Accepted Manuscript

Title: Investigation of the koala (*Phascolarctos cinereus*) hindgut microbiome via 16S pyrosequencing

Author: Christopher J. Barker Amber Gillett Adam Polkinghorne Peter Timms

PII: S0378-1135(13)00443-4
DOI: http://dx.doi.org/doi:10.1016/j.vetmic.2013.08.025
Reference: VETMIC 6338

To appear in: *VETMIC*

Received date: 14-6-2013
Revised date: 29-8-2013
Accepted date: 30-8-2013

Please cite this article as: Barker, C.J., Gillett, A., Polkinghorne, A., Timms, P., Investigation of the koala (*Phascolarctos cinereus*) hindgut microbiome via 16S pyrosequencing, *Veterinary Microbiology* (2013), http://dx.doi.org/10.1016/j.vetmic.2013.08.025

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

1 **Title:**

2 **Investigation of the koala (*Phascolarctos cinereus*) hindgut microbiome via 16S**
3 **pyrosequencing**

4

5 **Authors:**

6 **Christopher J. Barker[1], Amber Gillett[2], Adam Polkinghorne[1], and Peter Timms[1]**

7

8 **Author Affiliations:**

9 [1] Institute of Health and Biomedical Innovation (IHBI), 60 Musk Avenue, Kelvin Grove,
10 Queensland, 4059, Australia

11 [2] Australia Zoo Wildlife Hospital, Beerwah, Queensland, 4519, Australia

12

13 **Contact Details for Corresponding Author:**

14 **Postal Address:** Institute of Health and Biomedical Innovation (IHBI), 60 Musk Avenue, Kelvin
15 Grove, Queensland, 4059, Australia

16 **Phone no.:** +61 7 3138 6199

17 **Email:** p.timms@qut.edu.au

18

19

20

Accepted Manuscript

ACCEPTED MANUSCRIPT

1

2 **Abstract**

3 As a dietary source, the foliage of *Eucalyptus* spp. is low in available protein and carbohydrate
4 while containing polyphenolic compounds that interfere with enzymatic digestion. To overcome
5 this, the koala (*Phascolarctos cinereus*) has evolved a range of anatomical and physiological
6 adaptations to assist with digestion and absorption of nutrients from this food source. Microbial
7 fermentation of partially digested eucalyptus leaves is thought to be critical in this process,
8 however, little is known about the composition and diversity of microorganisms that are associated
9 with digestive health in this native species. In this study, we performed 16S rRNA gene
10 pyrosequencing of caecum, colon and faecal pellet samples from two wild, free ranging,
11 Queensland koalas. Our results reveal a highly complex and diverse ecosystem with considerable
12 intra-individual variation. Although samples were dominated by sequences from the Bacteroidetes
13 and Firmicutes phyla there was considerable variation at the genus level. This study is the first non-
14 culture based microbiota analysis, using 454-amplicon pyrosequencing, and provides preliminary
15 data to expand our understanding of the koala hindgut.

16

17 **Keywords**

18 16S; Koala; Microbiome; gut microbiome; 16SrRNA sequencing

19

20 **Introduction**

21 As the last surviving member of a family of prehistoric marsupials (*Phascolarctidae*), the koala
22 (*Phascolarctos cinereus*) is a biologically fascinating Australian arboreal marsupial. Part of the
23 uniqueness of this Australian marsupial comes from its unusual reliance on consumption of foliage
24 from species of *Eucalyptus* (Cork and Sanson, 1990). To other animals, the leaves of *Eucalyptus*
25 spp. are a nutrient-poor and toxic source of food as they are low in available proteins and
26 carbohydrates while having high concentrations of phenolic compounds (Cork and Foley, 1997).

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 The koala has evolved a range of anatomical and physiological adaptations allowing it to overcome
2 challenges associated with the consumption of *Eucalyptus* spp. leaf. Koalas possess a unique
3 gastrointestinal tract, comprised of a foregut and a greatly enlarged hindgut, including the proximal
4 colon and caecum (Cork and Sanson, 1990). The ratio of caecum size to body size of the koala is
5 the largest of any known mammal, with the mean retention time of matter passing through the
6 gastrointestinal tract thought to be one of the longest recorded for mammals (Krockenberger and
7 Hume, 2007).

8 The koala's hindgut is thought to be the primary location of microbial fermentation, an essential
9 process required for further breakdown of eucalypt leaf cells. The exact identity and role of
10 microorganisms in this process is currently unclear. However, the important role these organisms
11 play in koala digestive health is highlighted by the phenomenon of pap feeding. Pap is a thick faecal
12 paste made in the caecum of female koalas, and thought to be consumed by their young to assist in
13 inoculation of the gastrointestinal tract microflora (Osawa et al., 1993).

14 In an effort to characterise the role of microorganisms in the gastrointestinal tract of koalas, culture-
15 based studies have revealed a mixture of previously described and novel microorganisms (Osawa et
16 al., 1995; Osawa, 1990). Interestingly, a repeating theme of these studies was the identification of
17 tannin-degrading microorganisms including *Streptococcus bovis* (Osawa, 1990) and a new species
18 of bacteria, *Lonepinella koalarum*, within the family *Pasteurellaceae* (Osawa et al., 1995). These
19 bacteria were found in different locations of the koala gastrointestinal tract, including the bacterial
20 layer of the caecal wall, and in faecal specimens (Goel et al., 2005).

21 Despite these findings, the use of culture-based technologies alone has meant that attempts to
22 characterise the diversity and role of the koala gastrointestinal microbiota has been very limited.
23 Human studies have shown that 80% or more of phylotypes inhabiting the gut remain uncultured
24 (Zoetendal et al., 2008), suggesting that there are a significant number of potentially important
25 microorganisms that are yet to be identified in the koala gastrointestinal tract. Culture-independent
26 16S rRNA gene pyrosequencing has revolutionised our knowledge of the diversity and relationships

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 between gastrointestinal tract microflora in humans and animals. In this study we have utilised
2 high-throughput 16S rRNA gene pyrosequencing to perform a preliminary characterisation of the
3 diversity and distribution of bacteria within the caecum, colon and faecal pellet of two wild koalas.
4 Our results show a diverse and complex microbiota in the koala hindgut, and will serve as a basis
5 for further studies to understand the impact different microbial communities have on koala health
6 and disease.
7

8 **Methods**

9 ***Sample collection***

10 Caecum, colon and faecal pellet samples were obtained from two wild, free ranging Queensland
11 koalas (designated K1 and K2) admitted to the Australia Zoo Wildlife Hospital, Beerwah,
12 Queensland, Australia. Animal K1, a healthy 7 year old male, was admitted with severe head
13 trauma after being hit by a car. Physical evaluation resulted in the assignment of a body condition
14 score (BCS) of 6 out of 10 (considered within a 'healthy range' for wild koalas). Animal K2, an 8
15 year old female, was admitted to the hospital after being found severely emaciated with a BCS of 2
16 out of 10. Subsequent evaluation resulted in the diagnosis of cystitis (presumed to be due to a
17 chlamydial infection) and severe cachexia of unknown (or undetermined) cause. Both animals were
18 euthanized on welfare grounds and samples provided for microbiome analysis.
19 Caecum, colon and faecal pellet samples were obtained during post mortem shortly after euthanasia.
20 Caecal content was collected towards the blind end of the caecum, colonic content was collected
21 approximately 10cm preceding the rectum and a faecal pellet was milked from the rectum. Samples
22 were immediately stored at -20°C until processed and sequenced (no longer than 8 weeks after
23 collection).
24

25 ***DNA extraction***

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 Samples were thawed at room temperature and then processed for DNA extraction using a QIAamp
2 DNA Stool Mini Kit (QIAGEN, Valencia, CA). Briefly, 1.4 mL of ASL lysis buffer was added to
3 approximately 200 mg of sample in a tube containing ceramic beads for bead beading. The samples
4 were then homogenised using a bead-beater, for 5 minutes in one minute intervals. Subsequent steps
5 were followed according to the manufacturer's protocol for 'DNA isolation from stool for pathogen
6 detection'. DNA concentration and purity were determined using a NanoDrop spectrophotometer
7 (ND-1000, Nanodrop Technologies, Wilmington) at wavelengths of 230, 260 and 280 nm. DNA
8 preparations were standardised at 10ng/µL for each sample in accordance with sample preparation
9 guidelines for 454 pyrosequencing at the Australian Genome Research Facility.
10

11 ***Amplification and pyrosequencing of variable regions 1-3 of the 16S rRNA gene***
12 PCR amplification of variable regions 1-3 (V1-V3) from the 16S rRNA gene and subsequent 454
13 pyrosequencing were performed commercially by the Australian Genome Research Facility
14 (Brisbane, Australia). Briefly, PCR amplicons were generated using eubacterial 27F and 519R
15 primers along with AmpliTaq Gold 360 mastermix (Life Technologies, Australia). Thermal cycling
16 consisted of 3 minutes at 94°C, and 34 cycles consisting of 45 seconds at 94°C, 1 minute at 50°C,
17 and 1 minute at 72°C, followed by 7 minutes at 72°C. Resulting amplicons were measured by
18 fluorometry, normalised, measured by qPCR, normalised a second time, and then pooled in
19 equimolar ratios. This amplicon pool was then run on the GS-FLX platform using XLR70
20 chemistry (Roche, Australia).
21

22 ***Quality filtering of pyrosequencing reads***
23 Initially, multiple raw sequence files (sff files) were combined into a single file containing all
24 sequence data, and then fasta and quality files were extracted. This was done using Roche's 454
25 proprietary analysis software (sff tools). Barcodes were removed from the fasta sequences using the
26 Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso et al., 2010b) software package's

Accepted Manuscript

ACCEPTED MANUSCRIPT

split_library script. During this process, sequences were also filtered using quality parameters. Sequences >600 nucleotides in length or <200 nucleotides in length were discarded. Additionally, sequences with a mismatch in the primer sequence, an average quality score <25, a homopolymer run >6, or more than 6 ambiguous bases were also discarded. Resulting sequences had reverse primers removed using the truncate_reverse_primer script in the QIIME software package. The resulting fasta file was de-multiplexed using the split_fasta_on_sample_ids script in the QIIME software package. Individual fasta files were then multiplexed into caecum specific, colon specific, faecal pellet specific or a single fasta file containing all sequences, using Roche's sff tools software package. These files were used for downstream analysis via the QIIME software pipeline. During pre-processing of sequence files, quality control analysis of all samples was performed using the PRINSEQ-lite script (Schmieder and Edwards, 2011). Fasta and quality files were used to generate quality distribution graphs (data not shown) and sample statistics (average length, GC content, and occurrence of N) for sequences from each sample (Supplementary Table 1).

Analysis of quality filtered V1-V3 pyrosequencing reads

Pyrosequencing reads within a 97% similarity threshold were identified with UCLUST (Edgar, 2010). The most abundant sequence within each new or reference OTU was chosen as the OTU representative sequence. The PyNAST alignment algorithm (Caporaso et al., 2010a) was used to align OTU representative sequences against the Greengenes core set alignment (DeSantis et al., 2006) with a minimum alignment length of 365 and a minimum identity of 75% (177 OTU representative sequences failed to align). The OTU representative sequences were screened for chimeras with identified sequences being excluded from the downstream analysis. The PyNAST alignment of OTU representative sequences was used to infer a phylogenetic tree using FastTree software (Caporaso et al., 2010b). Taxonomy was assigned to each OTU representative sequence using a BLAST method along with the SILVA (108 release) reference database (Quast et al., 2013). OTU tables had chimeras and singletons excluded. An OTU table only containing OTUs with a

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 relative abundance >0.1% (most abundant OTU table) was created by calculating the relative
2 abundance for each sample and removing any OTUs with an abundance <0.1%. OTU tables were
3 summarised in QIIME to generate the relative abundance of taxonomic labels to the phylum and
4 genus level. Venn diagrams were generated using VENNY
5 (http://bioinfogp.cnb.csic.es/tools/venny/index.html). Heatmaps to represent the most abundant
6 OTUs from the caecum, colon and faecal pellet samples were constructed in iTOL (Supplementary
7 Figure 1.) (Letunic and Bork, 2011).
8

9 ***Sequence diversity analysis***

10 Microbial diversity was evaluated within samples (α diversity) or between samples (β diversity)
11 using QIIME. Alpha diversity was measured with the Chao1 (richness), Phylogenetic Diversity
12 (branch length-based diversity), Shannon diversity (OTU-based diversity), Shannon equitability,
13 and observed species metrics. Rarefaction, to a sub-sampling depth (determined by the minimum
14 number of sequences in a sample) of 6,736 reads/sample, was performed on all samples. Alpha
15 diversity means and corresponding error bars representing standard error of the mean (SEM) were
16 calculated using GraphPad PRISM software (GraphPad Software, CA, USA).

17 Beta diversity was evaluated with UniFrac, a community dissimilarity metric based on the fraction
18 of unique branch length observed in pairs of communities in a common phylogenetic tree
19 (Lozupone et al., 2007). Rarefaction, to a sub-sampling depth (determined by the minimum number
20 of sequences in a sample) of 7485 reads/sample, was performed on all samples. The phylogenetic
21 distance (UniFrac distance) is calculated as the fraction of unshared branch lengths between the pair
22 of communities. Unweighted UniFrac distances compare microbial communities within a
23 phylogenetic context based on the presence/absence of members, while weighted UniFrac also
24 incorporates relative abundance information (Lozupone et al., 2007). UniFrac-based Principal
25 Coordinate Analysis (PCoA) was used to map the UniFrac distance matrix onto a set of orthogonal
26 axes capturing the greatest amount of variation in all the samples tested. Distances between samples

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 on a PCoA plot reflect the corresponding dissimilarities in their community membership
2 (unweighted UniFrac) or community structure (weighted UniFrac).
3
4 **Results**
5 ***Characterisation of samples***
6 Caecum, colon and faecal pellet samples were collected from two wild, free ranging koalas (i.e.
7 surviving on a natural eucalypt diet). Animal K1 was admitted with a serious blunt force trauma
8 injury, but was otherwise considered healthy. Animal K2 was admitted after being found extremely
9 emaciated. It was subsequently diagnosed with severe cachexia of unknown (or undetermined)
10 cause and chlamydial cystitis. Both animals were euthanized on welfare grounds, due to the severity
11 of disease and injuries present.
12
13 ***Analysis of pyrosequencing data***
14 A total of 81,608 (47,345 for animal K1 and 34,263 for animal K2) sequences across all samples
15 were generated through pyrosequencing. The average length of these sequences was approximately
16 521 nucleotides, the GC content ranged from 26-67%, and there were 10,723 sequences containing
17 an ambiguous base call (N) (the maximum percentage of N's per sequence ranged from 9-51%).
18 Demultiplexing and quality trimming of the data set resulted in a total of 72,368 (41,801 for animal
19 K1 and 30,567 for animal K2) sequences with an average length of approximately 485 nucleotides,
20 an average GC content of approximately 52% and 2,563 containing an ambiguous base call (N) (the
21 maximum percentage of N's per sequence being 1%) (Supplementary Table 1).
22 The resulting data set was subsequently clustered into operational taxonomic units (OTUs) using
23 the algorithm UCLUST. From a total of 5277 OTUs, 1320 were identified as potential chimeras and
24 a further 2102 sequences were singletons. Removal of chimeras and singletons left 1855 OTUs for
25 downstream analysis (Supplementary Table 2).
26

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 ***Composition of hindgut microbial communities from a healthy koala (animal K1)***

2 Assignment of consensus taxonomy resulted in identification of 11 different phyla being
3 represented across the koala caecum, colon and faecal pellet samples. Bacteroidetes and Firmicutes
4 were found to be the most abundant pyhla with Bacteroidetes abundance ranging from 11.0% to
5 31.4% and Firmicutes abundance ranging from 40.8% to 64.8%. The abundance of Firmicutes was
6 lowest in the caecum and highest in the colon. In the caecum the ratio of Firmicutes:Bacteroidetes is
7 close to 1 (1.3:1) whereas in the colon and faecal pellet the ratio is significantly increased (5.9:1 and
8 3.18:1 respectively) (Supplementary Table 3). Cyanobacteria, Fusobacteria, Planctomycetes,
9 Proteobacteria and Synergistetes are also predominant phyla with abundances greater than 1%
10 across the different sample types. Cyanobacteria, Fusobacteria, Planctomycetes, and Proteobacteria
11 all have relative abundances that do not appear to change across the hindgut. In contrast, there is a
12 significant decrease in the abundance of Synergistetes in the faecal pellet sample (6.1%) compared
13 to the caecum (17.0%) and colon (14.8%) (Figure 1A and Supplementary Table 3).

14 Due to the sequencing depth being limited to approximately 12,000 reads per sample, initial genus
15 level analysis focused on the most abundant OTUs, defined as OTUs with a relative abundance
16 greater than 0.1% in a given sample. These OTUs were used for the generation of Figure 1B,
17 Supplementary Figure 1 and Supplementary Table 4. Genus level OTU classification resulted in
18 identification of 26 different genera having a relative abundance greater than 0.1% in any given
19 sample. *Alistipes* (12.7%), *Cloacibacillus* (10.9%), *Incertae sedis* (17.5%), *Ruminococcus* (23.4%),
20 and OTUs classified as 'Uncultured' (13.4%) were found to be the predominant OTUs across the
21 hindgut of the healthy animal. *Incertae sedis* is a term used to define a taxonomic group where its
22 broader relationships are unknown or undefined. Further investigation of the OTUs (from the OTU
23 table) classified as *Incertae sedis* found that they were either assigned to the order Clostridales or
24 Erysipelotrichales. OTUs classified to the order Clostridales accounted for over 99% of *Incertae*
25 *sedis* OTUs (data not shown). For the remainder of the manuscript *Incertae sedis* will be defined as
26 "*Incertae sedis* (unknown Clostridales)". OTUs classified as 'Uncultured' were further investigated

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 and found to contain sequences that were classified into five different phyla (Actinobacteria,
2 Chloroflexi, Firmicutes, Proteobacteria, and Synergistetes) and seven different orders
3 (Coriobacteriales, Anaerolineales, Clostridiales, Erysipelotrichales, Burkholderiales, Pasteurellales,
4 and Synergistales). However, the majority (>75%) were classified to the order Clostridiales
5 (Supplementary Table 5).
6 The predominant genera (>10% relative abundance) of the caecum was *Alistipes* (23.4%),
7 *Cloacibacillus* (15.9%), *Ruminococcus* (18.6%), and OTUs classified as 'Uncultured' (13.7%). The
8 predominant genera of the colon were *Cloacibacillus* (12.1%), *Incertae sedis* (unknown
9 Clostridales) (35.1%), *Ruminococcus* (16.6%), and OTUs classified as 'Uncultured' (14.1%).
10 Finally, the predominant genera of the faecal pellet were *Bacteroides* (10.1%), *Incertae sedis*
11 (unknown Clostridales) (13.9%), *Ruminococcus* (35.1%), and OTUs classified as 'Uncultured'
12 (12.3%) (Figure 1B and Supplementary Table 4). These results demonstrate a changing microbiome
13 across the hindgut of a healthy koala.
14
15 ***Comparative analysis of OTU composition across different sites of a healthy hindgut (animal K1)***
16 Comparative analysis of OTUs identified in the caecum, colon and faecal pellet of the healthy koala
17 (K1), found that the majority of OTUs were shared across all three gastrointestinal sites. Of the 54
18 different genera (including OTUs with <0.1% relative abundance) identified in the gastrointestinal
19 tract of animal K1, 26 were found to be shared across the caecum, colon and faecal pellet, five were
20 found to be unique to the caecum, eight unique to the colon, and seven unique to the faecal pellet.
21 Furthermore, two OTUs were found to be shared exclusively between the caecum and the colon,
22 three were shared exclusively between the colon and the faecal pellet and four were shared
23 exclusively between the caecum and faecal pellet (Figure 5).
24 In regards to abundance, OTUs unique to the caecum, colon, and faecal pellet, along with OTUs
25 shared exclusively between the different sites were found to make up a very small minority of the
26 microbiome (0.4-1%). The vast majority (99%) of OTUs are shared across the different sites,

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 however, the abundance of OTUs changes across the gastrointestinal tract. *Alistipes* is the dominant
2 genus of the caecum (21.9%) yet its abundance in the colon (4.7%) and faecal pellet (9.5%) is
3 significantly diminished. Relative abundance of the *Cloacibacillus* genus appears to gradually
4 decline across the gastrointestinal tract (caecum; 14.7%, colon; 11.6, faecal pellet; 4.2%). *Incertae*
5 *sedis* (unknown Clostridales) is the most dominant genera in the colon (33.8%), however it is
6 significantly lower in the caecum (4.2%) and faecal pellet (14.2%). In contrast, the *Ruminococcus*
7 genus makes up over a third of the faecal pellet microbiome (33.4%), and yet it is much less
8 abundant in the colon (16.5%) and caecum (18.2%). There are also examples of dominant genera
9 maintaining a similar level of abundance across the different sites. *Bacteroides*, *Bilophila*,
10 *Fusobacterium*, and OTUs classified as 'Uncultured' appear to have a consistent relative abundance
11 across the three different gastrointestinal tract sites (Figure 5).
12

13 ***Richness and diversity of OTUs identified from the koala hindgut***

14 Rarefaction curves indicate reasonable depth of sequence coverage, with curves beginning to level
15 off after 6000 reads (Figure 2A). Rank abundance curves, generated from complete OTU lists,
16 appear relatively steep, indicating low evenness across all samples with high ranking species having
17 much greater abundances than the low ranking species (Figure 3A-C). Evenness for the caecum and
18 pellet appear to be comparable, however colon samples show less evenness (Figure 3D). The alpha
19 diversity metrics, Chao1 (measuring richness), Phylogenetic diversity, Shannon equitability
20 (measuring evenness) and Shannon Diversity (measuring diversity whilst accounting for abundance
21 and evenness) showed no significant differences between different sample types, suggesting little
22 variability across the gastrointestinal tract samples (Figure 2B-E).

23 In unweighted UniFrac PCoA, the first principle coordinate (PC1) captured 52.8% of variation with
24 the second principle coordinate (PC2) capturing 17.3% of variation (Figure 4A & C). Although
25 there were not enough samples to undertake thorough statistical analysis, samples from animal K1
26 clustered together tightly. In contrast, samples from animal K2 were found to cluster away from the

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 K1 samples, suggesting that the microbiomes of the two animals are phylogenetically distinct. In
2 addition, samples from animal K2 did not cluster in relation to anatomical site, suggesting distinct
3 phylogenetic diversity across different anatomical sites of the hindgut. Samples from animal K1
4 showed tight clustering of the caecum, colon and faecal pellet, suggesting that phylogenetic
5 diversity of a healthy hindgut is comparable across the different sites. Both of these findings were
6 supported by the weighted UniFrac analysis where PC1 accounted for 59.6% variation and PC2 for
7 29.9% variation (Figure 4B & D).
8

9 ***Comparative analysis of hindgut OTU composition from a healthy (animal K1) versus diseased***
10 ***(animal K2) koala***

11 Animal K2 was severely emaciated and suffering from severe cachexia of unknown (or
12 undetermined) cause. Phylogenetic analysis of OTUs (with a relative abundance greater than 0.1%)
13 identified in the caecum, colon and faecal pellet demonstrate a clear difference in OTU presence
14 and abundance between the two animals (Supplementary Figure 1).

15 Comparative analysis revealed that from 60 caecal OTUs, classified to the genus level, 23 were
16 exclusive to animal K2, 11 were exclusive to animal K1 and 26 were shared (Supplementary Figure
17 2A). Of the 11 caecum OTUs that were exclusive to animal K1, there were three with a relative
18 abundance greater than 1% (*Bilophila, Cloacibacillus,* and *Fusobacterium*). The total abundance of
19 animal K1 caecum exclusive OTUs accounted for 7.7% of the total caecal microbiome. For animal
20 K2 caecum exclusive OTUs (23), there were only two with a relative abundance greater than 1%
21 (*Enterobacter* and *Candidatus Tammella*). OTUs found only in the caecum of animal K2 accounted
22 for 6.3% of the total caecal microbiome. Interestingly, this exclusive group of OTUs included
23 genera of known gastrointestinal pathogens (*Enterobacter* and *Escherichia-shigella*) at low levels
24 of abundance. Although *Escherichia* and *Shigella* are two separate genera the extremely high level
25 of genetic relatedness between the different species means that 16S classification databases are
26 unable to resolve which genus the 454 pyrosequencing reads fall into. Therefore, the genus level

Accepted Manuscript

ACCEPTED MANUSCRIPT

classification of OTUs is described as; '*Escherichia-shigella*'. Although the 26 shared OTUs make up the majority of the caecal microbiome (92-93%), there were distinct differences in relative abundance between the two animals. The relative abundance of *Alistipes*, *Cloacibacillus*, and *Ruminococcus* was much greater in animal K1, whereas the relative abundance of *Desulfovibrio* and *Parabacteroides* were much higher in animal K2 (Supplementary Table 6).

In the colon samples, 23 OTUs were shared between animals, 16 were exclusive to animal K1 and 19 exclusive to animal K2 (Supplementary Figure 2B). The OTUs exclusive to animal K1 make up 7.3% of the colon microbiome. Only two colon exclusive OTUs had a relative abundance greater than 1% (*Bilophila* and *Fusobacterium*). For animal K2, two colon exclusive OTUs were found to have an abundance greater than 1% (*Candidatus Tammella* and *Flavonifractor*), with the colon exclusive OTUs accounting for only 3.4% of the total microbiome. Interestingly, *Enterobacter* and *Escherichia-shigella* were not identified in the colon. In regard to the shared OTUs, similar to the caecum, the genera *Cloacibacillus* and *Ruminococcus* were found to be in greater abundance in animal K1, whereas the relative abundance of *Desulfovibrio* and *Parabacteroides* were higher in animal K2 (Supplementary Table 6).

In the faecal pellet samples, 26 OTUs were shared between the animals, 14 were exclusive to animal K1 and 16 exclusive to animal K2 (Supplementary Figure 2C). Exclusive OTUs make up 10.0% and 6.4% of the total faecal microbiomes for animals K1 and K2 respectively. *Bilophila*, *Cloacibacillus*, and *Fusobacterium* were found to be the most abundant (>1%), exclusive OTUs, in animal K1's faecal pellet. In contrast, only a single OTU, exclusive to animal K2 faecal pellet, was found to be greater than 1% (*Enterobacter*). In regards to OTUs shared between the two animals; *Alistipes*, *Bacteroides*, and *Ruminococcus* appear significantly more dominant in animal K1 whereas *Blautia*, *Incertae sedis* (unknown Clostridales), and *Parabacteroides* appear significantly more dominant in animal K2 (Supplementary Table 6).

Discussion

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 High throughput pyrosequencing of the 16S rRNA gene has revealed a diverse and complex
2 microbiota in the hindgut of the koala. Microbiota composition was found to vary across the
3 different anatomical sites and between the two different animals.
4 Analysis of microbial diversity found no significant differences when measuring the alpha diversity
5 of the different samples using several metrics (Phylogenetic diversity, Chao1, Equitability, and
6 Shannon) (Figure 2B-E) and a rank abundance analysis (Figure 3). Interestingly, UniFrac analysis
7 (beta diversity) found that samples from animal K1 clustered together tightly, suggesting
8 phylogenetic diversity is comparable across the different, hindgut, gastrointestinal sites of the
9 healthy koala (Figure 4). This is similar to findings from the human gastrointestinal tract (Stearns et
10 al., 2011) and the Brazilian Nelore steer (de Oliveira et al., 2013). In contrast, samples from animal
11 K2 clearly clustered away from the corresponding K1 samples and away from one another,
12 suggesting that the sick animal has a distinct phylogenetic diversity profile (Figure 4).
13 Taxonomic classification of OTUs led to the identification of 11 different phyla across the three
14 different anatomical sites. Only 0.3% of sequences were not classified to a known phylum. The
15 hindgut ecosystem of the koala is predominately colonised by two phyla, Bacteroidetes and
16 Firmicutes. These two phyla, which comprised 77% of our sequences, are also the most ubiquitous
17 phyla found across the vertebrate gut (Ley et al., 2008). The ratio of Firmicutes to Bacteroidetes
18 (FB ratio) has been associated with a number of different variables in humans, including age
19 (Mariat et al., 2009), obesity (Ley et al., 2006) and diet (De Filippo et al., 2010). Generally, a high
20 FB ratio has been associated with adults aged 20-45 years, obese people and a diet consisting of
21 animal protein and low levels of fibre (De Filippo et al., 2010; Ley et al., 2006; Mariat et al., 2009).
22 In the current study we found that the FB ratio changes considerably across the koala hindgut. In the
23 caecum the ratio is close to 1, whereas the ratio increases significantly in the colon (~6:1) and then
24 drops in the faecal pellet (~3:1) (Supplementary Table 3). Although the FB ratio has been found to
25 vary across different vertebrates, in the majority of species it is closer to 3 (Ley et al., 2008). Our
26 finding suggests that the bacterial fermentation process undertaken in the caecum may differ to that

ACCEPTED MANUSCRIPT

1 of the colon, with the caecum requiring a higher proportion of Bacteroidetes. Interestingly, a recent
2 study investigating the microbiota of the marsupial macropod (kangaroos and wallabies) also
3 identified a low FB ratio in the forestomach region where microbial fermentation of ingested plant
4 material occurs (Gulino et al., 2013).
5 At a deeper level of taxonomic classification, six genera were found to represent 85% of all
6 sequences. These genera include *Alistipes*, *Bacteroides*, *Cloacibacillus*, *Incertae sedis* (unknown
7 Clostridales), *Ruminococcus*, and OTUs classified as 'Uncultured' (Supplementary Table 5).
8 Although a variety of dominant genera were identified in the hindgut, there were significant
9 differences in the abundance of different microbial communities for the caecum, colon and faecal
10 pellet (Figure 1B, and 5 Supplementary Table 4 and Supplementary Figure 1). The caecum of
11 animal K1 (healthy) was dominated by *Alistipes* (23%), *Ruminococcus* (19%) and *Cloacibacillus*
12 (16%). The identification of *Ruminococcus* in the caecum is not surprising due to it containing
13 species associated with efficient degradation of plant cell walls (Flint et al., 2008). Unsurprisingly,
14 the highest abundance of *Ruminococcus* is found in the faecal pellet (which is indicative of the
15 lower section of the hind gut) where it can take advantage of dietary residue (plant cell wall) that
16 has evaded digestion in the upper gastrointestinal tract (Flint et al., 2008). *Cloacibacillus* is a
17 recently defined genus with a species isolated from the gastrointestinal tract of a pig, which is
18 capable of degrading mucin (Looft et al., 2012). However, little is known about this genus in
19 regards to its role in the digestive tract. Given the considerable abundance of this genus in both the
20 caecum and colon (12%), it may warrant further characterisation. In the lower hindgut (colon and
21 faecal pellet), there appears to be a considerable shift to the Firmicutes phylum (as observed
22 through the FB ratio), which is associated with a significant increase in the abundance of
23 *Ruminococcus* and other Firmicutes (OTUs classified as both *Incertae sedis* (unknown Clostridales)
24 and 'Uncultured'). Further analysis is required to clearly elucidate the dominant communities found
25 in the colon due to genus level classification not clearly resolving the Clostridiales identified at
26 these sites.

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 *Alistipes* is a recently defined genus created from the re-classification of *Bacteroides putredinis*
2 (Rautio et al., 2003). More recently it has been associated with an important role in the metabolism
3 of d-pinitol (Zhao et al., 2013). A study with *Bacillus subtilis* has found that this bacterium can use
4 pinitol as a sole carbon source, thought to be via the *myo*-inositol catabolism pathway. Pinitol is an
5 inositol-related compound, a 3-O-methyl of d-chiro-inositol. The *myo*-inositol catabolic pathway
6 involves step-wise reactions that ultimately yield dihydroxyacetone phosphate and acetyl-CoA, thus
7 acting as an energy source (Morinaga et al., 2006). Both pinitol and *myo*-inositol are cyclitols.
8 Significantly, cyclitols have been found to be key osmotica (a substance that acts to supplement
9 osmotic pressure) in higher plants, and may provide a putative link between adaptation of eucalypts
10 to aridity (Hasegawa et al., 2000). A metabolite profiling study of Eucalyptus trees has found that
11 there are numerous species (including species used as a food source by koalas (e.g. *Eucalyptus*
12 *saligna*)) containing high levels of *myo*-inositol (Merchant et al., 2006), suggesting a possible link
13 to the abundance of *Alistipes* in the koala caecum. Interestingly, the abundance of *Alistipes* is
14 significantly reduced in the colon and faecal pellet (caecum; 21.9%, colon; 4.7%, & faecal pellet;
15 9.5%). Although the benefit of *Alistipes* to the host is unclear, it was also found to be significantly
16 reduced (9%) in the caecum of animal K2 (diagnosed with severe cachexia of unknown cause).

17 Besides a significant decrease in caecal abundance of *Alistipes*, animal K2 also appeared to have a
18 significant reduction in both *Ruminococcus* (K1; 18.3%, K2; 0.47%) and *Cloacibacillus* (K1;
19 14.7%, K2; 0.08%). The reduction in *Ruminococcus* could be indicative of reduced bacterial
20 fermentation of cellulose, meaning reduced availability of essential nutrients. Furthermore, there
21 was also a significant increase in abundance of *Parabacteroides* and *Desulfovibrio*. Interestingly,
22 species from both genera are known to be opportunistic pathogens of humans (Goldstein et al.,
23 2003; Nakano et al., 2008), with the *Desulfovibrio* genus also containing species that are strict
24 animal pathogens (Lawson and Gebhart, 2000). As well as identifying potential opportunistic
25 pathogens, comparative analysis of animals K1 and K2 also identified the presence of known gut

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 pathogens (OTUs classified as either *Escherichia* or *Shigella*) that were exclusively found in the
2 gastrointestinal tract of animal K2 (Supplementary Table 6).
3 The comparative analysis of animals K1 and K2 also found that the genus *Akkermansia* is
4 completely absent from the caecum, colon and pellet of the emaciated animal, K2 (Supplementary
5 Table 6). A recent, elegant study by Everard *et al.* clearly defined a causal role for the species
6 *Akkermansia muciniphila* in controlling gut inflammation, the gut mucosal barrier and gut peptide
7 secretion (Everard et al., 2013). The abundance of the genus *Akkermansia* in the healthy animal
8 (K1) is much lower than that found in humans. However, this bacterium is normally found in the
9 mucus layer of the gut and thus the caecal content sampling technique may not have been extensive
10 enough to reveal the true abundance of this important bacterial community. Future studies should
11 incorporate sampling of the host tissue to ensure characterisation of important communities lining
12 the gut epithelium.
13 The koala's sole diet of eucalypt leaf is high in tannins which limit the availability of energy and
14 nutrients. Adaptation to this diet has resulted in the koala gastrointestinal tract harbouring tannin
15 degrading bacteria (Goel et al., 2005). Using culturing techniques, Osawa *et al.* identified the novel
16 tannin degrading bacterium, *Streptococcus bovis* and *Lonepinella koalarum*, from the koala caecum
17 and faeces (Osawa et al., 1995; Osawa, 1990). Surprisingly, both bacterial genera were not
18 identified during the initial analysis of this study. Further investigation, via BLAST alignment of
19 either *L. koalarum* or *S. bovis* 16S rRNA gene against an OTU fasta file, led to the identification of
20 OTUs that were highly homologous to both 16S sequences (data not shown). OTUs found to be
21 homologous to the *S. bovis* 16S sequence were generally classified to the Firmicutes phylum and as
22 *Incertae sedis* at the genus level. OTUs homologous to the *L. koalarum* 16S sequence were
23 classified to the Proteobacteria phylum, Pasteurellales order, and as 'Uncultured' at the genus level.
24 These findings suggest that both *S. bovis* and *L. koalarum* are not present in the SILVA database
25 used to assign classification. Interestingly, the abundance of these OTUs was very low (particularly
26 *L. koalarum*, <0.2%), highlighting the importance of rare or low abundant OTUs at a functional

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 level. Characterisation of less abundant OTUs is essential as abundance does not necessarily reflect
2 important and complex functionality (Arumugam et al., 2011).
3 Although this is a preliminary, observational study, it must be noted that limitations of sample size
4 make meaningful intra-animal and inter-animal comparisons very difficult. In addition, variations in
5 diet (based on differing geographical locations of the sampled animals) and sex may have also
6 impacted on the microbiome differences noted in this study. Future studies should attempt to
7 sample larger numbers of animals from multiple geographical locations and attempt to account for
8 factors such as sex, age and diet in order to define a core hindgut microbiome.
9 In conclusion, this study characterised the predominant bacterial communities across the hindgut of
10 the koala showing significant intra-individual variation, particularly in regards to abundance of
11 bacterial communities in the caecum versus the faecal pellet, and differences between a healthy and
12 diseased koala gastrointestinal tract microbiome. Future studies are needed to further define a core
13 'healthy' koala gastrointestinal tract microbiome at greater sequencing depth and functional
14 capacity in order to develop a complete picture of the koala microbiome and the role it may play in
15 health and disease.
16

17 **Acknowledgements**

18 This work was supported by an Australian Research Council (ARC) Linkage grant. We would like
19 to thank Claude Lacasse and hospital staff at the Australia Zoo Wildlife Hospital, for helping with
20 sample collection. We would also like to thank members of the Australian Genome Research
21 Facility (AGRF) sequencing team for their help with 454 pyrosequencing.
22

23 **References**

24 Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R.,
25 Tap, J., Bruls, T., Batto, J.M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L.,
26 Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F.,

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T.,
2 Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang, J., Guarner, F., Pedersen, O., de Vos,
3 W.M., Brunak, S., Dore, J., Meta, H.I.T.C., Antolin, M., Artiguenave, F., Blottiere, H.M.,
4 Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariaz, G.,
5 Dervyn, R., Foerstner, K.U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W.,
6 van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le
7 Roux, K., Maguin, E., Merieux, A., Melo Minardi, R., M'Rini, C., Muller, J., Oozeer, R.,
8 Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K.,
9 Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S.D.,
10 Bork, P., 2011. Enterotypes of the human gut microbiome. Nature 473, 174-180.
11 Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., Knight, R., 2010a.
12 PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26, 266-
13 267.
14 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer,
15 N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig,
16 J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J.,
17 Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J.,
18 Knight, R., 2010b. QIIME allows analysis of high-throughput community sequencing data.
19 Nature Methods 7, 335-336.
20 Cork, S.J., Foley, W.J. 1997. Digestive and metabolic adaptations of arboreal marsupials for dealing
21 with plant antinutrients and toxins., In: N.R., S., Hinds, L. (Eds.) Recent Advances in Marsupial
22 Biology UNSW Press, Sydney, 204-226.
23 Cork, S.J., Sanson, G.D. 1990. Digestion and nutrition in the koala: a review, In: Biology of the
24 Koala. Surrey Beatty, Sydney, 129-144.
25 De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., Collini, S.,
26 Pieraccini, G., Lionetti, P., 2010. Impact of diet in shaping gut microbiota revealed by a

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 comparative study in children from Europe and rural Africa. Proceedings of the National
2 Academy of Sciences of the United States of America 107, 14691-14696.
3 de Oliveira, M.N., Jewell, K.A., Freitas, F.S., Benjamin, L.A., Totola, M.R., Borges, A.C., Moraes,
4 C.A., Suen, G., 2013. Characterizing the microbiota across the gastrointestinal tract of a
5 Brazilian Nelore steer. Veterinary Microbiology 164, 307-314.
6 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi,
7 D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and
8 workbench compatible with ARB. Applied and Environmental Microbiology 72, 5069-5072.
9 Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics
10 26, 2460-2461.
11 Everard, A., Belzer, C., Geurts, L., Ouwerkerk, J.P., Druart, C., Bindels, L.B., Guiot, Y., Derrien,
12 M., Muccioli, G.G., Delzenne, N.M., de Vos, W.M., Cani, P.D., 2013. Cross-talk between
13 *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. Proceedings of
14 the National Academy of Sciences of the United States of America 110, 9066-9071.
15 Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R., White, B.A., 2008. Polysaccharide utilization by
16 gut bacteria: potential for new insights from genomic analysis. Nature Reviews. Microbiology 6,
17 121-131.
18 Goel, G., Puniya, A.K., Aguilar, C.N., Singh, K., 2005. Interaction of gut microflora with tannins in
19 feeds. Naturwissenschaften 92, 497-503.
20 Goldstein, E.J., Citron, D.M., Peraino, V.A., Cross, S.A., 2003. *Desulfovibrio desulfuricans*
21 bacteremia and review of human Desulfovibrio infections. Journal of Clinical Microbiology 41,
22 2752-2754.
23 Gulino, L.M., Ouwerkerk, D., Kang, A.Y., Maguire, A.J., Kienzle, M., Klieve, A.V., 2013.
24 Shedding light on the microbial community of the macropod foregut using 454-amplicon
25 pyrosequencing. PLoS One 8, e61463.

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 Hasegawa, P.M., Bressan, R.A., Zhu, J.K., Bohnert, H.J., 2000. Plant cellular and molecular
2 responses to high salinity. Annual Review of Plant Physiology and Plant Molecular Biology 51,
3 463-499.
4 Krockenberger, A.K., Hume, I.D., 2007. A flexible digestive strategy accommodates the nutritional
5 demands of reproduction in a free-living folivore, the Koala (*Phascolarctos cinereus*).
6 Functional Ecology. 21, 748-756.
7 Lawson, G.H., Gebhart, C.J., 2000. Proliferative enteropathy. Journal of Comparative Pathology
8 122, 77-100.
9 Letunic, I., Bork, P., 2011. Interactive Tree Of Life v2: online annotation and display of
10 phylogenetic trees made easy. Nucleic Acids Research 39, W475-478.
11 Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., Gordon, J.I., 2008. Worlds within worlds:
12 evolution of the vertebrate gut microbiota. Nature Reviews. Microbiology 6, 776-788.
13 Ley, R.E., Turnbaugh, P.J., Klein, S., Gordon, J.I., 2006. Microbial ecology: human gut microbes
14 associated with obesity. Nature 444, 1022-1023.
15 Looft, T., Levine, U.Y., Stanton, T.B., 2012. *Cloacibacillus porcorum* sp. nov., a mucin-degrading
16 bacterium from the swine intestinal tract and emended description of the genus *Cloacibacillus*.
17 International Journal of Systematic and Evolutionary Microbiology.
18 Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative beta
19 diversity measures lead to different insights into factors that structure microbial communities.
20 Applied and Environmental Microbiology. 73, 1576-1585.
21 Mariat, D., Firmesse, O., Levenez, F., Guimaraes, V., Sokol, H., Dore, J., Corthier, G., Furet, J.P.,
22 2009. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. BMC
23 Microbiology. 9, 123.
24 Merchant, A., Richter, A., Popp, M., Adams, M., 2006. Targeted metabolite profiling provides a
25 functional link among eucalypt taxonomy, physiology and evolution. Phytochemistry 67, 402-
26 408.

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 Morinaga, T., Yamaguchi, M., Makino, Y., Nanamiya, H., Takahashi, K., Yoshikawa, H.,
2 Kawamura, F., Ashida, H., Yoshida, K., 2006. Functional myo-inositol catabolic genes of
3 *Bacillus subtilis* Natto are involved in depletion of pinitol in Natto (fermented soybean).
4 Bioscience, Biotechnology, and Biochemistry 70, 1913-1920.
5 Nakano, V., Piazza, R.M., Cianciarullo, A.M., Bueris, V., Santos, M.F., Menezes, M.A., Mendes-
6 Ledesma, M.R., Szulczewski, V., Elias, W.P., Pumbwe, L., Wexler, H., Avila-Campos, M.J.,
7 2008. Adherence and invasion of Bacteroidales isolated from the human intestinal tract. Clinical
8 Microbiology and Infection 14, 955-963.
9 Osawa, R., Blanshard, W.H., Ocallaghan, P.G., 1993. Microbiological studies of the intestinal
10 microflora of the koala, *Phascolarctos cinereus* .2. Pap, a special maternal feces consumed by
11 juvenile koalas. Australian Journal of Zoology 41, 611-620.
12 Osawa, R., Rainey, F., Fujisawa, T., Lang, E., Busse, H.J., Walsh, T.P., Stackebrandt, E., 1995.
13 *Lonepinella koalarum* gen nov, sp nov, a new tannin protein complex degrading bacterium.
14 Systematic and Applied Microbiology 18, 368-373.
15 Osawa, R.O., 1990. Formation of a clear zone on tannin-treated brain heart infusion agar by a
16 Streptococcus Sp isolated from feces of koalas. Applied and Environmental Microbiology 56,
17 829-831.
18 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O.,
19 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-
20 based tools. Nucleic Acids Research 41, D590-596.
21 Rautio, M., Eerola, E., Vaisanen-Tunkelrott, M.L., Molitoris, D., Lawson, P., Collins, M.D.,
22 Jousimies-Somer, H., 2003. Reclassification of *Bacteroides putredinis* (Weinberg et al., 1937) in
23 a new genus *Alistipes* gen. nov., as *Alistipes putredinis* comb. nov., and description of *Alistipes*
24 *finegoldii* sp. nov., from human sources. Systematic and Applied Microbiology 26, 182-188.
25 Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets.
26 Bioinformatics 27, 863-864.

Accepted Manuscript

1 Stearns, J.C., Lynch, M.D., Senadheera, D.B., Tenenbaum, H.C., Goldberg, M.B., Cvitkovitch,
2 D.G., Croitoru, K., Moreno-Hagelsieb, G., Neufeld, J.D., 2011. Bacterial biogeography of the
3 human digestive tract. Scientific Reports 1, 170.
4 Zhao, Y., Wu, J., Li, J.V., Zhou, N.Y., Tang, H., Wang, Y., 2013. Gut microbiota composition
5 modifies fecal metabolic profiles in mice. Journal of Proteome Research 12, 2987-2999.
6 Zoetendal, E.G., Rajilic-Stojanovic, M., de Vos, W.M., 2008. High-throughput diversity and
7 functionality analysis of the gastrointestinal tract microbiota. Gut 57, 1605-1615.

8

9 **Figure Captions**

10 **Figure 1.**

11 Distribution of 16S rRNA gene sequences (Variable 1-3 regions) from koala hindgut samples. **A.**
12 Depicts the distribution of samples classified to the level of phyla. **B.** Depicts the distribution of
13 samples classified to the level of genus. A complete list of relative abundance counts for each phyla
14 and genera can be found in Supplementary tables 3 and 4.

15 **Figure 2.**

16 Alpha diversity based on 16S rRNA gene sequences (Variable 1-3 regions) from koala hindgut
17 samples. **A.** Rarefaction curves for the 6 different hindgut samples measured using the observed
18 species metric. **B.** Phylogenetic Diversity measurements (the minimum total branch length of the
19 phylogenetic tree that incorporates all OTUs in a sample). **C.** Chao 1, estimated richness
20 measurement for all hindgut samples. **D.** Shannon equitability index, a measure of evenness, for all
21 hindgut samples. **E.** Shannon diversity index, characterising species diversity by accounting for
22 abundance and evenness, for all hindgut samples.

23 **Figure 3.**

24 Rank abundance curves. **A.** Rank abundance curves for caecal OTUs. **B.** Rank abundance curves for
25 colon OTUs. **C.** Rank abundance curves for faecal pellet OTUs. **D.** Rank abundance curves for the
26 caecum, colon and faecal pellet. Averages were taken for each sample site, i.e. caecum rank

Accepted Manuscript

ACCEPTED MANUSCRIPT

1 abundance values were averaged from both animals to produce the caecum curve. For all graphs,
2 the y axis displays OTU relative abundance on a logarithmic scale, and the x axis lists the OTUs in
3 rank order of descending abundance.

4 **Figure 4.**

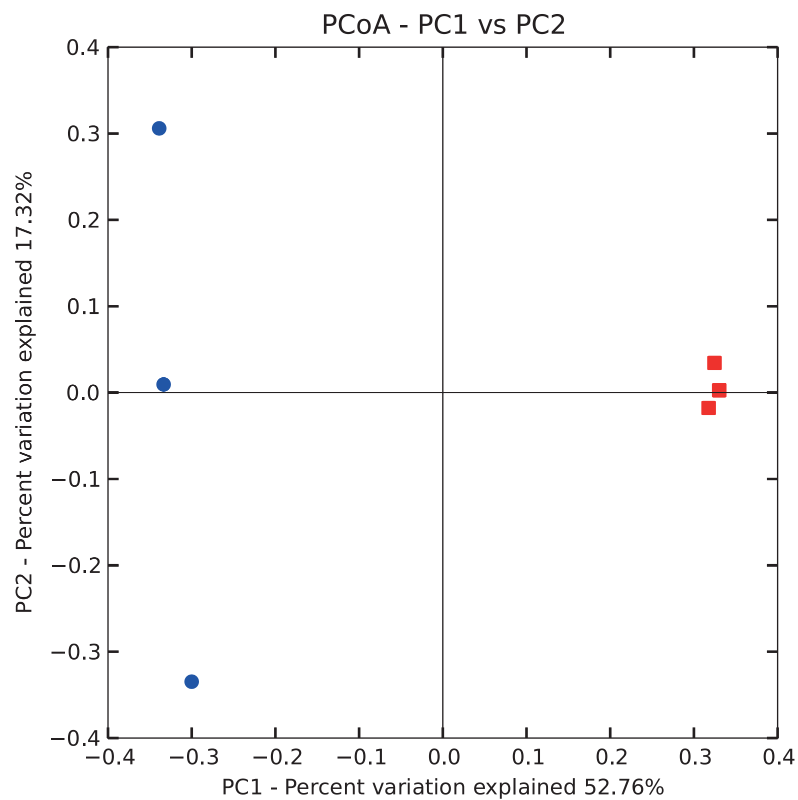
5 Principal Coordinate Analysis of unweighted and weighted UniFrac distances for 16S rRNA gene
6 sequences (Variable 1-3 regions) from koala hindgut samples. Beta diversity patterns were explored
7 using Principal Coordinate Analysis (PCoA). Samples were rarefied to 7485 reads/sample for all
8 figures. **A.** Depicts unweighted clustering of samples from animal K2 (sick animal, represented by
9 circular symbols) and animal K1 (healthy animal, represented by square symbols). **B.** Depicts
10 weighted clustering of samples from animal K2 (sick animal, represented by circular symbols) and
11 animal K1 (healthy animal, represented by square symbols). **C.** Depicts unweighted clustering of
12 samples from the caecum (represented by square symbols), colon (represented by triangle symbols)
13 and the faecal pellet (represented by circular symbols). **D.** Depicts weighted clustering of samples
14 from the caecum (represented by square symbols), colon (represented by triangle symbols) and the
15 faecal pellet (represented by circular symbols).

16 **Figure 5.**

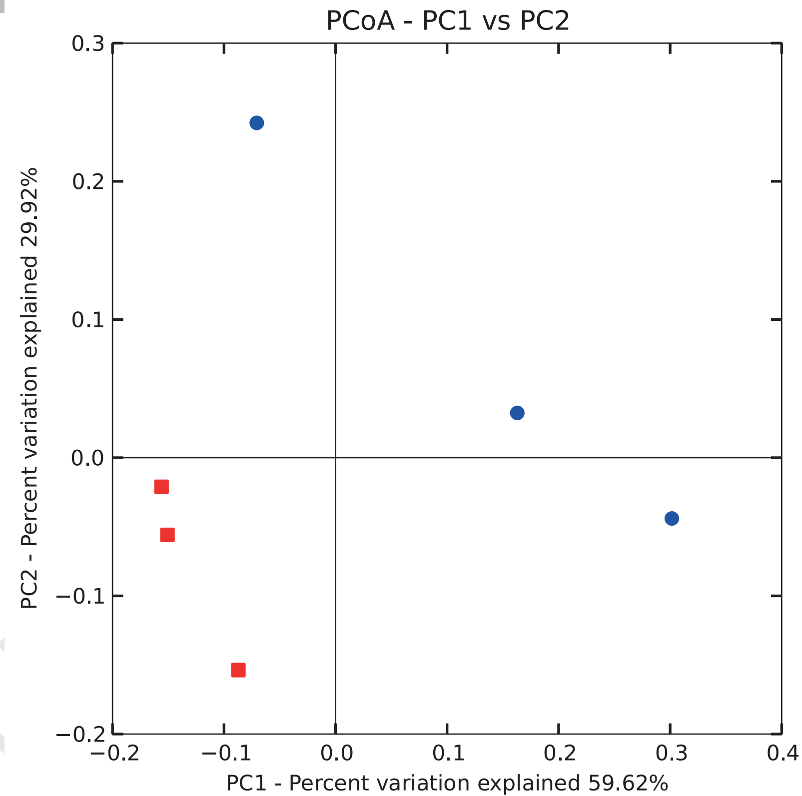
17 A venn diagram depicting the intra-individual variation of OTUs along the hindgut of a healthy
18 koala (K1). The corresponding table lists the relative abundance of OTUs at various sites.

Accepted Manuscript

A.



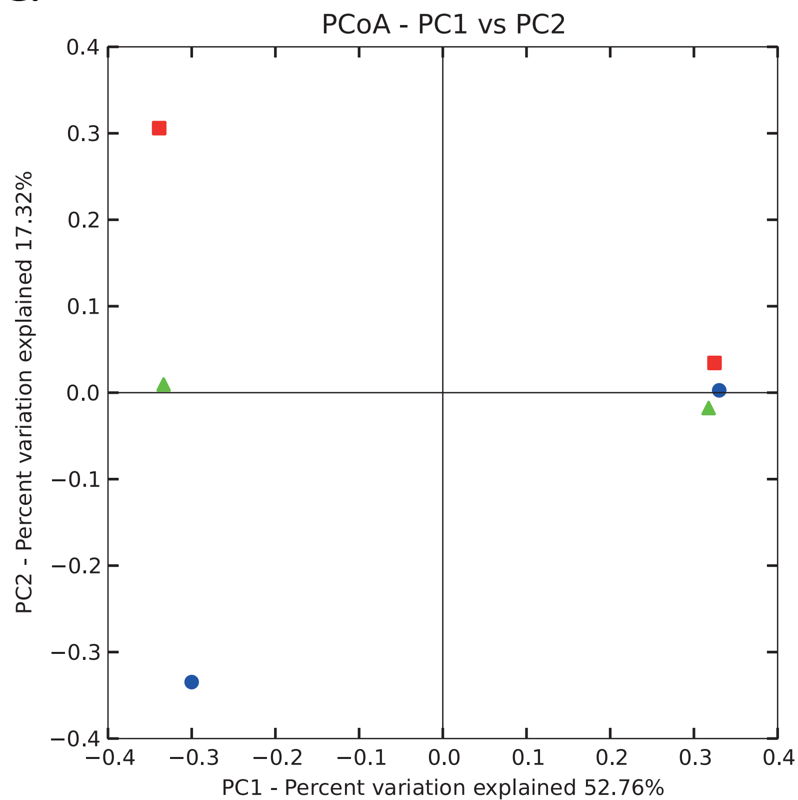
B.



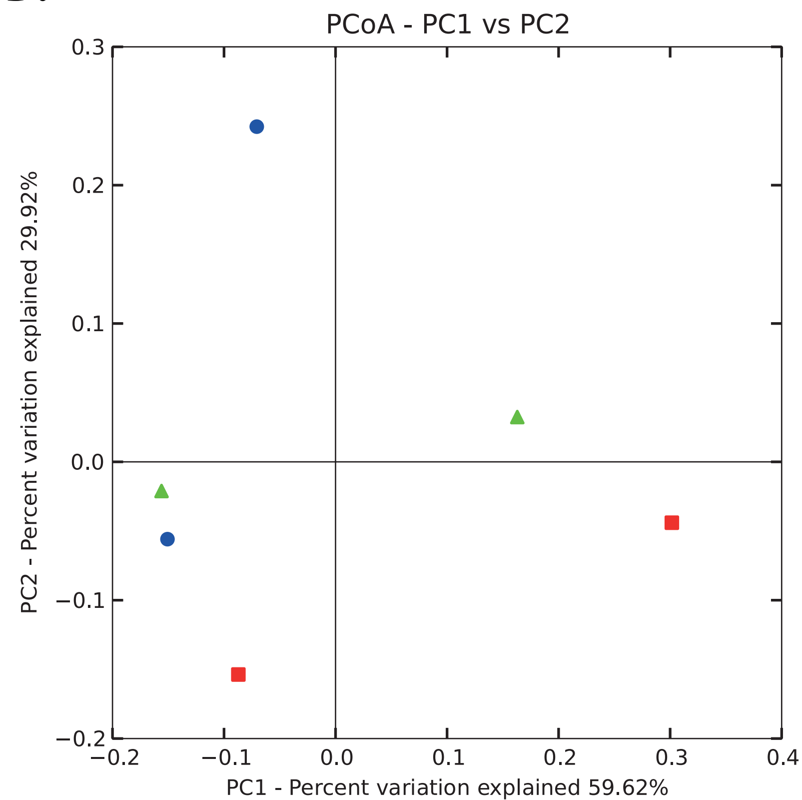
Key:

- K1 - Healthy
- K2 - Diseased

C.



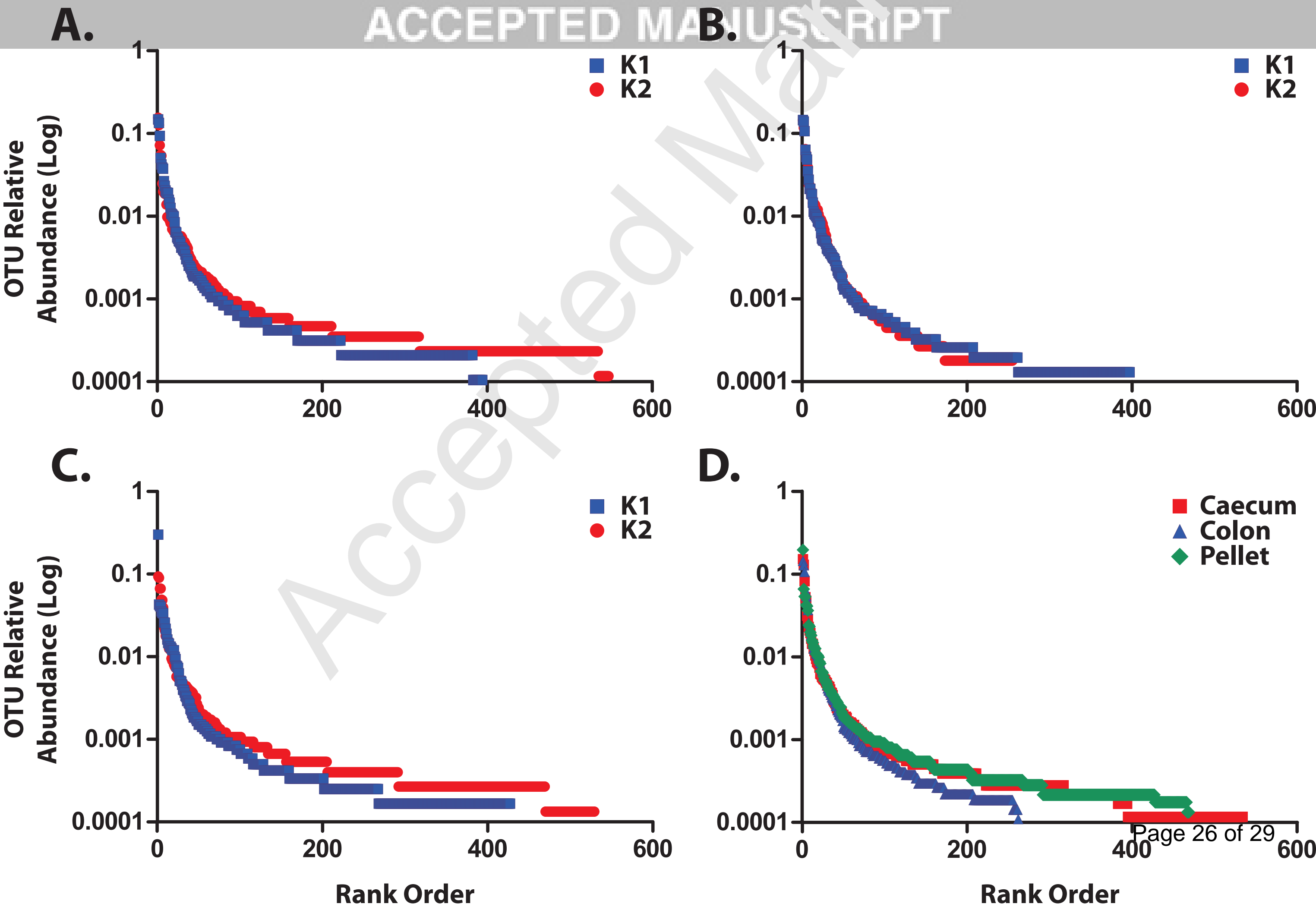
D.



Key:

- Caecum
- Colon
- Faecal Pellet

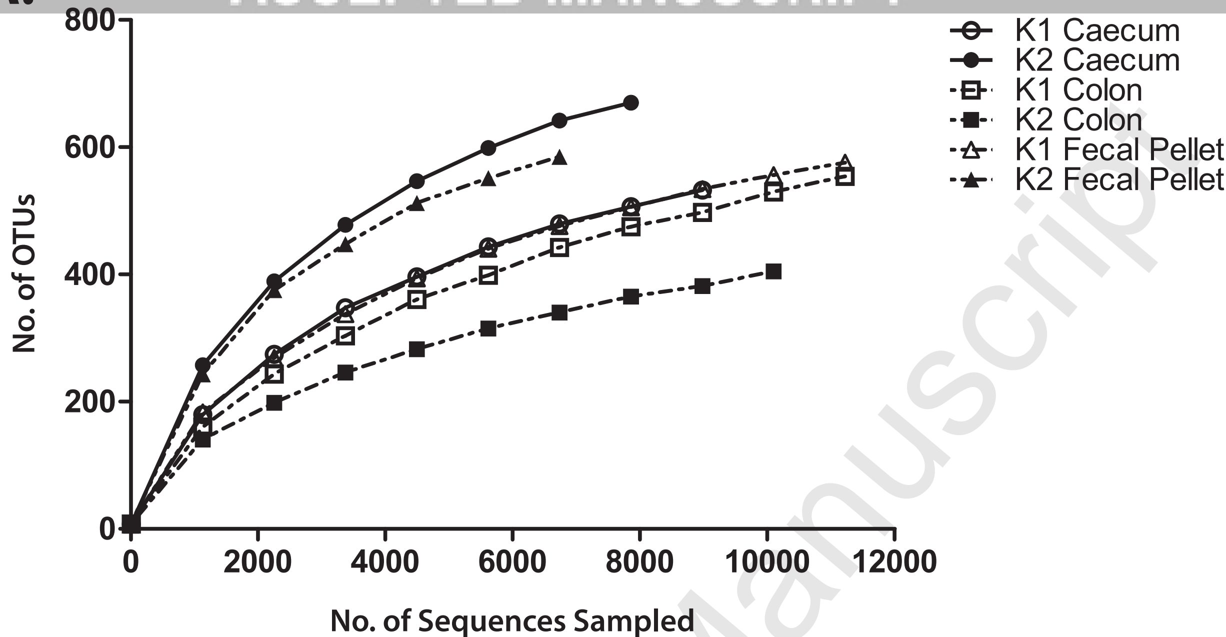
ACCEPTED MANUSCRIPT



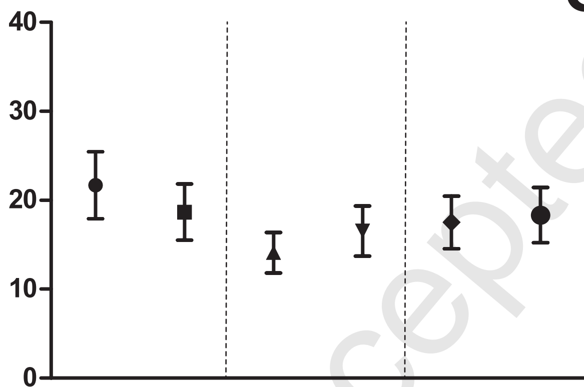
Accepted Manuscript

ACCEPTED MANUSCRIPT

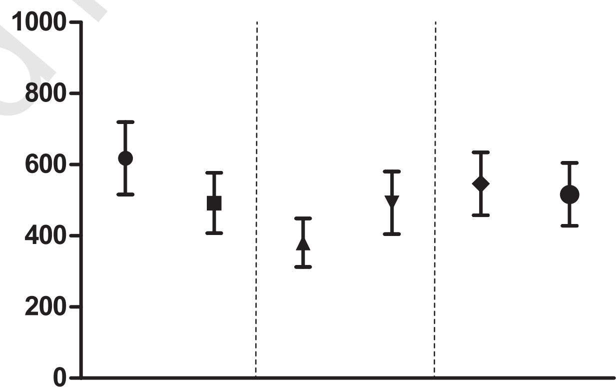
A.



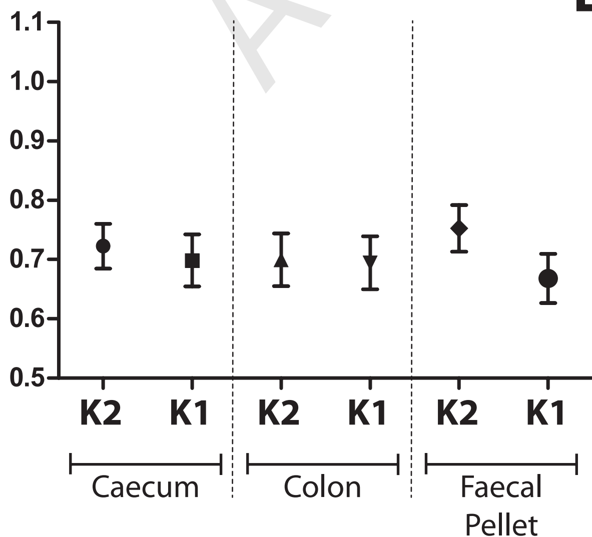
B.



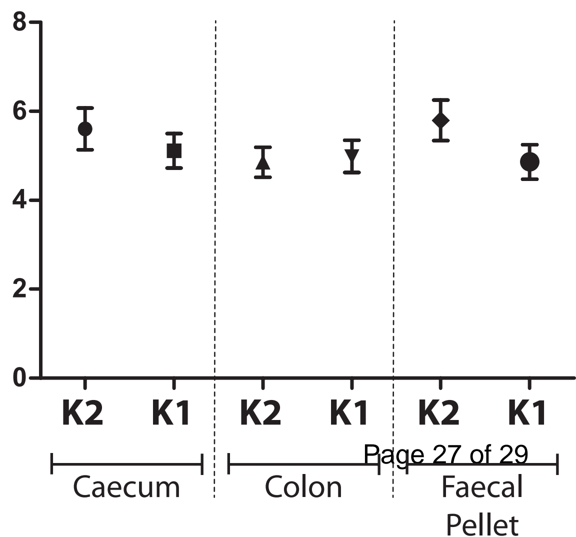
C.



D.

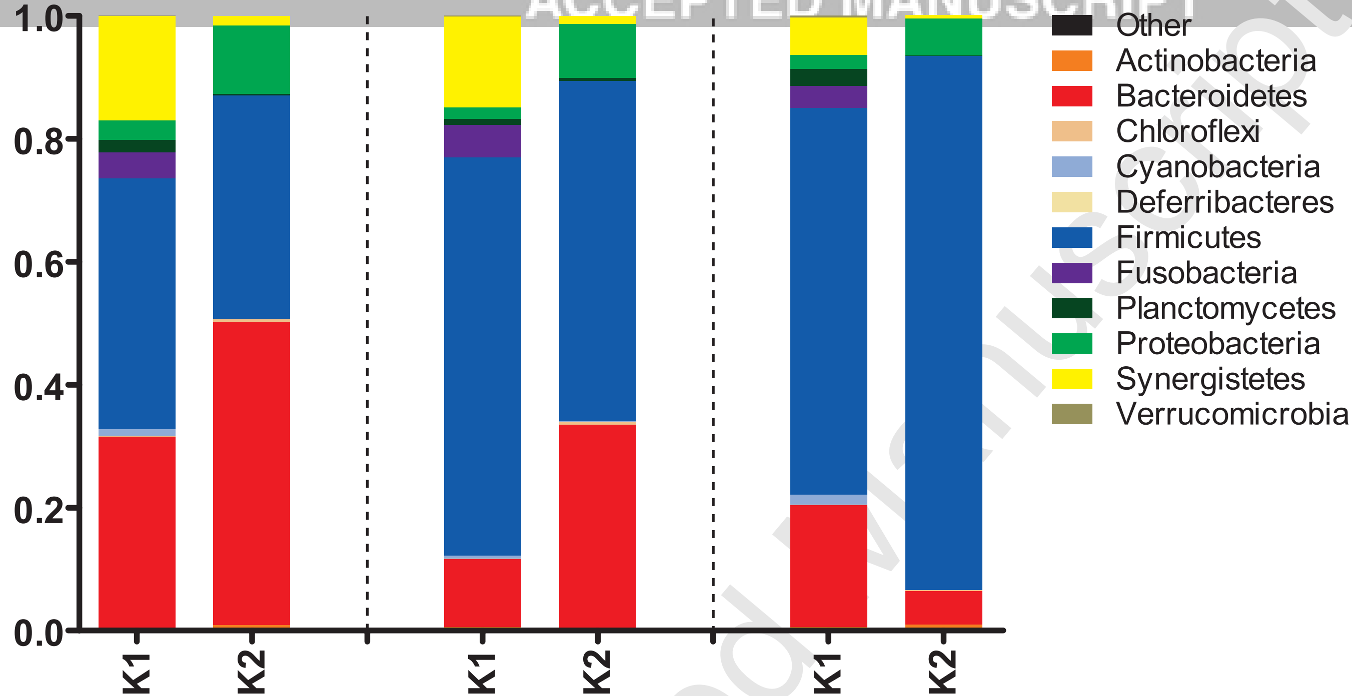


E.



Accepted Manuscript

ACCEPTED MANUSCRIPT

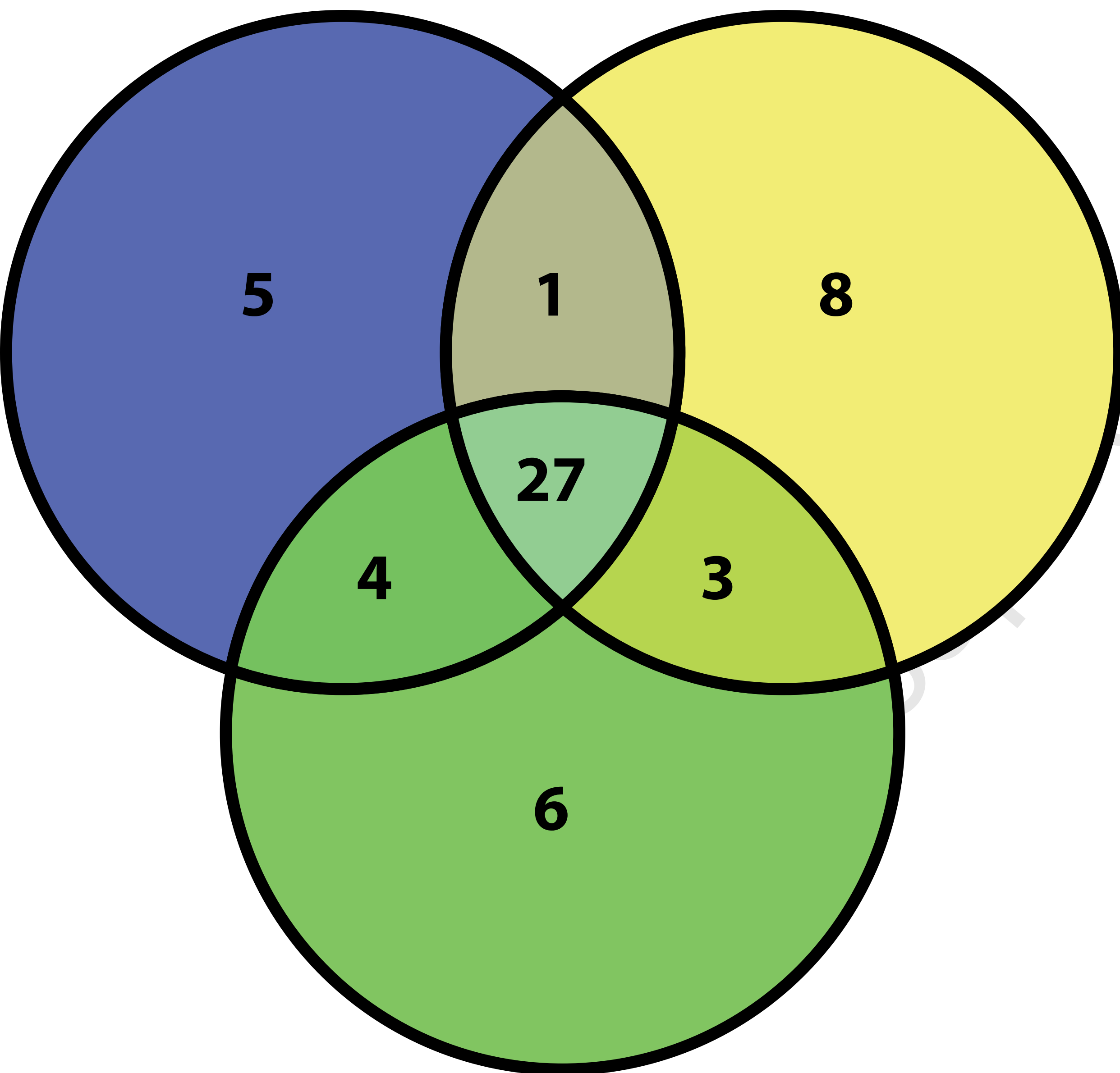


B.
Relative Abundance
1.0
0.8
0.6
0.4
0.2
0.0
K1
K2
K1
K2
K1
K2
Caecum
Colon
Faecal Pellet

Other
4c0d-2
Akkermansia
Alistipes
Alkaliflexus
Anaerovorax
Bacteroides
Bilophila
Blautia
Candidatus tammella
Cloacibacillus
Clostridium
Coprococcus
Desulfovibrio
Dysgonomonas
Enterobacter
Faecalibacterium
Flavonifractor
Fusobacterium
Gordonibacter
Incertae sedis
Megasphaera
Moryella
Nitrosomonas
Oxalobacter
Parabacteroides
Parasutterella
Phascolarctobacterium
Rikenella
Roseburia
Rs-k70 termite group
Ruminococcus
Shuttleworthia
Slackia
Uncultured
Vadinha49
Vc2.1 bac22

Accepted Manuscript

ACCEPTED MANUSCRIPT



Genus	Relative Abundance (%) CAECUM	COLON	PELLET	Site
Porphyromonas	0.042	0.020	NA	CAECUM & COLON
Actinobacillus	0.021	NA	0.008	CAECUM & PELLET
Barnesiella	0.021	NA	0.042	CAECUM & PELLET
Coprococcus	0.021	NA	0.050	CAECUM & PELLET
Saccharofermentans	0.094	NA	0.017	CAECUM & PELLET
Butyrivibrio	NA	0.026	0.017	COLON & PELLET
Clostridium	NA	0.026	0.092	COLON & PELLET
Faecalibacterium	NA	0.020	0.033	COLON & PELLET
Mucilaginibacter	0.031	NA	NA	CAECUM ONLY
Rikenella	0.063	NA	NA	CAECUM ONLY
S24-7	0.063	NA	NA	CAECUM ONLY
Sarcina	0.042	NA	NA	CAECUM ONLY
Vadinha17	0.021	NA	NA	CAECUM ONLY
Aerococcus	NA	0.078	NA	COLON ONLY
Anaerostipes	NA	0.046	NA	COLON ONLY
Cft112h7	NA	0.013	NA	COLON ONLY
Dorea	NA	0.065	NA	COLON ONLY
Enterobacter	NA	0.410	NA	COLON ONLY
Phocaeicola	NA	0.013	NA	COLON ONLY
Planomicrobium	NA	0.098	NA	COLON ONLY
Slackia	NA	0.202	NA	COLON ONLY
Chloroplast	NA	NA	0.017	PELLET ONLY
Enterorhabdus	NA	NA	0.067	PELLET ONLY
Geobacillus	NA	NA	0.017	PELLET ONLY
Gordonibacter	NA	NA	0.042	PELLET ONLY
Parasutterella	NA	NA	0.017	PELLET ONLY
Propionibacterium	NA	NA	0.017	PELLET ONLY

Genus	Relative Abundance (%) CAECUM	COLON	PELLET	Site
4c0d-2	1.132	0.489	1.578	SHARED
Akkermansia	0.042	0.065	0.259	SHARED
Alistipes	21.874	4.709	9.526	SHARED
Alkaliflexus	0.566	0.195	0.100	SHARED
Allisonella	0.136	0.013	0.017	SHARED
Anaerotruncus	0.063	0.033	0.134	SHARED
Anaerovorax	0.231	0.091	0.050	SHARED
Bacteroides	8.228	5.621	9.760	SHARED
Bilophila	2.316	1.264	1.570	SHARED
Blautia	0.304	0.046	0.267	SHARED
Cloacibacillus	14.726	11.587	4.199	SHARED
Desulfovibrio	0.440	0.137	0.334	SHARED
Dysgonomonas	0.105	0.130	0.109	SHARED
Fusobacterium	4.287	5.243	3.615	SHARED
Incertae sedis	4.161	33.837	14.176	SHARED
Megasphaera	0.063	0.254	0.058	SHARED
No BLAST Hit	0.157	0.410	0.301	SHARED
Oxalobacter	0.063	0.033	0.142	SHARED
Parabacteroides	0.346	0.221	0.225	SHARED
Phascolarctobacterium	4.570	0.879	2.738	SHARED
Pseudobutyrivibrio	0.084	0.085	0.050	SHARED
Roseburia	0.472	1.133	0.292	SHARED
Rs-k70 termite group	0.073	0.020	0.033	SHARED
Ruminococcus	18.248	16.453	33.353	SHARED
Uncultured	14.873	14.987	13.884	SHARED
Vadinha49	2.002	0.970	2.763	SHARED
Vc2.1 bac22	0.021	0.078	0.033	SHARED