

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
2.2.2 Beta Diversity	10
2.3 Data analysis using taxa abundance data	13
2.3.1 Stacked bar graphs of Taxa abundances at each level	14
A Appendix 1 (Taxa Abundance Tables)	30
B Appendix 2 (Taxa removed from filtered datasets at each level)	37
References:	43

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-2079

MMRRC strain ID: MMRRC_43535

Animal Information: The strain was donated to the MMRRC by Jan Carette at Stanford University. Fecal samples were obtained from animals housed under the care of Jan Carette at Stanford University.

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	Line	Sex
MMRRC.43535.1	AU040320-Hom	MMRRC.43535.1	Hom	MMRRC.43535	M
MMRRC.43535.2	AU040320-Hom	MMRRC.43535.2	Hom	MMRRC.43536	M
MMRRC.43535.3	AU040320-Hom	MMRRC.43535.3	Hom	MMRRC.43537	M
MMRRC.43535.4	AU040320-Hom	MMRRC.43535.4	Hom	MMRRC.43538	M
MMRRC.43535.5	AU040320-Hom	MMRRC.43535.5	Hom	MMRRC.43539	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 Mo-Bio (now 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: F319 (5'-ACTCCTACGGGAGGCAGCAGT-3') and R806 (5'-GGACTACNVGGGTWTCTAAT-3'). High-throughput sequencing was performed with Illumina MiSeq paired end 250-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 2018.4.0. 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 260 nts and reverse reads to 220 nts. Representative sequences were aligned using MAFFT (Kato 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 138354 ± 22973 (SD) individual sequencing reads per sample (Min= 104414; Max= 168937). After data processing the average number of sequences for each sample passing through to OTU classification was 58947 ± 12805 (SD). The average number of OTUs per sample was 258.

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 36000 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.

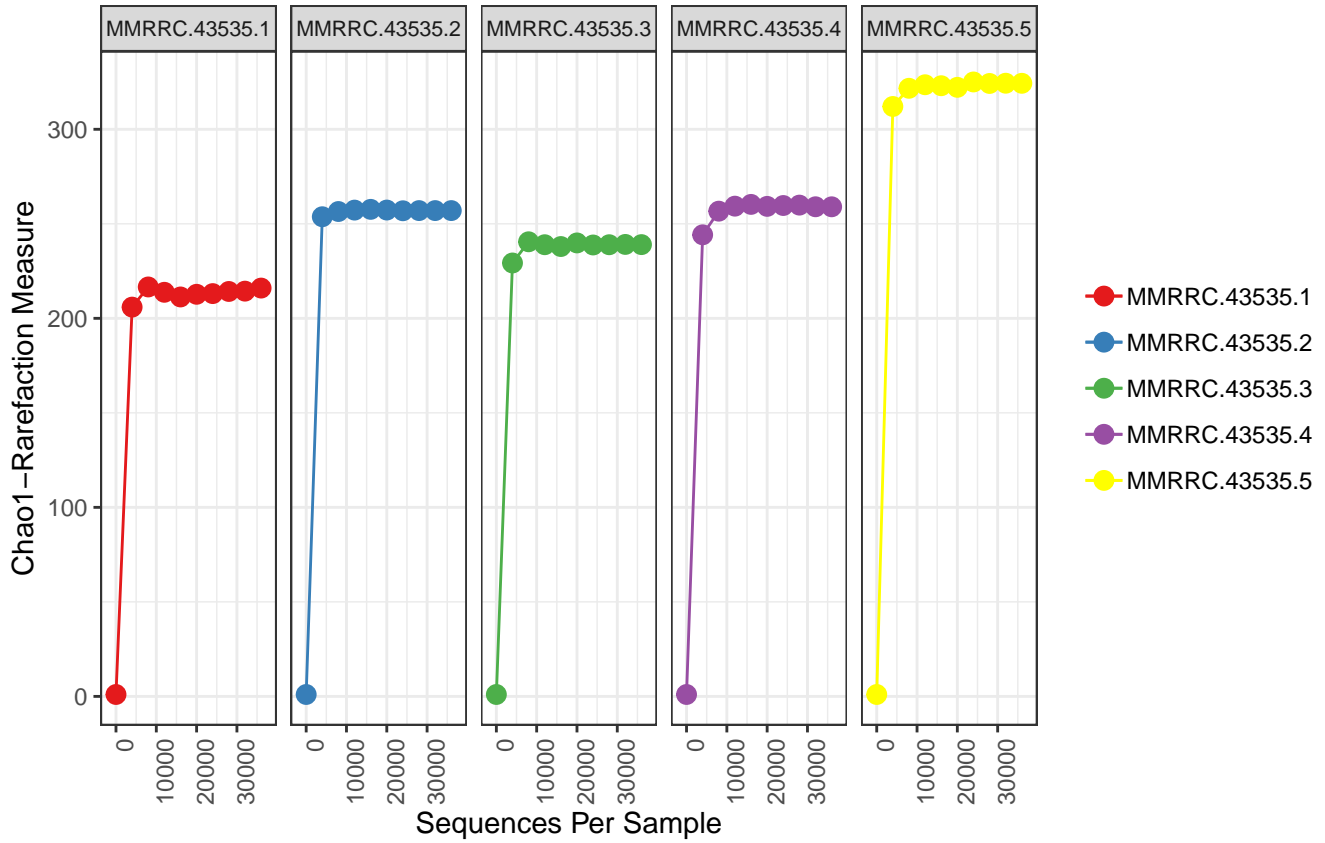


Figure 1: Chao1 Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

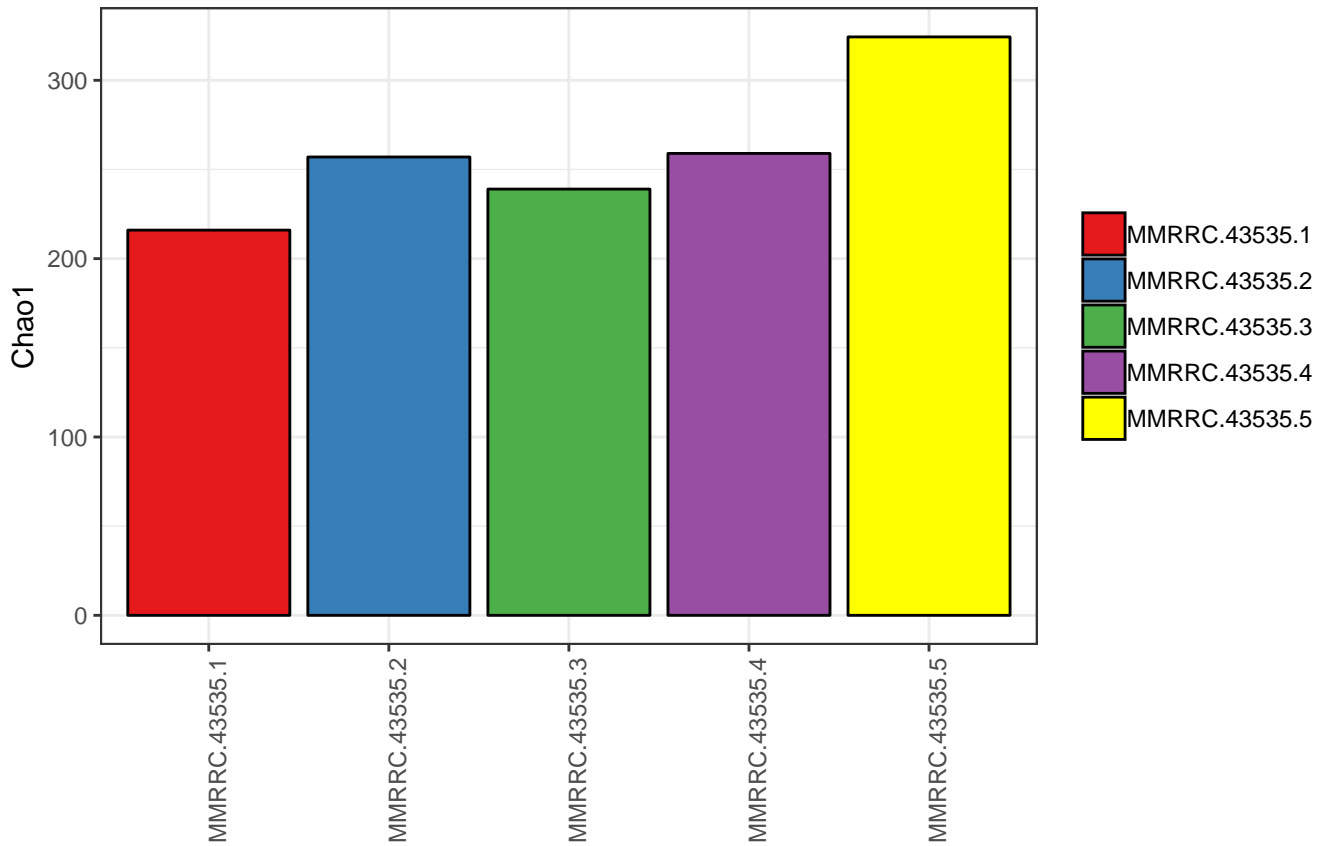


Figure 2: Chao1 Measure of Alpha Diversity- Boxplots-Animal_ID.

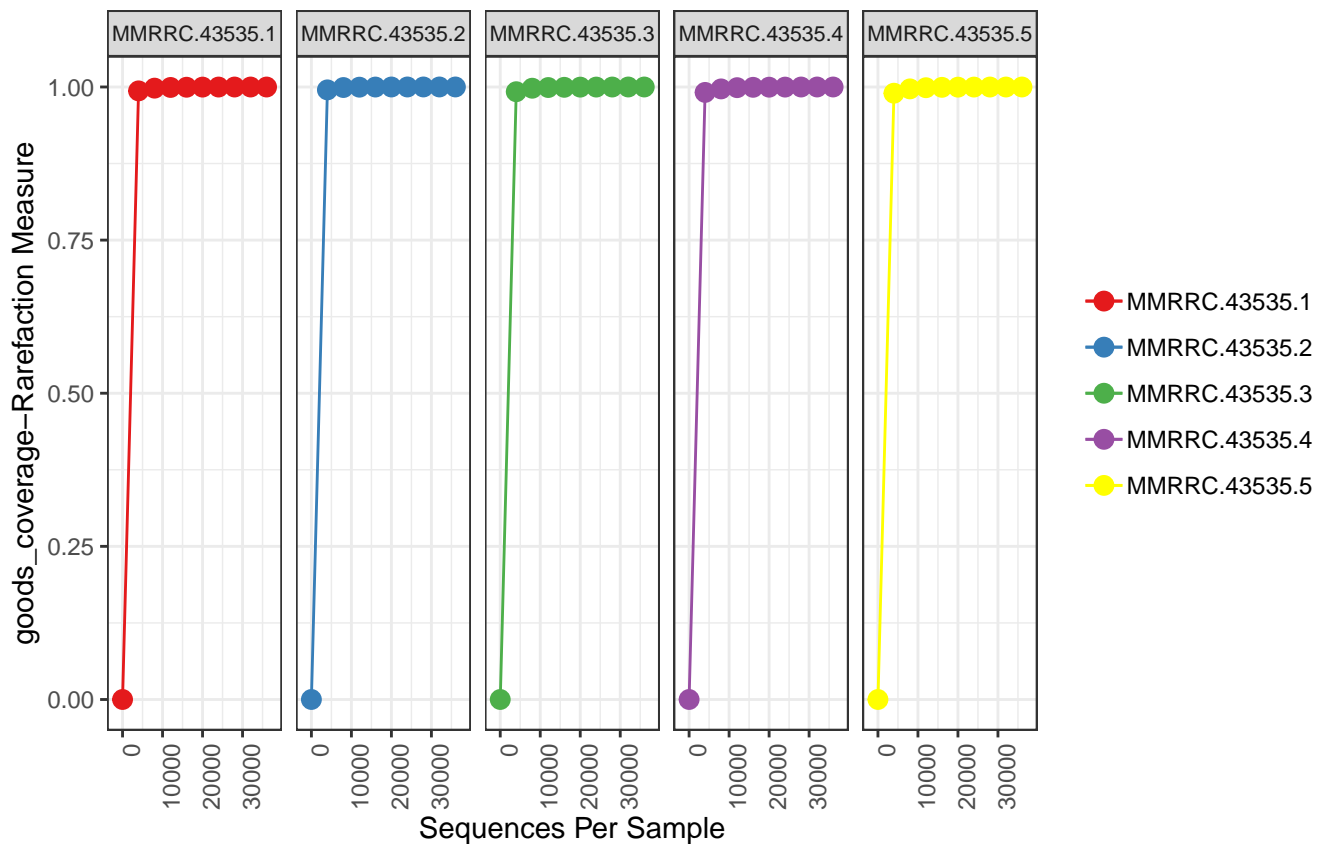


Figure 3: Goods Coverage Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

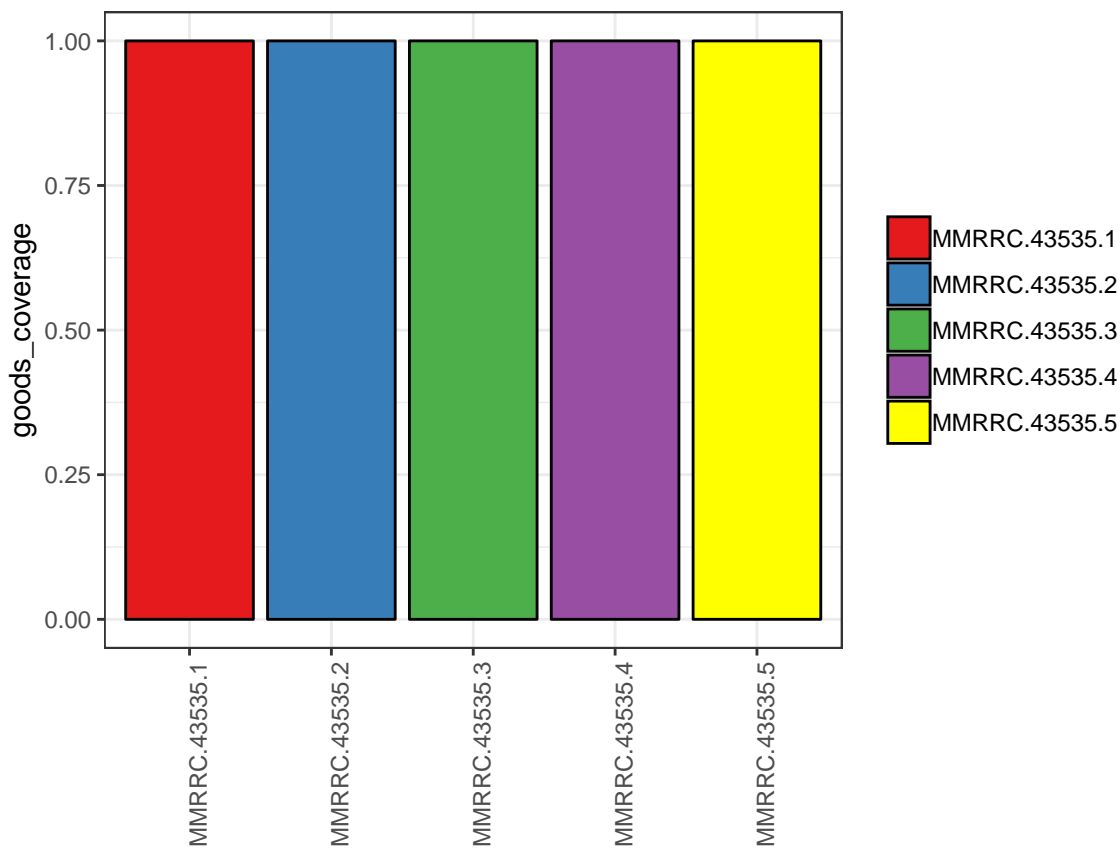


Figure 4: Goods Coverage Measure of Alpha Diversity- Boxplots-Animal_ID.

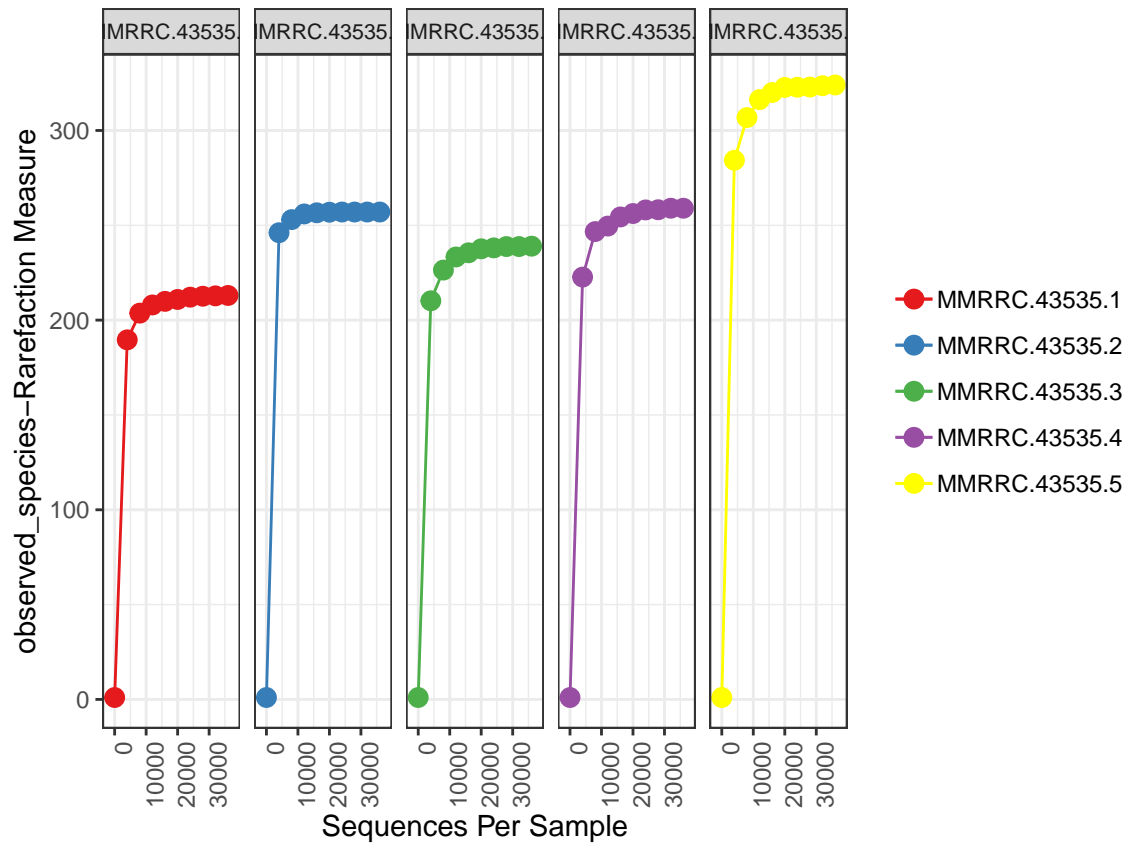


Figure 5: Observed Species Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

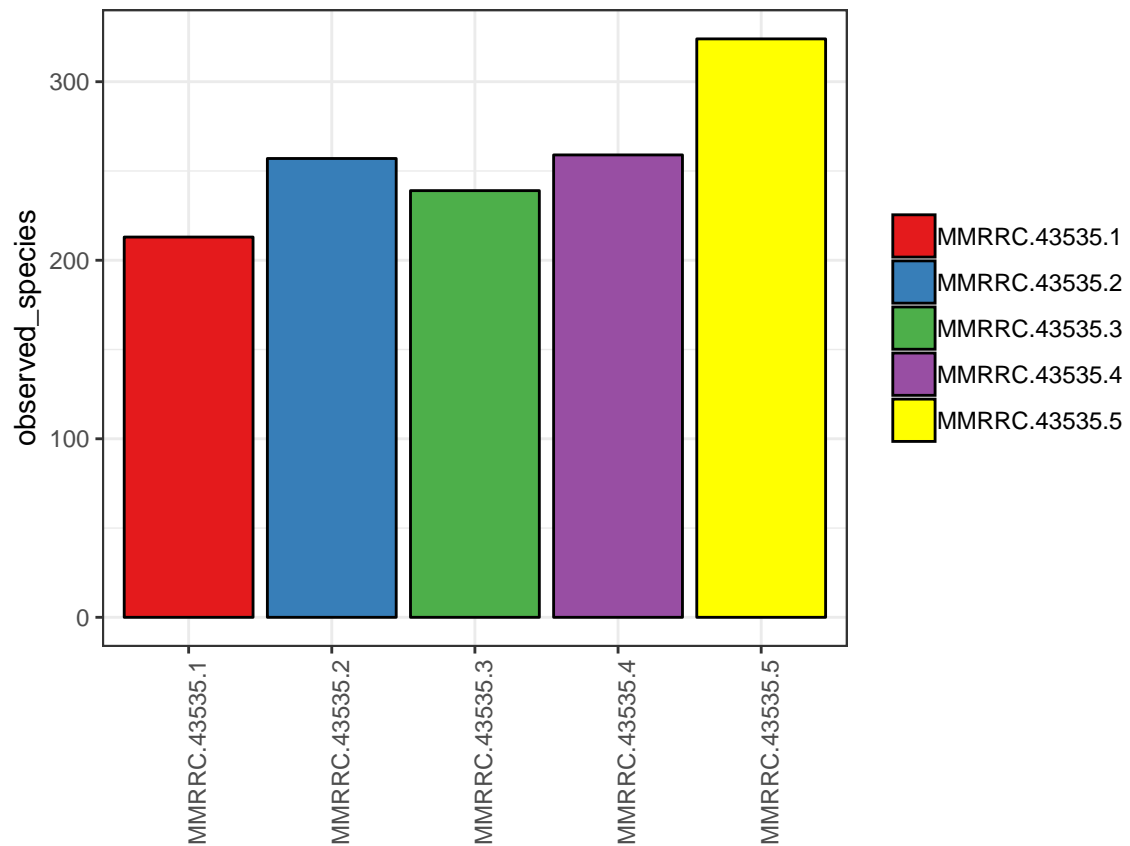


Figure 6: Observed Species Measure of Alpha Diversity- Boxplot-Animal_ID.

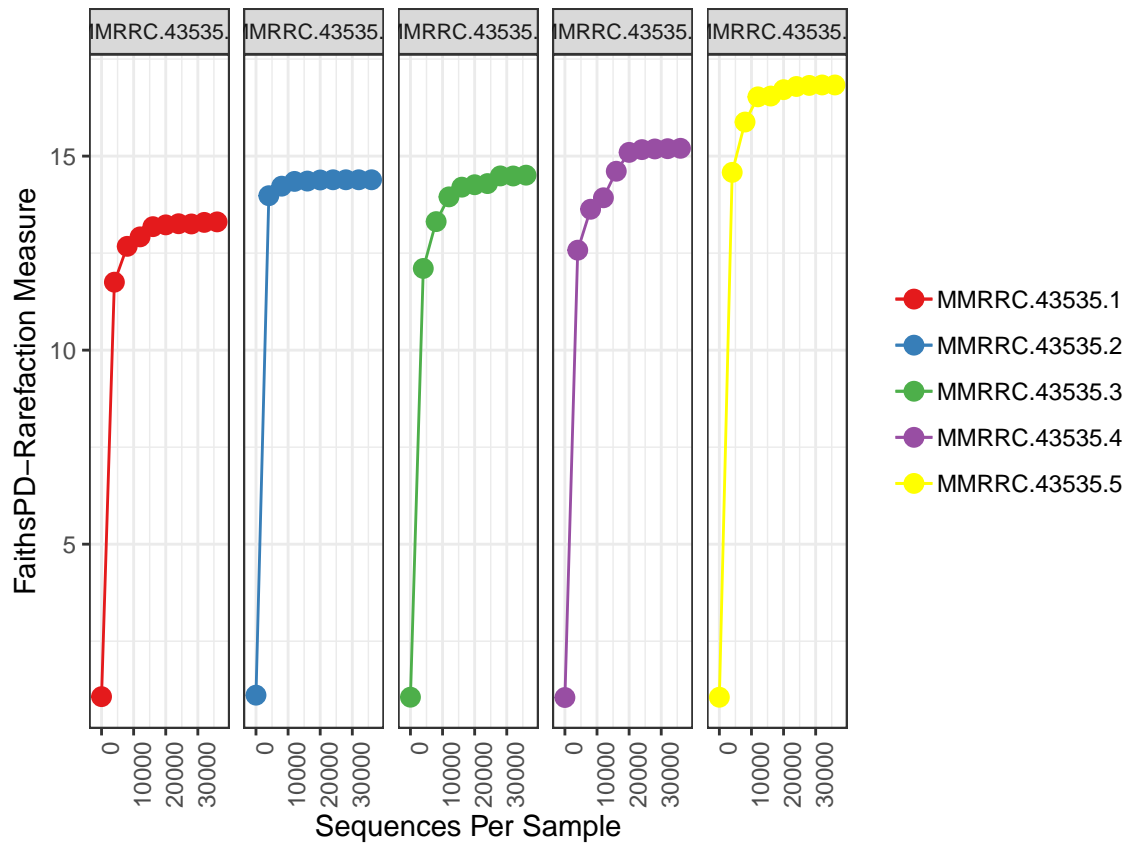


Figure 7: Faith's Phylogenetic Diversity Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

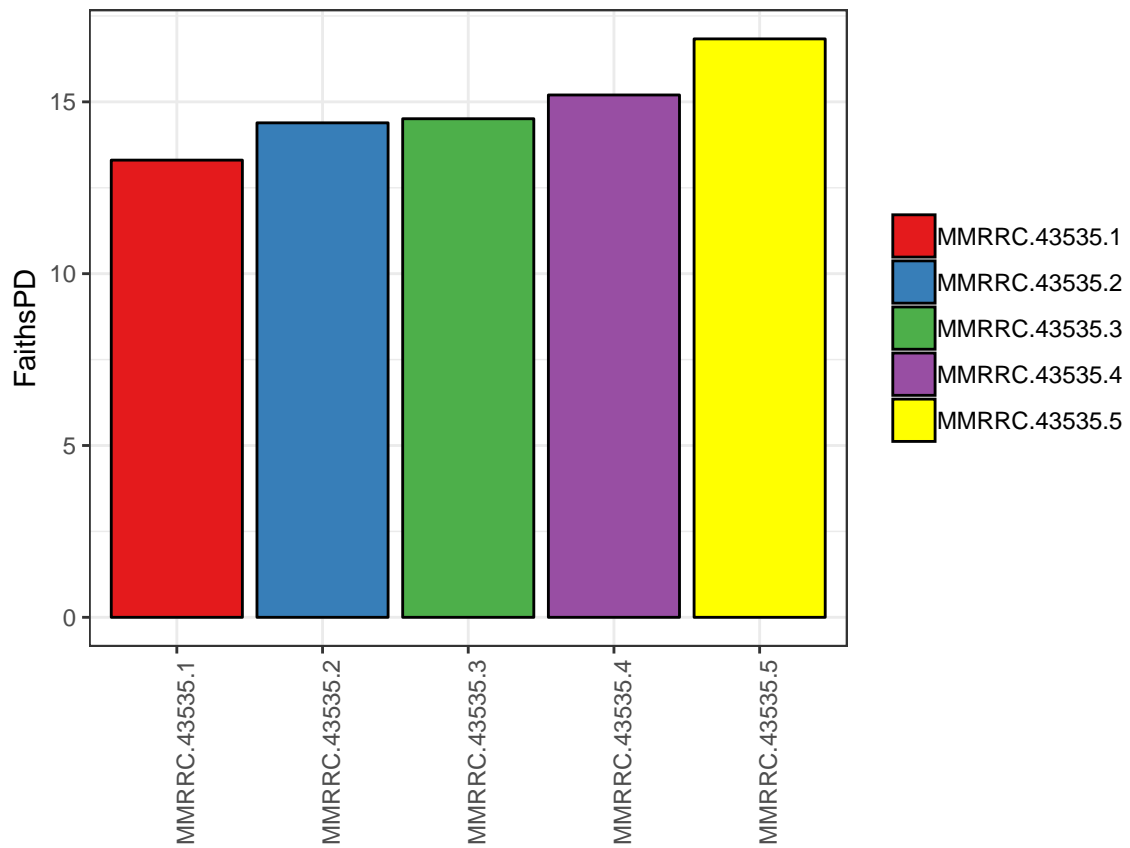


Figure 8: Faith's Phylogenetic Diversity Measure of Alpha Diversity- Boxplots-Animal_ID.

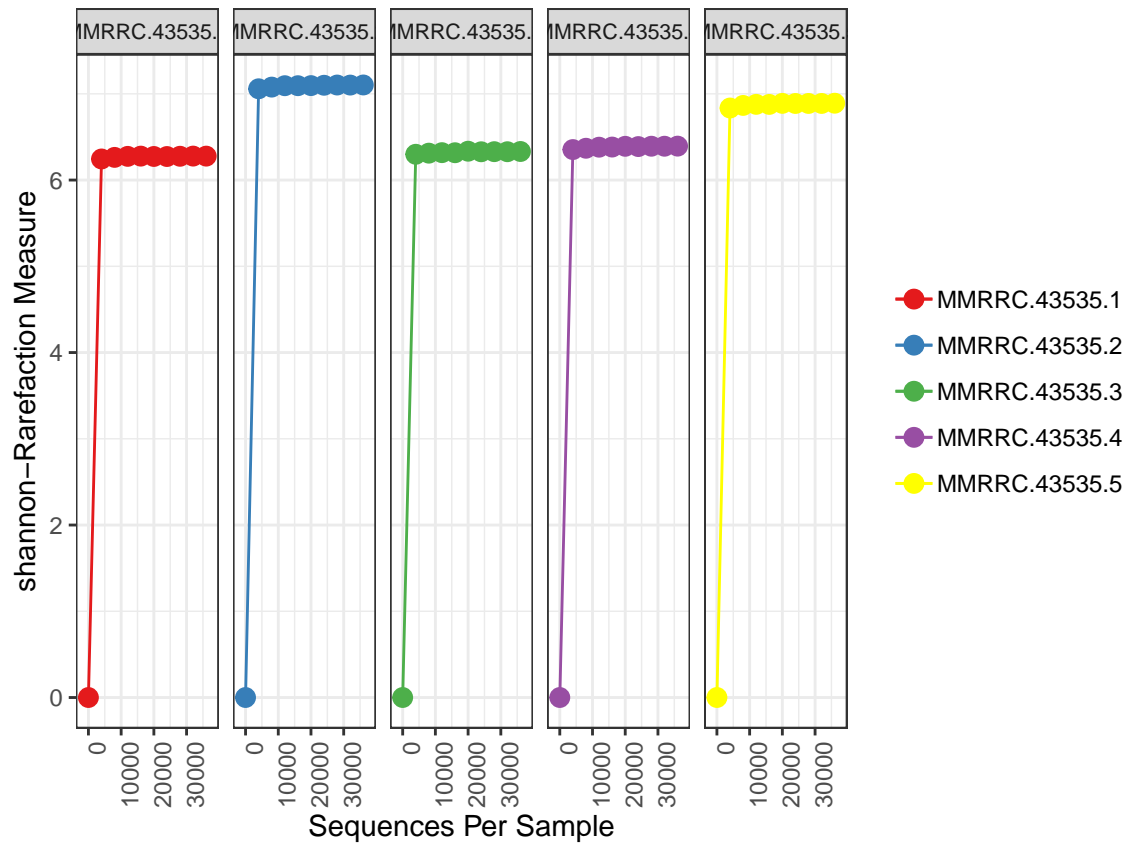


Figure 9: Shannon Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

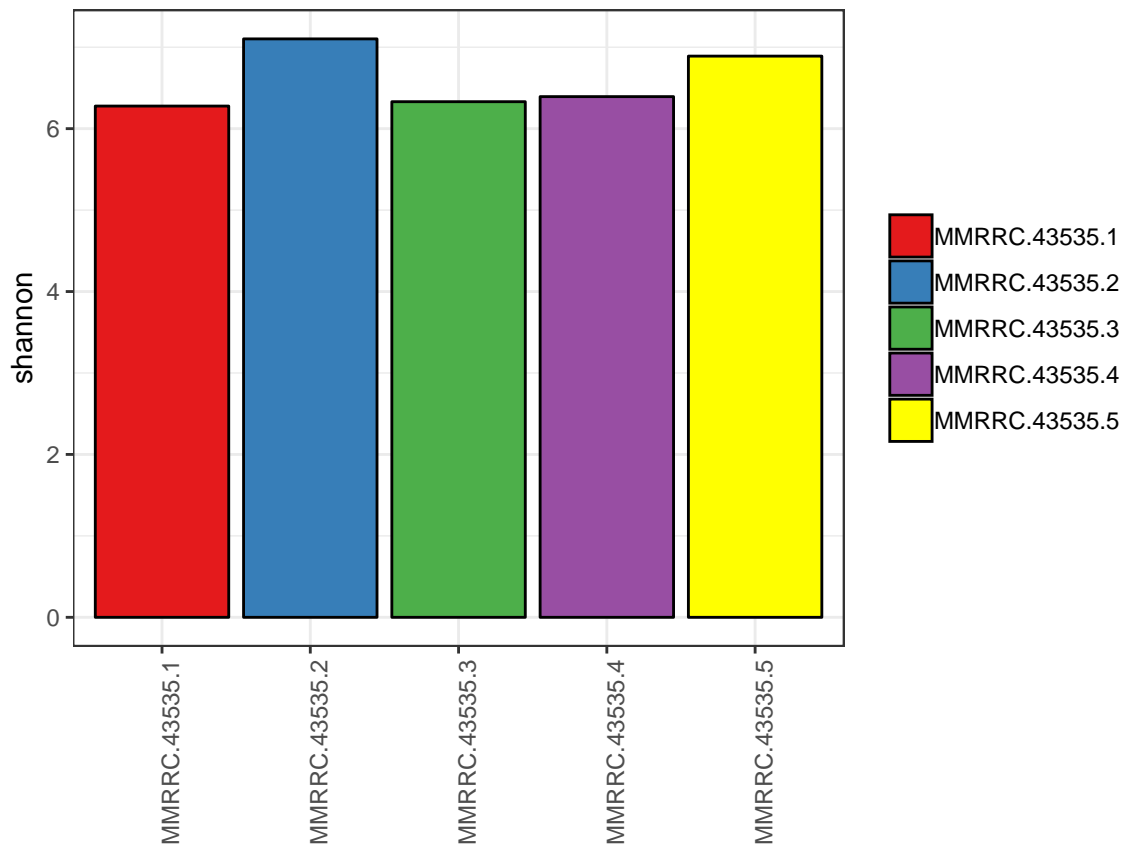


Figure 10: Shannon Measure of Alpha Diversity- Boxplots-Animal_ID.

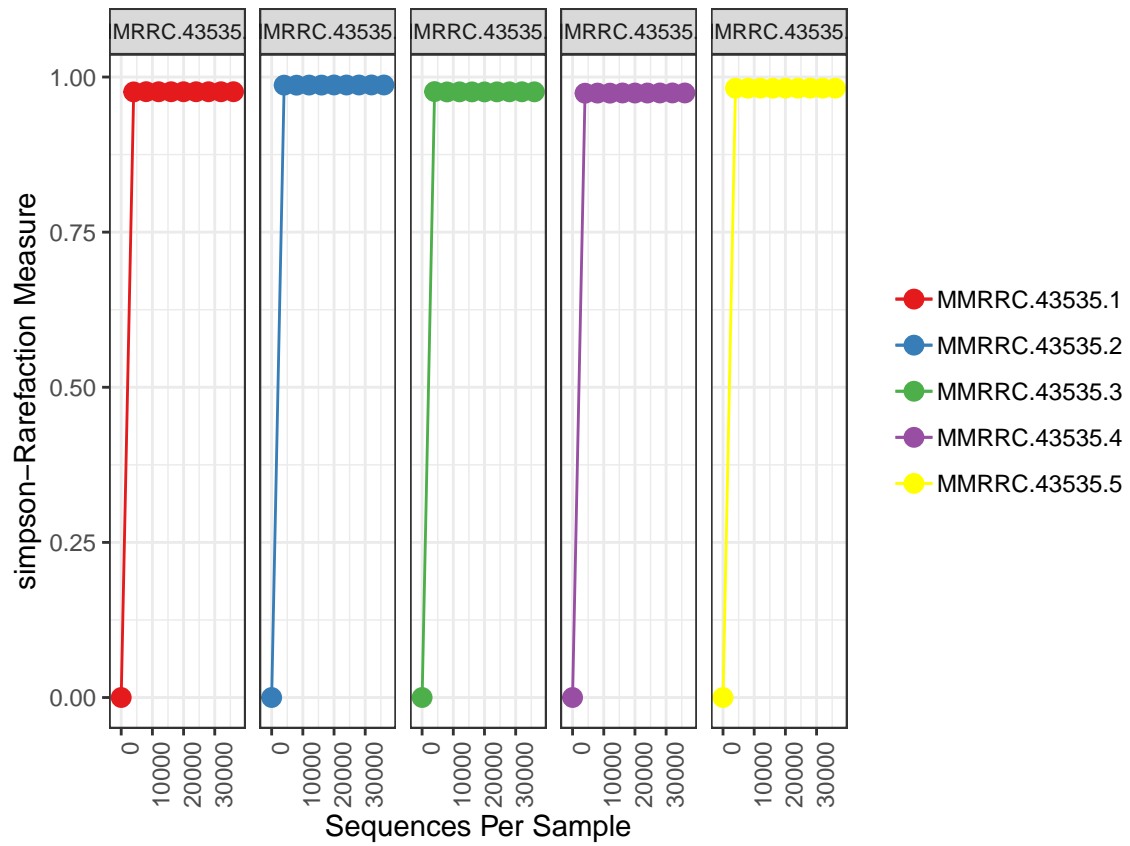


Figure 11: Simpson Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

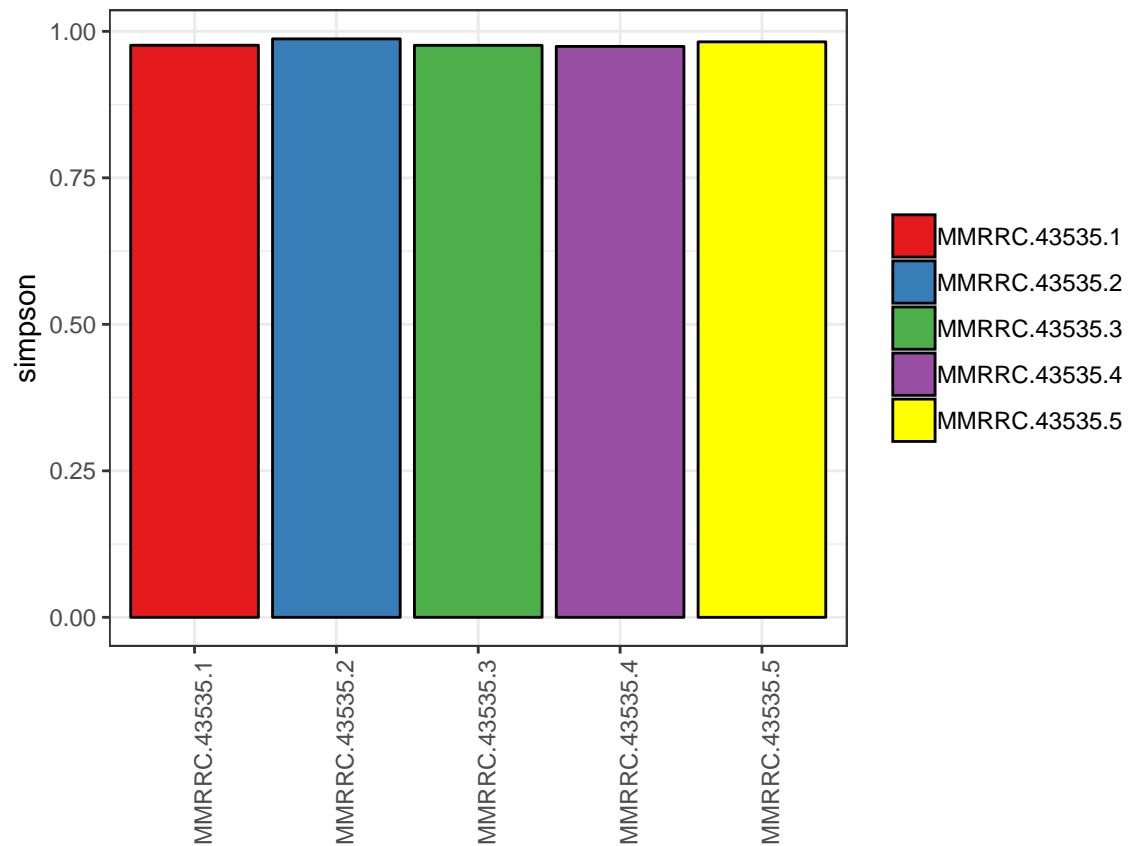


Figure 12: Simpson Measure of Alpha Diversity- Boxplots-Animal_ID.

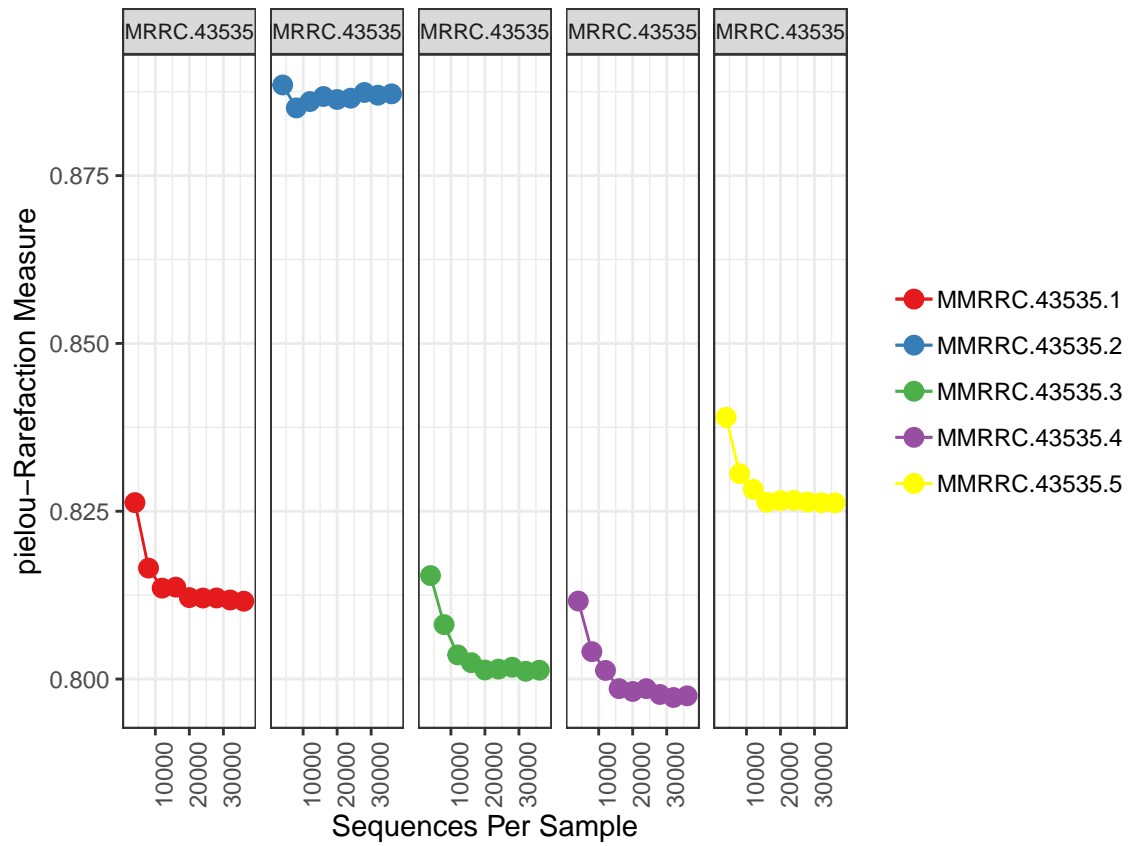


Figure 13: Pielou's Evenness Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

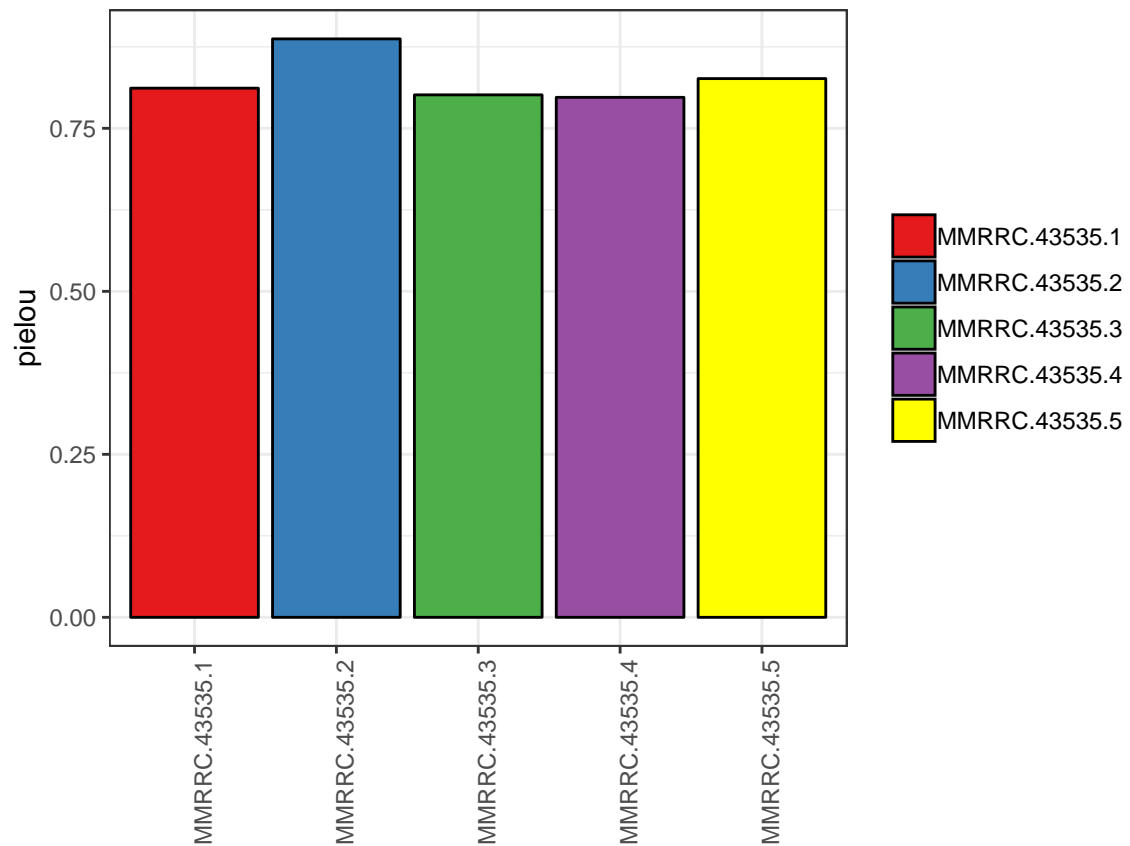


Figure 14: Pielou's Evenness Measure of Alpha Diversity- Boxplots-Animal_ID.

2.2.2 Beta Diversity

Beta diversity is the diversity between different treatment groups. Beta diversity analyses include the phylogenetic measure of unweighted and weighted UniFrac distances and the non-phylogenetic Bray-Curtis and Jaccard methods. Principal coordinates analysis (PCoA) of the first 2 coordinates of the distance matrices are plotted. The shaded ellipse depicts the 95% confidence interval. For more information on the UniFrac distance metric for community comparisons, please see (Lozupone et al., 2011). Comparing the sample beta diversity was performed with PERMANOVA (anderson).

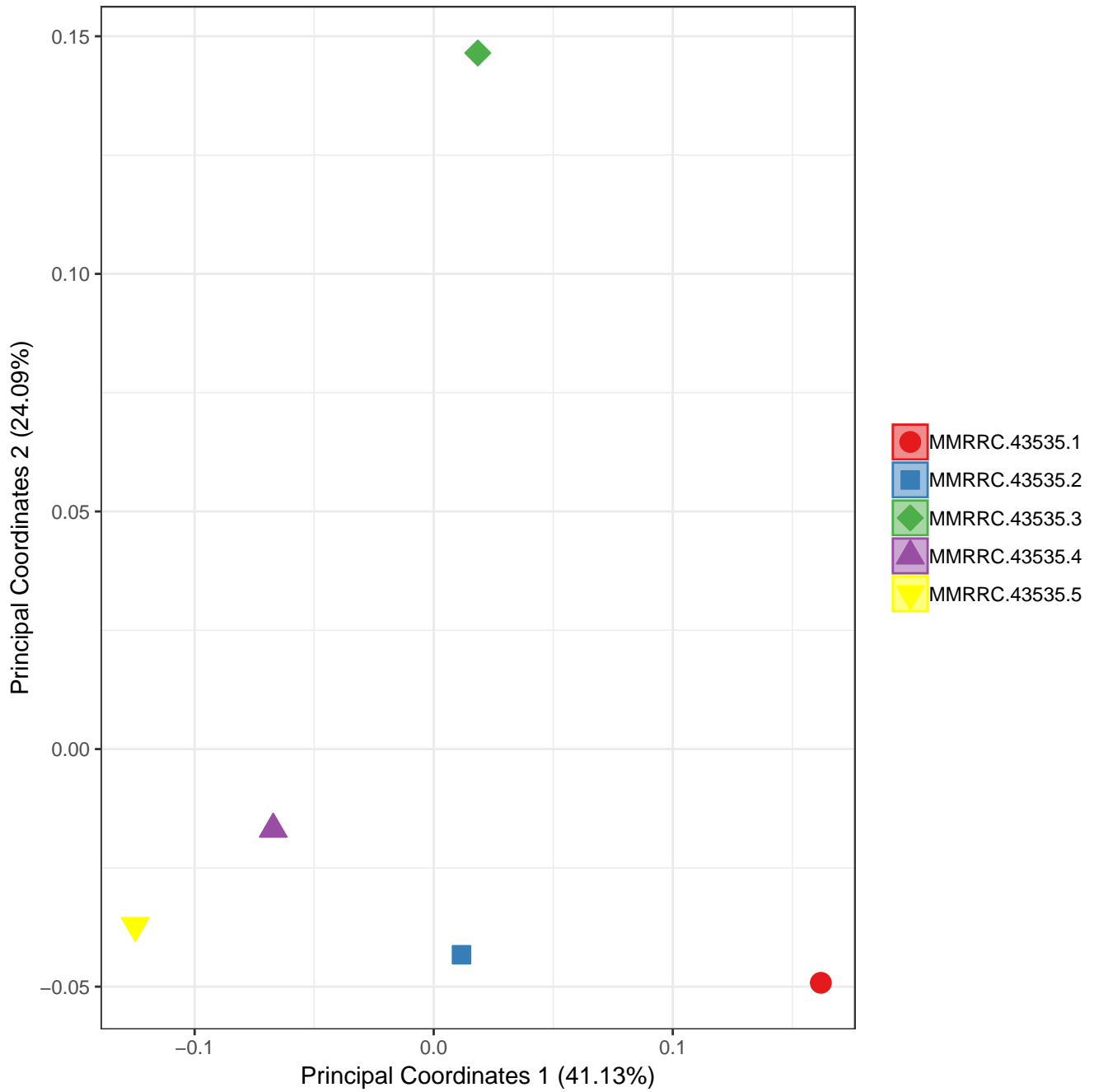


Figure 15: Unweighted UniFrac Measure of Beta Diversity-Animal_ID.

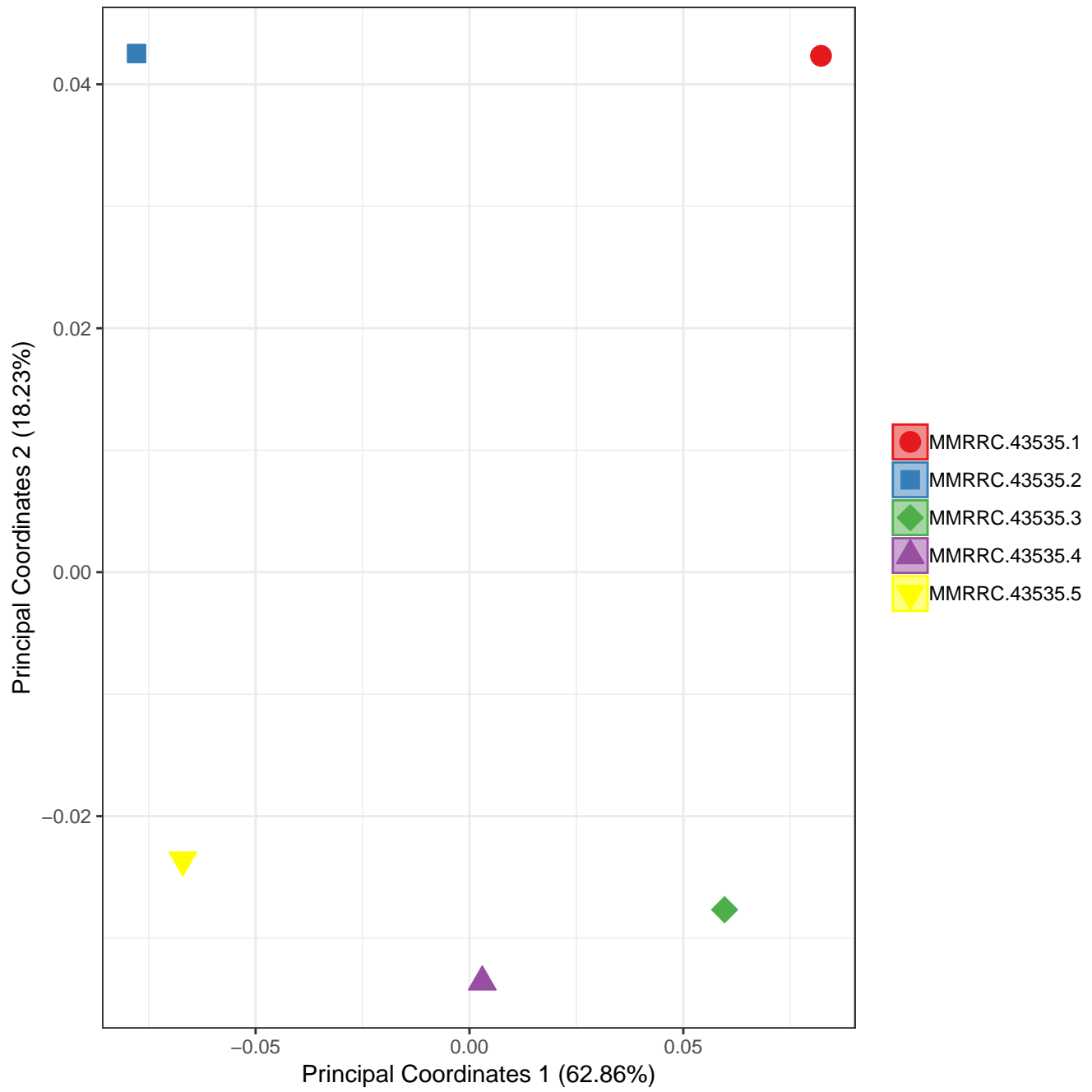


Figure 16: Weighted Unifrac Measure of Beta Diversity-Animal_ID.

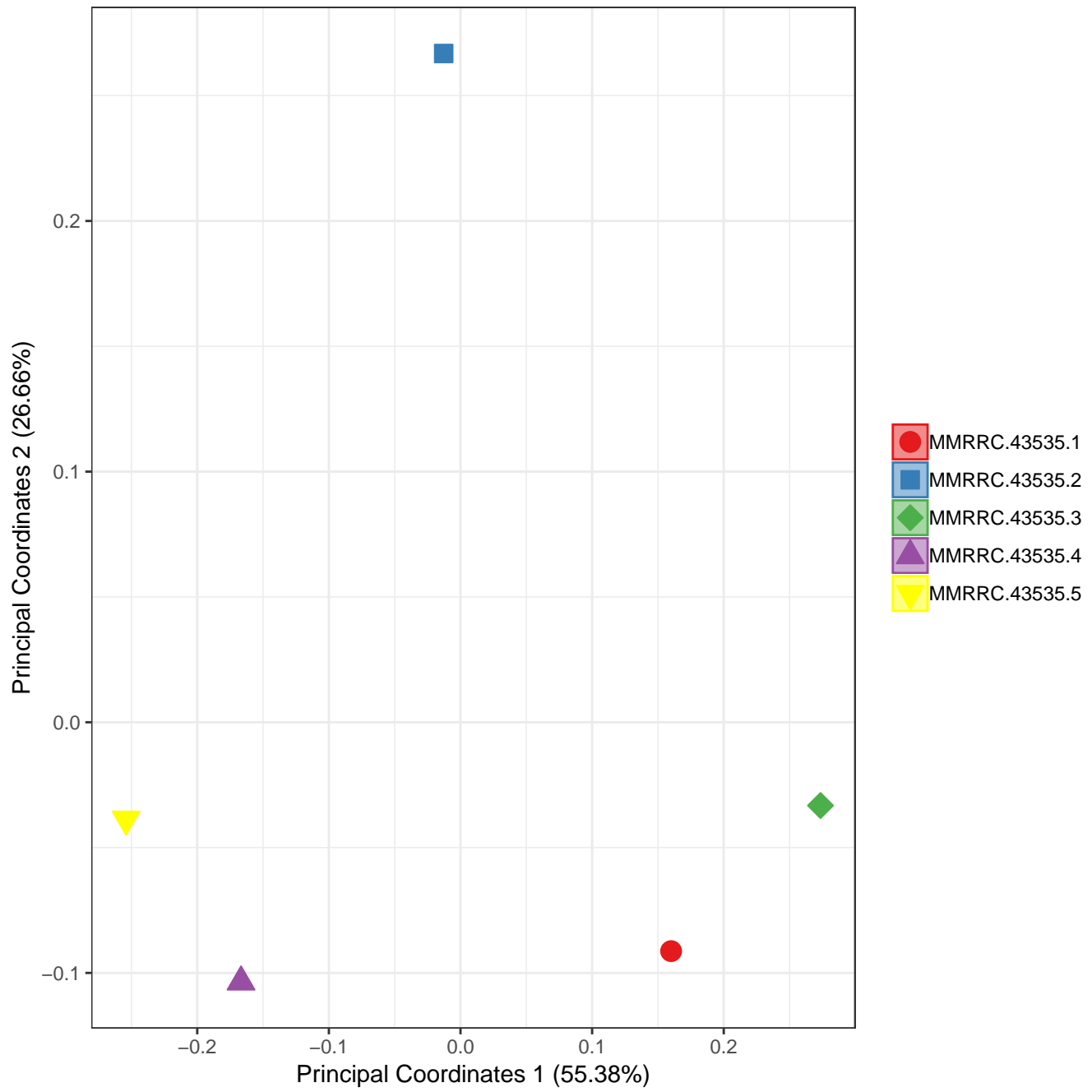


Figure 17: Bray-Curtis Measure of Beta Diversity-Animal_ID.

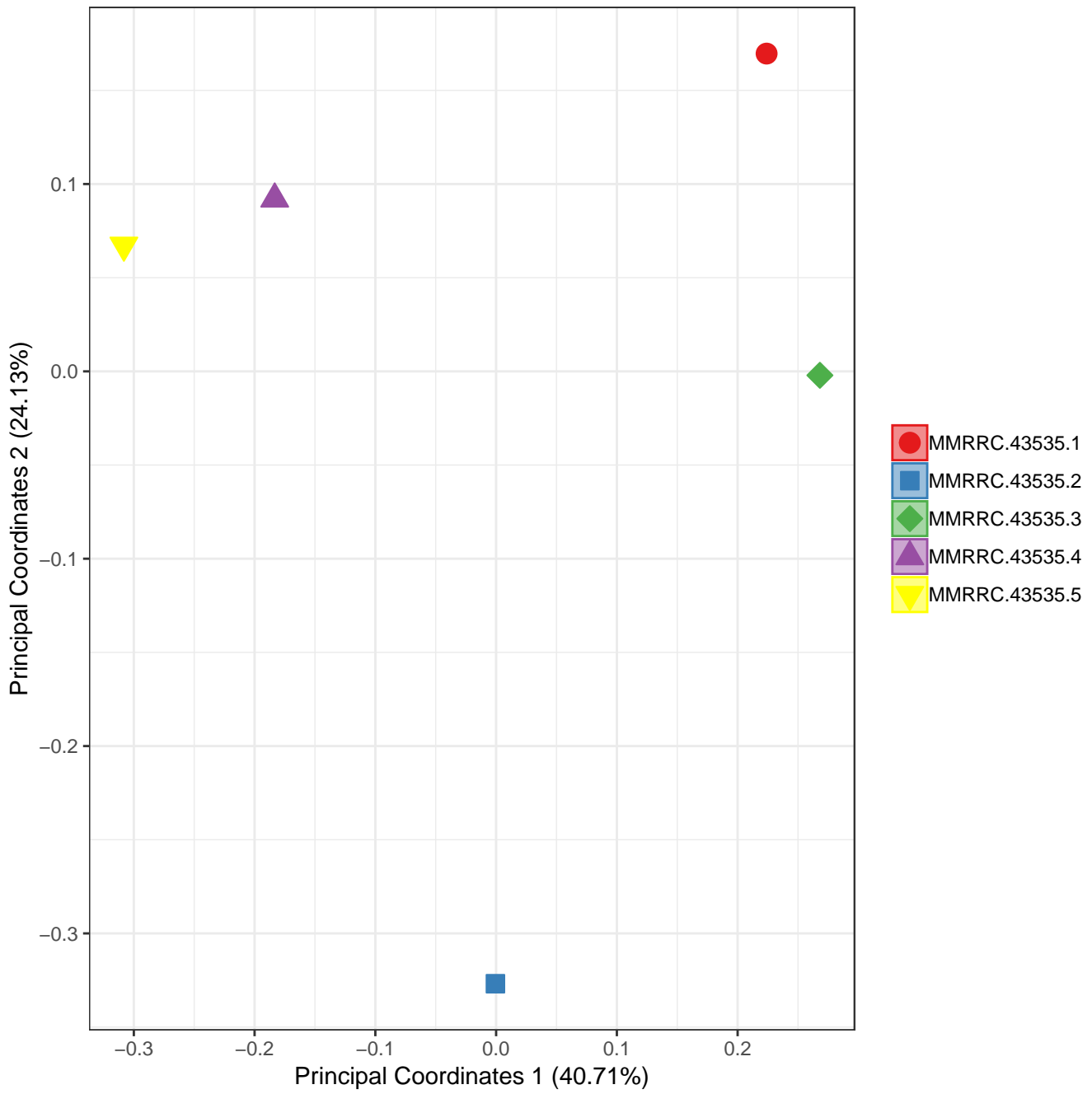


Figure 18: Jaccard Measure of Beta Diversity-Animal_ID.

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.

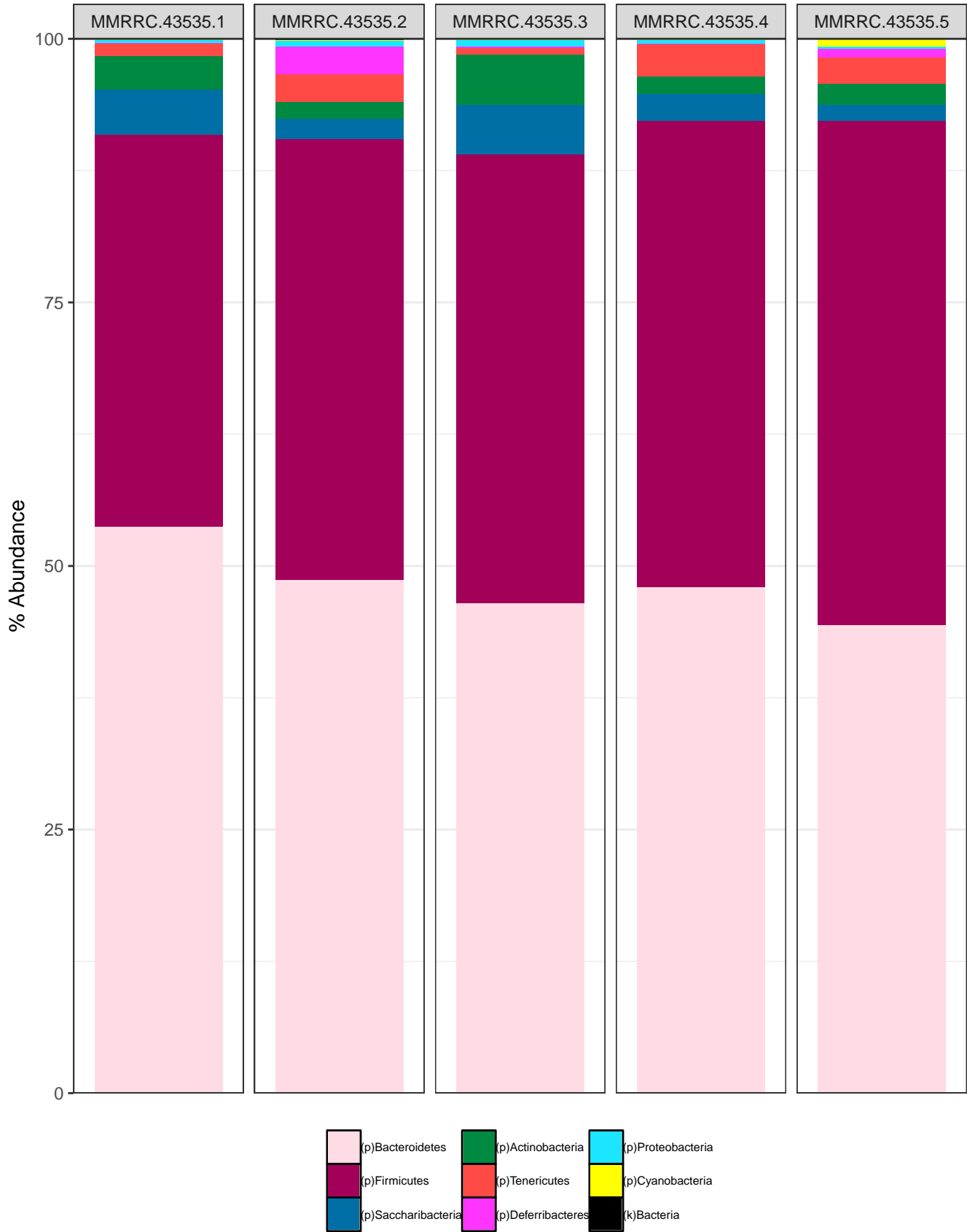


Figure 19: Phylum level abundances-UNFILTERED

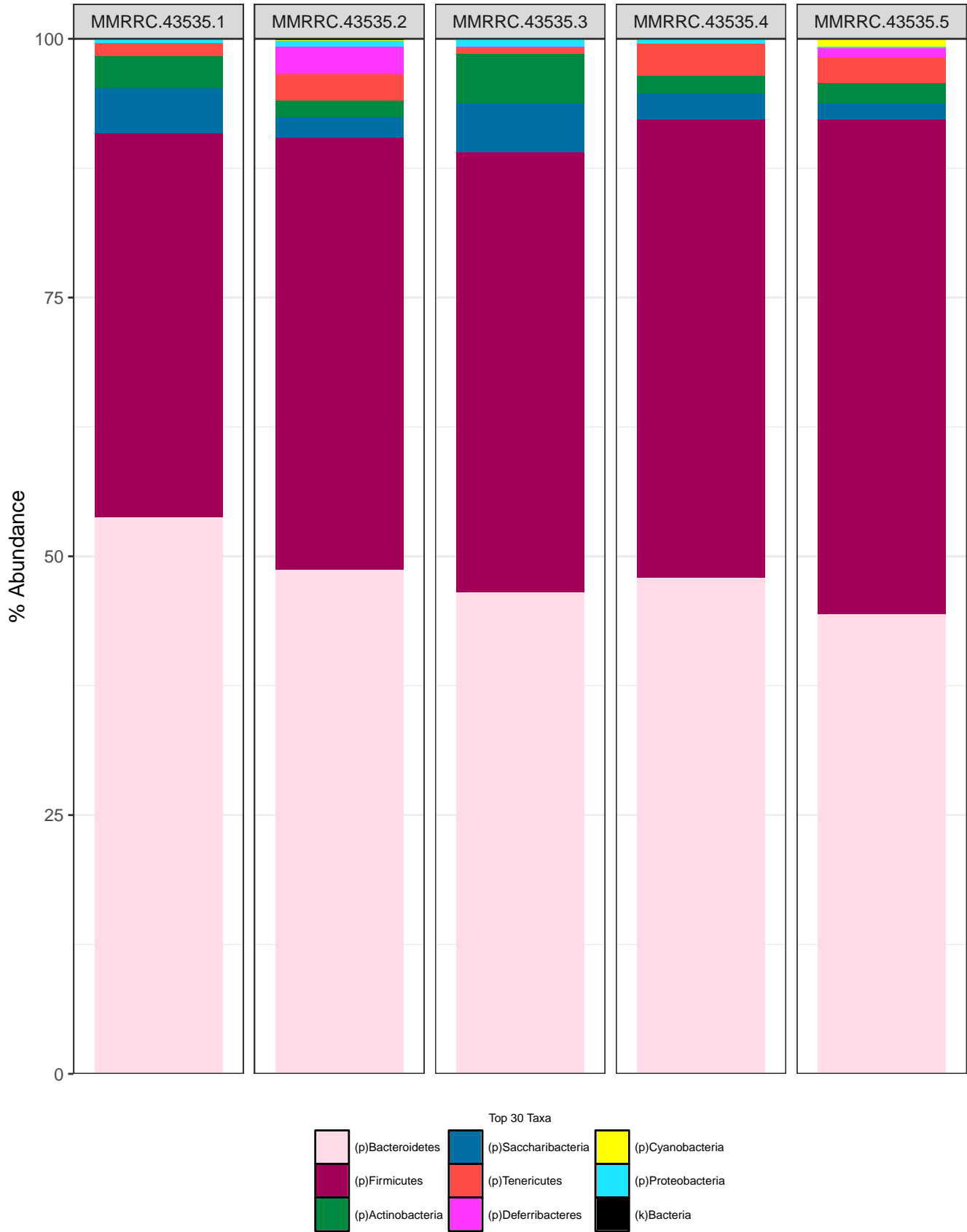


Figure 20: Phylum level abundances-FILTERED: 0 phylum level taxa removed

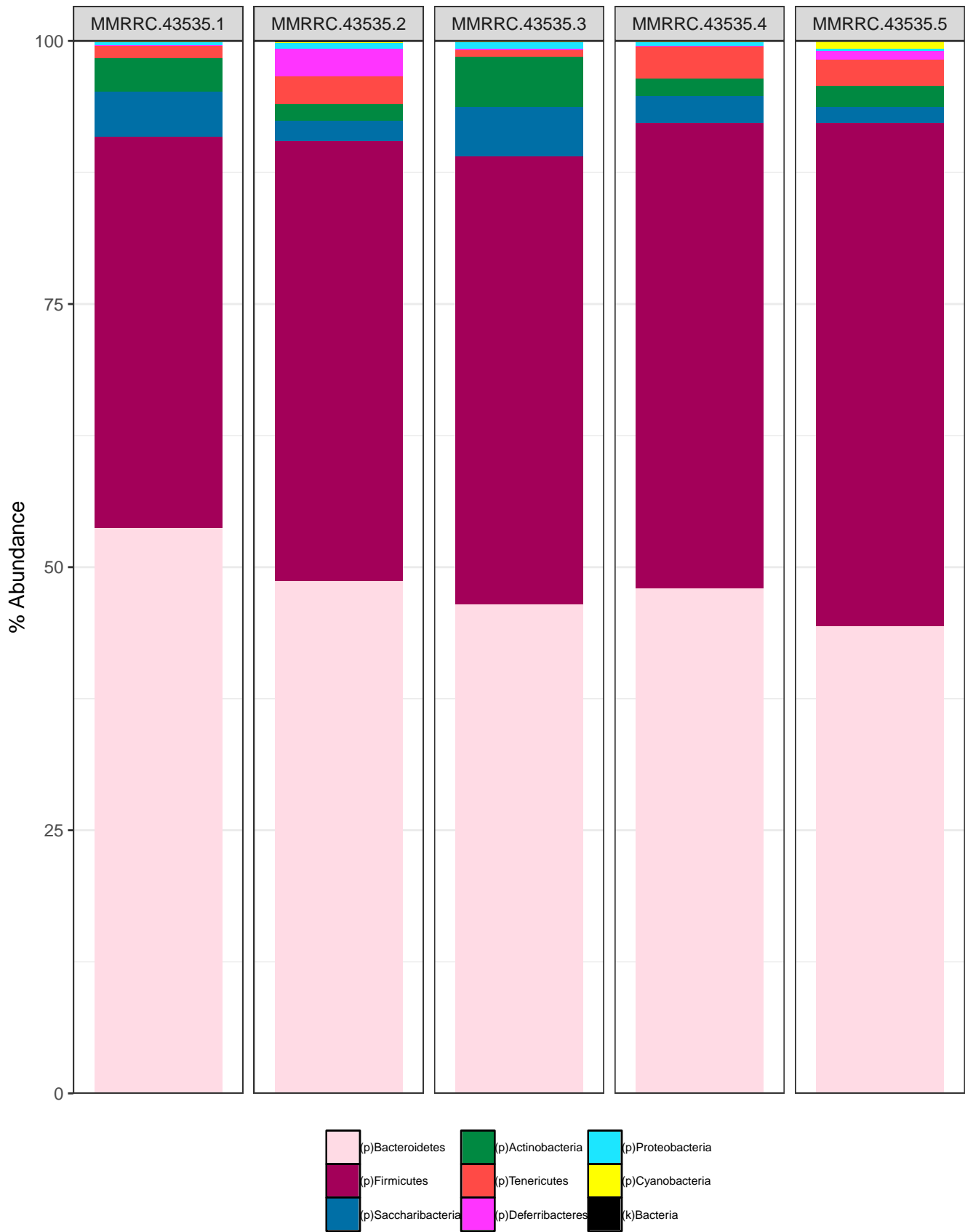


Figure 21: Phylum level abundances-FILTERED & RESCALED

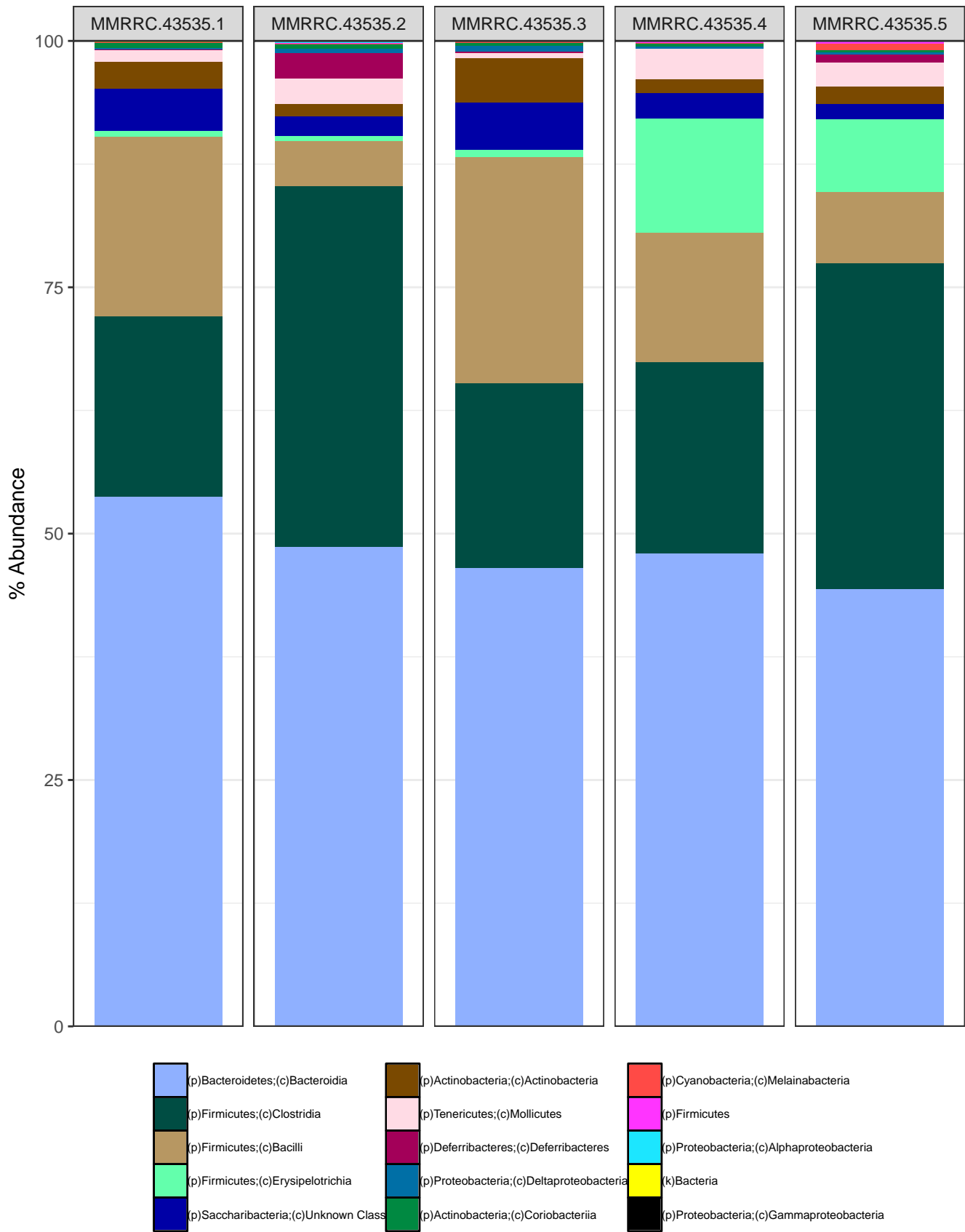


Figure 22: Class level abundances-UNFILTERED

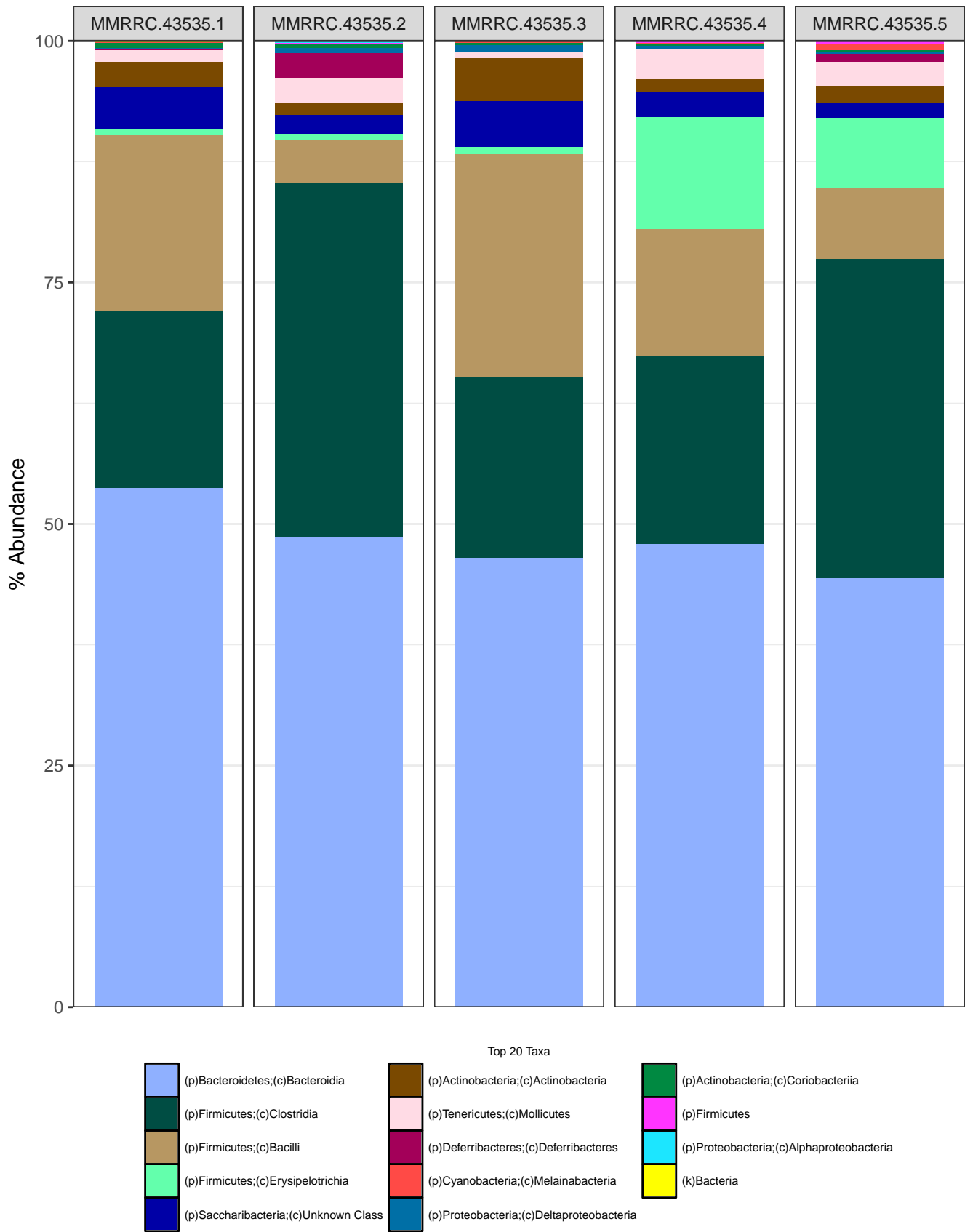


Figure 23: Class level abundances-FILTERED: 1 class level taxa removed

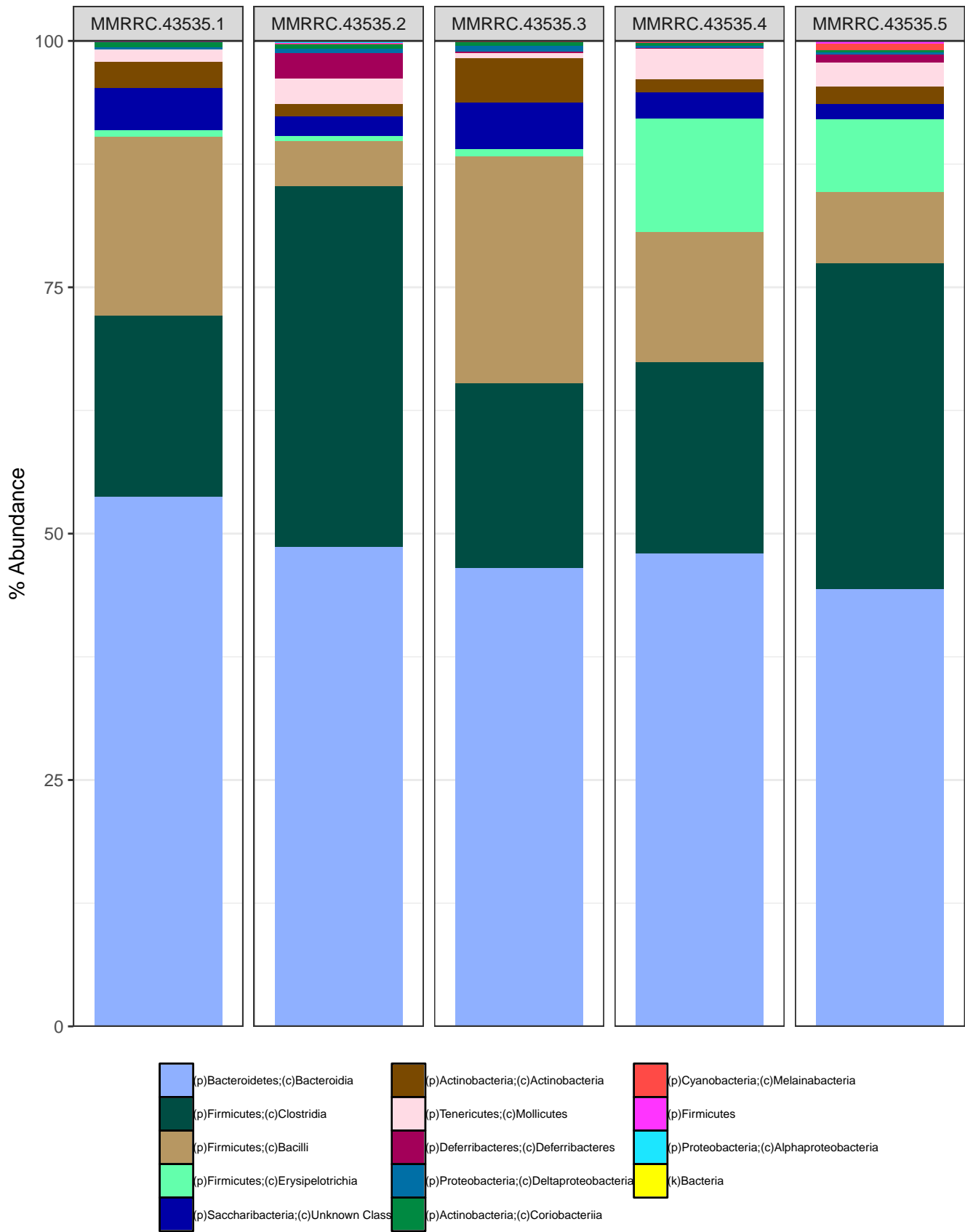


Figure 24: Class level abundances-FILTERED & RESCALED

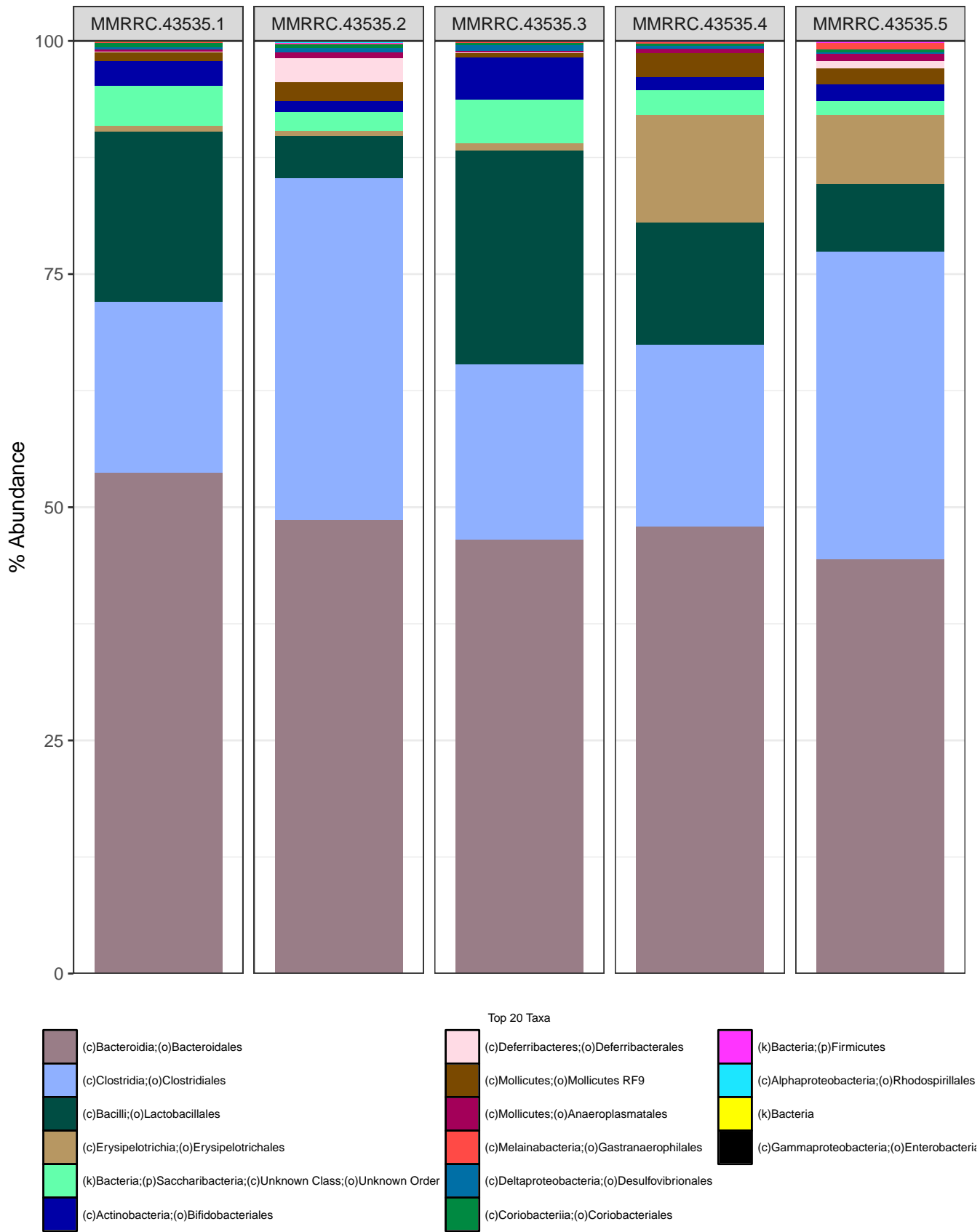


Figure 25: Order level abundances-UNFILTERED

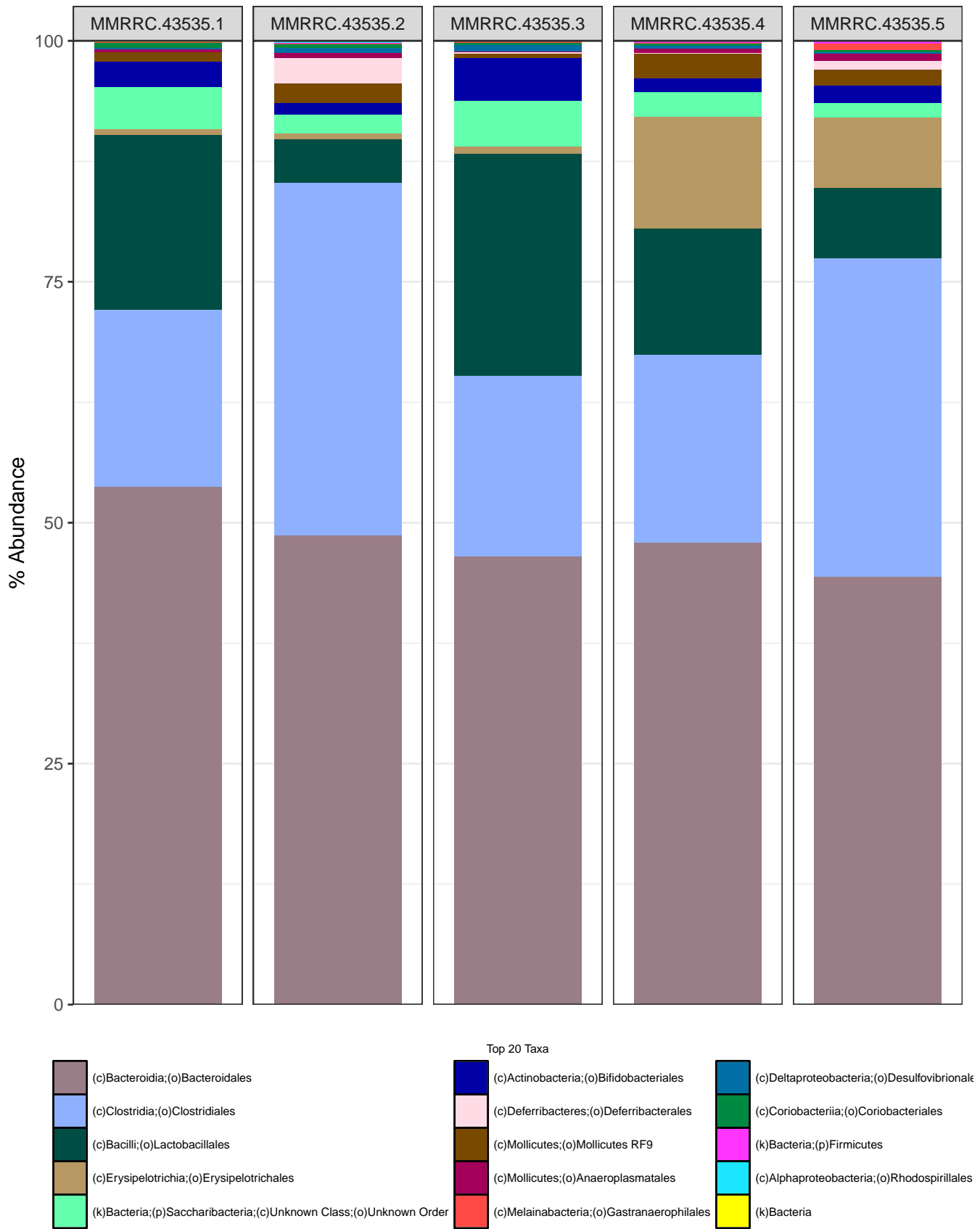


Figure 26: Order level abundances-FILTERED: 1 order level taxa removed

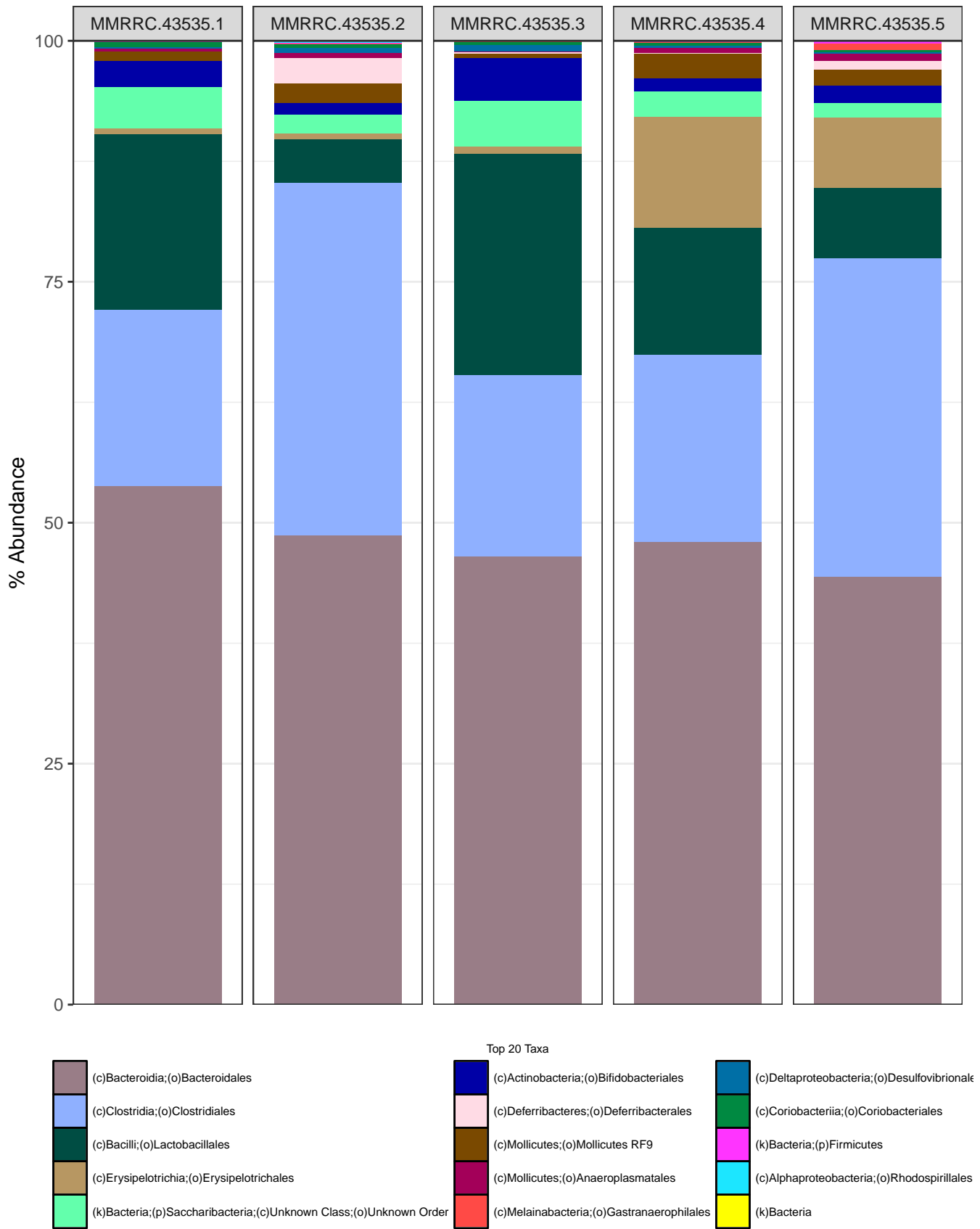


Figure 27: Order level abundances-FILTERED & RESCALED

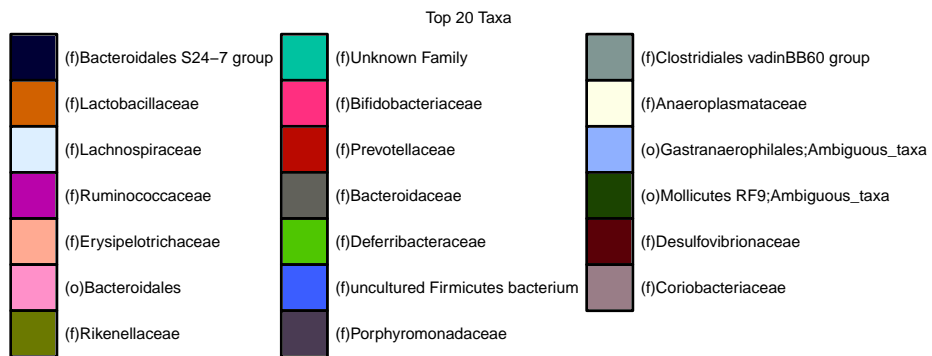
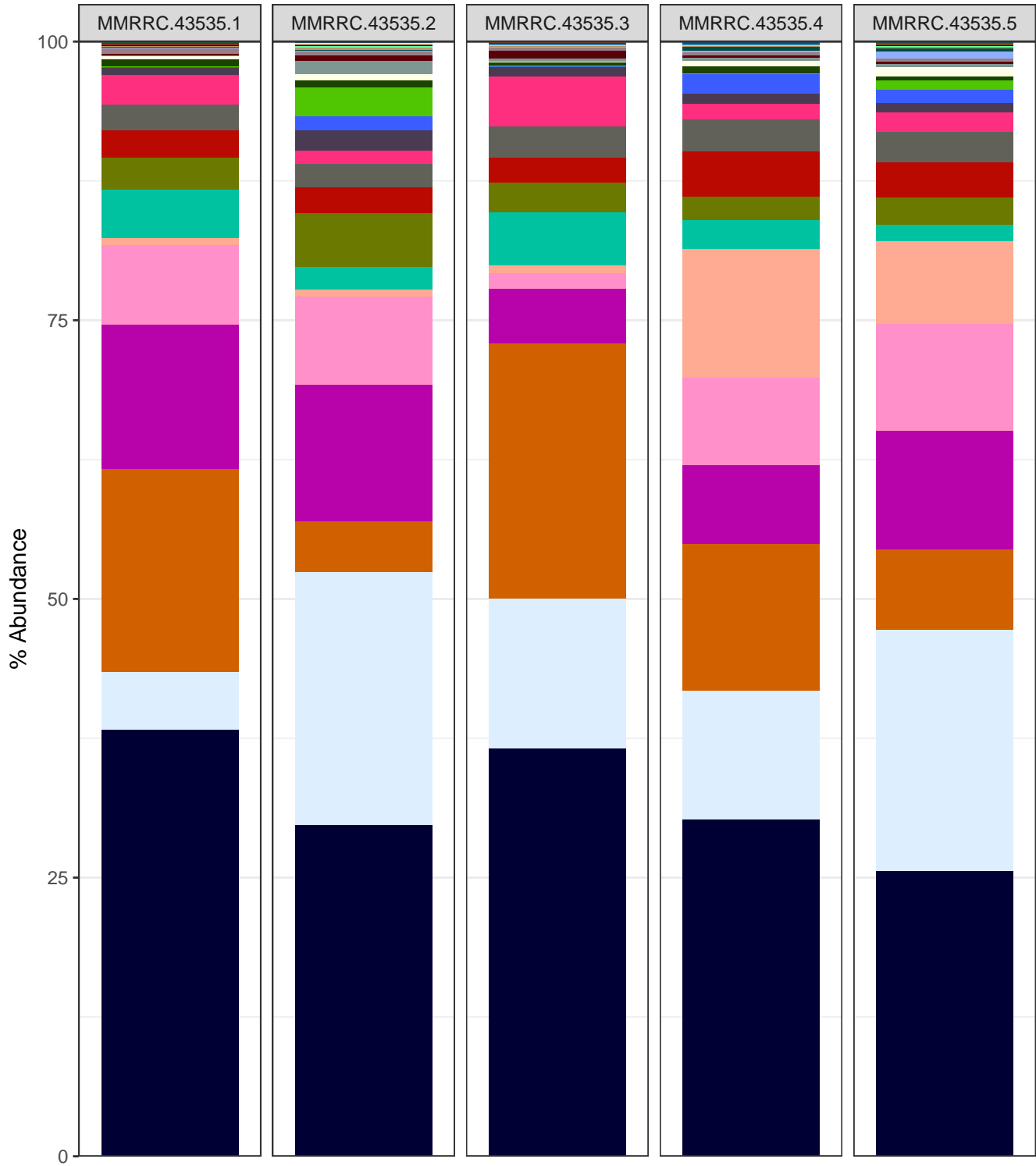


Figure 28: Family level abundances-UNFILTERED

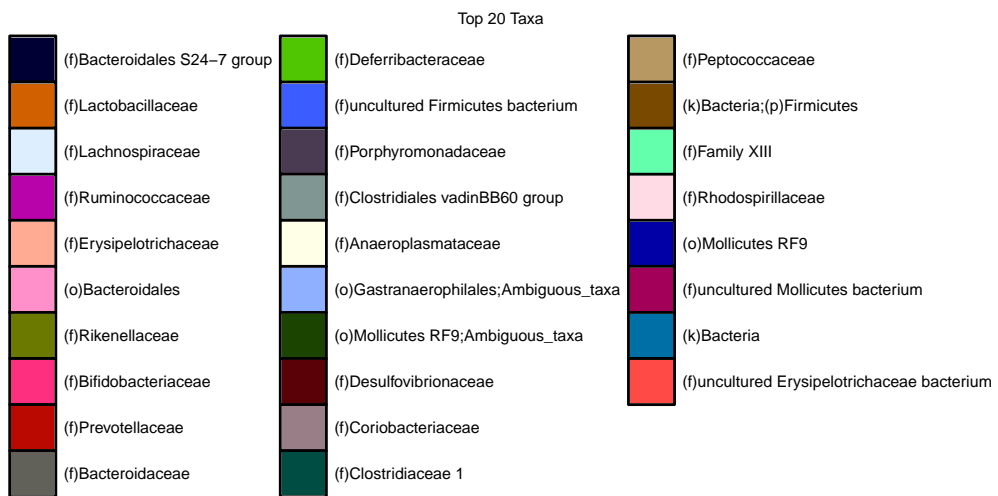
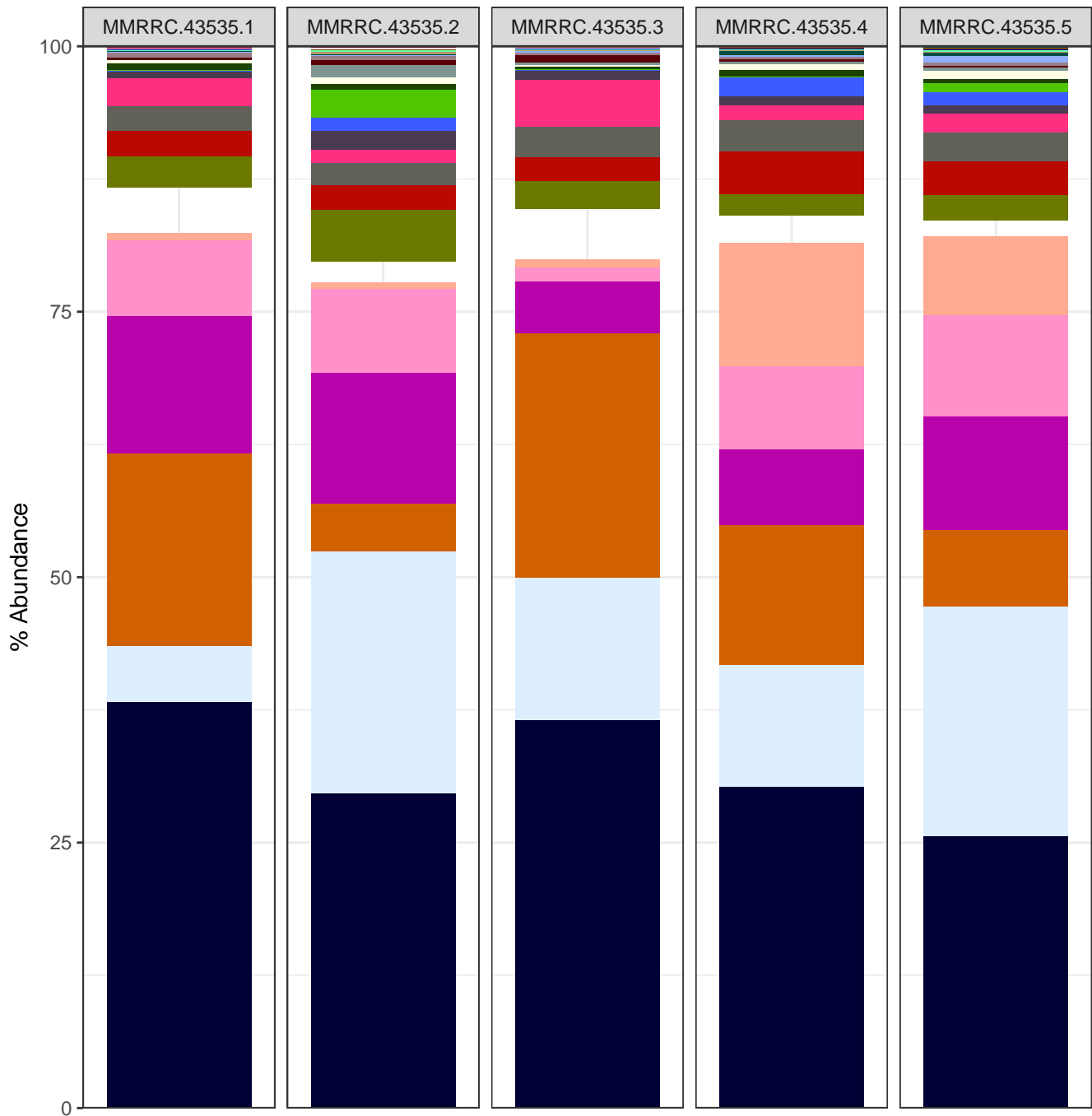


Figure 29: Family level abundances-FILTERED: 6 family level taxa removed

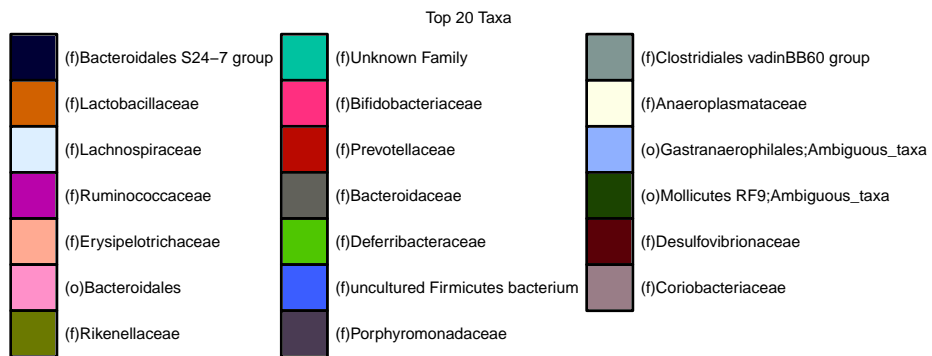
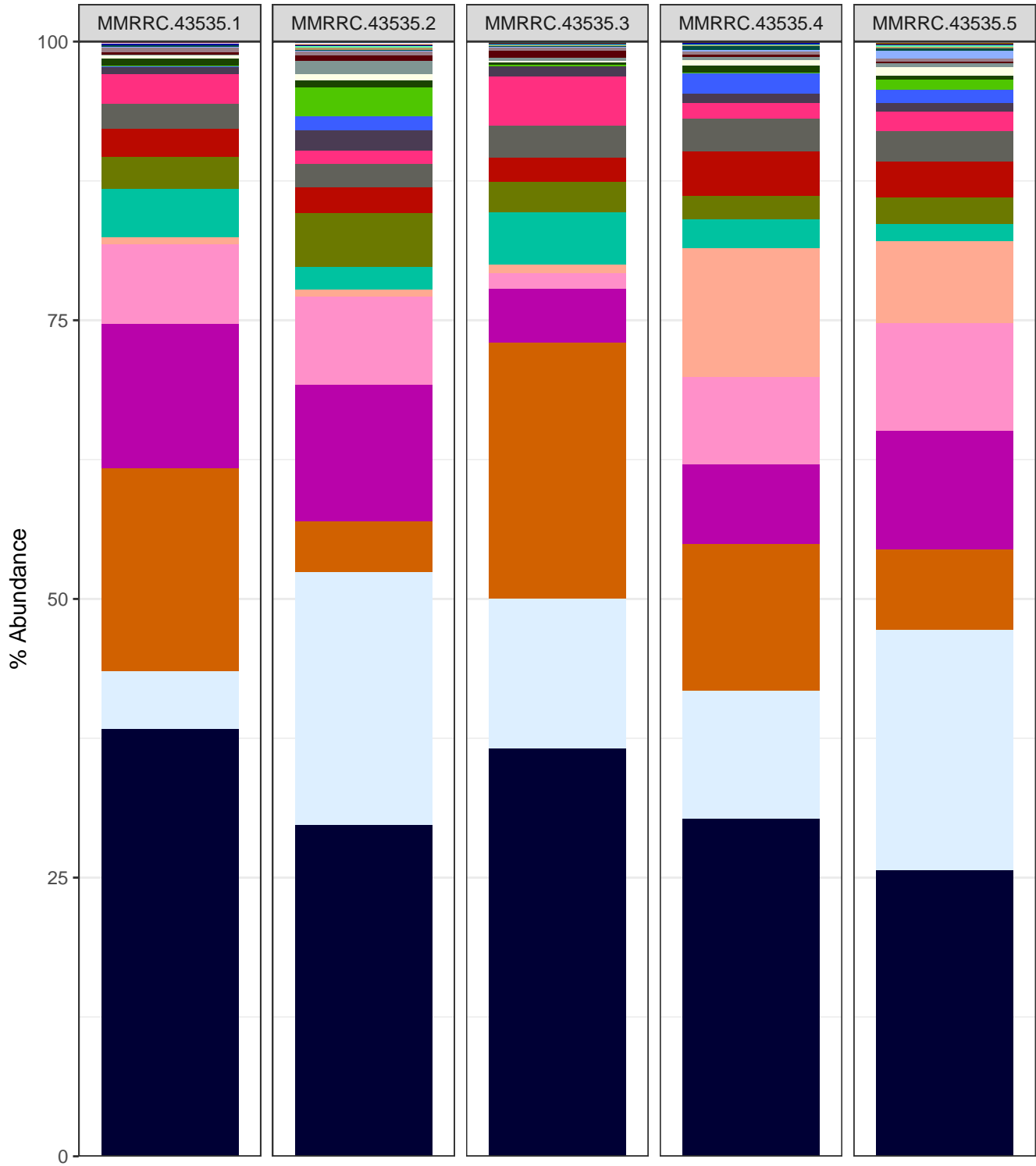


Figure 30: Family level abundances-FILTERED & RESCALED

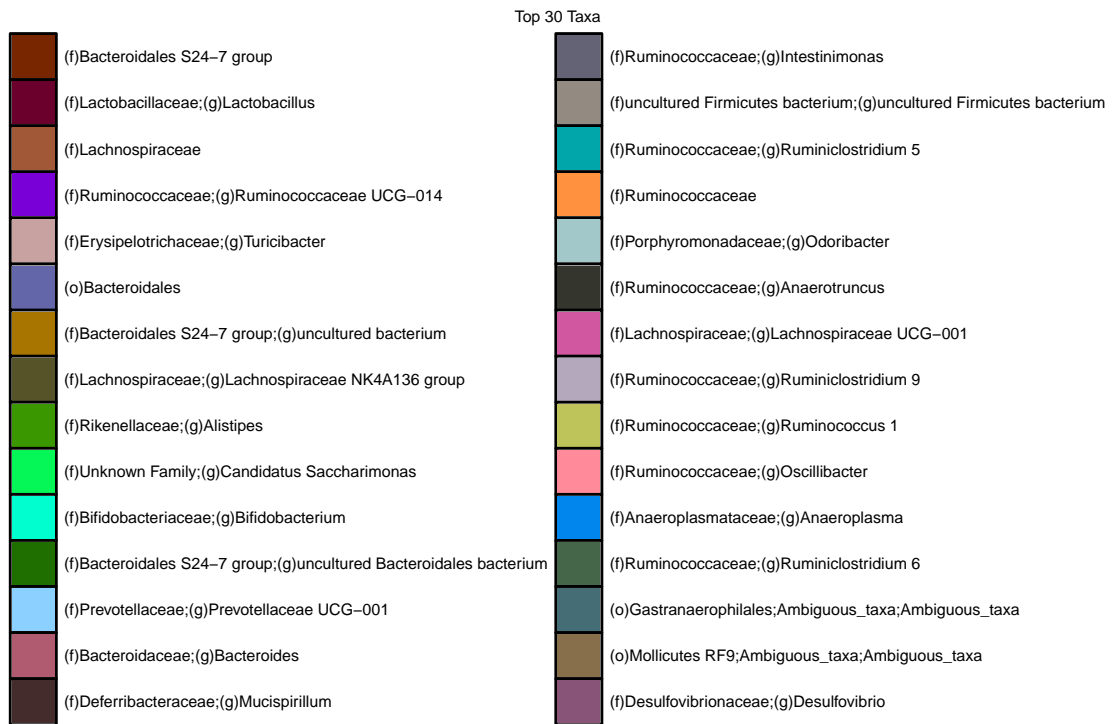
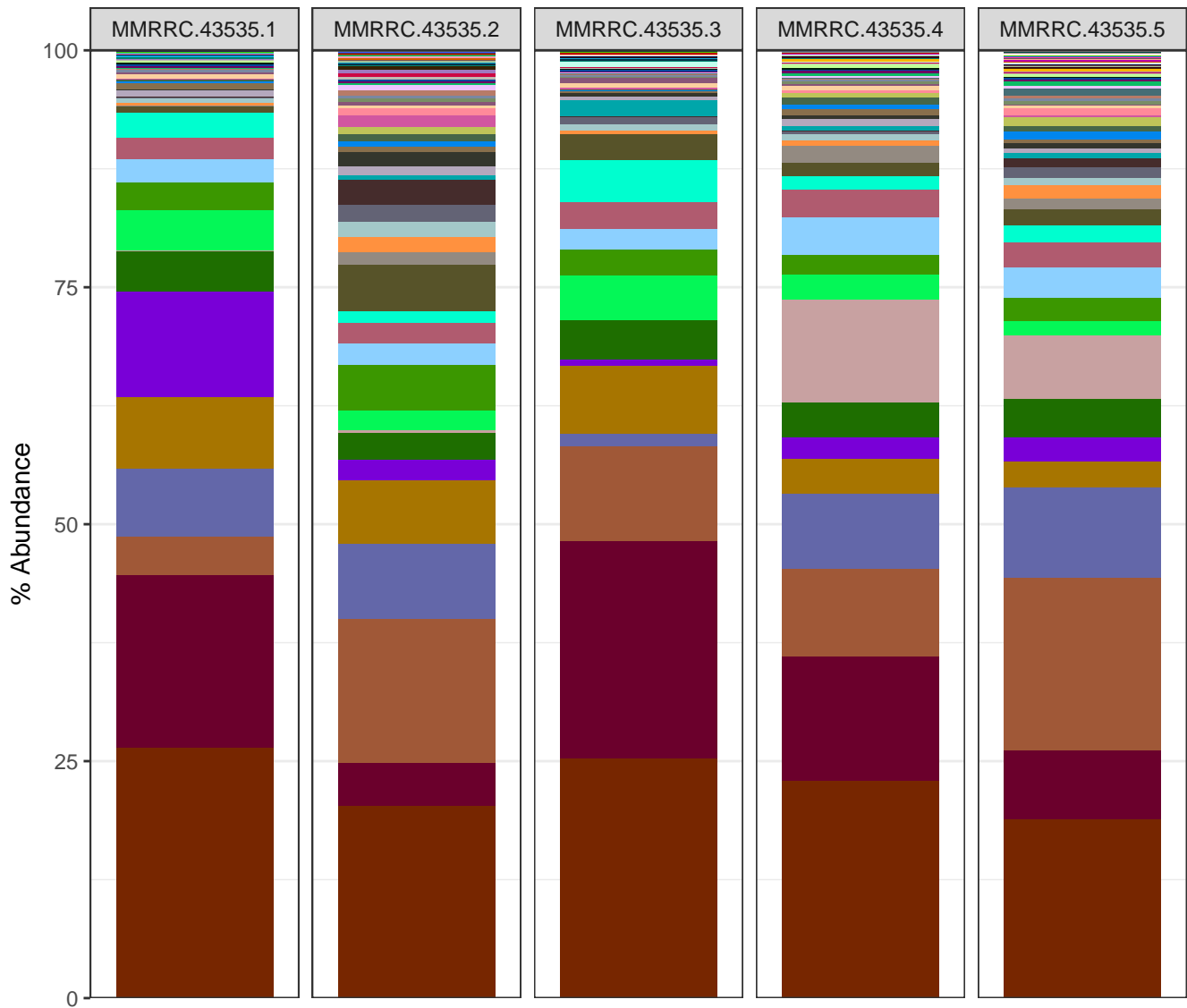
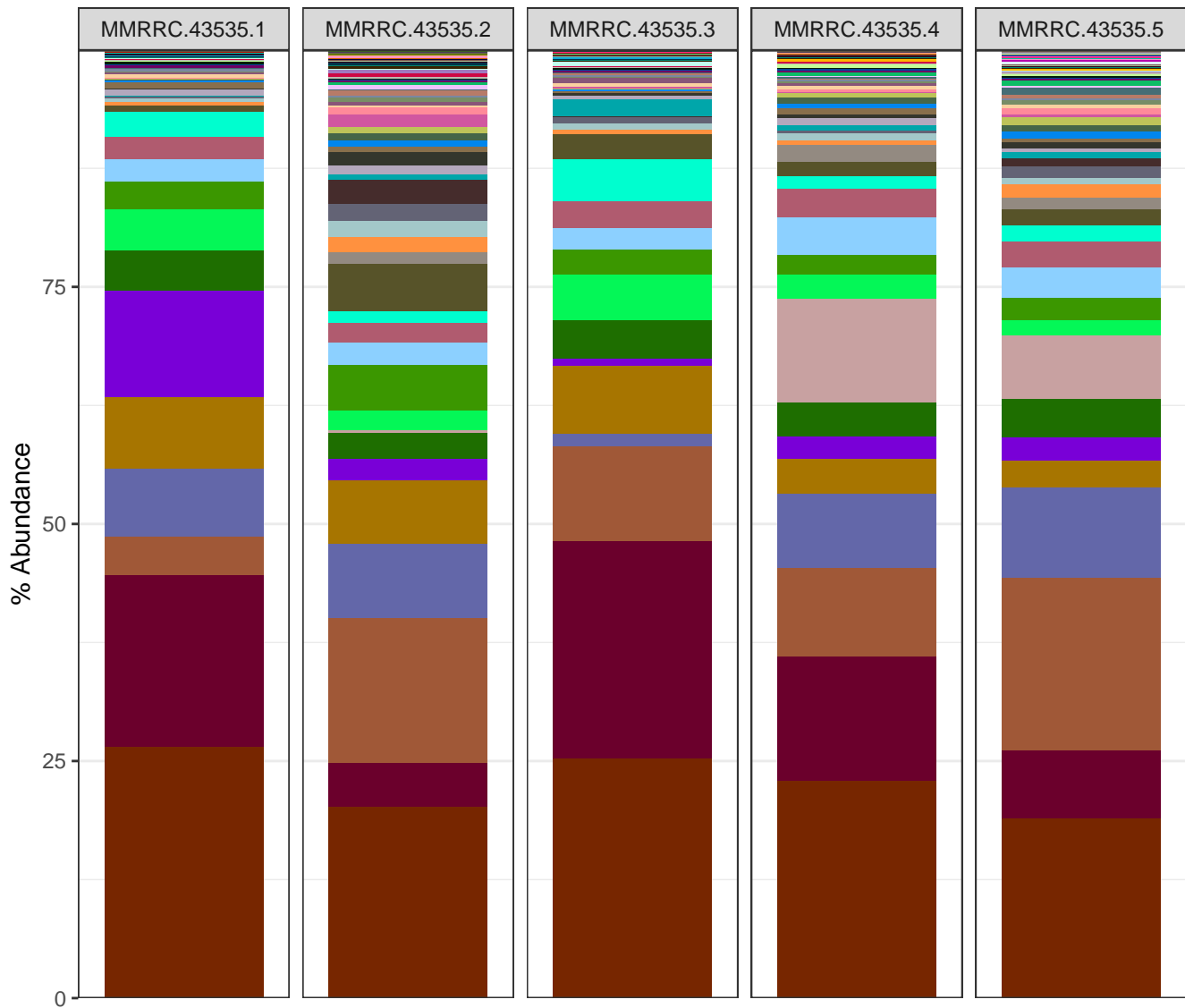


Figure 31: Genus level abundances-UNFILTERED



- (f)Bacteroidales S24-7 group
- (f)Lactobacillaceae;(g)Lactobacillus
- (f)Lachnospiraceae
- (f)Ruminococcaceae;(g)Ruminococcaceae UCG-014
- (f)Erysipelotrichaceae;(g)Turicibacter
- (o)Bacteroidales
- (f)Bacteroidales S24-7 group;(g)uncultured bacterium
- (f)Lachnospiraceae;(g)Lachnospiraceae NK4A136 group
- (f)Rikenellaceae;(g)Alistipes
- (f)Unknown Family;(g)Candidatus Saccharimonas
- (f)Bifidobacteriaceae;(g)Bifidobacterium
- (f)Bacteroidales S24-7 group;(g)uncultured Bacteroidales bacterium
- (f)Prevotellaceae;(g)Prevotellaceae UCG-001
- (f)Bacteroidaceae;(g)Bacteroides
- (f)Deferribacteraceae;(g)Mucispirillum

Top 30 Taxa

- (f)Ruminococcaceae;(g)Intestinimonas
- (f)uncultured Firmicutes bacterium;(g)uncultured Firmicutes bacterium
- (f)Ruminococcaceae;(g)Ruminiclostridium 5
- (f)Ruminococcaceae
- (f)Porphyromonadaceae;(g)Odoribacter
- (f)Ruminococcaceae;(g)Anaerotruncus
- (f)Lachnospiraceae;(g)Lachnospiraceae UCG-001
- (f)Ruminococcaceae;(g)Ruminiclostridium 9
- (f)Ruminococcaceae;(g)Ruminococcus 1
- (f)Ruminococcaceae;(g)Oscillibacter
- (f)Anaeroplasmataceae;(g)Anaeroplasma
- (f)Ruminococcaceae;(g)Ruminiclostridium 6
- (o)Gastranaerophilales;Ambiguous_taxa;Ambiguous_taxa
- (o)Mollicutes RF9;Ambiguous_taxa;Ambiguous_taxa
- (f)Desulfovibrionaceae;(g)Desulfovibrio

Figure 32: Genus level abundances-FILTERED: 14 genus level taxa removed

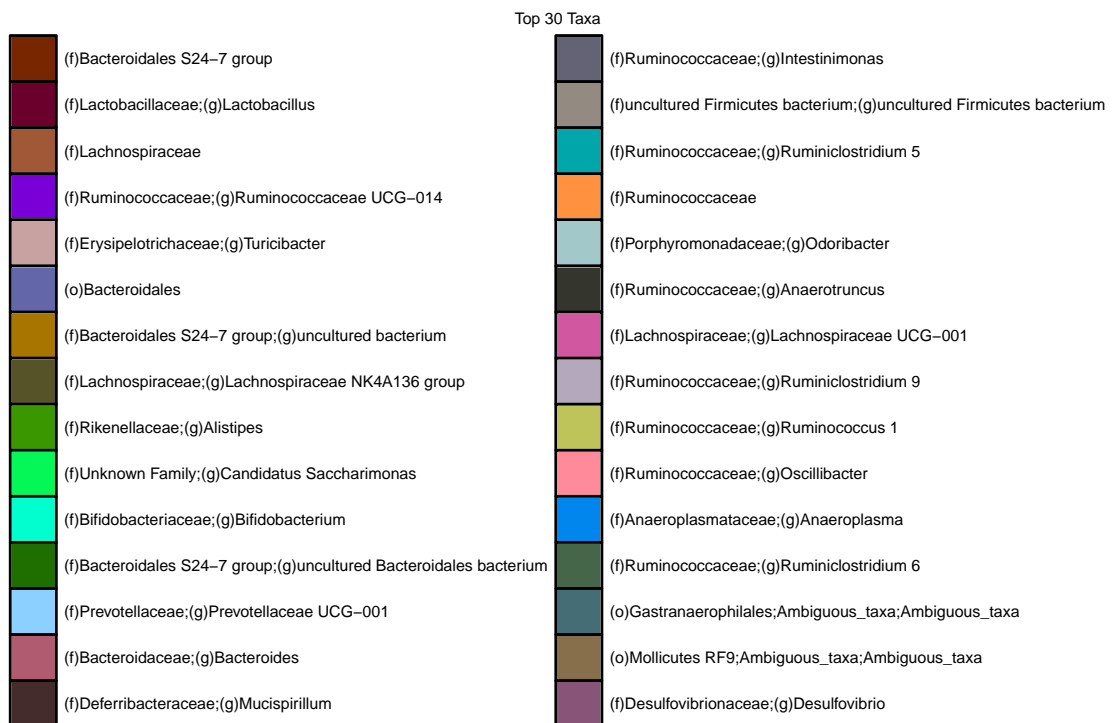
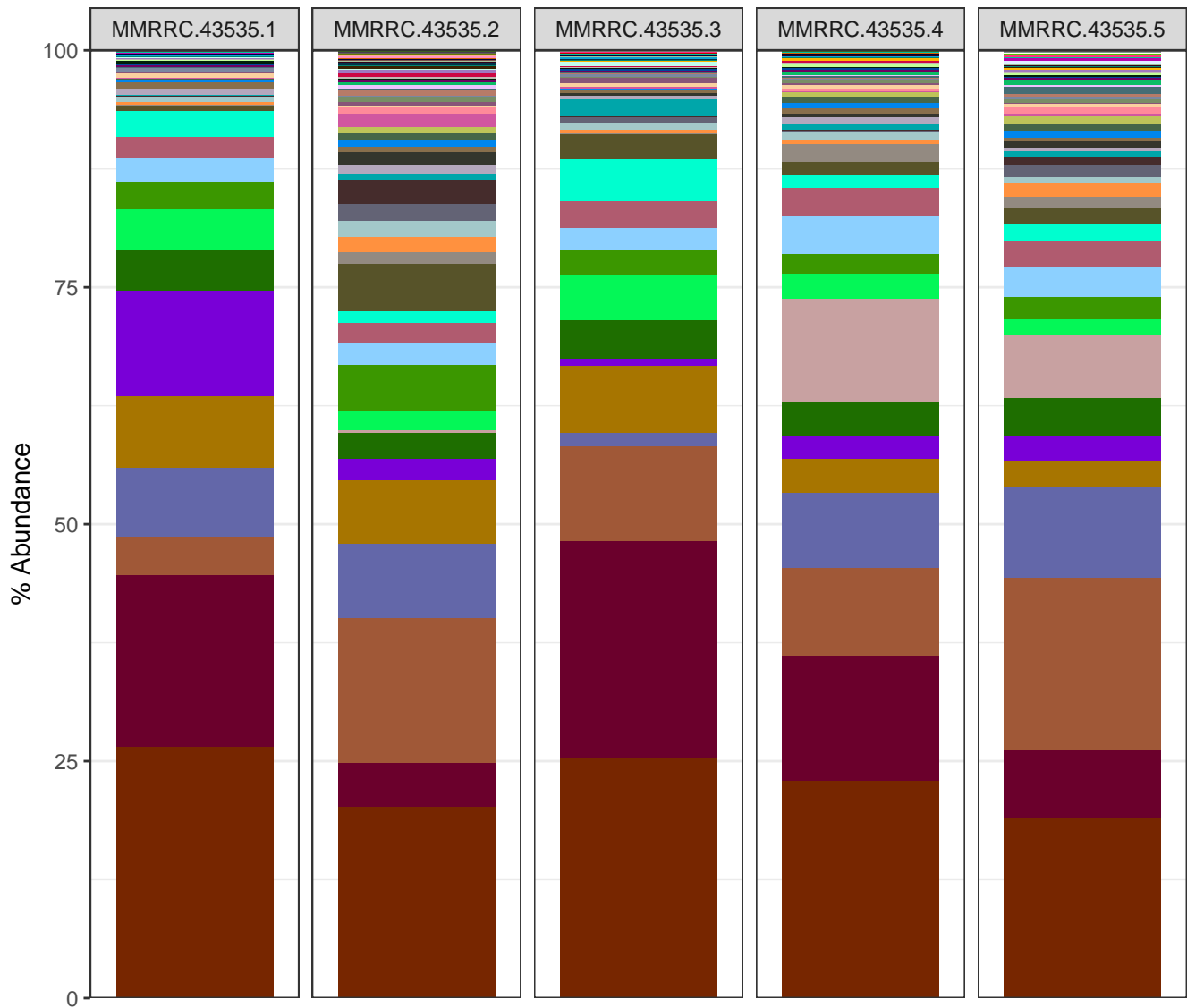


Figure 33: Genus level abundances-FILTERED & RESCALED

A Appendix 1 (Taxa Abundance Tables)

Table 2: Phylum (Filtered) level taxa group mean percent abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(k)Bacteria	0.03%	0.07%	0.01%	0.01%	0.03%
(p)Actinobacteria	3.15%	1.61%	4.77%	1.64%	1.99%
(p)Bacteroidetes	53.76%	48.7%	46.53%	47.99%	44.39%
(p)Cyanobacteria	0.07%	0.08%	0.08%	0.13%	0.69%
(p)Deferribacteres	0.05%	2.61%	0.13%	0.04%	0.85%
(p)Firmicutes	37.15%	41.75%	42.51%	44.22%	47.84%
(p)Proteobacteria	0.29%	0.61%	0.67%	0.31%	0.19%
(p)Saccharibacteria	4.29%	2%	4.74%	2.59%	1.54%
(p)Tenericutes	1.21%	2.57%	0.56%	3.08%	2.48%
NA	100	100	100	100	100

Table 3: Class (Filtered) level taxa group mean percent abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(k)Bacteria	0.03%	0.07%	0.01%	0.01%	0.03%
(p)Actinobacteria;(c)Actinobacteria	2.67%	1.21%	4.48%	1.39%	1.74%
(p)Actinobacteria;(c)Coriobacteriia	0.48%	0.4%	0.3%	0.25%	0.26%
(p)Bacteroidetes;(c)Bacteroidia	53.76%	48.7%	46.53%	47.99%	44.39%
(p)Cyanobacteria;(c)Melainabacteria	0.07%	0.08%	0.08%	0.13%	0.69%
(p)Deferribacteres;(c)Deferribacteres	0.05%	2.61%	0.13%	0.04%	0.85%
(p)Firmicutes	0%	0.06%	0.02%	0.08%	0.15%
(p)Firmicutes;(c)Bacilli	18.19%	4.53%	22.96%	13.15%	7.26%
(p)Firmicutes;(c)Clostridia	18.33%	36.58%	18.79%	19.43%	33.06%
(p)Firmicutes;(c)Erysipelotrichia	0.63%	0.58%	0.74%	11.56%	7.37%
(p)Proteobacteria;(c)Alphaproteobacteria	0.05%	0.13%	0.02%	0.01%	0.03%
(p)Proteobacteria;(c)Deltaproteobacteria	0.19%	0.48%	0.64%	0.26%	0.16%
(p)Saccharibacteria;(c)Unknown Class	4.29%	2%	4.74%	2.59%	1.54%
(p)Tenericutes;(c)Mollicutes	1.21%	2.57%	0.56%	3.08%	2.48%
NA	99.951	100	99.988	99.96	100

Table 4: Order (Filtered) level taxa group mean percent abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(c)Actinobacteria;(o)Bifidobacteriales	2.67%	1.21%	4.48%	1.39%	1.74%
(c)Alphaproteobacteria;(o)Rhodospirillales	0.05%	0.13%	0.02%	0.01%	0.03%
(c)Bacilli;(o)Lactobacillales	18.19%	4.53%	22.96%	13.15%	7.26%
(c)Bacteroidia;(o)Bacteroidales	53.76%	48.7%	46.53%	47.99%	44.39%
(c)Clostridia;(o)Clostridiales	18.33%	36.58%	18.79%	19.43%	33.06%
(c)Coriobacteriia;(o)Coriobacteriales	0.48%	0.4%	0.3%	0.25%	0.26%
(c)Deferribacteres;(o)Deferribacterales	0.05%	2.61%	0.13%	0.04%	0.85%
(c)Deltaproteobacteria;(o)Desulfovibrionales	0.19%	0.48%	0.64%	0.26%	0.16%
(c)Erysipelotrichia;(o)Erysipelotrichales	0.63%	0.58%	0.74%	11.56%	7.37%
(c)Melainabacteria;(o)Gastranaerophilales	0.07%	0.08%	0.08%	0.13%	0.69%
(c)Mollicutes;(o)Anaeroplasmatales	0.27%	0.59%	0.11%	0.51%	0.78%
(c)Mollicutes;(o)Mollicutes RF9	0.94%	1.97%	0.45%	2.57%	1.7%
(k)Bacteria	0.03%	0.07%	0.01%	0.01%	0.03%
(k)Bacteria;(p)Firmicutes	0%	0.06%	0.02%	0.08%	0.15%
(k)Bacteria;(p)Saccharibacteria;(c)Unknown Class;(o)Unknown Order	4.29%	2%	4.74%	2.59%	1.54%
NA	99.951	100	99.988	99.96	100

Table 5: Family (Filtered) level taxa group mean percent abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(f)Anaeroplasmataceae	0.27%	0.59%	0.11%	0.51%	0.78%
(f)Bacteroidaceae	2.28%	2.11%	2.82%	2.94%	2.71%
(f)Bacteroidales S24-7 group	38.31%	29.7%	36.58%	30.26%	25.65%
(f)Bifidobacteriaceae	2.67%	1.21%	4.48%	1.39%	1.74%
(f)Clostridiaceae 1	0.1%	0%	0%	0.39%	0.23%
(f)Clostridiales vadinBB60 group	0.06%	1.19%	0.29%	0.22%	0.33%
(f)Coriobacteriaceae	0.48%	0.4%	0.3%	0.25%	0.26%
(f)Deferribacteraceae	0.05%	2.61%	0.13%	0.04%	0.85%
(f)Desulfovibrionaceae	0.19%	0.48%	0.64%	0.26%	0.16%
(f)Erysipelotrichaceae	0.63%	0.58%	0.74%	11.56%	7.37%
(f)Family XIII	0.02%	0.14%	0.08%	0.09%	0.12%
(f)Lachnospiraceae	5.2%	22.73%	13.43%	11.51%	21.57%
(f)Lactobacillaceae	18.15%	4.53%	22.95%	13.14%	7.25%
(f)Peptococcaceae	0.01%	0.23%	0.1%	0.08%	0.13%
(f)Porphyromonadaceae	0.61%	1.78%	0.86%	0.87%	0.84%
(f)Prevotellaceae	2.46%	2.28%	2.22%	4%	3.2%
(f)Rhodospirillaceae	0.05%	0.13%	0.02%	0.01%	0.03%
(f)Rikenellaceae	2.91%	4.89%	2.68%	2.08%	2.38%
(f)Ruminococcaceae	12.93%	12.29%	4.87%	7.12%	10.64%
(f)uncultured Erysipelotrichaceae bacterium	0.01%	0%	0.07%	0%	0%
(f)uncultured Firmicutes bacterium	0.08%	1.27%	0.02%	1.81%	1.21%
(f)uncultured Mollicutes bacterium	0.1%	0.06%	0.03%	0.01%	0.02%
(k)Bacteria	0.03%	0.07%	0.01%	0.01%	0.03%
(k)Bacteria;(p)Firmicutes	0%	0.06%	0.02%	0.08%	0.15%
(k)Bacteria;(p)Saccharibacteria;(c)Unknown Class;(o)Unknown Order;(f)Unknown Family	4.29%	2%	4.74%	2.59%	1.54%
(o)Bacteroidales	7.2%	7.92%	1.38%	7.84%	9.61%
(o)Gastranaerophilales;Ambiguous_taxa	0.07%	0.08%	0.08%	0.13%	0.69%
(o)Mollicutes RF9	0.1%	0.07%	0.09%	0.1%	0.09%
(o)Mollicutes RF9;Ambiguous_taxa	0.65%	0.57%	0.25%	0.65%	0.37%
NA	99.911	100	99.965	99.934	99.958

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

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(f)Anaeroplasmataceae;(g)Anaeroplasma	0.27%	0.59%	0.11%	0.51%	0.78%
(f)Bacteroidaceae;(g)Bacteroides	2.28%	2.11%	2.82%	2.94%	2.71%
(f)Bacteroidales S24-7 group	26.49%	20.27%	25.31%	22.94%	18.92%
(f)Bacteroidales S24-7 group;(g)uncultured bacterium	7.56%	6.64%	7.13%	3.66%	2.75%
(f)Bacteroidales S24-7 group;(g)uncultured Bacteroidales bacterium	4.26%	2.8%	4.14%	3.66%	3.98%
(f)Bifidobacteriaceae;(g)Bifidobacterium	2.67%	1.21%	4.48%	1.39%	1.74%
(f)Clostridiaceae 1;(g)Candidatus Arthromitus	0.1%	0%	0%	0.39%	0.23%
(f)Clostridiales vadinBB60 group	0%	0.3%	0.04%	0.06%	0.07%
(f)Clostridiales vadinBB60 group;(g)uncultured bacterium	0%	0.08%	0.06%	0.02%	0%
(f)Clostridiales vadinBB60 group;(g)uncultured organism	0.02%	0.28%	0.11%	0%	0%
(f)Clostridiales vadinBB60 group;Ambiguous_taxa	0.04%	0.53%	0.08%	0.14%	0.26%
(f)Coriobacteriaceae	0.1%	0.08%	0.1%	0%	0.02%
(f)Coriobacteriaceae;(g)Enterorhabdus	0.33%	0.24%	0.2%	0.19%	0.21%
(f)Coriobacteriaceae;(g)Parvibacter	0.05%	0.06%	0%	0.03%	0.02%
(f)Deferribacteraceae;(g)Mucispirillum	0.05%	2.61%	0.13%	0.04%	0.85%
(f)Desulfovibrionaceae;(g)Bilophila	0%	0.11%	0.06%	0.04%	0.08%
(f)Desulfovibrionaceae;(g)Desulfovibrio	0.19%	0.36%	0.58%	0.22%	0.08%
(f)Erysipelotrichaceae	0.28%	0.04%	0.13%	0.15%	0.18%
(f)Erysipelotrichaceae;(g)Erysipelatoclostridium	0.21%	0.19%	0.11%	0.23%	0.18%
(f)Erysipelotrichaceae;(g)Faecalibaculum	0.02%	0.05%	0%	0.27%	0.23%
(f)Erysipelotrichaceae;(g)Turicibacter	0.02%	0.3%	0%	10.9%	6.77%
(f)Erysipelotrichaceae;(g)uncultured	0.09%	0%	0.49%	0%	0%
(f)Family XIII;(g)Family XIII AD3011 group	0.02%	0.01%	0.02%	0.03%	0.05%
(f)Family XIII;(g)Family XIII UCG-001	0%	0.09%	0.04%	0.04%	0.06%
(f)Lachnospiraceae	4.06%	15.24%	9.94%	9.3%	18.14%
(f)Lachnospiraceae;(g)[Ruminococcus] gnavus group	0.03%	0.41%	0%	0.04%	0.12%
(f)Lachnospiraceae;(g)Acetatifactor	0%	0.05%	0%	0%	0.06%
(f)Lachnospiraceae;(g)Blautia	0%	0.03%	0.01%	0.04%	0.22%
(f)Lachnospiraceae;(g)Coprococcus 1	0%	0.52%	0.1%	0.12%	0.36%
(f)Lachnospiraceae;(g)Lachnospiraceae NK4A136 group	0.63%	4.96%	2.67%	1.45%	1.68%
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-001	0.06%	1.27%	0.19%	0.06%	0.28%
(f)Lachnospiraceae;(g)Marvinbryantia	0.43%	0.19%	0.39%	0.45%	0.29%
(f)Lachnospiraceae;(g)Moryella	0%	0.08%	0.1%	0%	0%
(f)Lachnospiraceae;(g)Tyzzerella 3	0%	0%	0.01%	0.05%	0.2%
(f)Lachnospiraceae;(g)uncultured	0%	0%	0.02%	0%	0.14%
(f)Lactobacillaceae;(g)Lactobacillus	18.15%	4.53%	22.95%	13.14%	7.25%
(f)Peptococcaceae;(g)Peptococcus	0%	0.1%	0.05%	0.02%	0.04%
(f)Peptococcaceae;(g)uncultured	0.01%	0.13%	0.06%	0.06%	0.08%
(f)Porphyromonadaceae;(g)Odoribacter	0.44%	1.63%	0.68%	0.7%	0.69%
(f)Porphyromonadaceae;(g)Parabacteroides	0.16%	0.15%	0.18%	0.16%	0.14%
(f)Prevotellaceae;(g)Prevotellaceae UCG-001	2.45%	2.28%	2.22%	4%	3.2%
(f)Rhodospirillaceae;(g)uncultured	0.05%	0.13%	0.02%	0.01%	0.03%
(f)Rikenellaceae;(g)Alistipes	2.91%	4.89%	2.68%	2.08%	2.38%
(f)Ruminococcaceae	0.29%	1.63%	0.39%	0.49%	1.43%
(f)Ruminococcaceae;(g)[Eubacterium] coprostanoligenes group	0%	0.22%	0%	0.31%	0.51%
(f)Ruminococcaceae;(g)Anaerotruncus	0.08%	1.49%	0.29%	0.38%	0.64%
(f)Ruminococcaceae;(g)Butyricicoccus	0%	0.04%	0%	0%	0.12%
(f)Ruminococcaceae;(g)Hydrogenoanaerobacterium	0%	0%	0%	0%	0.08%
(f)Ruminococcaceae;(g)Intestinimonas	0.14%	1.81%	0.68%	0.28%	1.19%
(f)Ruminococcaceae;(g)Oscillibacter	0.04%	0.8%	0.08%	0.23%	0.68%

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

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(f)Ruminococcaceae;(g)Ruminiclostridium	0.05%	0.46%	0.1%	0.25%	0.46%
(f)Ruminococcaceae;(g)Ruminiclostridium 5	0.06%	0.49%	1.72%	0.54%	0.63%
(f)Ruminococcaceae;(g)Ruminiclostridium 6	0.1%	0.71%	0.05%	0.69%	0.65%
(f)Ruminococcaceae;(g)Ruminiclostridium 9	0.63%	0.93%	0.4%	0.75%	0.41%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-003	0%	0.38%	0.11%	0.11%	0.08%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-004	0.04%	0.08%	0.04%	0.14%	0.05%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-009	0%	0%	0.01%	0%	0.12%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-010	0.13%	0.24%	0.1%	0.11%	0.11%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-013	0.17%	0.01%	0.18%	0.02%	0.02%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-014	11.14%	2.26%	0.72%	2.31%	2.54%
(f)Ruminococcaceae;(g)Ruminococcus 1	0.04%	0.74%	0%	0.48%	0.88%
(f)Ruminococcaceae;(g)Subdoligranulum	0%	0%	0%	0.05%	0.05%
(f)uncultured Erysipelotrichaceae bacterium;(g)uncultured Erysipelotrichaceae bacterium	0.01%	0%	0.07%	0%	0%
(f)uncultured Firmicutes bacterium;(g)uncultured Firmicutes bacterium	0.08%	1.27%	0.02%	1.81%	1.21%
(f)uncultured Mollicutes bacterium;(g)uncultured Mollicutes bacterium	0.1%	0.06%	0.03%	0.01%	0.02%
(f)Unknown Family;(g)Candidatus Saccharimonas	4.29%	2%	4.74%	2.59%	1.54%
(k)Bacteria	0.03%	0.07%	0.01%	0.01%	0.03%
(k)Bacteria;(p)Firmicutes	0%	0.06%	0.02%	0.08%	0.15%
(o)Bacteroidales	7.2%	7.92%	1.38%	7.84%	9.61%
(o)Gastranaerophilales;Ambiguous_taxa;Ambiguous_taxa	0.07%	0.08%	0.08%	0.13%	0.69%
(o)Mollicutes RF9	0.1%	0.07%	0.09%	0.1%	0.09%
(o)Mollicutes RF9;Ambiguous_taxa;Ambiguous_taxa	0.65%	0.57%	0.25%	0.65%	0.37%
NA	99.874	99.945	99.935	99.875	99.836

B Appendix 2 (Taxa removed from filtered datasets at each level)

Table 7: Phylum level taxa removed via filtering- group mean abundances

X.Total...Removed.	X.0.	X.0..1	X.0..2	X.0..3	X.0..4
Total % Removed	0	0	0	0	0

Table 8: Class level taxa removed via filtering- group mean abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(p)Proteobacteria;(c)Gammaproteobacteria	0.0491	0	0.0123	0.04	0
Total % Removed	0.0491	0	0.0123	0.04	0

Table 9: Order level taxa removed via filtering- group mean abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(c)Gammaproteobacteria;(o)Enterobacteriales	0.0491	0	0.0123	0.04	0
Total % Removed	0.0491	0	0.0123	0.04	0

Table 10: Family level taxa removed via filtering- group mean abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(f)Christensenellaceae	0.0000	0	0.0092	0.0192	0.0076
(f)Enterobacteriaceae	0.0491	0	0.0123	0.0400	0.0000
(f)Enterococcaceae	0.0399	0	0.0138	0.0064	0.0000
(f)Streptococcaceae	0.0000	0	0.0000	0.0000	0.0076
(o)Clostridiales	0.0000	0	0.0000	0.0000	0.0273
Total % Removed	0.0890	0	0.0353	0.0657	0.0425

Table 11: Genus level taxa removed via filtering- group mean abundances

Taxa	MMRRC.43535.1	MMRRC.43535.2	MMRRC.43535.3	MMRRC.43535.4	MMRRC.43535.5
(f)Christensenellaceae;(g)uncultured	0.0000	0.0000	0.0092	0.0192	0.0076
(f)Coriobacteriaceae;(g)Gordonibacter	0.0000	0.0221	0.0000	0.0272	0.0167
(f)Enterobacteriaceae	0.0491	0.0000	0.0123	0.0400	0.0000
(f)Enterococcaceae	0.0000	0.0000	0.0138	0.0000	0.0000
(f)Enterococcaceae;(g)Enterococcus	0.0399	0.0000	0.0000	0.0064	0.0000
(f)Erysipelotrichaceae;(g)uncultured bacterium	0.0169	0.0000	0.0000	0.0128	0.0152
(f)Family XIII;(g)[Eubacterium] brachy group	0.0000	0.0332	0.0184	0.0192	0.0167
(f)Lachnospiraceae;(g)Lachnospiraceae FCS020 group	0.0000	0.0000	0.0000	0.0000	0.0212
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-010	0.0000	0.0000	0.0000	0.0000	0.0061
(f)Lachnospiraceae;(g)Tyzzerella	0.0000	0.0000	0.0108	0.0000	0.0455
(f)Prevotellaceae	0.0107	0.0000	0.0000	0.0000	0.0000
(f)Ruminococcaceae;(g)Ruminiclostridium 1	0.0092	0.0000	0.0000	0.0000	0.0000
(f)Streptococcaceae;(g)Streptococcus	0.0000	0.0000	0.0000	0.0000	0.0076
(o)Clostridiales	0.0000	0.0000	0.0000	0.0000	0.0273
Total % Removed	0.1259	0.0553	0.0645	0.1249	0.1639

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. doi: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. doi: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. doi: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. doi: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. doi:10.1093/nar/gks1219.