion of assembled metagenomic reads binned to a phylogenetic order based on association to 16S gene assignment.

Open reading frames that define genes were identified using FragGeneScan and annotated against the protein family A (PfamA) database as a method for identification of the functional capacity of the microbiome. A specific focus of these bioinformatic analyses was in the identification of the genetic potential underpinning the uptake of transition state metals (e.g. cobalt, nickel) which are critical for key fermentation schemes linked with methane, acetate and propionate formation. Forty five Pfam classes were chosen for an initial screening including families critical to methanogenesis that target cobalt uptake (Appendix 6). Fifty four percent of the variance in the data allowed for the separation of the low and high methane phenotypes as the first component of a PCA analysis (Figure 11). The data variance was mainly driven by a decrease in the abundance of ABC transporters responsible for the uptake of metals into the microbial cell. In addition a decrease in the presence of tetrahydromethanopterin S-methyltransferase subunits also was responsible for separating the low methane phenotype. Tetrahydromethanopterin S-methyltransferase is part of a complex that catalyzes the formation of methyl-coenzyme M and tetrahydromethanopterin from coenzyme M and methyl-tetrahydromethanopterin. This is an energy-conserving, sodium-ion translocating step and is essential to all methanogenic pathways.

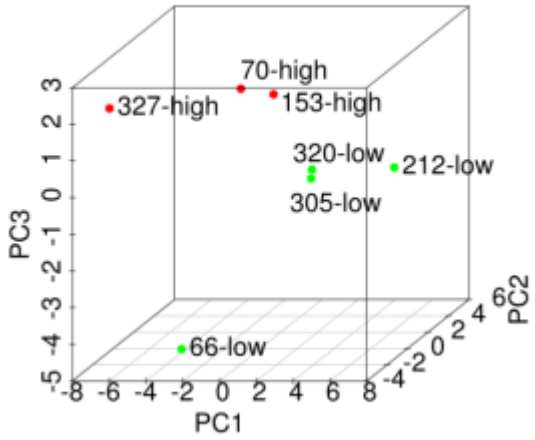


Figure 11. Principal component analysis of Pfam classes associated with transition metal binding and uptake for high methane phenotype animals (red) and low methane phenotype animals (green). The first component explains 54.4%, second component 19.4 and third component 12.9% of the variance in the data.

Simulating low and high methane phenotypes through culturing

Several combinations of bacterial species and metahnogenic species were trialled before the development of consistent consortia that exhibited either high or low methane yields. Final fermentation studies were performed using six bacterial species plus *Methanobrevibacter smithii* that resulted in a high methane yield (Table 3). The low methane yield consortium included the same species with the inclusion of an additional acetogenic bacterial species (Table 3). Initial studies were performed using these consortia in a defined anaerobic rumen media with the addition of 0.8% (w/v) of ground lucerne (1mm screened). Further experiments followed with 0.8% (w/v) of ground Rhodes grass (1mm screened).

Table 3. Microbial species included in fermentation consortium

Organism	Strain	Media	Genome	Consortium
Blautia Hydrogenotrophica	**DSM 10507**	**M8C**	**DSM 10507**	**High/Low**
Butyrivibrio fibrisolvens	**AcTF2**	**M8C**	**16**	**High/Low**
Fibrobacter succinogenes	**S85**	**M8C Lucerne**	**S85**	**High/Low**
Lachnospiraceae sp.	**A4**	**M8C**	**A4**	**Low**
Prevotella ruminicola	**23**	**M8C**	**23**	**High/Low**
Ruminococcus albus	**AR67/ Ra8**	**M8C Lucerne**	**Ra8**	**High/Low**
Ruminococcus flavefaciens	**Y1**	**M8C Lucerne**	**FD1 & 17**	**High/Low**
Methanobrevibacter smithii	**SM9**	**BRN**	**SM9**	**High/Low**

Measures of gas fermentation parameters for lucerne fermentations varied between the low and high methane systems with the low methane consortium on average across all time points producing 22% less methane. Free hydrogen levels were also significantly different ($p \leq 0.01$) between the cultures with an average decrease of 47% in the low methane cultures (Figure 12a). There was no significant difference ($p = 0.56$) measured for fibre digestion between the low and high cultures 0.241 ± 0.003, 0.251 ± 0.161 mg of fibre degraded (average of triplicate measures ± SEM). Rhodes grass fermentations also produced significant variance ($p \leq 0.01$) between the low and high methane cultures with a 23% and 38% reduction at 24 and 30 hours respectively. Hydrogen levels in the low methane cultures from Rhodes grass never exceeded the 140 mmoles detected at zero hours and were at levels that were 80% less compared to the high methane cultures that reached similar levels of hydrogen production to the Lucerne fermentations (Figure 12b).

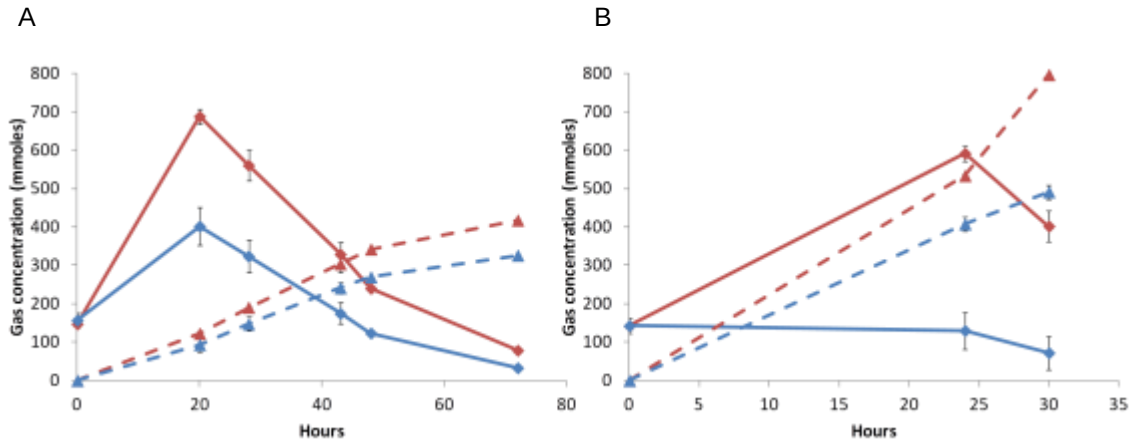


Figure 12. (A) Lucerne substrate; (B) Rhodes grass substrate. Hydrogen (solid line diamond markers) and methane (dashed line triangle marker) concentrations (mM) for microbial fermentations on Lucerne substrate for low (blue lines and markers) and high (red lines and markers) yield cultures.

Microbial fermentation end products in the form of short chain fatty acids showed significant ($p \leq 0.01$) increases in the production of acetate for the low methane cultures at all-time points when compared to the high culture for both Lucerne and Rhodes grass (Table 4 and 5). While in general all other SCFA were similar when compared at each time point, however, a significant increase between time points (20-72 hrs and 24-30 hrs) as an indication of continued microbial fermentation was evident. The acetate to propionate ratio remained higher for the low methane cultures across all time points as a result of increased acetate production. In addition levels of formate utilization by Rhodes grass cultures were 12.29 ± 1.31 mM for the high methane culture and 17.03 ± 0.10 mM for the low methane cultures at 24hrs for Rhodes grass fermentations.

Table 4. Volatile fatty acid concentrations (mM) for microbial fermentations on Lucerne substrate

Ferment time (h)	20		43		72	
Consortium	Low	High	Low	High	Low	High
Acetate	27.34 ± 1.14[a]	22.02 ± 0.46[b]	29.72 ± 0.87[c]	25.81 ± 0.7[a]	36.27 ± 0.56[d]	29.55 ± 0.36[c]
Propionate	7.16 ± 0.28[a]	7.1 ± 0.49[a]	10.25 ± 0.36[b]	10.95 ± 0.31[b,c]	11.38 ± 0.1[c]	11.43 ± 0.36[c]
Butyrate	14.78 ± 0.13[a]	15.41 ± 0.17[a,b]	15.81 ± 0.38[b]	17.4 ± 0.42[c]	18.33 ± 0.29[d]	19.04 ± 0.04[d]
iso-Butyrate	1.19 ± 0.02[a]	1.03 ± 0.06[b]	1.39 ± 0.04[c]	1.47 ± 0.05[c]	1.43 ± 0.02[c]	1.4 ± 0.01[c]
Valerate	0.14 ± 0.00[a]	0.24 ± 0.05[a,b]	0.13 ± 0.03[a]	0.37 ± 0.03[b]	0.09 ± 0.02[a]	0.31 ± 0.05[b]
iso-Valerate	1.94 ± 0.02[a]	1.88 ± 0.03[a]	2.36 ± 0.02[b]	2.43 ± 0.06[b]	2.87 ± 0.03[c]	2.89 ± 0.03[c]
A:P	3.82 ± 0.03[a]	3.13 ± 0.26[b]	2.90 ± 0.03[b,d]	2.36 ± 0.13[c]	3.18 ± 0.02[b]	2.59 ± 0.12[c,d]

Values with the same row with different superscripts are significantly different $p < 0.05$

Table 5. Volatile fatty acid concentrations (mM) for microbial fermentations on Rhodes grass substrate

Ferment time (h)	24		30	
Consortium	Low	High	Low	High
Acetate	21.52 ± 0.37[a]	12.26 ±0.28[b]	23.24 ± 0.17[c]	14.38 ± 0.95[d]
Propionate	6.00 ± 0.06[a]	5.72 ± 0.19[a]	6.48 ± 0.17[b]	6.68 ± 0.12[b]
Butyrate	12.02 ± 0.12[a]	12.01 ± 0.18[a]	12.28 ± 0.11[a,b]	12.47 ± 0.09[b]
iso-Butyrate	0.89 ± 0.01[a]	0.40 ± 0.06[b]	0.86 ± 0.04[a]	0.72 ± 0.03[c]
Valerate	0.23 ± 0.00[a]	0.23 ± 0.04[a]	0.10 ± 0.05[b]	0.12 ± 0.04[b]
iso-Valerate	1.47 ± 0.04[a]	0.67 ± 0.04[b]	1.70 ± 0.03[c]	1.10 ± 0.03[d]
A:P	3.59 ± 0.09[a]	2.14 ± 0.04[b]	3.59 ± 0.12[a]	2.15 ± 0.15[b]

Values with the same row with different superscripts are significantly different $p < 0.05$

Metatrascriptomic analysis of "low" and "high" consortia fermentations

Extraction of microbial total RNA (all RNA types) was performed on Rhodes grass fermentations for "low" and "high" methane phenotypes at 24 and 30 hrs. Yields and quality of the total RNA as ascertained by measuring on the Bioanalyzer RNA nano-chip revealed RNA concentrations and quality suitable for metatranscriptomic sequencing (Appendix 7). Sequence read data generated as paired ends of a 250 bp library resulted in just over 815 million reads (407 million paired ends) for the 24 libraries. Depletion of ribosomal RNA reads using Ribozero resulted in the near complete removal of rRNA reads, with on average only 2.8% of the data being identified as rRNA. Mapping of transcript data to complete or draft reference genomes allowed for the mapping of approximately 80% of the data (Figure 13). The level of transcripts for each species revealed a rank contribution of *Prevotella ruminicola* > *Butyrivibrio fibrisolvens* > *Blautia Hydrogenotrophica* > *Ruminococcus albus* >

Methanobrevibacter smithii > *Fibrobacter succinogenes* > *Ruminococcus flavefaciens* for the high methane cultures. Similar ranks were seen for the low cultures with the addition of the *Lachnospiraceae*

sp. TWA4, which contributed approximately 1% of the transcript data, with a reduction in the *M. smithii* transcripts by 50% to only 2% of the data. Only the fibre degrading species *Ruminococcus albus* and *Fibrobacter succinogenes* increased transcript levels between 24 and 30 hours. Differential gene expression for *B. hydrogenotrophica* and *M. smithii* showed an increase in the number of genes up regulated for the high methane cultures while all the remaining species had a greater number of genes up regulated for the low methane cultures (Appendix 8).

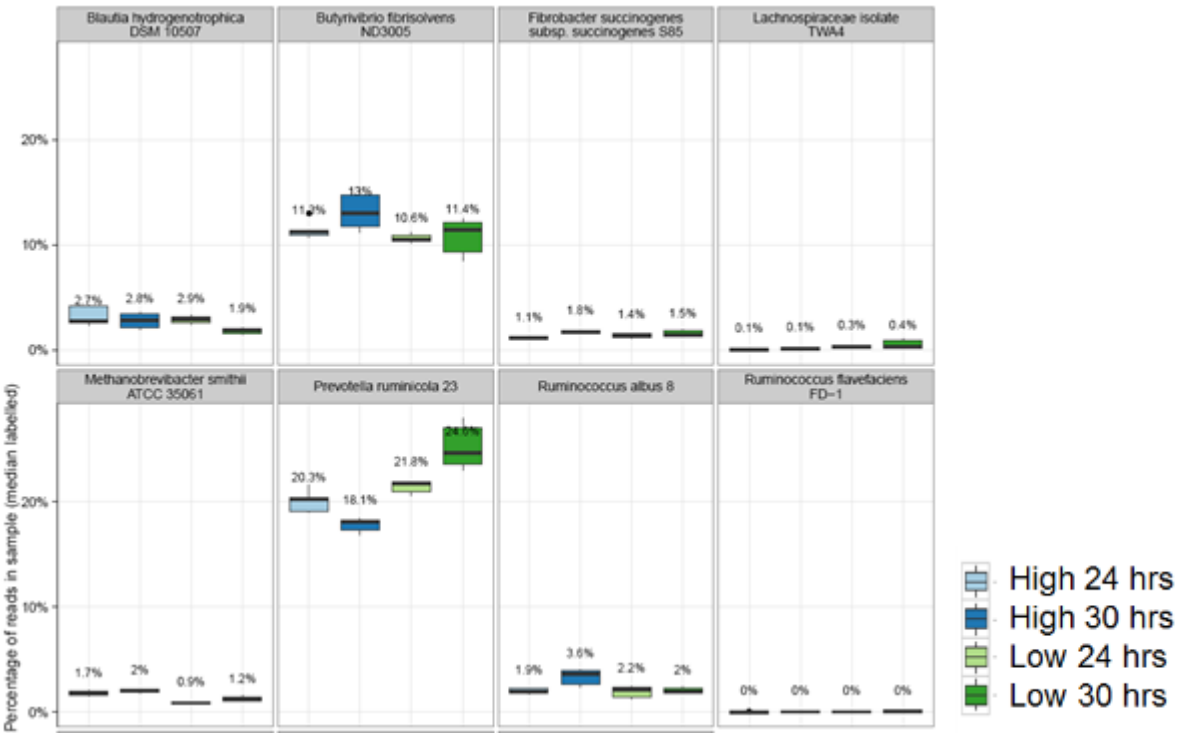


Figure 13. Quantity of RNAseq read data that mapped to each reference microbial genome for high and low methane cultures at 24 and 30 hr time points. Box plots are for 3 replicates at each time point and phenotype.

Transcripts assigned as carbohydrate active enzymes (CAZy) that are likely to be involved in the breakdown of plant fibre including glycoside hydrolases (GH), glycosyl transferases (GT), carbohydrate binding modules (CBM), polysaccharide lyases (PE), carbohydrate esterases (CE), dockerin and cohesion proteins showed a similar expression profile for each cellulolytic species based on time of sample and did not vary between the low or high methane phenotypes (Appendix 9).

Transcripts mapped to KEGG fermentative pathways that were up regulated in high methane cultures at 24 hours showed that the dominant species *P. ruminicola* was responsible for the production of propionate, while *F. succinogenes, B. fibrisolvens* and *R. albus* all produced actetate (Figure 14A). *F. succinogenes* also expressed genes for the production of succinate and *B. fibrisolvens* had active gene pathways for butyrate formation. *R. albus* was the only species that expressed genes at high levels for the production of ethanol in the high methane culture.

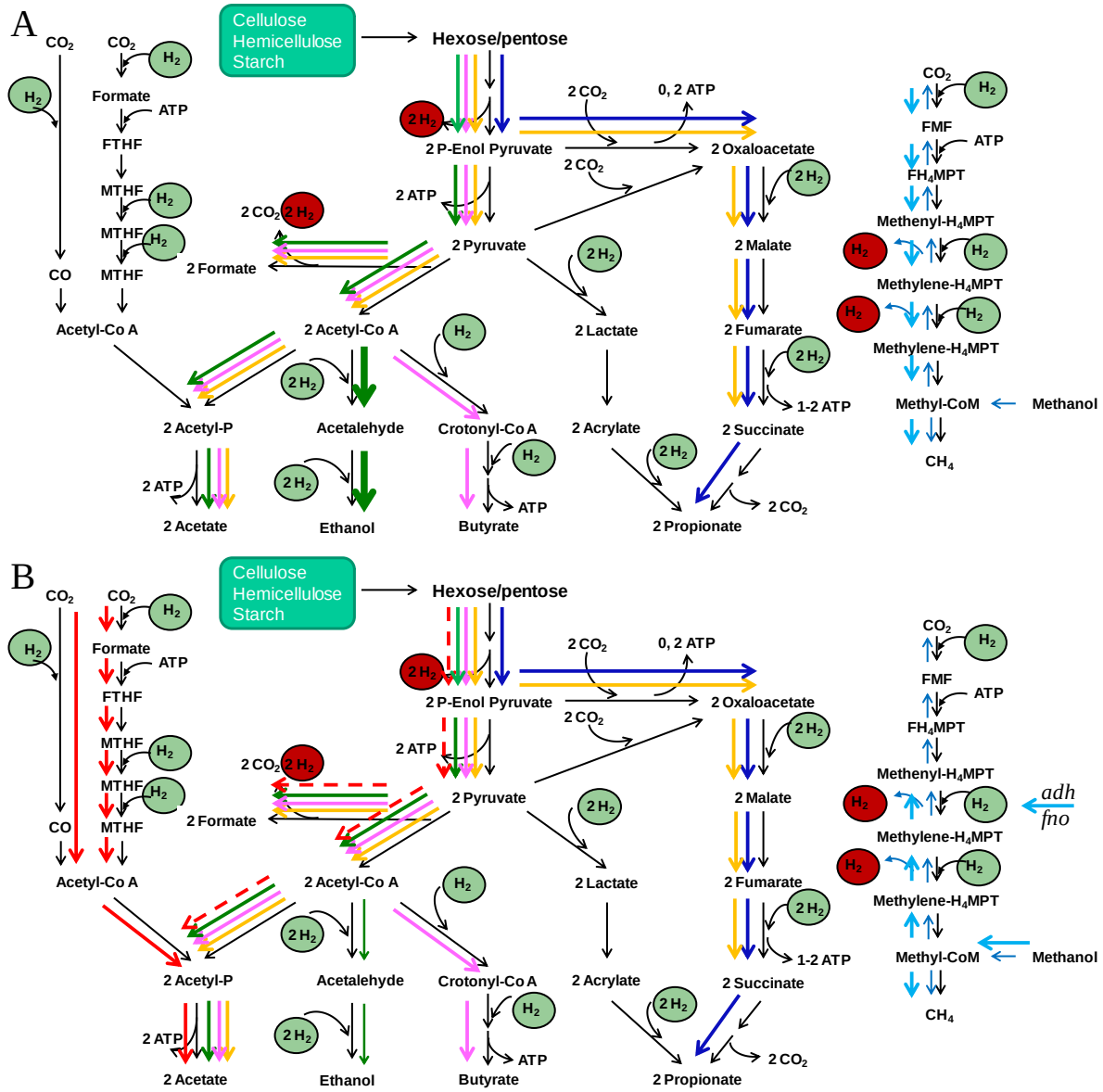


Figure 14. Transcripts mapping to fermentative pathways that are up regulated in A) high methane and B) low methane cultures. Coloured line represent species *P. ruminicola* (dark blue), *F. succinogenes* (yellow), *B. fibrisolvens* (pink), *R. albus* (green), *M. smithii* (light blue) and TWA4 (red).Dashed red line indicates heterotrophic acetate pathway for TWA4.

Methanogenesis by *M. smithii* followed the hydrogenotrophic pathway of oxidizing CO_2 while using H_2 as the reducing agent to drive electron transport. Differential gene expression for 24 hours in the low methane cultures were hydrogen was depleted showed similar fermentative pathways for *P. ruminicola, F. succinogenes* and *B. fibrisolvens* while *R. albus* redirected fermentation away from ethanol into acetate production (Figure 14B). *M. smithii* produced an increase in expression for the methanol specific methyl-transferases of the methylotrophic methanogenesis pathway and an increase in a NADP-dependent alcohol (*adh*) dehydrogenase and a F_{420}-dependent NADP oxidoreductase (*fno*). This coincided with a decrease in expression of the methanofuran and methanopterin generating "upper steps" of the autotrophic pathway (Figure 14 & 15).

The highest expressed genes for the acetogenic bacterial species TWA4 were those associated with the reductive acetogenic H_2 consuming Wood-Ljungdahl pathway and the uptake of glycerol and conversion to acetate via the glycolysis pathway (Figure 14B).

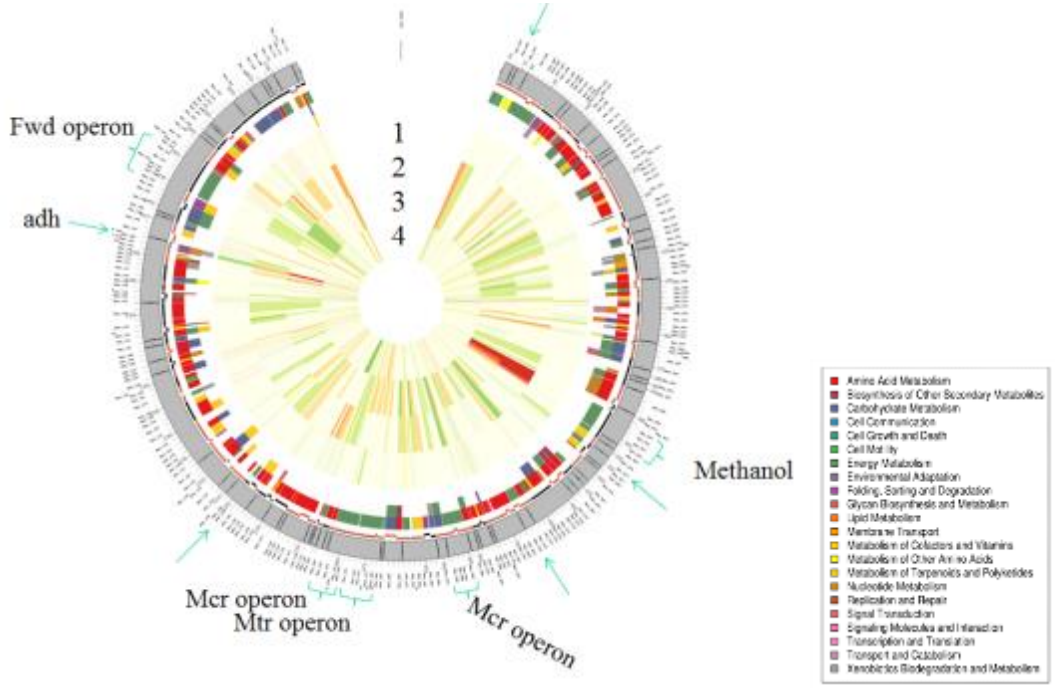


Figure 15: Up-regulated genes involved in methanogenesis for *M. smithii*. Rings 1-4 represent significant changes in gene expression levels (low green to high red) (1) 30 hrs fermentation high methane versus low methane, (2) 24 hrs high methane versus low methane fermentation, (3) high methane system 30 hrs versus 24 hrs and (4) low methane system 30 hrs versus 24 hrs respectively. Operon structures within the genome are indicated by parenthesis and gene id. Arrows indicate single up regulated genes involved in methanogenesis.

Development of qPCR assay for the monitoring of specific fermentative steps.

Genes encoding for enzymes considered to be critical for the determination of pathways that contribute to a low methane phenotype were either identified within the literature or designed based on analysis of the metagenomics/transcriptomic data (Table 6). Gene alignments of these key functional gens using data recovered from cattle rumen metagenomic data sets along with full length gene sequences from representative genomes and metagenomics datasets was used to ascertain appropriate sites for the design of qPCR primers for the purpose of assessing the transcriptional levels of these genes in low methane systems. Designed primers for acetyl-coA synthase and methanol transfer were found to be specific for their respective target gene as judged by sequence data of cloned products (data not shown). The efficiency of amplification of the gene targets by qPCR assay ranged from (1.8 – 2.1), indicating close to maximal efficiency of amplification.

Table 6: Key functional genes identified for monitoring low methane microbial fermentation

Pathway	Step	Gene	Reference
Methanogenesis	methyl reduction	mcrA	(Denman *et al.*, 2007)
		mcrA/mtrA	(Shi *et al.*, 2014)
	methanol transfer	mtaB	This study
Acetogenesis	acetyl coA synthase	acs	This study
Propionate	methylmalonyl-CoA decarboxylase	mmdA	(Reichardt *et al.*, 2014)

4. Discussion

Microbial ecology

Methane produced from cattle is an end by-product aimed at the removal of excess hydrogen produced from the bacterial conversion of plant polysaccharides into volatile fatty acids for the animal. Previous studies have studied the link between low RFI to reductions in methane emissions (Nkrumah *et al.*, 2006, Hegarty *et al.*, 2007). Nkrumah and colleagues speculated that the greater dry matter intake by high RFI animals was partly attributed to the low rumen metabolizability of the feed by the animals. Data here would suggest that this can be attributed at least in part to the changes in the observed microbial populations between the low RFI (low methane) and high RFI (high methane). In cattle that were producing “Higher” than or “Lower” than predicted methane outputs based on residual feed intake (RFI) there was found to be a difference in bacterial composition. However the methanogen populations were not seen to be significantly different in their compositions. With no significant change to the identified methanogen populations it is likely that in this situation the low methane phenotype is bacterially driven and would suggest changes in abundance to bacterial populations that produce/consume hydrogen. There was on a broad level a decrease in Bacteroidetes and a concurrent increase in Firmicutes, Proteobacteria and Cyanobacteria phyla. Closer examination revealed that 44 specific bacterial species were linked to the low methane animal phenotype, of which the majority were from the *Prevotella* genus. Previous work with Japanese goats fed bromochloromethane (BCM) to inhibit methanogens, also found a direct link between these species and a low methane state and increases in propionate production (Mitsumori *et al.*, 2011). An increase in propionate production by *Prevotella* species would consume H_2 and lower the rumen H_2 partial pressure accordingly aiding in re-oxidation of NADH and maintaining efficient rumen fermentation. Further studies of the goat rumen microbiome using 16S amplicon analysis also confirmed an increase in the two dominant *Prevotella* OTUs (Denman et al 2015 submitted). Unlike the low cattle however, the Bacteroidetes phylum was observed to increase with higher doses of BCM with a concomitant decrease in Firmicutes, Synergistetes and Lentisphaera. Bromochloromethane is an halogenated methane analogue that is believed to inhibit methane production by reacting with reduced vitamin B12 and inhibiting the cobamide-dependent methyl transferase step of methanogenesis (Wood *et al.*, 1968, Chalupa, 1977). This step is immediately prior to the terminal reductive reaction and is responsible for the synthesis of methyl coenzyme M (Wood *et al.*, 1982). The specificity of BCM to only target methanogens is not known and it is likely that due to the similarity in enzyme structure and co factor requirements for reductive acetogens, that these populations may also be affected. Functional gene studies for reductive acetogenesis in the rumen of goats treated with BCM clearly showed changes in diversity for these populations and 16S amplicon data showed a reduction in their abundance (Denman et al 2015 submitted). One of the most abundant groups of bacteria identified as core member of the low methane cattle phenotype were from the Lachnospiraceae family and were closely related to species identified in the foregut of the low methane producing Tammar Wallaby as known acetogenic bacteria. Bacterial 16S rDNA clone libraries generated from the Tammar Wallaby have previously shown little overlap with cattle rumen samples, indicating very divergent populations that contribute to the low methane phenotype (Pope *et al.*, 2010). The cattle used for comparison with the Tammar Wallabies were only limited and were focused on the microbial populations that attached to the biofuel substrate switch grass with no consideration for a methane phenotype. The contribution of these species to the low methane cattle phenotype is not understood but warrants further investigation based on their abundance in the low methane cattle and other low methane environments. An isolate from this group has recently been described from the Tammar Wallaby as an acetogenic species that can switch between mixotrophic growth (concomitant consumption of hydrogen during heterotrophic growth) with glycerol and autotrophic growth (Gagen *et al.*, 2014). Under heterotrophic growth conditions, isolate TWA4 did not significantly stimulate methanogenesis in a co-culture with *M. smithii,* contrary to the expectation for fermentatively growing organisms. This isolate has the capability to redirect the hydrogen that it produces into acetate production and not release it into the gut environment where methanogens could utilise it as an energy source. This strategy may underpin part of the explanation for a low methane phenotype in these animals.

A further synergy with the low methane rumen and Tammar Wallabies was the increase in the abundance of the Succinivibrionaceae family with a 30 times increase in the rumen of low methane animals. A Succinivibrionaceae isolate WG-1, from the Tammar Wallaby has been isolated and studied and is a known succinate producer from starch diets (Pope *et al.*, 2011). It was concluded that promotion of this

species in the rumen could contribute to substrate oxidation and reductions with little H_2 being released for methane formation. The increase in succinate formation would be available to the animal directly or further transformed by *Prevotella* species to propionate.

An interesting observation was the increase in bacteria associated with the phylum Cyanobacteria in low methane animals. Cyanobacteria are responsible for the production of oxygen as a by-product of photosynthesis. This group has recently been reclassified to explain the presence of Cyanobacteria in anaerobic "dark" environments (Di Rienzi *et al.*, 2013). It is hypothesised from genome reconstruction that these bacteria are obligate anaerobic fermentative bacteria that produce hydrogen, lactate, acetate and formate as end products (Di Rienzi *et al.*, 2013). The identification of a FeFe hydrogenase in the reconstructed genomes suggests the production of high molar levels of H_2 and would typically require close association with a H_2 consuming species to maintain low partial H_2 pressures. For the numbers of this group to increase in the low methane animals would require syntrophy with methanogens, sulphate reducers or acetogenic bacteria. Methanogen diversity was not altered between phenotypes, with *Methanobrevibacter* dominating both phenotypes, suggesting that the changes in the bacterial diversity and increase in acetogenic associated bacteria may be linked with these novel bacteria. A close relationship between these bacteria would allow for efficient interspecies H_2 transfer and benefit both populations.

Samples made available through the NLMP network allowed access to 320 rumen samples that were collected from animals as part of the genetic technologies to reduce methane emissions from Australian beef cattle project (Robert Herd NSW DPI: 01200.044; B.CCH.6310). This encompassed paired rumen and faecal samples for each animal. Due to the nature of the trail from which the samples were collected it was observed the the major influence on the rumen microbiome data for both the bacterial and archaeal populations was the cohort within year (which accounted for changes in sex, age, location, weight and time of the year). The use of a linear model to account for these factors could not clearly separate microbiomes based on methane emissions alone. The nature of these sequencing methods can generate anomalies in the data analysis if factors such as changes in the range of species diversity and sample coverage are not accounted for. The substantial difference between sample diversity and abundance metrics likely reduces the sensitivity of the linear model.

A new methodology to appropriately handle the library size and diversity differences was adapted using the DESeq2 modelling package. This approach calculated the mean OTU value as a quantity proportional to the total sequencing effort present in the sample, scaled by a normalisation factor which adjusted for library size differences across samples. When the samples were divided into groups based on their methane yields there were 299 differentially abundant OTUs between the high and low treatments (FDR corrected, $q < 0.05$). Again the low methane OTU's in the rumen were dominated by *Prevotella* species, but also identified increases in the fibrolytic *Fibrobacter* species. This is in agreement with the 16S amplicon data from the low methane animals described above and that for the goats treated with BCM. The common population changes appear to be increases to specific *Prevotella* and *Fibrobacter* species with decreases to *Ruminococcus* species. Changes to Synergistaceae and Veillonaceae seem to be more variable across the trials and may reflect more the intervention method (BCM) or the feed offered.

The majority of methane production from cattle is known to be produced from the rumen with only a minor contribution from lower gut fermentation (Lassey *et al.*, 1997). Faecal samples are easily collected from animals and would provide a rapid analysis tool for defining microbial population changes between animals. Unfortunately the limited differences detected between the faecal microbiota of low and high methane animals would conclude that it is not possible to develop a methodology for determining the difference in animal methane phenotype from faecal sample microbiome composition.

Functional genomics (metagenomics)

Phylogenetic assignment of marker genes against curated datasets accurately identifies to the species level only a small percentage of the microbiota. However, key methanogen populations were observed at varying levels in the "low" and "high" methane animals with reads assigned to *M ruminantium,* being observed 50% less in the low methane animals. Methods that identified the ribosomal sequences within the metagenomic data were more robust and were in general agreement with the previous amplified 16S specific amplicon data. Similar population shifts as a contribution to the metagenomics DNA pool were

associated with an increase in certain *Prevotella*, and *Fibrobacter* species and Lachnospiraceae, Porphyromonadaceae and Succinivibrionaceae family members. This provides further evidence to the importance of these populations in low methane animals. Although the Ruminococcaceae family was increased in the low methane animals, sequences most closely associated with the fibrolytic *Ruminococcus* species were decreased in a manner similar to the 16S amplicon data.

The generation of contiguous sequences from within a metagenomic sample can indicate the relative abundance of the species from which genomic material originated, implying that those microbes contribute a greater quantity of nucleic material to the sample and are more likely to be sequenced to a greater depth. The use of sequence compositional based algorithms to more accurately classify and group sequences to phylogenetic taxonomy was only achieved to the order level, but could account for 95% of the data and is a reflection of the high level of diversity within the cattle rumen and novelty of the composition compared to currently available reference datasets. However, functional characterization of the microbiome through the annotation of the metagenomic sequence data, especially those genes involved in transitional metal uptake and methanogenesis produced a clear separation based on the methane phenotype. Of greatest interest was the marked decline in the number of sequences associated with the abundance of ABC transporters responsible for the uptake of metals into the microbial cells and of read abundance associated with tetrahydromethanopterin S-methyltransferase subunits. Tetrahydromethanopterin S-methyltransferase is part of a complex that catalyzes the formation of methyl-coenzyme M and tetrahydromethanopterin from coenzyme M and methyl-tetrahydromethanopterin. This is an energy-conserving, sodium-ion translocating step and is essential to all methanogenic pathways. In a recent study, metagenomic and metatranscriptomic analysis of microbes from sheep fed a pelleted lucerne diet selected for low and high methane genetic phenotypes did not show an altered abundance for these genes at the genomic level based on methane phenotype (Shi *et al.*, 2014). However they did find highly significant changes at the transcriptomic levels revealing lower gene transcripts for these energy yielding steps in low methane animals. Variation in animal species and diet offered may explain the lower abundance of these genes in low methane cattle. However, metatranscriptomic testing of the rumen should be performed to determine if gene regulation is also lower in low methane cattle.

Simulating low methane phenotypes to study gene transcriptomics

Microbial ecology data and functional analysis of metagenomic data provide evidence of the key species that define a low methane phenotype in cattle and goats. Fermentation cultures containing some of these key functional species were established to elucidate changes in functional pathways associated with lower methane in order to better understand these mechanisms. The introduction of TWA4 a known acetogenic bacterium, closely related to species detected to be increased in the rumen of low methane cattle was responsible for directly competing with the methanogenic species for hydrogen in the cultures. Regardless of the plant substrate provided both fermentations containing TWA4 resulted in a decrease in methane production with a concurrent increase in acetate for these systems. Lucerne fermentations produced greater quantities of free hydrogen for low methane cultures compared to the Rhodes grass cultures and is most likely a consequence of higher levels of soluble and easily fermentable sugars in Lucerne hay. In this instance, the bacterial species capable of utilizing hydrogen were not required to compete for hydrogen with the methanogens, as partial pressures of H_2 were maintained at higher levels. Generally methanogens have a competitive advantage over bacterial species when competing for available hydrogen in a system, due to the lower thermodynamic potential of methanogenesis (Cord-Ruwisch *et al.*, 1988). With cultures grown on Rhodes grass, where less fermentable material is readily available, the low methane cultures maintained a low partial pressure of H_2. Although this reduction in H_2 partial pressure would favour the thermodynamics of methane production the consortia in the low methane culture resulted in a much greater reduction of methane compared to lucerne (38% compared to 22% respectively).

Sequencing of the mRNA transcripts using RNAseq allowed for the discrete mapping of read data onto the relevant annotated microbial genomes. The proportion of RNA transcripts assigned to each species mirrored the abundance of the species within the rumen with *P. ruminicola* being the most dominant followed by *B. fibrisolvens* and then the dominant fibre degraders *R. albus* and *F. succinogenes. R.* albus was actively expressing the gene encoding for albusin B, which has been identified as the *R. flavefaciens* specific bacteriocin that allows *R. albus* to dominate in the rumen. *M. smithii* as the representative of the

methanogen populations contributed up to 4% of the data in the high methane cultures which are at the high levels of what is to be expected in the rumen (Janssen & Kirs, 2008). The *M. smithii* strain used for the fermentations is one of the most efficient and fastest growing of the isolates available within our laboratory and the high transcript levels likely reflects its superior growth characteristics. The decrease in transcripts and differential gene expression in particular to genes associated with methanogenesis observed in *M. smithii* reflects its contribution to the fermentation parameters, when methane production was decreased by 23 to 38% (24 hr and 30 hr respectively) there was a concurrent decrease in *M. smithii* transcripts with a 50% reduction to 2% of the total transcripts sequenced. Under optimal conditions for methane production in the high methane cultures, the most expressed genes for *M. smithi* were those associated with methanogenesis using the hydrogenotrophic pathway. Within the low methane cultures *M. smithii* has altered the pathway to compensate for the lower hydrogen concentrations. Methanol or ethanol as a product of plant fibre fermentation can be oxidised by the NADP-dependent alcohol dehydrogenase and the resulting reduced H_2 used to form $F_{420}H_2$ by using the NADPH-dependent F_{420}-H_2 oxidoreductase. Both of these enzymes were significantly up regulated in the low methane cultures and would allow for the production of $F_{420}H_2$ under H_2 limiting conditions observed in the low methane cultures. In addition, three methanol:cobalamin methyltransferases which transfer the methyl group from methanol to coenzyme-M to produce methyl-CoM which is subsequently reduced to methane were the most highly expressed genes in *M. smithii* for the low methane cultures. Although it is known that ethanol and methanol can enhance the growth of *M. ruminantium* in the presence of H_2 (Leahy *et al.*, 2010), true methylotrophic methanogens all belong to the order Methanosarcinales, Methanoplasmatales and the genus *Methanosphaera,*. Methylotrophic methanogens generate the electrons required for the reduction of the methyl group to methane from the production of CO_2 through the stepwise generation of reduction equivalents by reversing the hydrogenotrophic methanogenesis pathway (Thauer, 1998, Fricke *et al.*, 2006) . This “disproportionate” pathway results in the production of 3 moles of methane from 4 moles of methanol. The transcriptome for *M. smithii* reveals an incomplete disproportionate pathway and is therefore only capable of converting 4 moles of methanol/ethanol to 2 moles of methane. This implies that at lower hydrogen levels, *M. smithii* has the ability to adapt by switching to a less efficient methane production and obtain hydrogen from methanol and or ethanol at two steps in the methanogenic pathway. Methanol and ethanol are known fermentation products of plant fermentation and in both fermentation systems studied here *R. albus* was contributing ethanol as a final fermentation end product.

Metatranscriptomic data generated from low and high methane emitting sheep has recently been published (Shi *et al.*, 2014). The authors were able to show links with reductions at the transcriptional level in the methanogenesis pathway between the two animal phenotypes. The transcript changes found in the low methane fermenters described here concur with these findings, but the animal data does not describe the fate of the methanol:cobalamin methyltransferases. Further analysis of the animal dataset will need to be undertaken to determine if changes to this system are evident in animals.

The acetogen TWA4 was responsible for the increased acetate production in the low methane fermentations and actively competed and sequestered available hydrogen in direct competition with *M. smithii*. Species TWA4 achieved this by performing mixotrophic growth, utilising the reductive acetogenic pathway as well as converting glycerol to acetate via the glycolysis pathway ensuring net consumption of hydrogen. The transcript levels suggest that the abundance of TWA4 did not reach dominant numbers within the fermentation at 24 hours. Further studies revealed that the lysis method used to extract RNA did not efficiently lyse the TWA4 species (data not shown). A more extensive lysis system did effectively lyse TWA4 and a qPCR assay revealed TWA4 to be present at approximately 3% of the population.

Measurements of H_2 concentrations in cattle have shown the rate of H_2 production immediately after feeding leads to a transient overloading of methanogenic archaea capacity to use H_2, resulting in peaks in H_2 emissions from microbial fermentation (Rooke *et al.*, 2014). Hydrogen concentrations in the high methane fermenters followed this pattern with increased concentration of H_2 being detected in cultures that contained *M. smithii* as the only hydrogen utilising species. Increased hydrogen concentrations are known to affect *Ruminococcus* species and their fermentative pathways. Although higher hydrogen concentrations altered the fermentative pathways and end products of fermentation for *R. abus*, the changes did not alter the expression of enzymes involved in plant hydrolysis. In fact the CAZy enzymes for the three major fibre degraders were not altered between the low and high methane fermenters. Clustering of expression values for each enzyme showed that the level of hydrogen/methane had less of

an effect then the time at which the sample was collected. The differences in gene expression between the time points would indicate changes in the substrate available as fibre break down progresses, require the use of different enzymes. However, hydrogen concentrations did alter the fermentation profile for *R. albus* in the high methane/hydrogen fermentations where there was up regulation of genes for the production of ethanol and FeFe hydrogenases involved in the redirection of energy pathways to regenerate NAD^+ which is inhibited at high hydrogen concentrations (Zheng *et al.*, 2014). *R. albus* was able to redirect fermentation to the more energy yielding acetate production in the low methane cultures were hydrogen was maintained at lower concentrations by TWA4 and *M. smithii*.

Formate is also an electron donor source for methanogenesis and measures tracking formate also indicated that there was an increased level of utilization of formate in the low methane cultures. Pure culture experiments and the genome sequence of *Lachnospiraceae sp.* TWA4 have shown that this species can use mixo and heterotrophic acetogenic pathways including the use of formate for the production of acetate. The transcriptomic analysis of the high and low cultures on Rhodes grass indicates expression of genes for TWA4 in utilising formate both in the reductive acetogenic pathway and for a pyruvate formate-lyase gene that converts pyruvate to acetyl-coA with the release of formate that could then be redirected to the reductive acetogenic pathway rather then released as an end product. *M. smithii* did not show a significant up regulation of the formate utilising genes that could allow for formate to be used as an electron donor.

Researchers have previously investigated the feasibility of promoting acteogenic pathways in the rumen or in *in vitro* rumen fermentations by firstly reducing methanogens to promote a suitable environment for acetogensis, but with little success (Demeyer *et al.*, 1996, Immig *et al.*, 1996, Nollet *et al.*, 1998, Lopez *et al.*, 1999). In most cases the rumen microbiome has adapted to the anti-methanogenic strategy, rendering its potency inadequate and allowing methanogen levels to return to normal. Furthermore, acetogenic bacteria possess varying levels of competiveness to effectively compete for hydrogen, with researches noting that only two of the six strains tested could sustain a reduced methane state in rumen *in vitro* fermentations (Lopez *et al.*, 1999). These previous studies have primarily focused on monitoring the effects of the treatments by changes in rumen fermentation parameters; later researchers began to track specific populations using qPCR to more accurately quantify changes in the populations (Watanabe *et al.*, 2010). However, none of these studies could monitor specific species and assign their functional contribution to the rumen system or how the microbiome had overcome the intervention strategy. With the advent of high throughput sequencing and metatranscriptomic workflows it is now possible to clearly quantify the function and level of activity a species is contributing and which energy pathways are altered during these microbiome fluxes. The biochemical gas production and SCFA measures all indicate that changes to fermentation pathways have been influenced by the bacterial composition between the low and high consortia and the introduction of TWA4 and its effect on hydrogen concentrations. Transcriptomic analysis correlates with these biochemical measures showing the major fermentation pathways for each species and the changes in these to hydrogen inhibition for some species. TWA4 is seen to possess active pathways that are capable of sequestering H_2 away from *M. smithii,* but also competitively removing formate as an alternative electron donor source from *M. smithii.* Due to the competition for electron donors *M. smithii* is required to activate a less efficient mechanism for generating methane through methyl transferases and alcohol dehydrogenases, resulting in decreased methane production and cell numbers.

Metatransciptomics of complex microbial populations, such as those within the rumen, across large numbers of animals, although feasible, would involve a large expense and generation of more relevant metagenomic and genomic scaffolds to aid in transcript mapping. Quantitative PCR primers have been designed that target multiple genes in three major fermentative pathways observed to have altered expression levels in low methane systems. Primers have been designed and sourced from the literature to target these pathways. All primer sets were found to have acceptable amplification efficiencies and were found to be specific and only amplified target genes ensuring their robustness as part of a qPCR assay. Appropriately collected rumen samples for the purpose of RNA extraction will need to be sourced for the testing of these primers under real conditions. The qPCR assay will allow rapid monitoring of changes to these key fermentative pathways and aid in further refinements of current understanding of these pathways in the cattle rumen.

5. Conclusion

The rumen microbiota of cattle and sheep contain a complex mix of microbiota that are responsible for the bioconversion of plant material into energy yielding products for the animal. This is achieved through hydrolysis of the plant polysaccharides and then pathways involving both hydrogen-producing and hydrogen-consuming steps into various fermentation end products such as short chain fatty acids (SCFA). Methanogenic populations have evolved in this system. They have a niche role to consume metabolic [H] and produce methane, which accounts for a 2-12% loss of the metabolic energy from feed. Redirection of H_2 into more energy yielding end products that the animal can utilize will not only lead to reductions in methane emissions but efficiency improvements in animal livestock production.

At the outset of this project, our understandings of the rumen microbiota and the functional roles that different populations perform in the rumen were limited largely to studies of laboratory isolates. Although this strategy has led to the identification and association of microbial groups to key functions in the rumen, the isolates represent only a small proportion of the potential that resides in the rumen. The use of improved sequencing and molecular methods has allowed for a greater depth of analysis of the rumen under varying rumen methane phenotypes. Findings from this research have identified differences in rumen microbes which may explain some of the difference between low and high RFI as it pertains to lower methane production. Interestingly, the microbiota of low RFI animals were found to harbor similar dominant populations to those that were found in the gut of the low methane producing Tammar Wallabies. These populations are responsible for efficient turnover of soluble sugars without the release of H_2 into the gut environment. Fibrolytic populations are important primary colonizes of plant material, and any reduction in these populations is likely to affect rumen fermentation. Consistent with data from low methane goats in which increases in rumen H_2 partial pressure led to reduction in the *Ruminococcus* species and increase in *F. succinogenes*, the rumen of low methane cattle were observed to follow a similar strategy. Although rumen H_2 partial pressure concentrations were not measured in the cattle studied here, it would indicate that fiber digestion is not compromised in low methane rumen systems.

Monitoring of gene transcripts in low methane fermenters found an acetogen species responsible for redirection of H_2 from methanogenesis to acetate production. This is the first evidence of a successful competition in a complex system of a reductive acetogen competing with a methanogen for the available H_2. The methaogen species was found to alter its normal methanogenic pathway to attempt to offset this competition by upregulating gene involved in the use of alcohols as alternative H_2 sources. Which are likely to be less efficient at providing energy to the methanogenic species and explains the decrease in methane production.

Strategies that promote these species within cattle are likely to result in lower methane emissions and will revolve around increased starch and soluble sugar content in diets. In addition, these same substrates are likely to increase the abundance of the other important Tammar wallaby gut Succinivibrionaceae species in the cattle rumen. Isolates of these two important species have been obtained from the Tammar gut and could be used as part of an inoculum strategy, although cattle isolates would be more preferable and attempts to isolate cattle isolates should be undertaken.

6. Future research needs

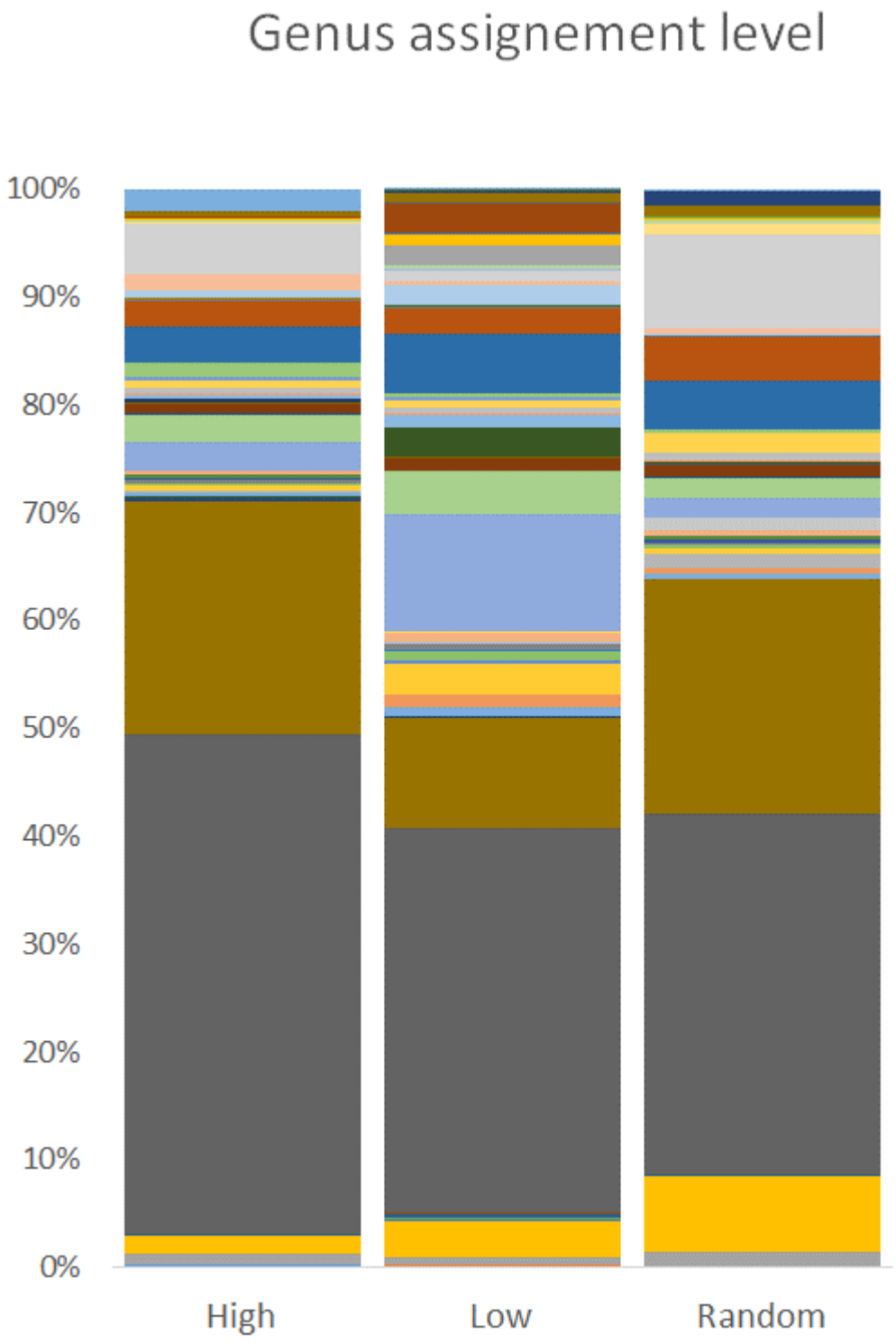
The findings from this project have identified microbial species and their functional pathways that aid in the reduction of methane in cattle. Strategies that promote these species within cattle are likely to result in lower methane emissions and will focus on dietary manipulation. The introduction of novel species into the rumen will require the conditioning of the bacteria and the provision of the correct substrates for those introduced species to establish and maintain sufficient numbers to provide impact in the rumen. Further work to define these optimum conditions should be undertaken as should efforts to isolate cattle representatives from these important groups.

7. Publications

1. Denman SE. “Metagenome analysis, community structure” at the Joint RuminOmics/Rumen Microbial Genomics Network Workshop in conjunction with greenhouse gasses and animal agriculture 2013
2. Morrison M, Rosewarne, C Denman SE. “Making more of metagenomics - or -- when everything old is new again?” Gut microbiology: from sequence to function. Rowett – INRA 2014
3. Denman, S, Bragg, L, Smith, W, McSweeney, C and Morrison,M. “Metabolism Pathway Shifts by Bacteria Competing for Hydrogen in Invitro Mixed Rumen Microbial Fermentations.” CGIF 2015
4. Gagen, E. J., et al. (2014). "Investigation of a new acetogen isolated from an enrichment of the tammar wallaby forestomach." BMC Microbiology 14(1): 314.
5. Denman et al (2015) “Metagenomic analysis of the rumen microbial community following inhibition of methane formation by a halogenated methane analogue”. Submitted to Frontiers in Microbiology

8. Appendices

Appendix 1

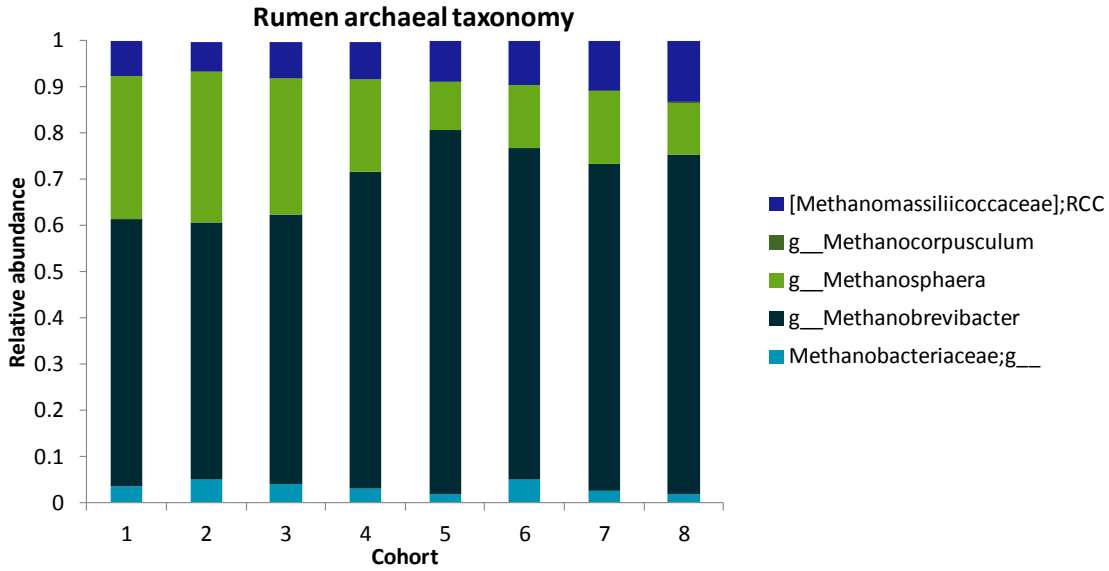


- k__Bacteria;p__Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__RFN20
- k__Bacteria;p__Synergistetes;c__Synergistia;o__Synergistales;f__Synergistaceae;Other
- k__Bacteria;p__Synergistetes;c__Synergistia;o__Synergistales;f__Dethiosulfovibrionaceae;g__Pyramidobacter
- k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
- k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Sphaerochaetales;f__Sphaerochaetaceae;g__Sphaerochaeta
- k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__
- k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
- k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfobacterales;f__Desulfobulbaceae;g__Desulfobulbus
- k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
- k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
- k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__[Coprobacillaceae];g__Sharpea
- k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5
- k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__L7A_E11
- k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Coriobacteriales;f__Coriobacteriaceae;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Coriobacteriales;f__;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Succiniclasticum
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Selenomonas
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Mitsuokella
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Acidaminococcus
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Shuttleworthia
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Moryella
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Catonella
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Mogibacterium
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;Other;Other
- k__Bacteria;p__Firmicutes;c__Clostridia;o__;f__;g__
- k__Bacteria;p__Firmicutes;c__Clostridia;Other;Other;Other
- k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
- k__Bacteria;p__Firmicutes;Other;Other;Other;Other
- k__Bacteria;p__Fibrobacteres;c__Fibrobacteria;o__Fibrobacterales;f__Fibrobacteraceae;g__Fibrobacter
- k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
- k__Bacteria;p__Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
- k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Anaerolineales;f__Anaerolinaceae;g__SHD-231
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__p-2534-18B5;g__
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__YRC22
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;Other
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Paludibacter
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__BS11;g__
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
- k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;Other;Other
- k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Pseudonocardiaceae;g__Saccharopolyspora
- k__Bacteria;Other;Other;Other;Other;Other

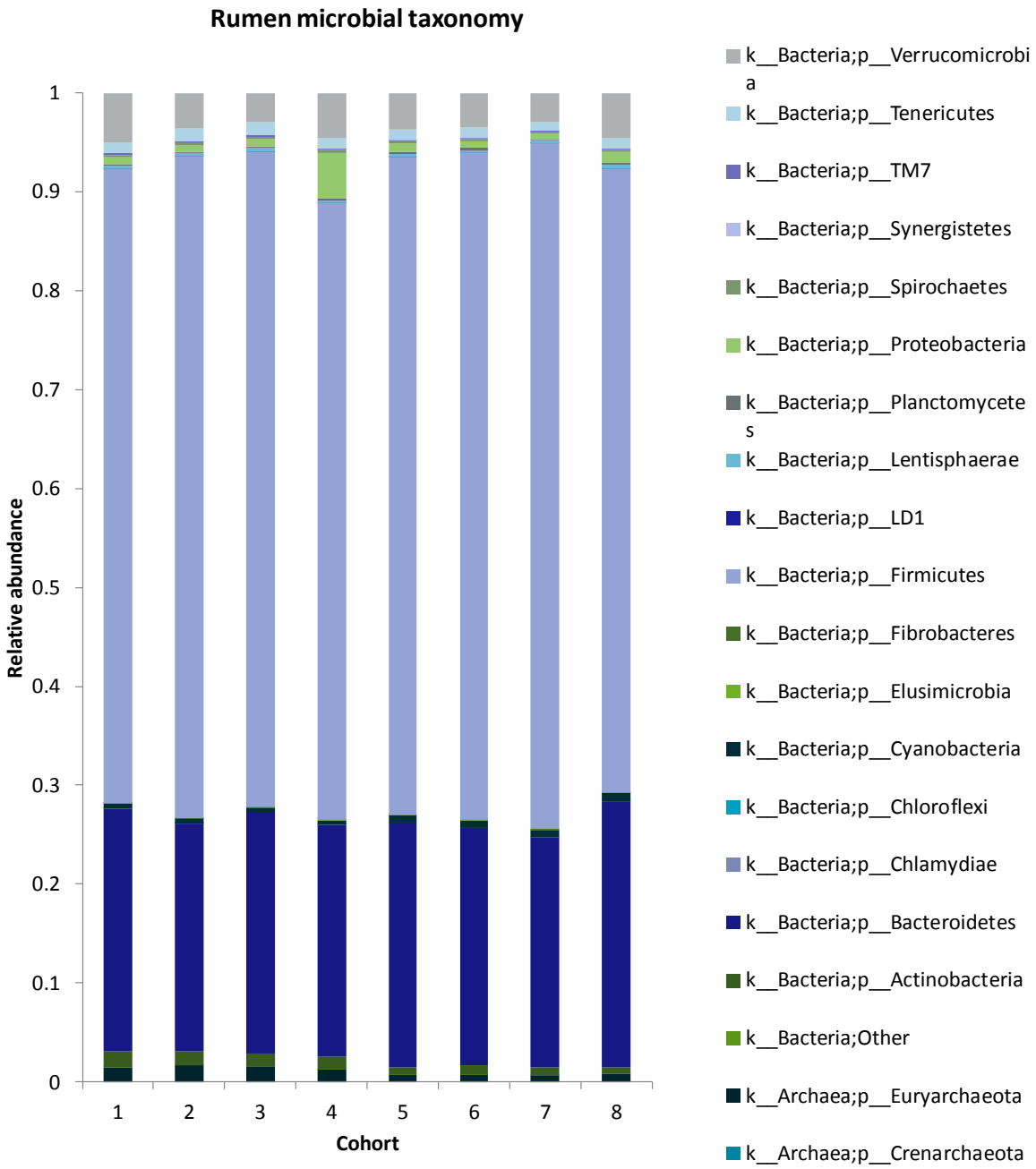
Relative abundance of microbial sequence data assigned to the genus level for low and high methane animals

Appendix 2

Rumen bacterial and archaeal taxonomy abundance for animal cohorts within year



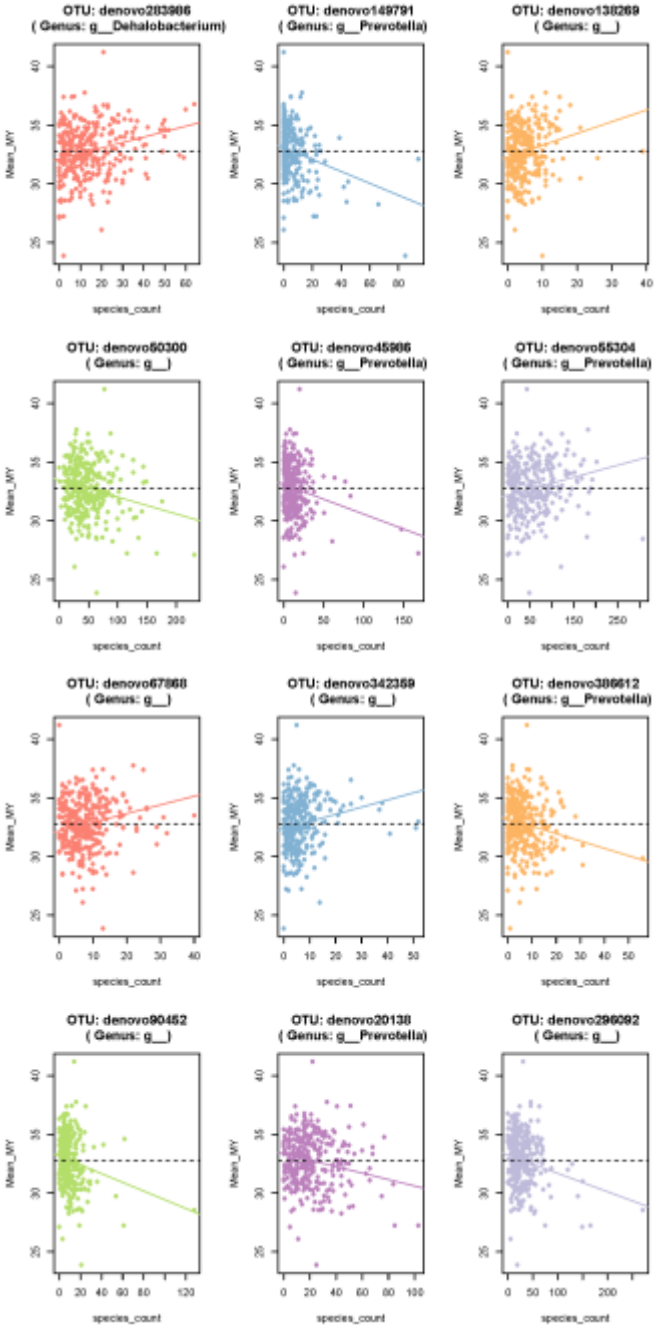
Relative abundance of archaeal sequence data assigned to the genus level for animals grouped as cohorts within year Trangie male (1&2), Trangie female (3&4), Glen Innes male (5&6) and Glen Innes female (7&8)

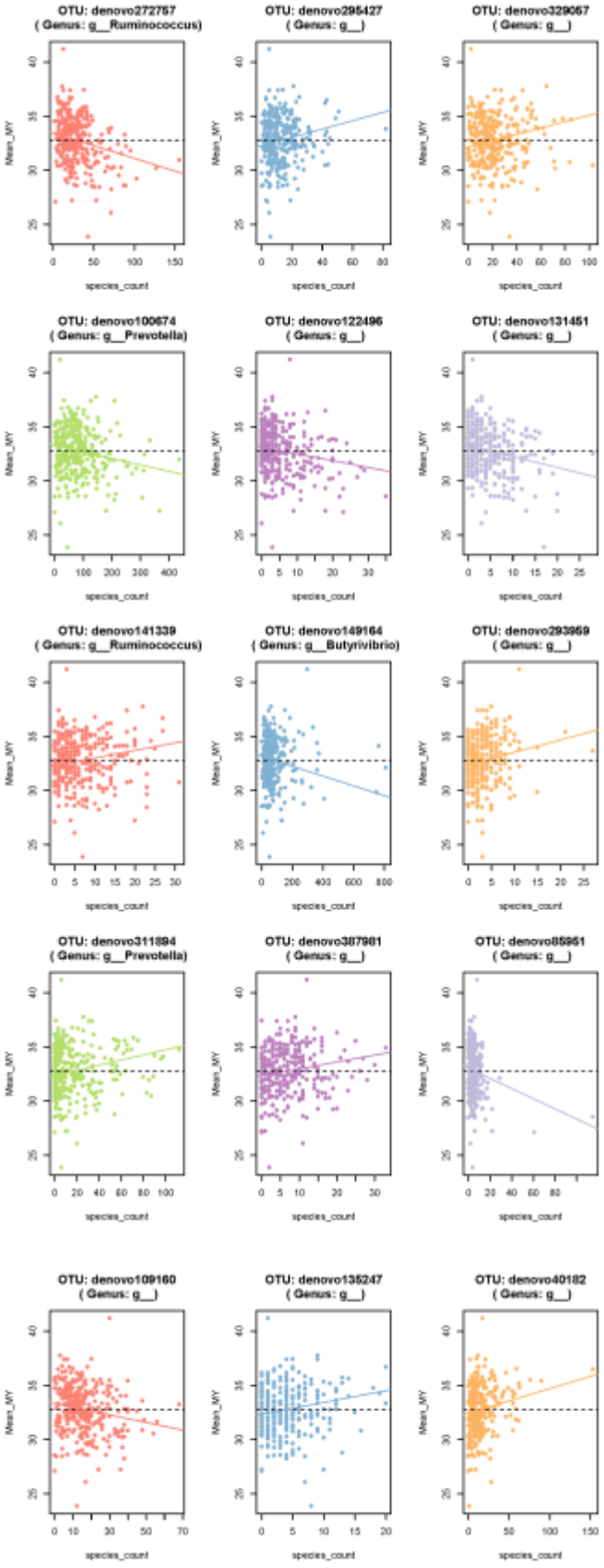


Relative abundance of microbial sequence data assigned to the phylum level for animals grouped as cohorts within year Trangie male (1&2), Trangie female (3&4), Glen Innes male (5&6) and Glen Innes female (7&8)

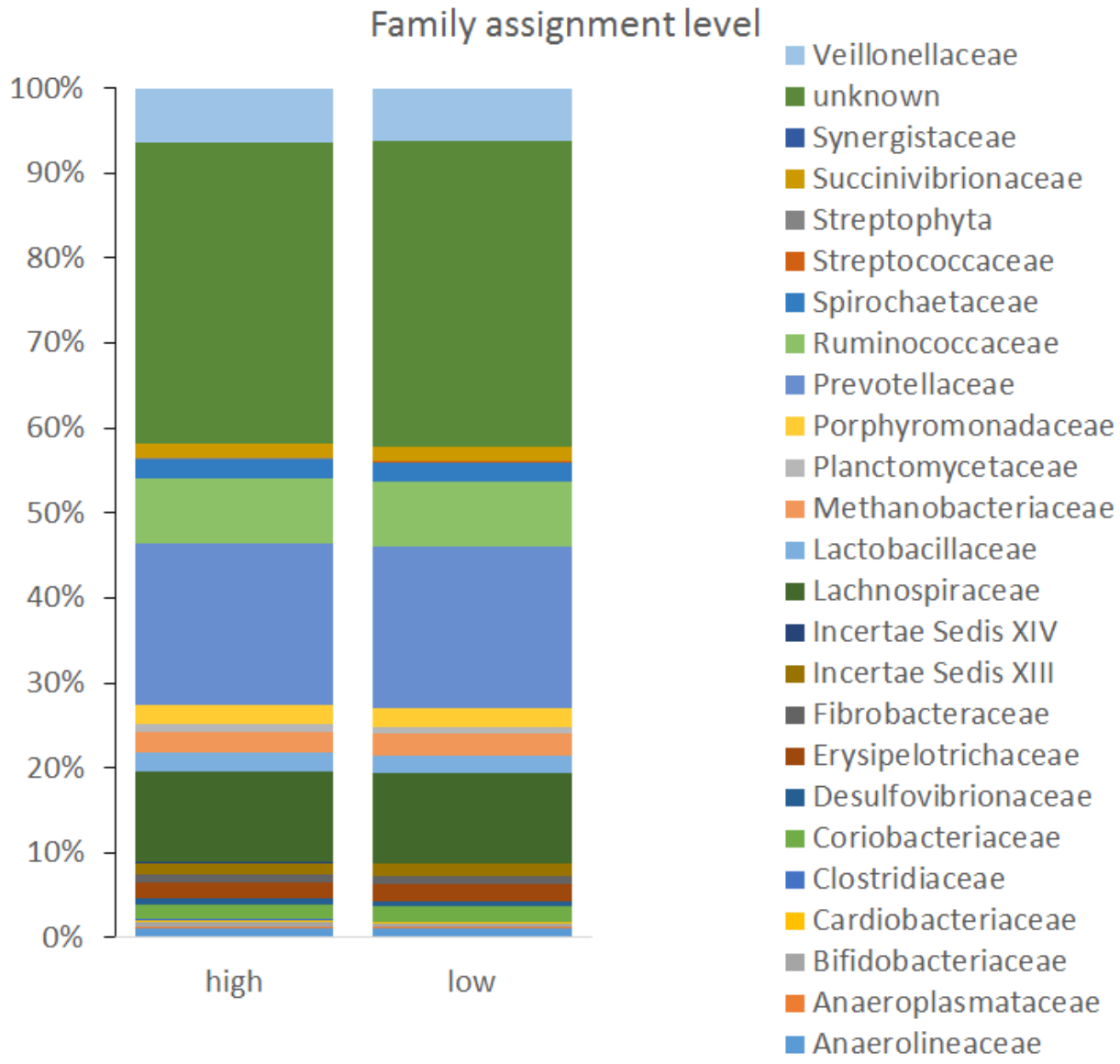
Appendix 3

Significantly altered OTU's from linear model for rumen samples



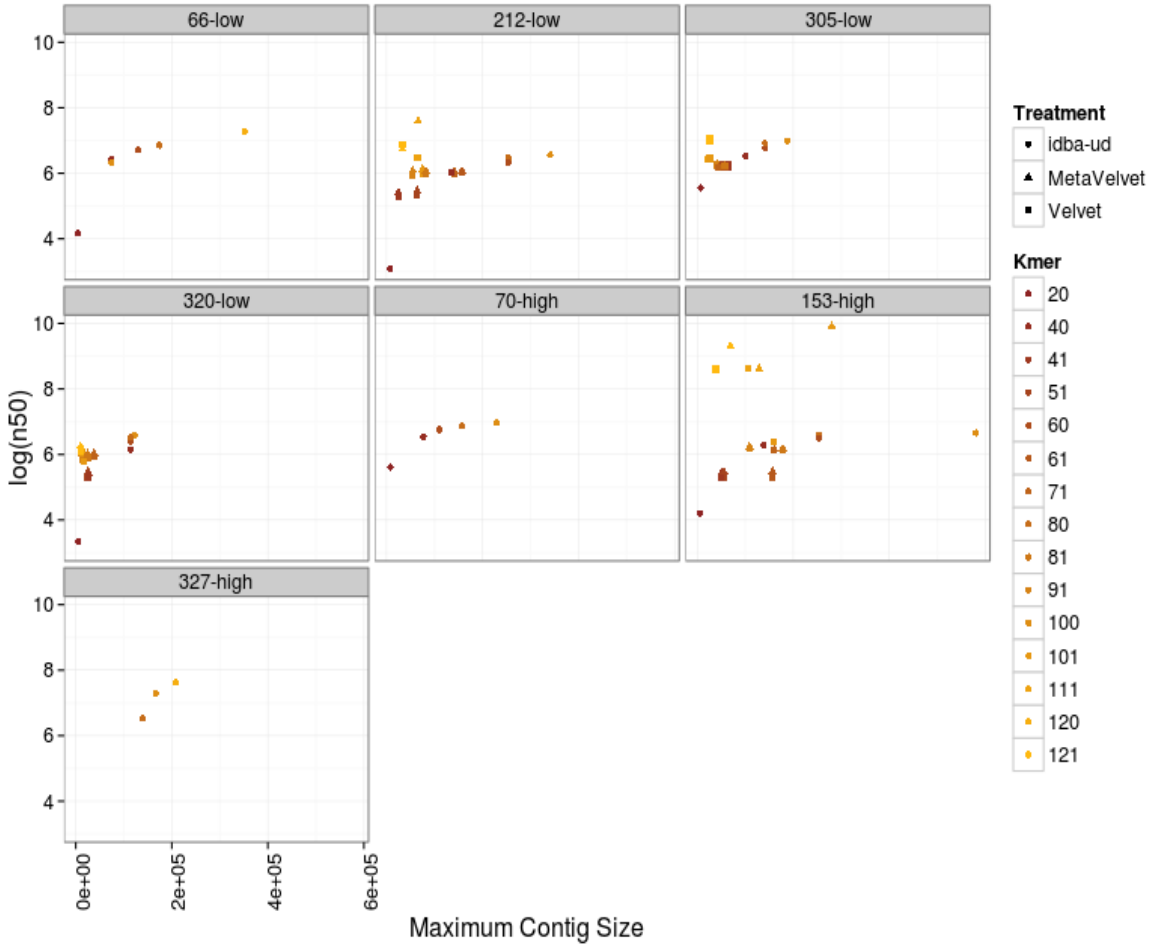

OTU: denovo272757
(Genus: g__Ruminococcus)
OTU: denovo295427
(Genus: g__)
OTU: denovo329057
(Genus: g__)
OTU: denovo100674
(Genus: g__Prevotella)
OTU: denovo122496
(Genus: g__)
OTU: denovo131451
(Genus: g__)
OTU: denovo141339
(Genus: g__Ruminococcus)
OTU: denovo149164
(Genus: g__Butyrivibrio)
OTU: denovo293959
(Genus: g__)
OTU: denovo311894
(Genus: g__Prevotella)
OTU: denovo387981
(Genus: g__)
OTU: denovo85961
(Genus: g__)
OTU: denovo109160
(Genus: g__)
OTU: denovo135247
(Genus: g__)
OTU: denovo40182
(Genus: g__)
Mean_MY
species_count

Appendix 4



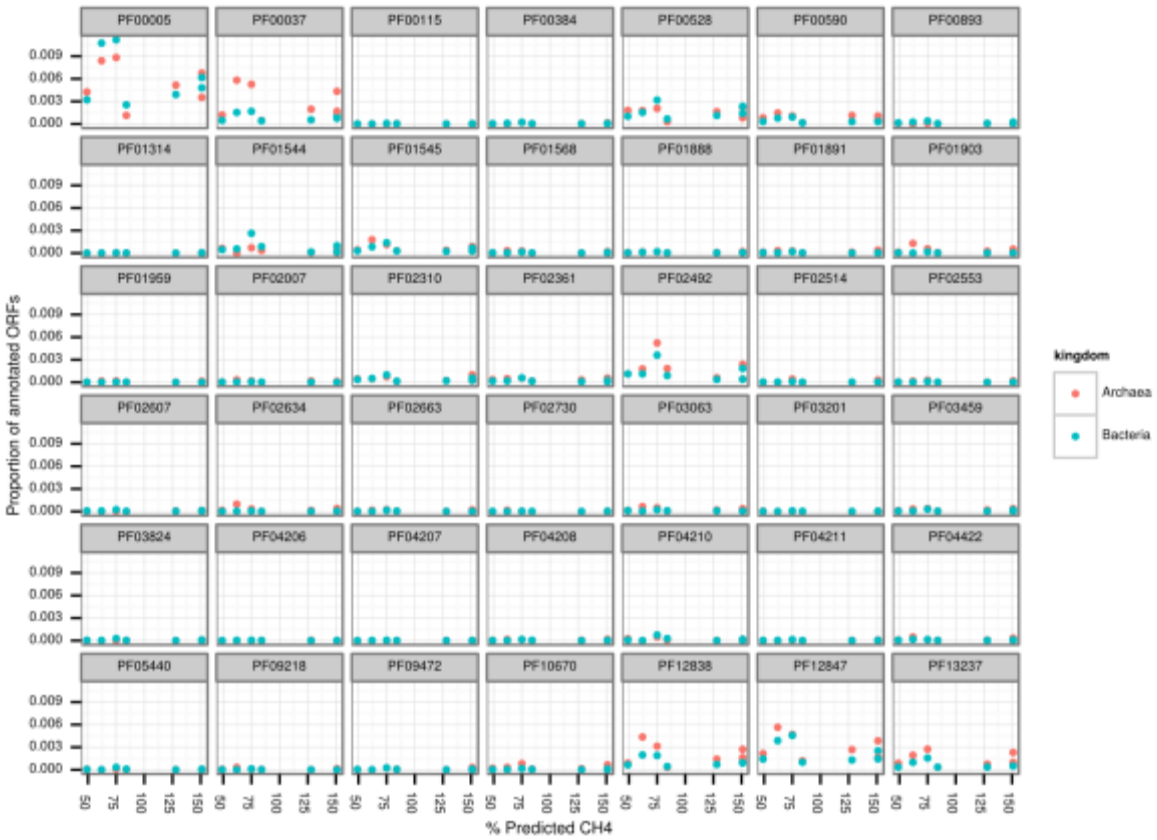
Relative abundance of microbial sequence data assigned to the family level for low and high methane animals

Appendix 5

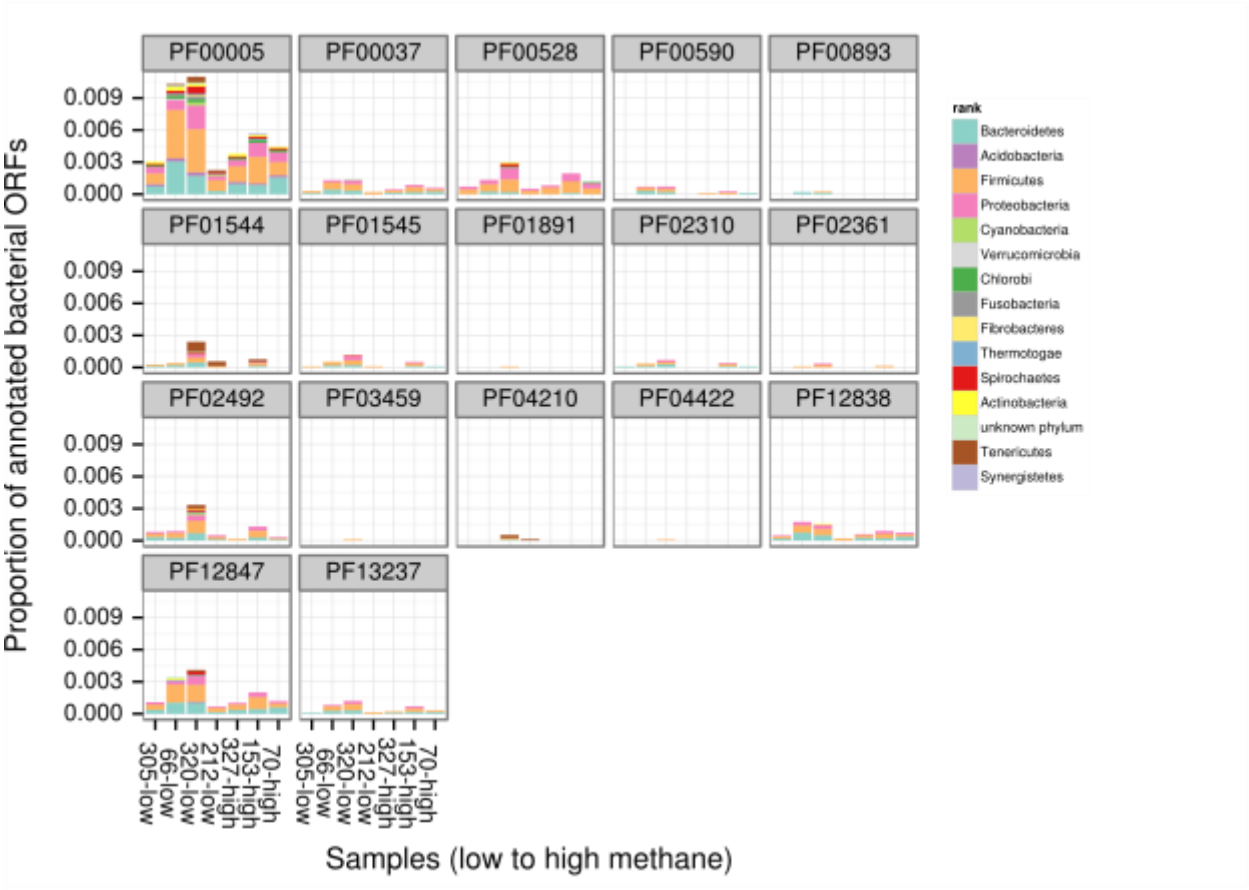


Plots showing the N50 values versus the contig size for all seven samples using the three assembly algorithms (velvet ■, metavelvet ▲and IDBA-UD ♦) at varying kmer settings. Samples 66-low, 70-high and 327-high were only assembled using IBDA-UD.

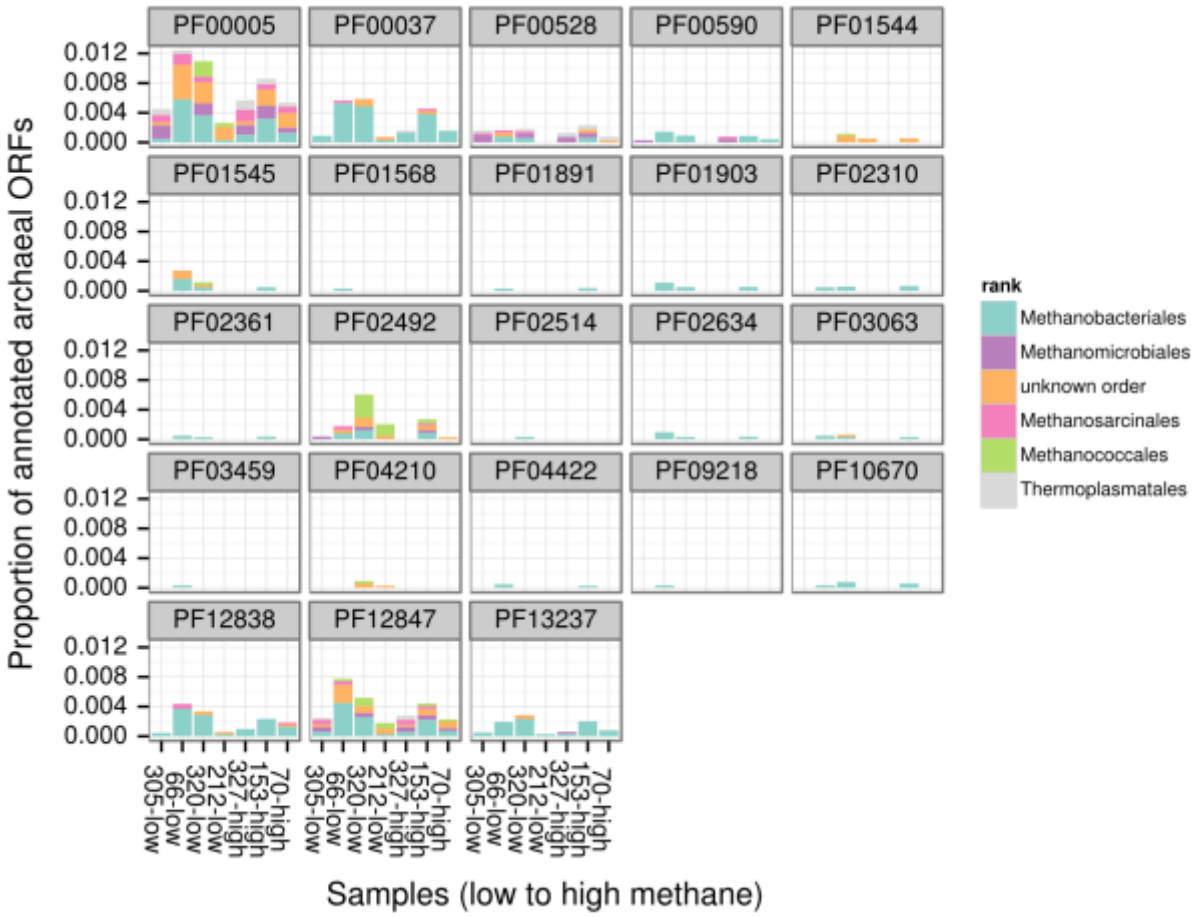
Appendix 6



Proportion of annotated metagenomic reads assigned to pFams for both bacterial and archaeal focused on uptake of transition state metals and methanogenesis pathways.



Proportion of annotated metagenomic reads assigned to pFams for bacterial phyla focused on significantly altered uptake families of transition state metals and methanogenesis pathways.



Proportion of annotated metagenomic reads assigned to pFams for archaeal order focused on significantly altered uptake families of transition state metals and methanogenesis pathways.

Description of families of genes associate with Pfam's focused on uptake of transition state metals and methanogenesis pathways.

Pfam_id	Family	Description
PF00005	ABC_tran	ABC transporter
PF00037	Fer4	4Fe-4S binding domain
PF00115	COX1	Cytochrome C and Quinol oxidase polypeptide I
PF00384	Molybdopterin	Molybdopterin oxidoreductase
PF00528	BPD_transp_1	Binding-protein-dependent transport system inner membrane component
PF00590	TP_methylase	Tetrapyrrole (Corrin/Porphyrin) Methylases
PF00893	Multi_Drug_Res	Small Multidrug Resistance protein
PF01314	AFOR_C	Aldehyde ferredoxin oxidoreductase, domains 2 & 3
PF01544	CorA	CorA-like Mg2+ transporter protein
PF01545	Cation_efflux	Cation efflux family
PF01568	Molydop_binding	Molydopterin dinucleotide binding domain
PF01888	CbiD	CbiD
PF01891	CbiM	Cobalt uptake substrate-specific transmembrane region
PF01903	CbiX	CbiX
PF01959	DHQS	3-dehydroquinate synthase (EC 4.6.1.3)
PF02007	MtrH	Tetrahydromethanopterin S-methyltransferase MtrH subunit
PF02310	B12-binding	B12 binding domain
PF02361	CbiQ	Cobalt transport protein
PF02492	cobW	CobW/HypB/UreG, nucleotide-binding domain
PF02514	CobN-Mg_chel	CobN/Magnesium Chelatase
PF02553	CbiN	Cobalt transport protein component CbiN
PF02607	B12-binding_2	B12 binding domain
PF02634	FdhD-NarQ	FdhD/NarQ family
PF02663	FmdE	FmdE, Molybdenum formylmethanofuran dehydrogenase operon
PF02730	AFOR_N	Aldehyde ferredoxin oxidoreductase, N-terminal domain
PF03063	Prismane	Prismane/CO dehydrogenase family
PF03201	HMD	H2-forming N5,N10-methylene-tetrahydromethanopterin dehydrogenase
PF03459	TOBE	TOBE domain
PF03824	NicO	High-affinity nickel-transport protein
PF04206	MtrE	Tetrahydromethanopterin S-methyltransferase, subunit E
PF04207	MtrD	Tetrahydromethanopterin S-methyltransferase, subunit D
PF04208	MtrA	Tetrahydromethanopterin S-methyltransferase, subunit A
PF04210	MtrG	Tetrahydromethanopterin S-methyltransferase, subunit G
PF04211	MtrC	Tetrahydromethanopterin S-methyltransferase, subunit C
PF04422	FrhB_FdhB_N	Coenzyme F420 hydrogenase/dehydrogenase, beta subunit N-term
PF05440	MtrB	Tetrahydromethanopterin S-methyltransferase subunit B
PF09218	DUF1959	Domain of unknown function (DUF1959)
PF09472	MtrF	Tetrahydromethanopterin S-methyltransferase, F subunit (MtrF)
PF10670	DUF4198	Domain of unknown function (DUF4198)
PF12838	Fer4_7	4Fe-4S dicluster domain
PF12847	Methyltransf_18	Methyltransferase domain
PF13237	Fer4_10	4Fe-4S dicluster domain
PF13620	CarboxypepD_reg	Carboxypeptidase regulatory-like domain

Appendix 7

Total RNA yields and quality measures for culture fermentation on Rhodes grass

Sample	RNA Conc. (ng/µl)	rRNA Ratio (23s / 16s)	RNA Integrity Number (RIN)
Rhodes T24_High_1	815.8	1.3	9.3
Rhodes T24_High_2	846.8	1.3	9.5
Rhodes T24_High_3	517.3	1.2	9.5
Rhodes T24_Low_4	1059.4	1.4	9.6
Rhodes T24_Low_5	540.9	1.6	9.9
Rhodes T24_Low_6	1463.7	1.7	9.8
Rhodes T30_High_1	1144.3	1.3	9.6
Rhodes T30_High_2	426.8	1.5	9.7
Rhodes T30_High_3	1177.6	1.5	9.6
Rhodes T30_Low_4	1371.9	1.5	9.6
Rhodes T30_Low_5	1025.7	1.6	9.7
Rhodes T30_Low_6	1200.1	1.5	9.5

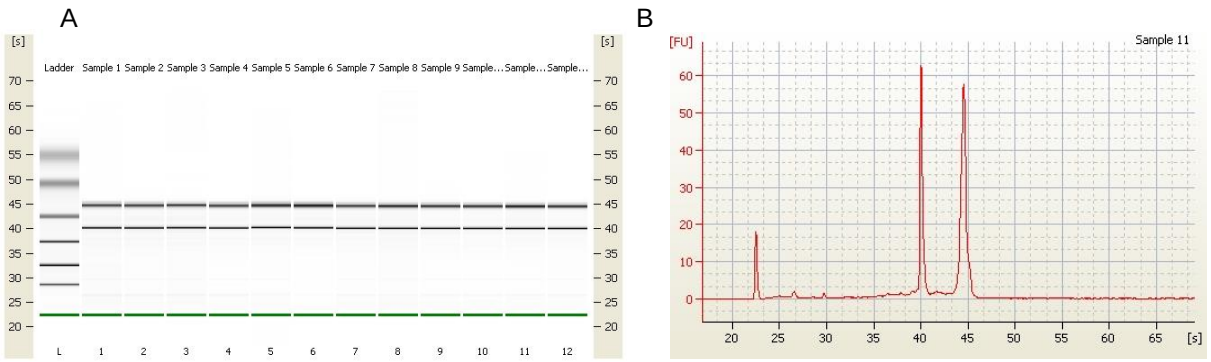
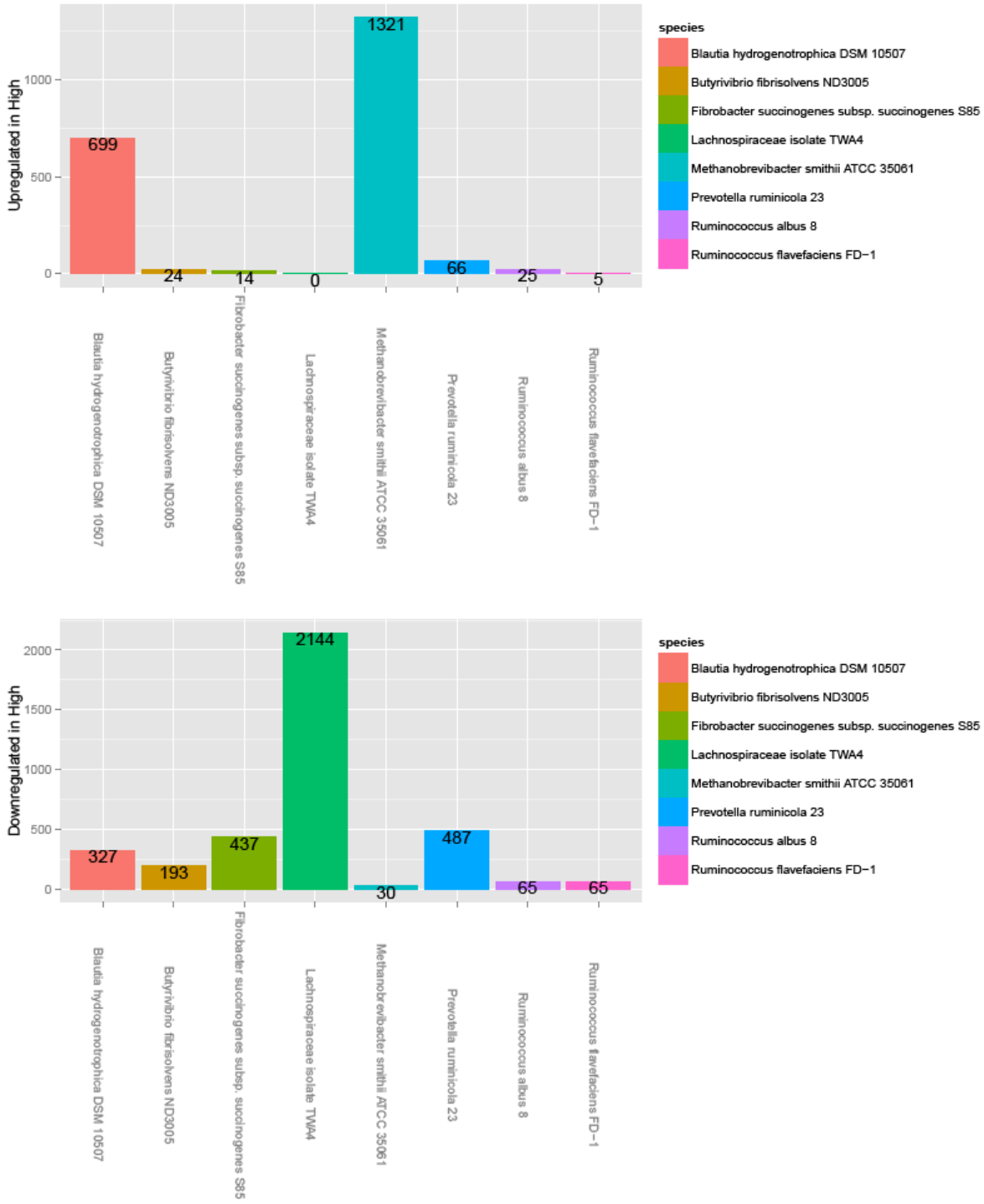


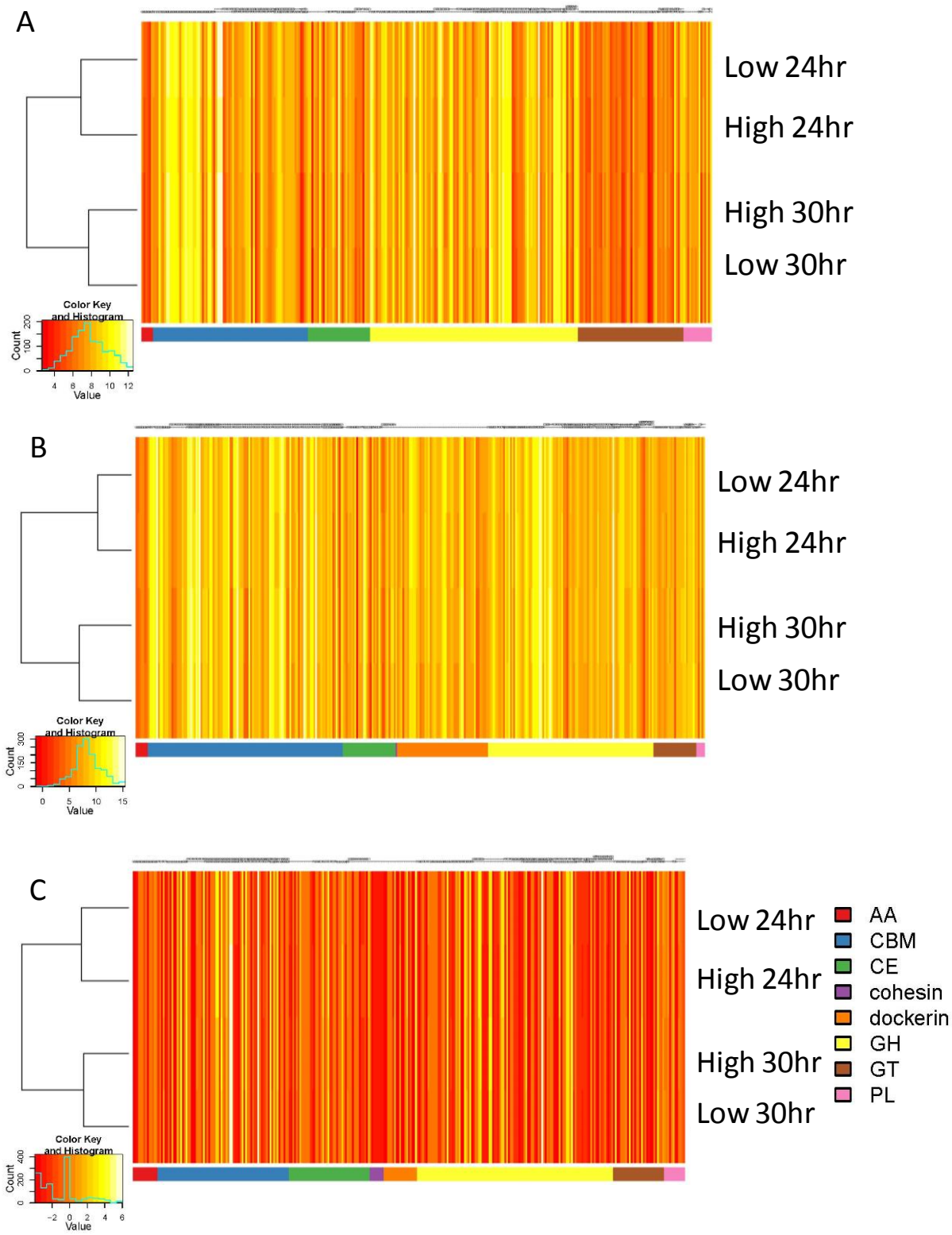
Figure: (A) Bioanalyzer gel image of total RNA extracted from 24 hr (samples 1-6) and 30 hr (samples 7-12) for high and low methane fementations. (B) Bioanalyzer chromatogram of a total RNA sample peaks indicate internal standard, 16S and 23S rRNA (left to right respectively).

Appendix 8



Counts for up regulated gene transcripts for high (upper panel) and low methane cultures (lower panel) assigned to each individual species.

Appendix 9



Clustered heat maps for the abundance of transcripts mapped to each CAZy gene for A) *F. succinogenes,* B) *R. albus* and C) *R. flavefaciens* are represented in columns (lighter colours represent up regulation) for low and high methame cultures at 24 and 30 hours. Auxiliary activities (AA), carbohydrate binding modules (CBM), carbohydrate esterases (CE), cohesion, dockerin, glycoside hydrolases (GH), glycosyl transferases (GT) and polysaccharide lyases (PL).

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW & Noble WS (2009) MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202-W208.

Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bryant MP (1972) Commentary on the Hungate technique for culture of anaerobic bacteria. *Am J Clin Nutr* **25**: 1324-1328.

Caporaso JG, Kuczynski J, Stombaugh J*, et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Chalupa W (1977) Manipulating Rumen Fermentation. *J Anim Sci* **45**: 585-599.

Chessel D, Dufour AB & Thioulouse J (2004) The ade4 package - I : One-table methods. *R News* **4**: 5-10.

Cord-Ruwisch R, Seitz H-Jr & Conrad R (1988) The capacity of hydrogenotrophic anaerobic bacteria to compete for traces of hydrogen depends on the redox potential of the terminal electron acceptor. *Arch Microbiol* **149**: 350-357.

Demeyer DI, Fiedler D & DeGraeve KG (1996) Attempted induction of reductive acetogenesis into the rumen fermentation in vitro. *Reprod Nutr Dev* **36**: 233-240.

Denman SE, Tomkins NW & McSweeney CS (2007) Quantitation and diversity analysis of ruminal methanogenic populations in response to the antimethanogenic compound bromochloromethane. *FEMS Microbiol Ecol* **62**: 313-322.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P & Andersen GL (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069-5072.

Di Rienzi SC, Sharon I, Wrighton KC*, et al.* (2013) The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *eLife* **2**: e01102.

Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G & Thauer RK (2006) The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. *J Bacteriol* **188**: 642-658.

Gagen EJ, Wang J, Padmanabha J*, et al.* (2014) Investigation of a new acetogen isolated from an enrichment of the tammar wallaby forestomach. *BMC Microbiol* **14**: 314.

Hegarty RS, Goopy JP, Herd RM & McCorkell B (2007) Cattle selected for lower residual feed intake have reduced daily methane production. *J Anim Sci* **85**: 1479-1486.

Huang Y, Gilna P & Li W (2009) Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25**: 1338-1340.

Hungate RE (1969) Chapter IV A Roll Tube Method for Cultivation of Strict Anaerobes. *Methods in Microbiology,* Vol. Volume 3, Part B (Norris JR & Ribbons DW, eds.), p.^pp. 117-132. Academic Press.

Immig I, Demeyer D, Fiedler D, Van Nevel C & Mbanzamihigo L (1996) Attempts to induce reductive acetogenesis into a sheep rumen. *Arch Tierernahr* **49**: 363-370.

Janssen PH & Kirs M (2008) Structure of the archaeal community of the rumen. *Appl Environ Microbiol* **74**: 3619-3625.

Lassey KR, Ulyatt MJ, Martin RJ, Walker CF & David Shelton I (1997) Methane emissions measured directly from grazing livestock in New Zealand. *Atmos Environ* **31**: 2905-2914.

Leahy SC, Kelly WJ, Altermann E*, et al.* (2010) The Genome Sequence of the Rumen Methanogen <italic>Methanobrevibacter ruminantium</italic> Reveals New Possibilities for Controlling Ruminant Methane Emissions. *PLoS ONE* **5**: e8926.

Lopez S, McIntosh E, Wallace RJ & Newbold CJ (1999) Effect of adding acetogenic bacteria on methane production by mixed rumen microorganisms. *Anim Feed Sci Technol* **78**: 1-9.

Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Ludwig W, Strunk O, Westram R*, et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363-1371.

McMurdie PJ & Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

Mitsumori M, Shinkai T, Takenaka A, Enishi O, Higuchi K, Kobayashi Y, Nonaka I, Asanuma N, Denman SE & McSweeney CS (2011) Responses in digestion, rumen fermentation and microbial populations to inhibition of methane formation by a halogenated methane analogue. *Br J Nutr* 1-10.

Nkrumah JD, Okine EK, Mathison GW, Schmid K, Li C, Basarab JA, Price MA, Wang Z & Moore SS (2006) Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J Anim Sci* **84**: 145-153.

Nollet L, Mbanzamihigo L, Demeyer D & Verstraete W (1998) Effect of the addition of Peptostreptococcus productus ATCC 35244 on reductive acetogenesis in the ruminal ecosystem after inhibition of methanogenesis by cell-free supernatant of Lactobacillus plantarum 80. *Anim Feed Sci Technol* **71**: 49-66.

Playne MJ (1985) Determination of ethanol, volatile fatty acids, lactic and succinic acids in fermentation liquids by gas chromatography. *J Sci Food Agric* **36**: 638-644.

Pope PB, Smith W, Denman SE, Tringe SG, Barry K, Hugenholtz P, McSweeney CS, McHardy AC & Morrison M (2011) Isolation of Succinivibrionaceae implicated in low methane emissions from Tammar wallabies. *Science* **333**: 646-648.

Pope PB, Denman SE, Jones M*, et al.* (2010) Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. *Proc Natl Acad Sci U S A* **107**: 14793-14798.

Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, Flint HJ & Louis P (2014) Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* **8**: 1323-1335.

Rho M, Tang H & Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* **38**: e191.

Rooke JA, Wallace RJ, Duthie CA, McKain N, de Souza SM, Hyslop JJ, Ross DW, Waterhouse T & Roehe R (2014) Hydrogen and methane emissions from beef cattle and their rumen microbial community vary with diet, time after feeding and genotype. *Br J Nutr* 1-10.

Rost B (1999) Twilight zone of protein sequence alignments. *Protein Engineering* **12**: 85-94.

Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O & Huttenhower C (2012) Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* **9**: 811-814.

Shi W, Moon CD, Leahy SC*, et al.* (2014) Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. *Genome Res* **24**: 1517-1525.

Thauer RK (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144 (Pt 9)**: 2377-2406.

Watanabe Y, Suzuki R, Koike S, Nagashima K, Mochizuki M, Forster RJ & Kobayashi Y (2010) In vitro evaluation of cashew nut shell liquid as a methane-inhibiting and propionate-enhancing agent for ruminants. *J Dairy Sci* **93**: 5258-5267.

Wood JM, Kennedy FS & Wolfe RS (1968) The reaction of multihalogenated hydrocarbons with free and bound reduced vitamin B 12. *Biochemistry* **7**: 1707-1713.

Wood JM, Moura I, Moura JJ, Santos MH, Xavier AV, LeGall J & Scandellari M (1982) Role of vitamin B12 in methyl transfer for methane biosynthesis by Methanosarcina barkeri. *Science* **216**: 303-305.

Yu Z & Morrison M (2004) Improved extraction of PCR-quality community DNA from digesta and fecal samples. *BioTechniques* **36**: 808-812.

Zheng Y, Kahnt J, Kwon IH, Mackie RI & Thauer RK (2014) Hydrogen formation and its regulation in Ruminococcus albus: involvement of an electron-bifurcating [FeFe]-hydrogenase, of a non-electron-bifurcating [FeFe]-hydrogenase, and of a putative hydrogen-sensing [FeFe]-hydrogenase. *J Bacteriol* **196**: 3840-3852.