



D4006- Gut Microbiota Analysis

UC Davis MMPC - Microbiome & Host Response Core

Contents

1 Methods:	1
1.1 Sequencing	1
1.2 Data processing	1
2 Summary of Findings:	2
2.1 Sequencing analysis	2
2.2 Microbial diversity	2
2.2.1 Alpha Diversity	2
2.3 Data analysis using taxa abundance data	10
2.3.1 Stacked bar graphs of Taxa abundances at each level	11
A Appendix 1 (Taxa Abundance Tables)	17
References:	24

Core Contacts:

Helen E. Raybould, Ph.D., Core Leader (heraybould@ucdavis.edu)
 Trina A. Knotts, Ph.D., Core Co-Leader (taknotts@ucdavis.edu)
 Michael L. Goodson, Ph.D., Core Scientist (mlgoodson@ucdavis.edu)

Client(s): Kent Lloyd, DVM PhD ;MMRRC; UC Davis

Project #: MBP-3995

MMRRC strain ID: MMRRC.50622

Animal Information: The strain was donated to the MMRRC by Jyothi Arikath at University of Nebraska Medical Center. Fecal samples were obtained from animals housed under the care of Jyothi Arikath at University of Nebraska Medical Center consuming Teklad 8656.

1 Methods:

Brief Project Description: MMRRC strains are often contributed to the MMRRC to fulfill the resource sharing aspects of NIH grants. Since transporting mice to another facility often causes a microbiota shift, having a record of the original fecal microbiota from the donor institution where the original phenotyping or testing was performed may prove helpful if a phenotype is lost after transfer. Several MMRRC mouse lines were selected for fecal microbiota profiling of the microbiota.

Table 1: Animal-Strain Information

X.SampleID	TreatmentGroup	Animal_ID	Genotype	MMRRC.ID	Sex
MMRRC.50622.1	MMRRC.050622_Hemi_1	1	Hemi	MMRRC.050622	M

1.1 Sequencing

Frozen fecal or regional gut samples were shipped on dry ice to UC Davis MMPC and Host Microbe Systems Biology Core. Total DNA was extracted using Qiagen PowerFecal kit. Sample libraries were prepared and analyzed by barcoded amplicon sequencing. In brief, the purified DNA was amplified on the V4 region of the 16S rRNA genes via PCR using the following primers: F341 (5'-CCTACGGGNGGCWGCAG-3') and R806 (5'-GGACTACNVTGGGTWTCTAAT-3'). High-throughput sequencing was performed with Illumina MiSeq paired end 300-bp run.

1.2 Data processing

The data derived from sequencing was processed using QIIME2 for 16S based microbiota analyses (QIIME 2 Development Team (2017)). Demultiplexed paired end sequences that already had barcodes and adapters removed were analyzed using QIIME 2 version 2021.4.0. For quality filtering and feature (OTU) prediction, we used DADA2 (Callahan et al. (2016)). Upon reviewing the sequence quality data, we trimmed 0 nucleotides (nts) from the 5' end of the forward and 0 nts from the reverse reads. Forward reads were truncated to 270 nts and reverse reads to 220 nts. Representative sequences were aligned using MAFFT (Katoh and Standley (2013)). A phylogenetic tree of the aligned sequences was made using FastTree 2 (Price, Dehal, and Arkin (2010)). OTUs/features were taxonomically classified using a pre-trained Naive Bayes taxonomy classifier. The classifier was trained using the Silva 128 97% OTUs (Quast et al. (2013)) for the 319F-806R region. Tables of taxonomic counts and percentage (relative frequency) were generated. Diversity analyses were run on the resulting OTU/feature .biom tables to provide both phylogenetic and non-phylogenetic metrics of alpha and beta diversity (Lozupone et al. (2011)). Additional data analysis (PLS-DA) and statistics were performed with R . Visualization files (.qzv) can be viewed through <http://view.qiime2.org> . The most relevant files for browsing in this manner are the feature tables (table.qzv (seq count data by taxa) and relative_freq_table.qzv (% abundance by taxa)), and taxa bar plots (taxa-bar-plots.qzv).

2 Summary of Findings:

2.1 Sequencing analysis

We obtained a mean of $40200 \pm NA$ (SD) individual sequencing reads per sample (Min= 40200; Max= 40200). After data processing the average number of sequences for each sample passing through to OTU classification was $27370 \pm NA$ (SD). The average number of OTUs per sample was 242.

2.2 Microbial diversity

Alpha and beta diversity measures were calculated using the QIIME pipeline (QIIME 2). Rarefaction was used to sample the same number of random reads from each sample for the diversity analyses. The sampling depth was set at 27300 sequences per sample.

2.2.1 Alpha Diversity

Alpha diversity is a measure of a specific sample's richness or diversity. Six different measures of alpha diversity were analyzed (Chao1, Good's coverage, Observed species, Faith's phylogenetic diversity (PD)), Shannon index, and Simpson's index. To account for differences in the number of reads in a sample, we performed rarefaction to collect a random set of sequences on a set number of sequences to rule out sample depth bias.



Figure 1: Chao1 Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

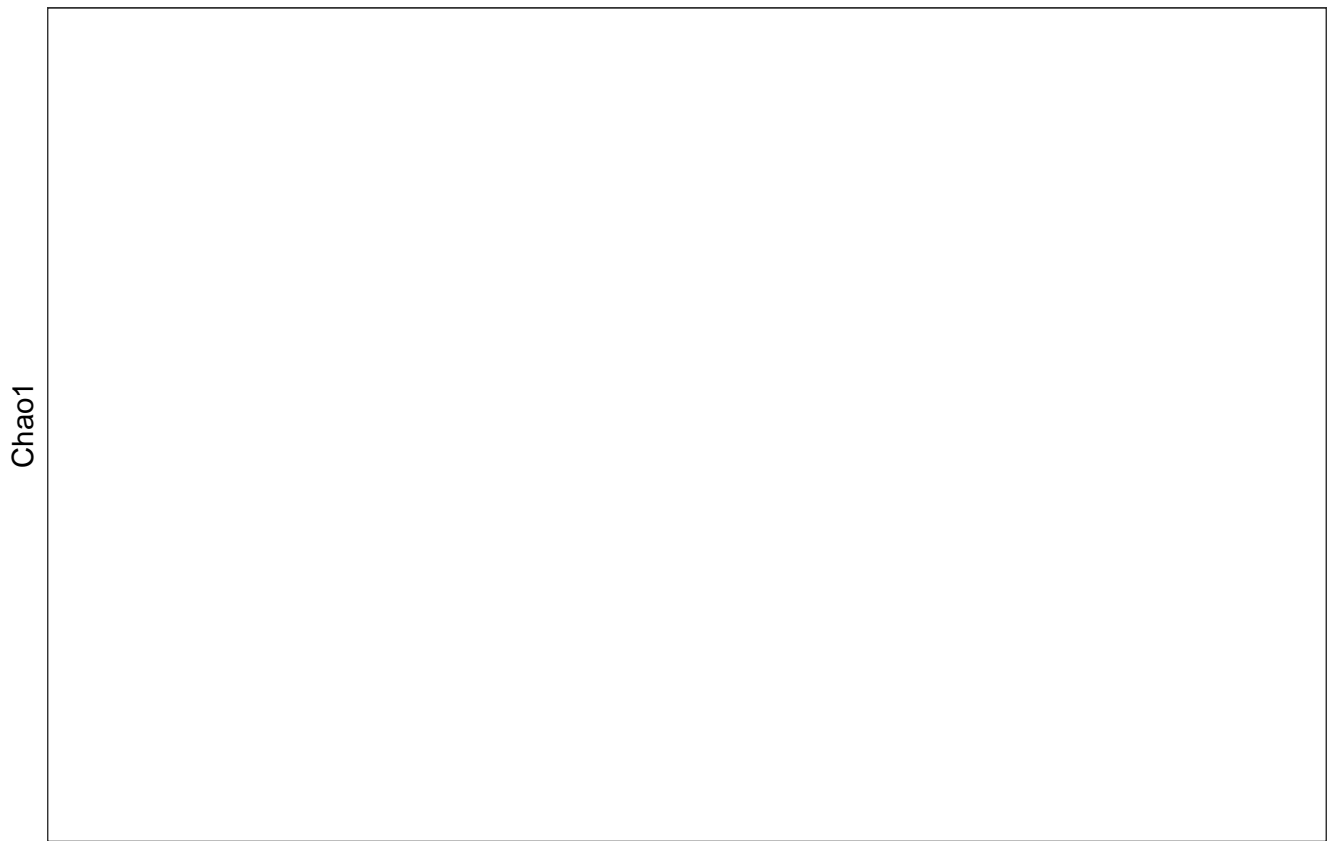


Figure 2: Chao1 Measure of Alpha Diversity- Boxplots-Animal_ID.



Figure 3: Goods Coverage Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

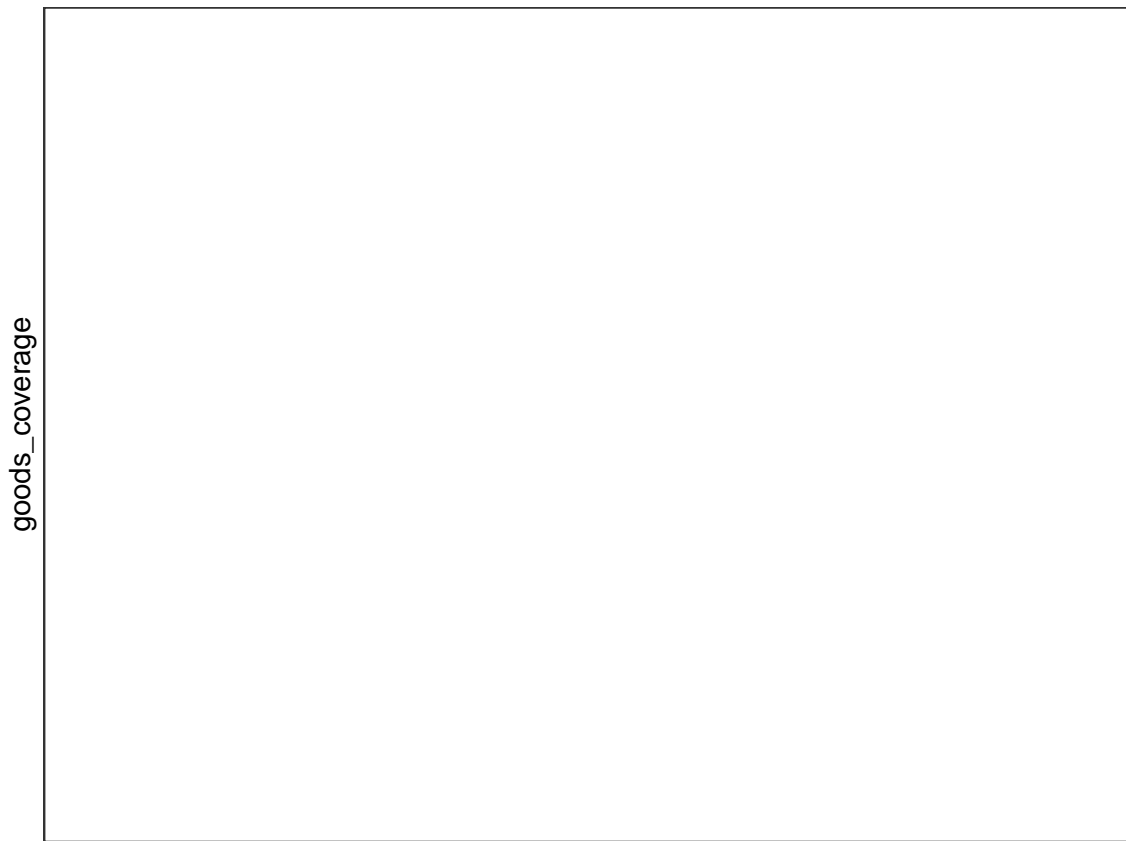


Figure 4: Goods Coverage Measure of Alpha Diversity- Boxplots-Animal_ID.



Figure 5: Observed Species Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

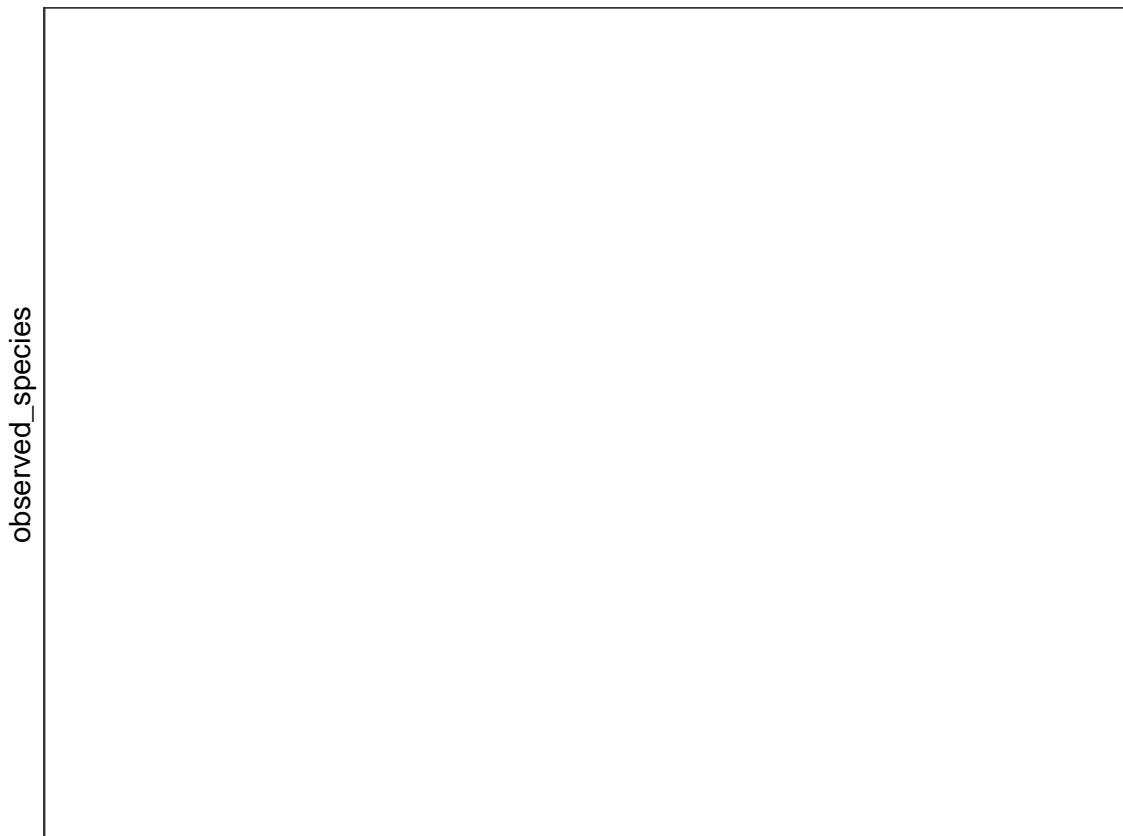


Figure 6: Observed Species Measure of Alpha Diversity- Boxplot-Animal_ID.



Figure 7: Faith's Phylogenetic Diversity Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

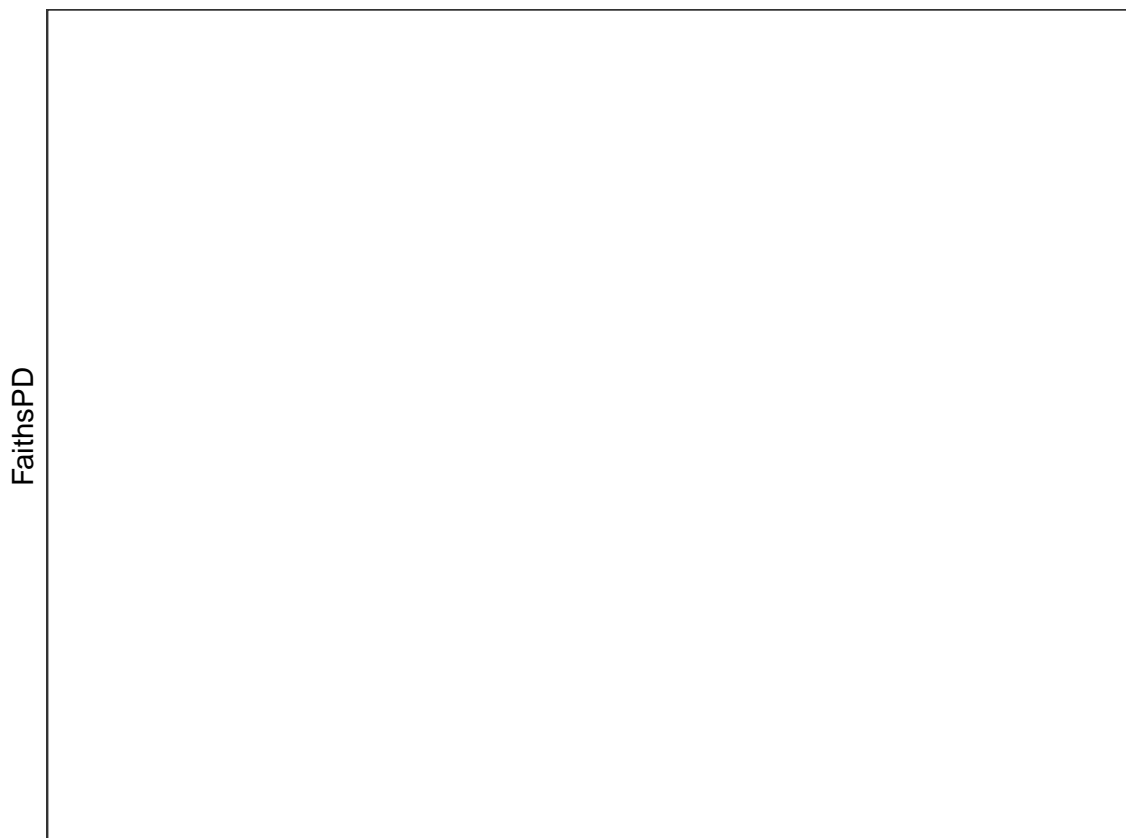


Figure 8: Faith's Phylogenetic Diversity Measure of Alpha Diversity- Boxplots-Animal_ID.



Figure 9: Shannon Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

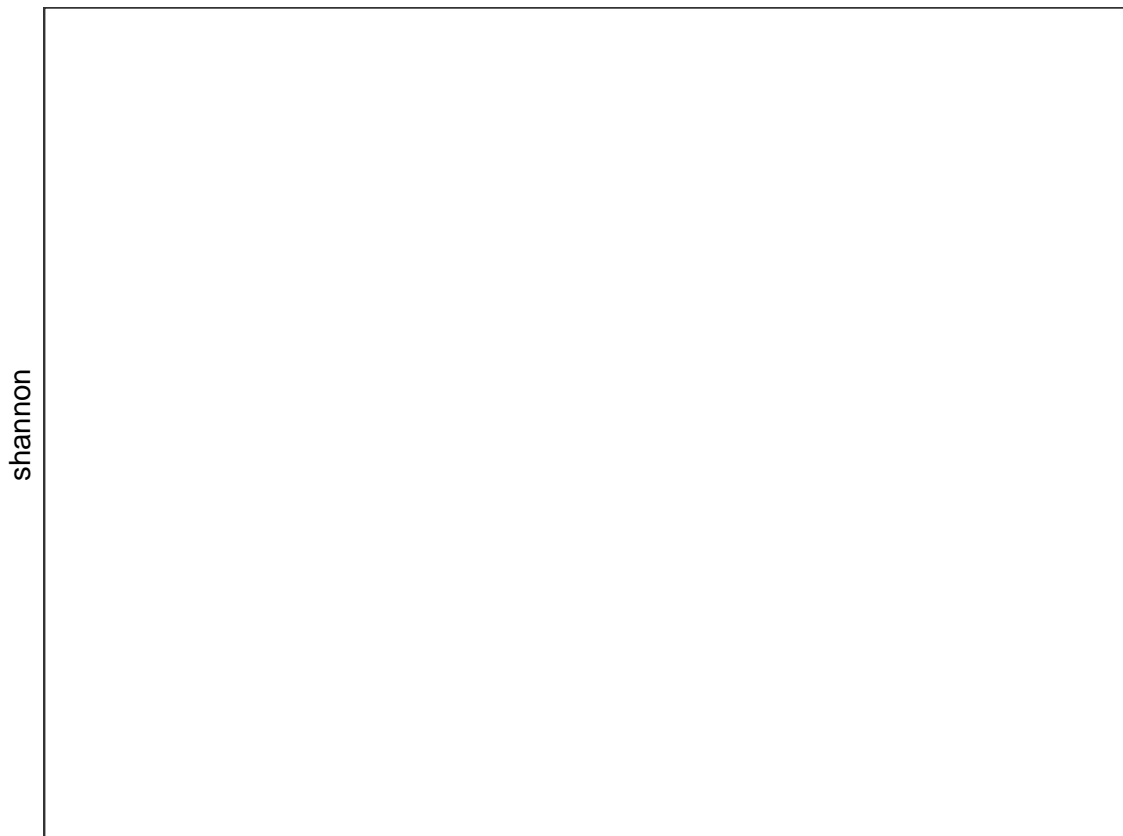


Figure 10: Shannon Measure of Alpha Diversity- Boxplots-Animal_ID.



Figure 11: Simpson Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

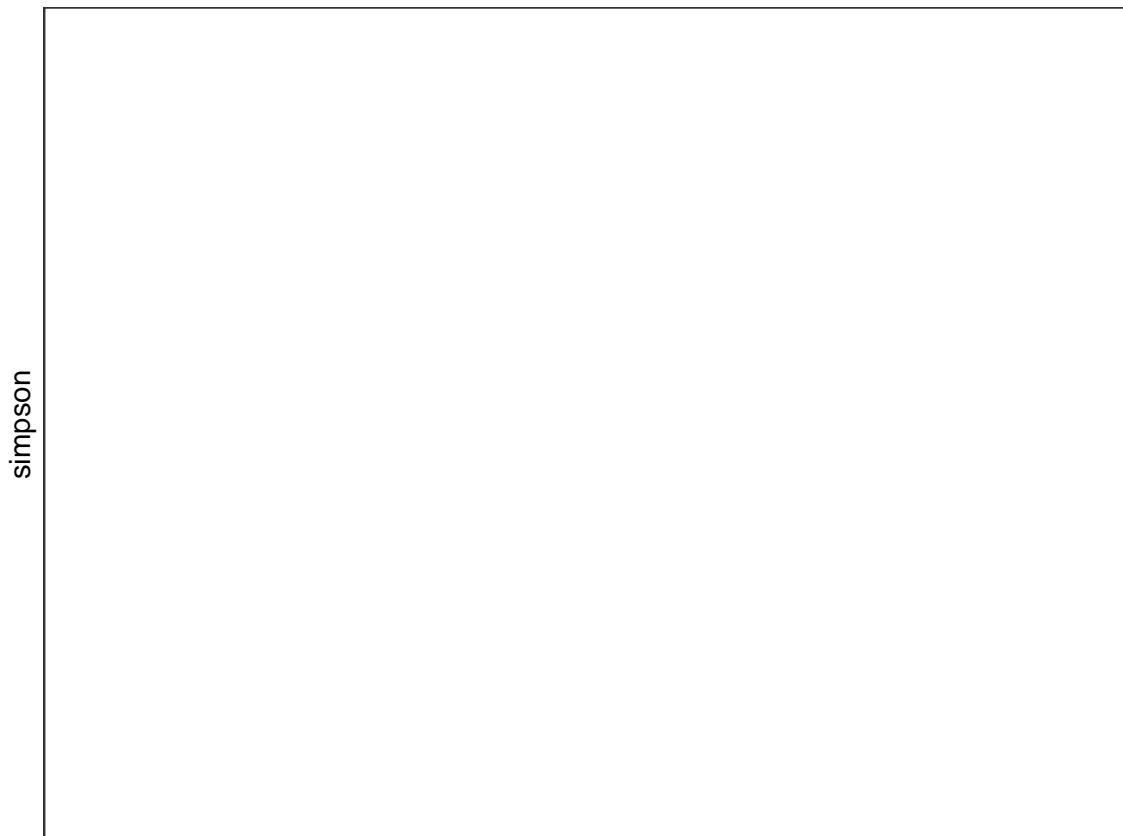


Figure 12: Simpson Measure of Alpha Diversity- Boxplots-Animal_ID.



Figure 13: Pielou's Evenness Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

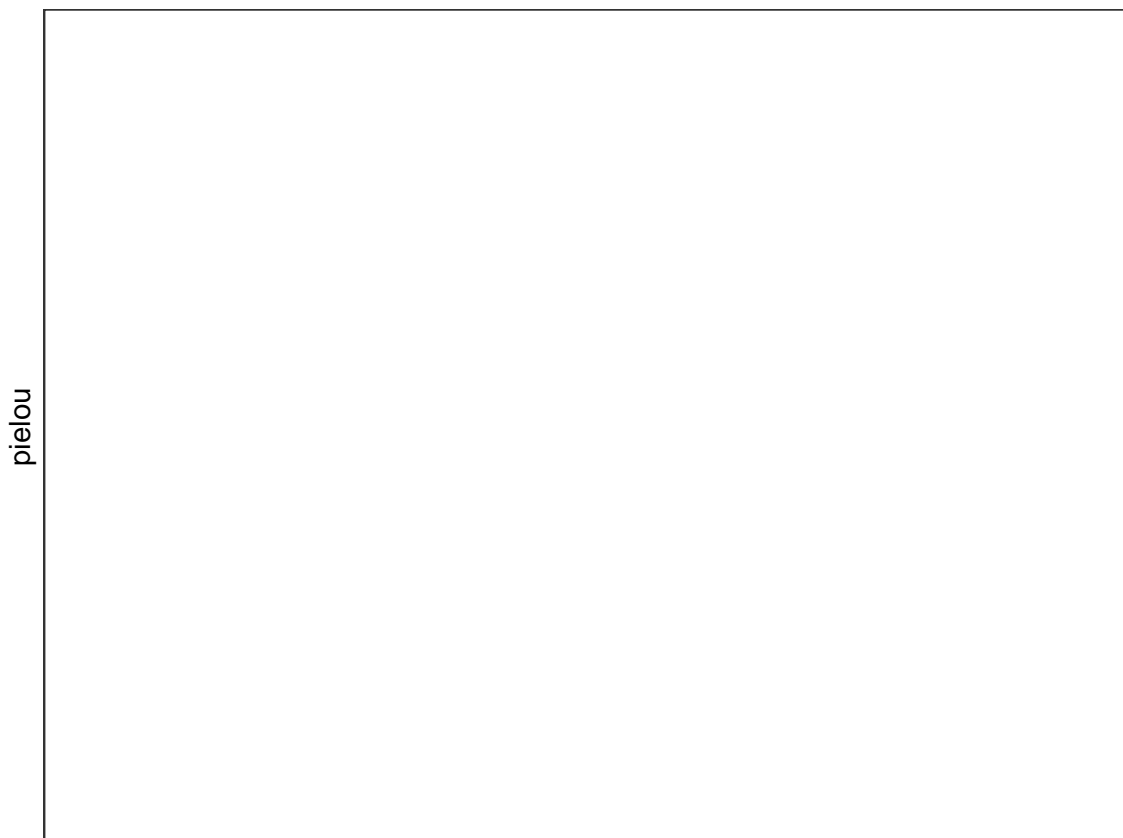


Figure 14: Pielou's Evenness Measure of Alpha Diversity- Boxplots-Animal_ID.

2.3 Data analysis using taxa abundance data

2.3.1 Stacked bar graphs of Taxa abundances at each level

Sequence counts are converted to relative abundance (% abundance). Taxa abundance data is presented in 3 ways: unfiltered, filtered, and filtered/rescaled. "Unfiltered" is the unmanipulated taxa abundance data. "Filtered" data removes low abundance taxa (mean of any group < 0.05%) and genera not present in at least 50% of samples within an experimental group. Filtering out taxa will then reduce the overall abundance to less than 100%. "Filtered /Rescaled" renormalizes the filtered data back to 100% after the low abundance taxa are removed.

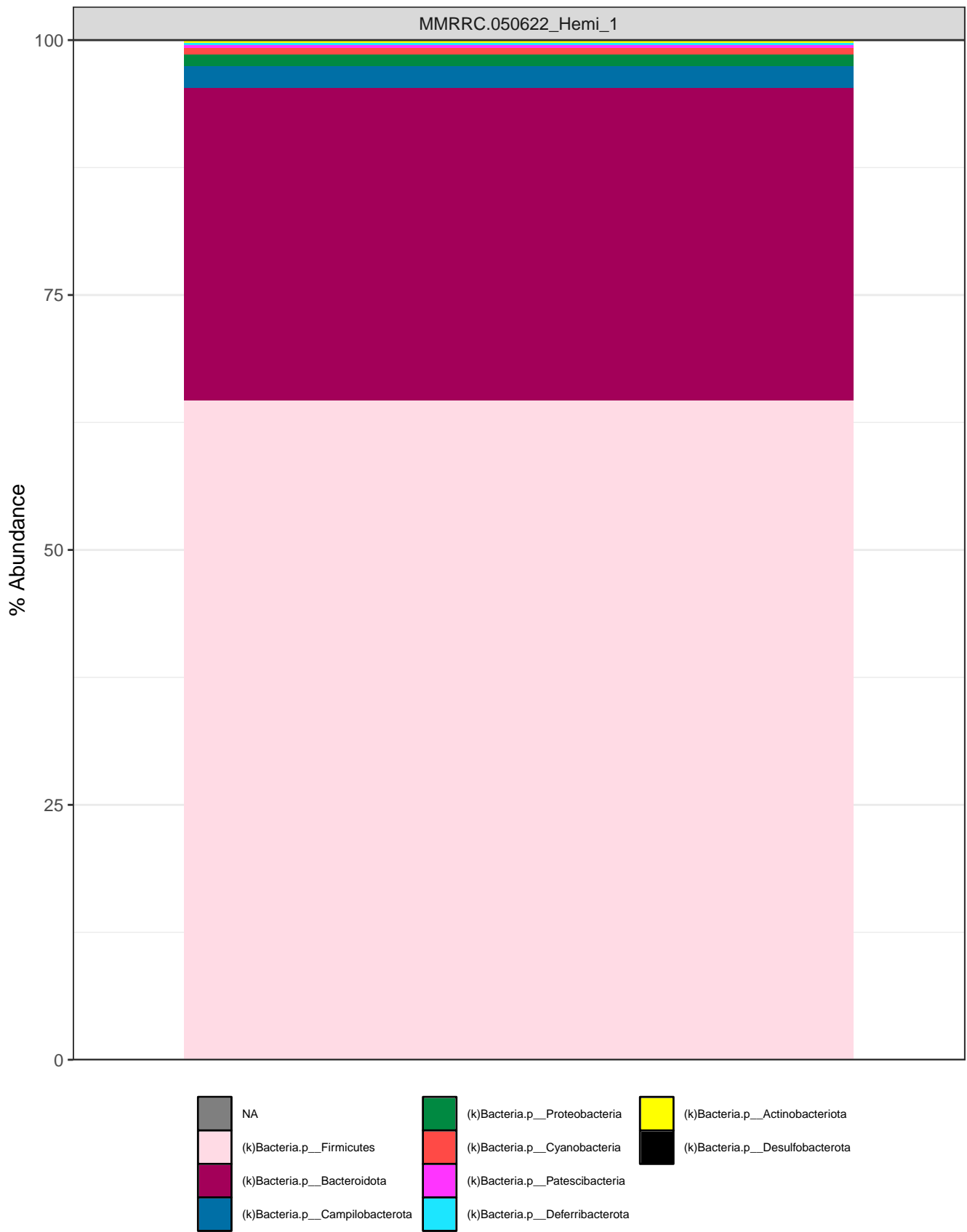


Figure 15: Phylum level abundances-UNFILTERED

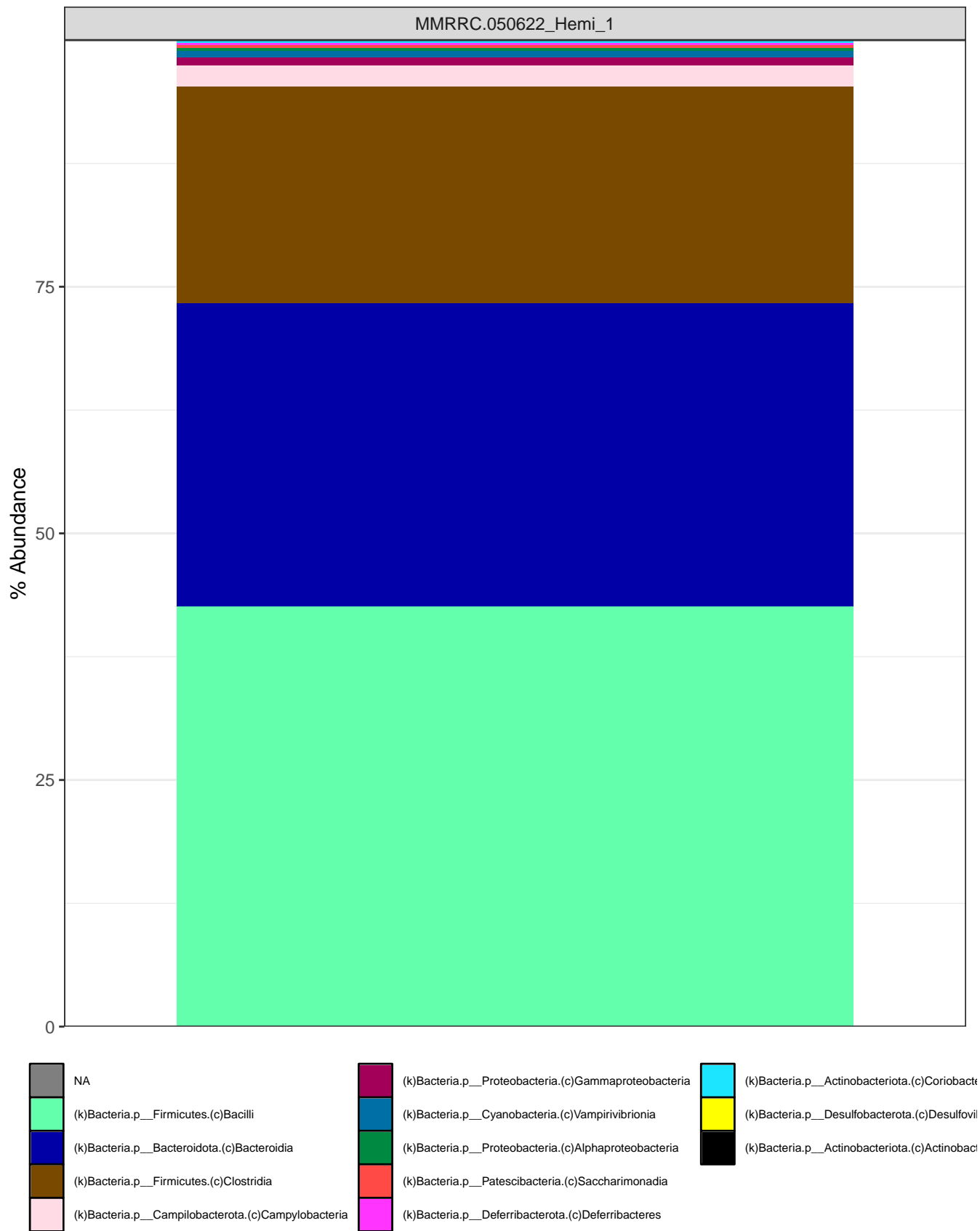


Figure 16: Class level abundances-UNFILTERED

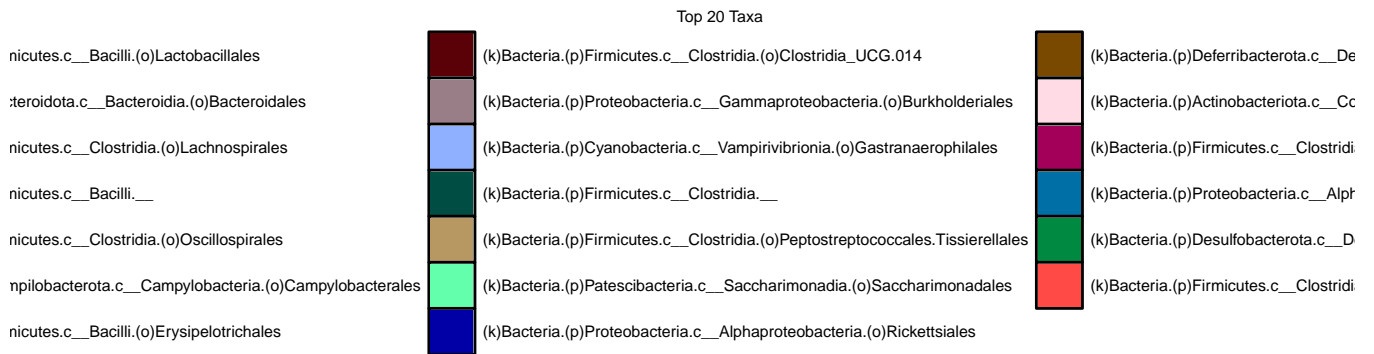
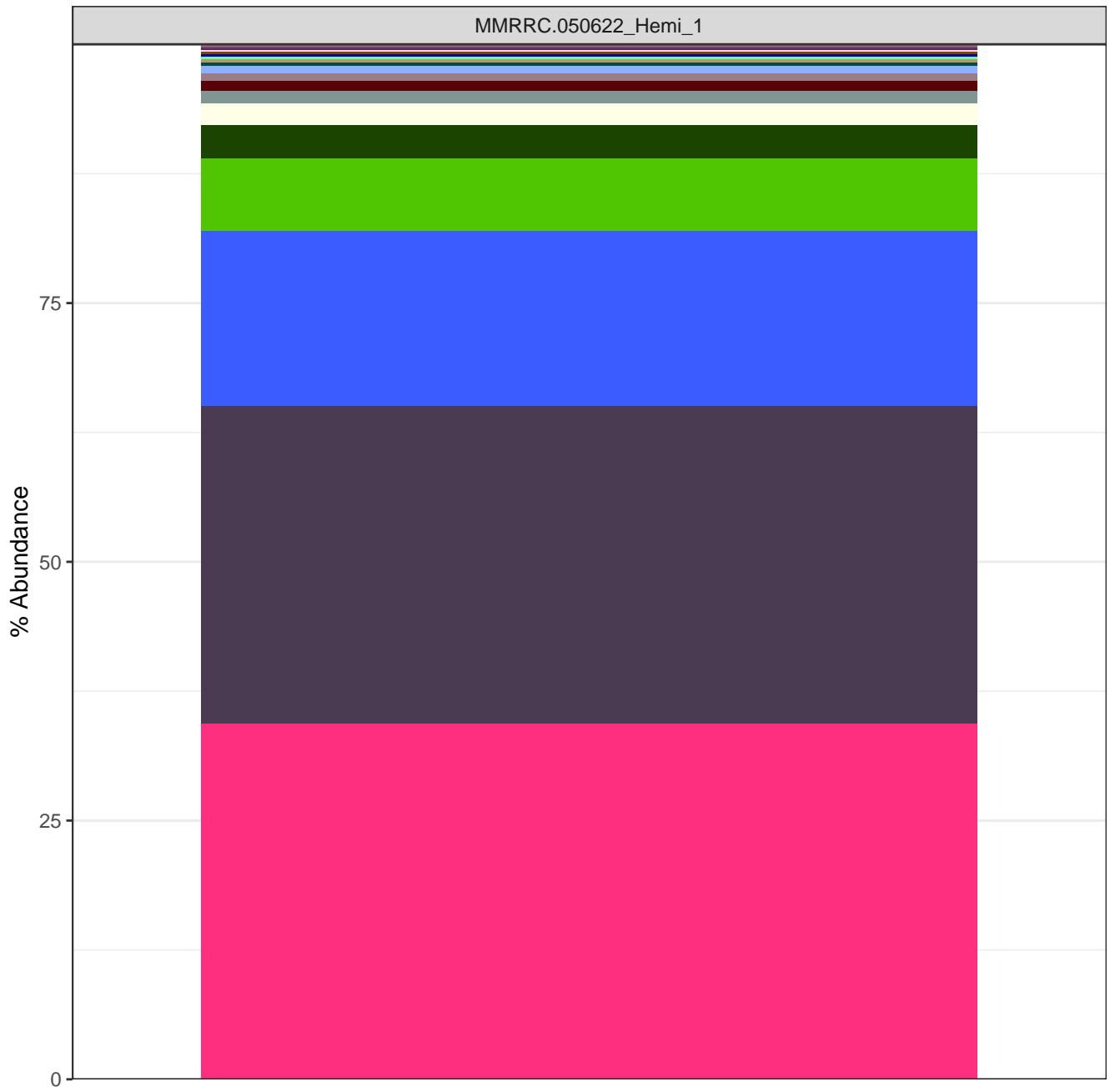


Figure 17: Order level abundances-UNFILTERED

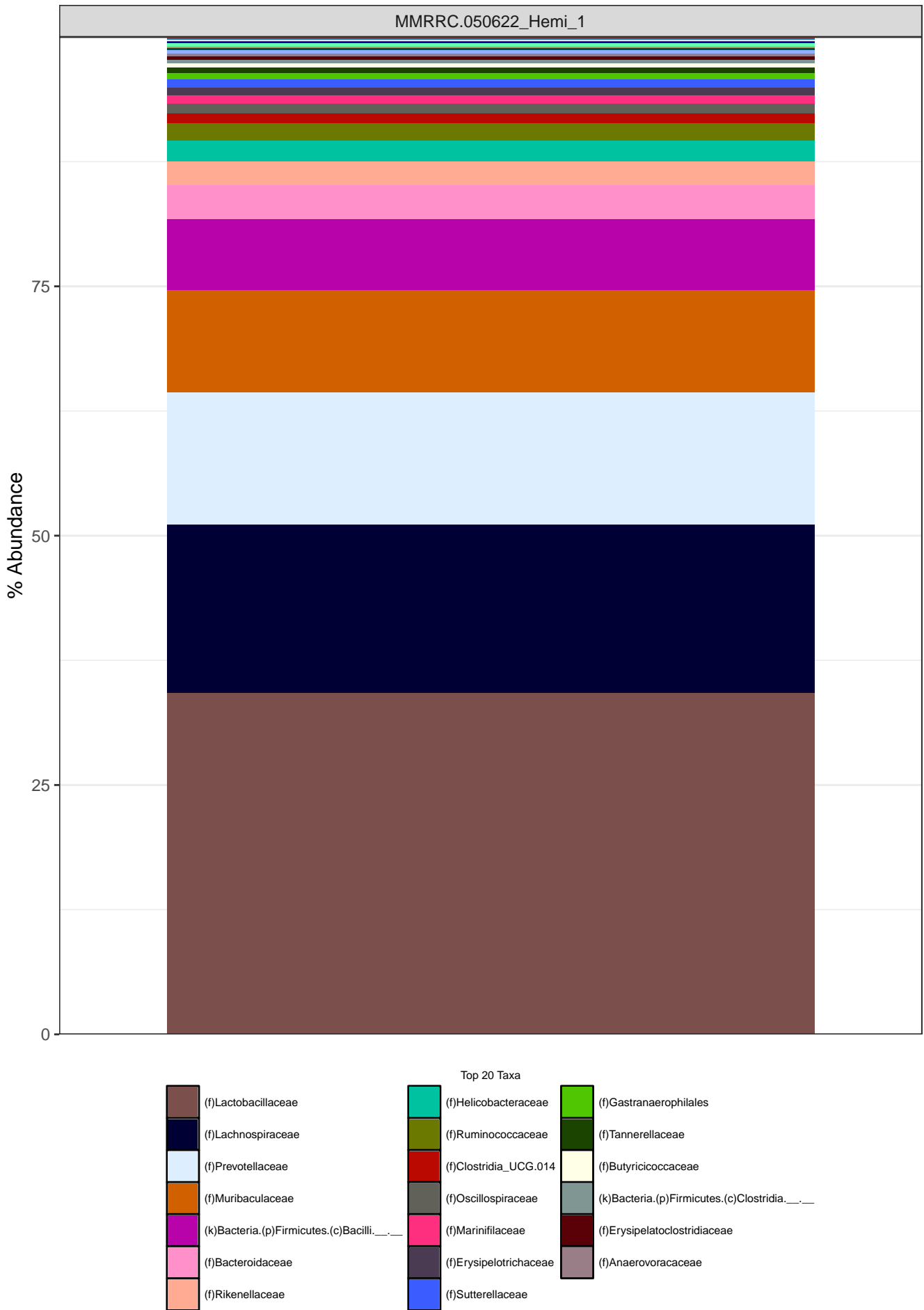


Figure 18: Family level abundances-UNFILTERED

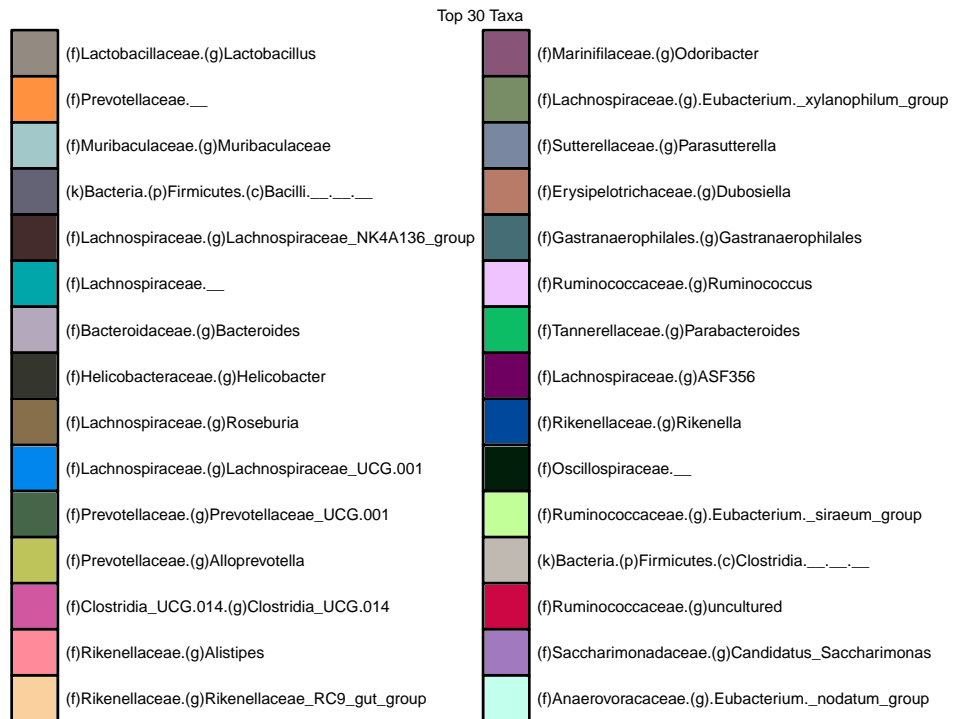
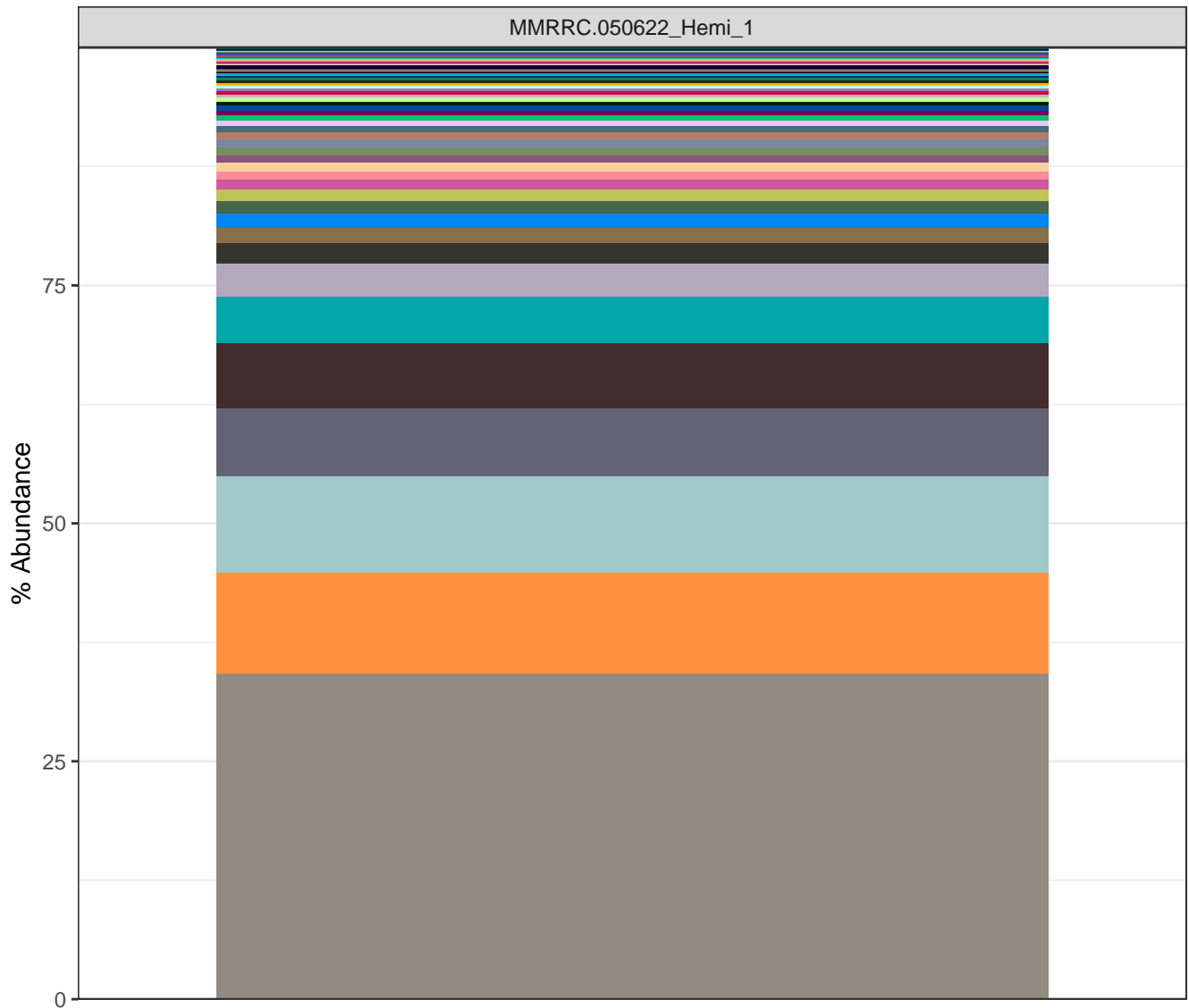


Figure 19: Genus level abundances-UNFILTERED

A Appendix 1 (Taxa Abundance Tables)

Table 2: Phylum level taxa group mean percent abundances

Taxa	MMRRC.050622_Hemi_1
(k)Bacteria.p__Desulfobacterota	0.08%
(k)Bacteria.p__Actinobacteriota	0.21%
(k)Bacteria.p__Deferribacterota	0.22%
(k)Bacteria.p__Patescibacteria	0.3%
(k)Bacteria.p__Cyanobacteria	0.65%
(k)Bacteria.p__Proteobacteria	1.11%
(k)Bacteria.p__Campilobacterota	2.13%
(k)Bacteria.p__Bacteroidota	30.72%
(k)Bacteria.p__Firmicutes	64.59%

Table 3: Class level taxa group mean percent abundances

Taxa	MMRRC.050622_Hemi_1
(k)Bacteria.p__Actinobacteriota.(c)Actinobacteria	0.03%
(k)Bacteria.p__Desulfobacterota.(c)Desulfovibrionia	0.08%
(k)Bacteria.p__Actinobacteriota.(c)Coriobacteriia	0.18%
(k)Bacteria.p__Deferribacterota.(c)Deferribacteres	0.22%
(k)Bacteria.p__Patescibacteria.(c)Saccharimonadia	0.3%
(k)Bacteria.p__Proteobacteria.(c)Alphaproteobacteria	0.32%
(k)Bacteria.p__Cyanobacteria.(c)Vampirivibrionia	0.65%
(k)Bacteria.p__Proteobacteria.(c)Gammaproteobacteria	0.79%
(k)Bacteria.p__Campilobacterota.(c)Campylobacteria	2.13%
(k)Bacteria.p__Firmicutes.(c)Clostridia	21.99%
(k)Bacteria.p__Bacteroidota.(c)Bacteroidia	30.72%
(k)Bacteria.p__Firmicutes.(c)Bacilli	42.6%

Table 4: Order level taxa group mean percent abundances

Taxa	MMRRC.050622_Hemi_1
(k)Bacteria.(p)Actinobacteriota.c__Actinobacteria.(o)Bifidobacteriales	0.03%
(k)Bacteria.(p)Firmicutes.c__Bacilli.(o)RF39	0.03%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Peptococcales	0.04%
(k)Bacteria.(p)Firmicutes.c__Bacilli.(o)Mycoplasmatales	0.06%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Monoglobales	0.07%
(k)Bacteria.(p)Desulfobacterota.c__Desulfovibrionia.(o)Desulfovibrionales	0.08%
(k)Bacteria.(p)Proteobacteria.c__Alphaproteobacteria.(o)Rhodospirillales	0.09%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Clostridia__vadinBB60__group	0.15%
(k)Bacteria.(p)Actinobacteriota.c__Coriobacteriia.(o)Coriobacteriales	0.18%
(k)Bacteria.(p)Deferribacterota.c__Deferribacteres.(o)Deferribacterales	0.22%
(k)Bacteria.(p)Proteobacteria.c__Alphaproteobacteria.(o)Rickettsiales	0.23%
(k)Bacteria.(p)Patescibacteria.c__Saccharimonadia.(o)Saccharimonadales	0.3%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Peptostreptococcales.Tissierellales	0.31%
(k)Bacteria.(p)Firmicutes.c__Clostridia.___	0.37%
(k)Bacteria.(p)Cyanobacteria.c__Vampirivibrionia.(o)Gastranaerophilales	0.65%
(k)Bacteria.(p)Proteobacteria.c__Gammaproteobacteria.(o)Burkholderiales	0.79%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Clostridia_UCG.014	0.97%
(k)Bacteria.(p)Firmicutes.c__Bacilli.(o)Erysipelotrichales	1.14%
(k)Bacteria.(p)Campilobacterota.c__Campylobacteria.(o)Campylobacterales	2.13%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Oscillospirales	3.18%
(k)Bacteria.(p)Firmicutes.c__Bacilli.___	7.05%
(k)Bacteria.(p)Firmicutes.c__Clostridia.(o)Lachnospirales	16.91%
(k)Bacteria.(p)Bacteroidota.c__Bacteroidia.(o)Bacteroidales	30.72%
(k)Bacteria.(p)Firmicutes.c__Bacilli.(o)Lactobacillales	34.31%

Table 5: Family level taxa group mean percent abundances

Taxa	MMRRC.050622_Hemi_1
(f)Streptococcaceae	1%
(f).Eubacterium._coprostanoligenes_group	1%
(f)Bifidobacteriaceae	1%
(f)RF39	1%
(f)Peptococcaceae	1%
(f)UCG.010	1%
(f)Mycoplasmataceae	1%
(f)Monoglobaceae	1%
(f)Desulfovibrionaceae	1%
(f)Enterococcaceae	1%
(f)Clostridia_vadinBB60_group	1%
(f)uncultured	2%
(f)Eggerthellaceae	1%
(f)Deferribacteraceae	1%
(f)Saccharimonadaceae	1%
(f)Anaerovoracaceae	1%
(f)Erysipelatoclostridiaceae	1%
(k)Bacteria.(p)Firmicutes.(c)Clostridia.____	1%
(f)Butyricocccaceae	1%
(f)Tannerellaceae	1%
(f)Gastranaerophilales	1%
(f)Sutterellaceae	1%
(f)Erysipelotrichaceae	1%
(f)Marinifilaceae	1%
(f)Oscillospiraceae	1%
(f)Clostridia_UCG.014	1%
(f)Ruminococcaceae	1%
(f)Helicobacteraceae	1%
(f)Rikenellaceae	1%
(f)Bacteroidaceae	1%
(k)Bacteria.(p)Firmicutes.(c)Bacilli.____	1%
(f)Muribaculaceae	1%
(f)Prevotellaceae	1%
(f)Lachnospiraceae	1%
(f)Lactobacillaceae	1%

Table 6: Genus level taxa group mean percent abundances

Taxa	MMRRC.050622_Hemi_1
(f) Anaerovoracaceae.(g) Family_XIII_AD3011_group	1%
(f) Peptococcaceae.(g) uncultured	1%
(f) Lachnospiraceae.(g) Acetatifactor	1%
(f) Streptococcaceae.(g) Lactococcus	1%
(f) Eggerthellaceae.____	1%
(f) Peptococcaceae.(g) Peptococcus	1%
(f) Eubacterium._coprostanoligenes_group.(g) Eubacterium._coprostanoligenes_group	1%
(f) Bifidobacteriaceae.(g) Bifidobacterium	1%
(f) Lachnospiraceae.(g) Lachnospiraceae_UCG.006	1%
(f) Ruminococcaceae.(g) Ruminococcaceae	1%
(f) Erysipelatoclostridiaceae.(g) Erysipelatoclostridium	1%
(f) RF39.(g) RF39	1%
(f) Lachnospiraceae.(g) Lachnospiraceae_FCS020_group	1%
(f) Lachnospiraceae.(g) Tyzzerella	1%
(f) Lachnospiraceae.(g) uncultured	1%
(f) UCG.010.(g) UCG.010	1%
(f) Oscillospiraceae.(g) NK4A214_group	1%
(f) Ruminococcaceae.(g) Negativibacillus	1%
(f) Lachnospiraceae.(g) Marvinbryantia	1%
(f) Lachnospiraceae.(g) Tuzzerella	1%
(f) Mycoplasmataceae.(g) Mycoplasma	1%
(f) Lachnospiraceae.(g) GCA.900066575	1%
(f) Monoglobaceae.(g) Monoglobus	1%
(f) Erysipelotrichaceae.(g) uncultured	1%
(f) Desulfovibrionaceae.(g) Desulfovibrio	1%
(f) Lachnospiraceae.(g) A2	1%
(f) Oscillospiraceae.(g) Colidextribacter	1%
(f) Prevotellaceae.(g) Paraprevotella	1%
(f) Enterococcaceae.(g) Enterococcus	1%
(f) Muribaculaceae.(g) Muribaculum	1%
(f) Butyricocccaceae.(g) UCG.009	1%
(f) Clostridia_vadinBB60_group.(g) Clostridia_vadinBB60_group	1%
(f) Eggerthellaceae.(g) Enterorhabdus	1%
(f) uncultured.(g) uncultured	2%
(f) Oscillospiraceae.(g) Oscillibacter	1%
(f) Ruminococcaceae.(g) Incertae_Sedis	1%
(f) Oscillospiraceae.(g) uncultured	1%
(f) Deferribacteraceae.(g) Mucispirillum	1%
(f) Lachnospiraceae.(g) Lachnoclostridium	1%
(f) Ruminococcaceae.(g) Anaerotruncus	1%
(f) Erysipelatoclostridiaceae.(g) Candidatus_Stoquefichus	1%
(f) Butyricocccaceae.(g) Butyricococcus	1%
(f) Anaerovoracaceae.(g) Eubacterium._nodatum_group	1%
(f) Saccharimonadaceae.(g) Candidatus_Saccharimonas	1%
(f) Ruminococcaceae.(g) uncultured	1%
(k) Bacteria.(p) Firmicutes.(c) Clostridia.____.____	1%
(f) Ruminococcaceae.(g) Eubacterium._siraeum_group	1%
(f) Oscillospiraceae.____	1%
(f) Rikenellaceae.(g) Rikenella	1%
(f) Lachnospiraceae.(g) ASF356	1%

Table 6: Genus level taxa group mean percent abundances (*continued*)

Taxa	MMRRC.050622_Hemi_1
(f)Tannerellaceae.(g)Parabacteroides	1%
(f)Ruminococcaceae.(g)Ruminococcus	1%
(f)Gastranaerophilales.(g)Gastranaerophilales	1%
(f)Erysipelotrichaceae.(g)Dubosiella	1%
(f)Sutterellaceae.(g)Parasutterella	1%
(f)Lachnospiraceae.(g).Eubacterium._xylanophilum_group	1%
(f)Marinifilaceae.(g)Odoribacter	1%
(f)Rikenellaceae.(g)Rikenellaceae_RC9_gut_group	1%
(f)Rikenellaceae.(g)Alistipes	1%
(f)Clostridia_UCG.014.(g)Clostridia_UCG.014	1%
(f)Prevotellaceae.(g)Alloprevotella	1%
(f)Prevotellaceae.(g)Prevotellaceae_UCG.001	1%
(f)Lachnospiraceae.(g)Lachnospiraceae_UCG.001	1%
(f)Lachnospiraceae.(g)Roseburia	1%
(f)Helicobacteraceae.(g)Helicobacter	1%
(f)Bacteroidaceae.(g)Bacteroides	1%
(f)Lachnospiraceae.____	1%
(f)Lachnospiraceae.(g)Lachnospiraceae_NK4A136_group	1%
(k)Bacteria.(p)Firmicutes.(c)Bacilli.____.____	1%
(f)Muribaculaceae.(g)Muribaculaceae	1%
(f)Prevotellaceae.____	1%
(f)Lactobacillaceae.(g)Lactobacillus	1%

Important: Please remember the policy to acknowledge the UC Davis MMPC in your publications or presentations of the work that was generated using UC Davis MMPC services. Please include a statement in the acknowledgement section: Research was supported by NIH grant U24-DK092993 (MMPC-University of California Davis Microbiome and Host Response Core, RRID:SCR_015361).

References:

- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. Johnson, and S. P. Holmes. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” Journal Article. *Nat Methods* 13 (7): 581–3. <https://doi.org/10.1038/nmeth.3869>.
- Katoh, K., and D. M. Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” Journal Article. *Mol Biol Evol* 30 (4): 772–80. <https://doi.org/10.1093/molbev/mst010>.
- Lozupone, C., M. E. Lladser, D. Knights, J. Stombaugh, and R. Knight. 2011. “UniFrac: An Effective Distance Metric for Microbial Community Comparison.” Journal Article. *ISME J* 5 (2): 169–72. <https://doi.org/10.1038/ismej.2010.133>.
- Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. “FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments.” Journal Article. *PLoS One* 5 (3): e9490. <https://doi.org/10.1371/journal.pone.0009490>.
- QIIME 2 Development Team. 2017. “QIIME 2 Docs.” Web Page. <https://docs.qiime2.org/2017.9/tutorials/moving-pictures/>.
- Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. “The Silva Ribosomal Rna Gene Database Project: Improved Data Processing and Web-Based Tools.” Journal Article. *Nucleic Acids Res* 41: D590–6. <https://doi.org/10.1093/nar/gks1219>.