D4006- Gut Microbiota Analysis

UC Davis MMPC - Microbiome & Host Response Core

Contents

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_043567

Animal Information: The strain was donated to the MMRRC by Paul Burgoyne at MRC NIMR. Fecal samples were obtained from animals housed under the care of Arthur Arnold at UCLA.

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 facilty 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.043567.M56564	MMRRC.043567__Tg.plus_M	M56564	Tg.plus	MMRRC.043567	M
MMRRC.043567.M94895	MMRRC.043567__Tg.plus_M	M94895	Tg.plus	MMRRC.043567	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 2019.1.0. 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 200 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 70610 ± 18179 (SD) individual sequencing reads per sample (Min= 57755; Max= 83464). After data processing the average number of sequences for each sample passing through to OTU classification was 34520 ± 16600 (SD). The average number of OTUs per sample was 142.

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 22700 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.
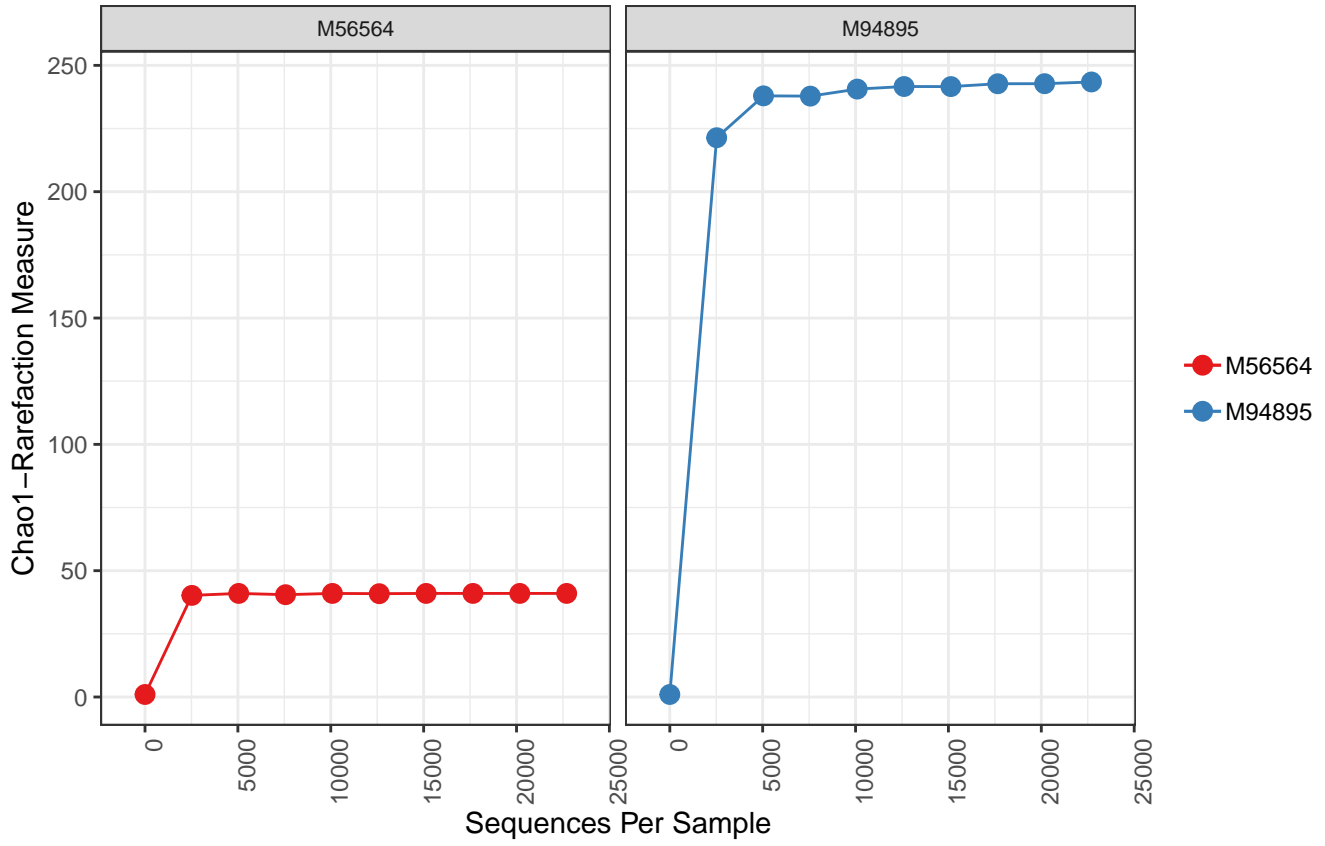
Figure 1: Chao1 Measure of Alpha Diversity- Rarefaction curve-Animal_ID.
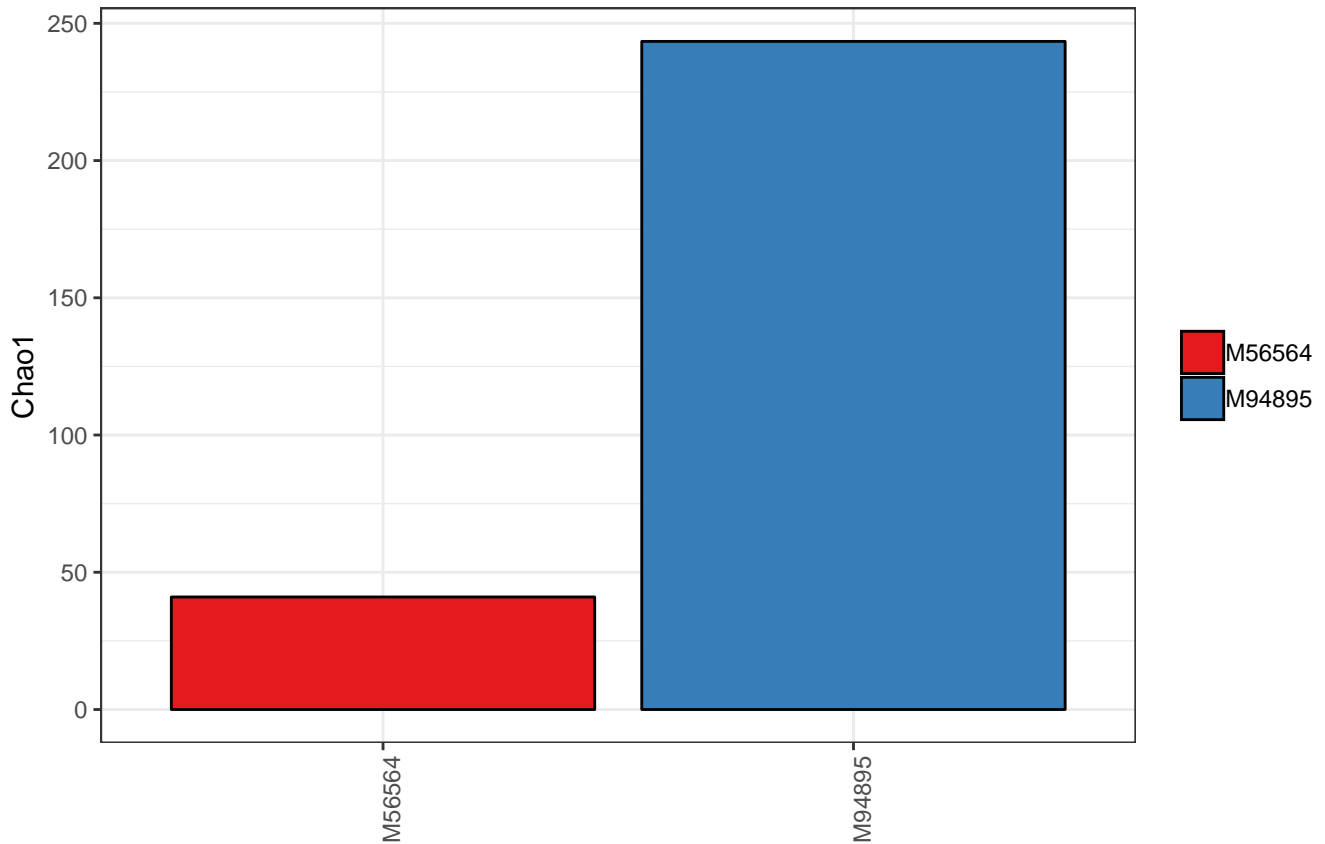

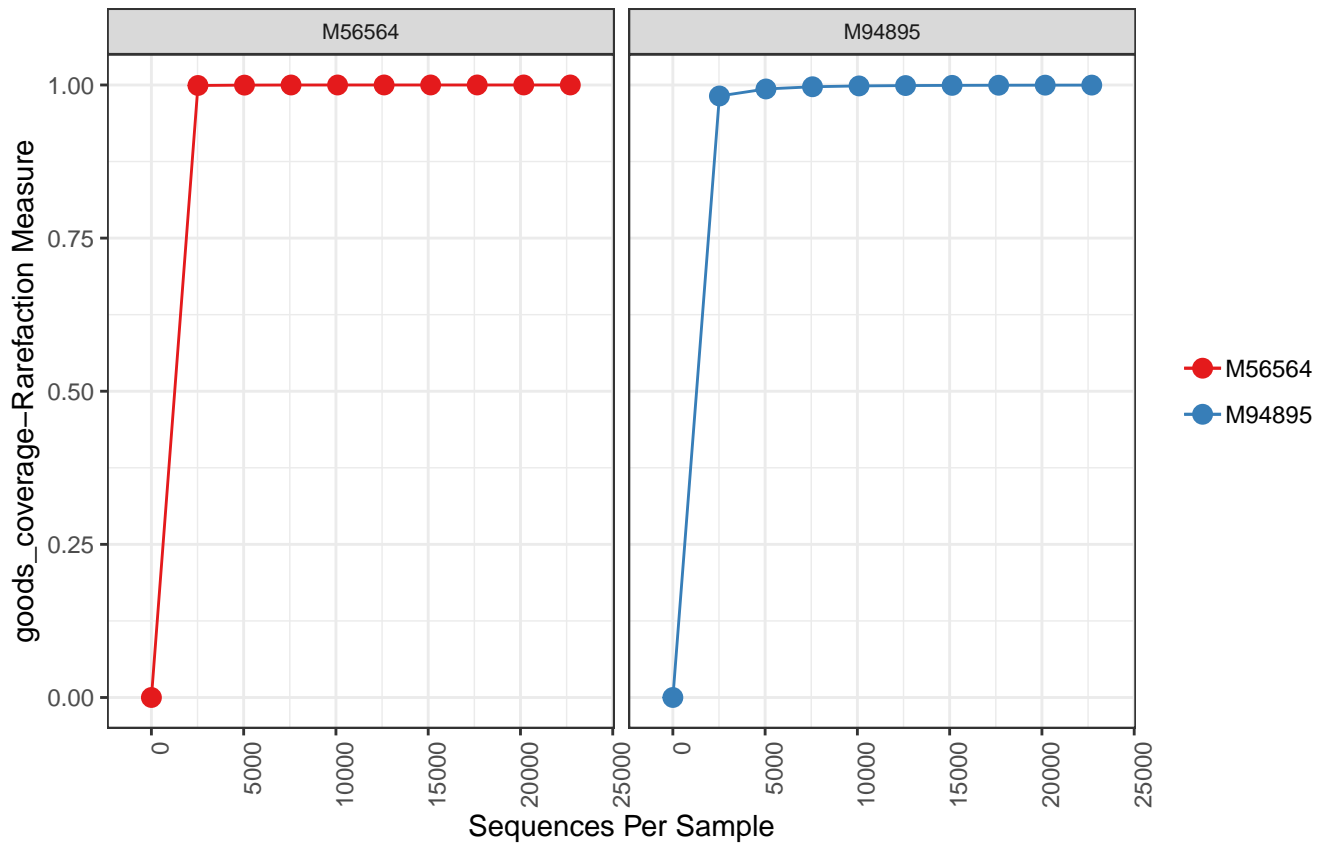Figure 2: Chao1 Measure of Alpha Diversity- Boxplots-Animal_ID.

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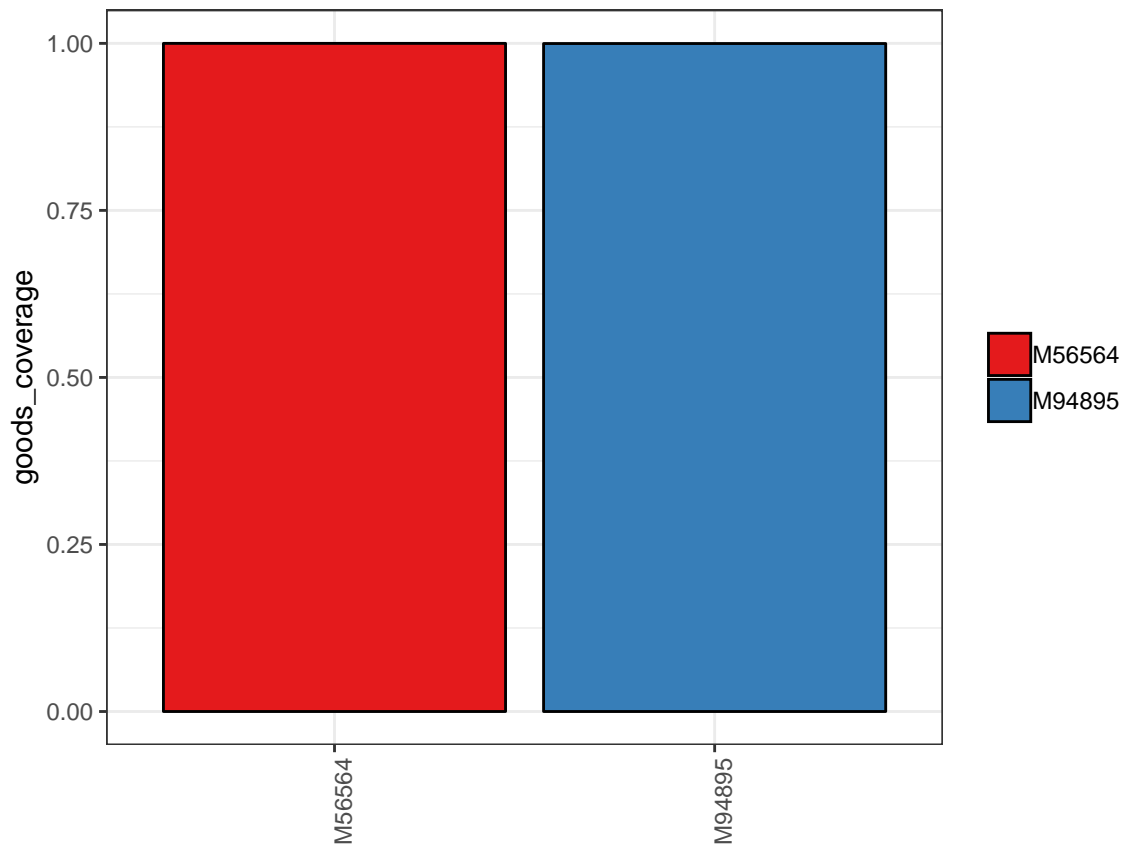

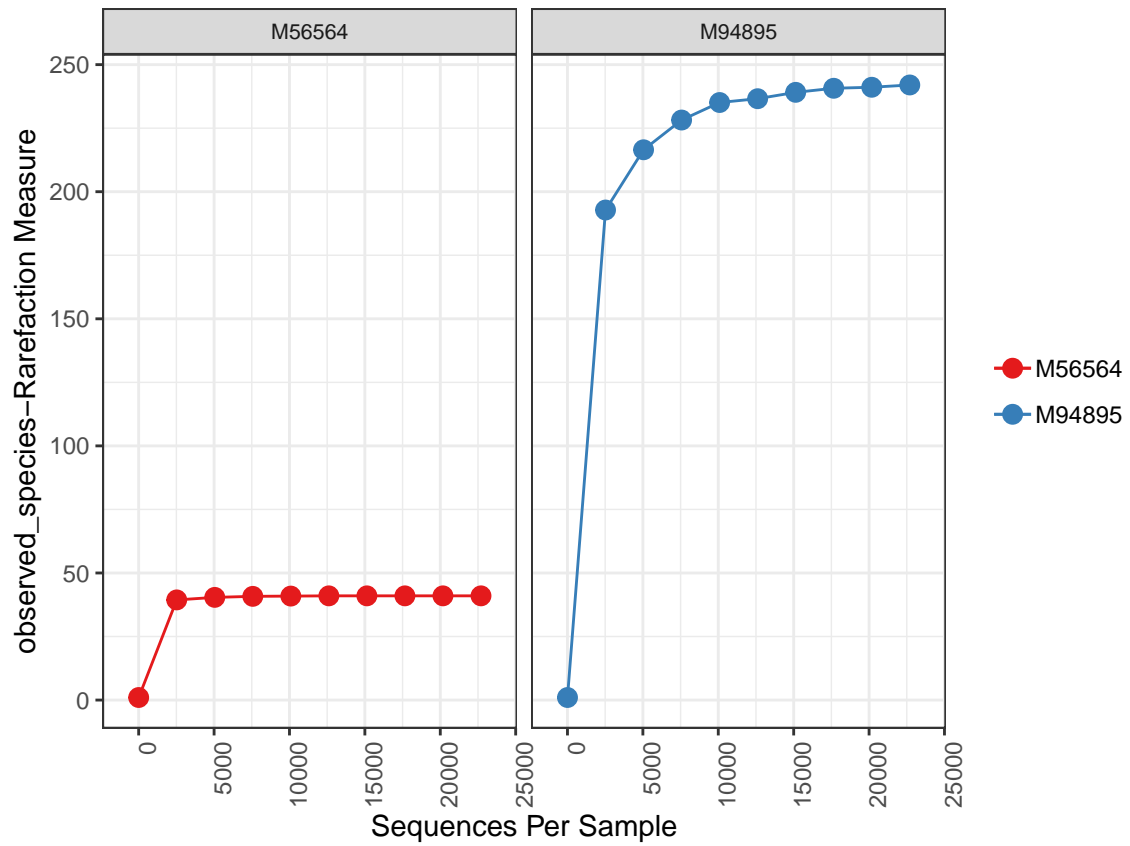Figure 4: Goods Coverage Measure of Alpha Diversity- Boxplots-Animal_ID.

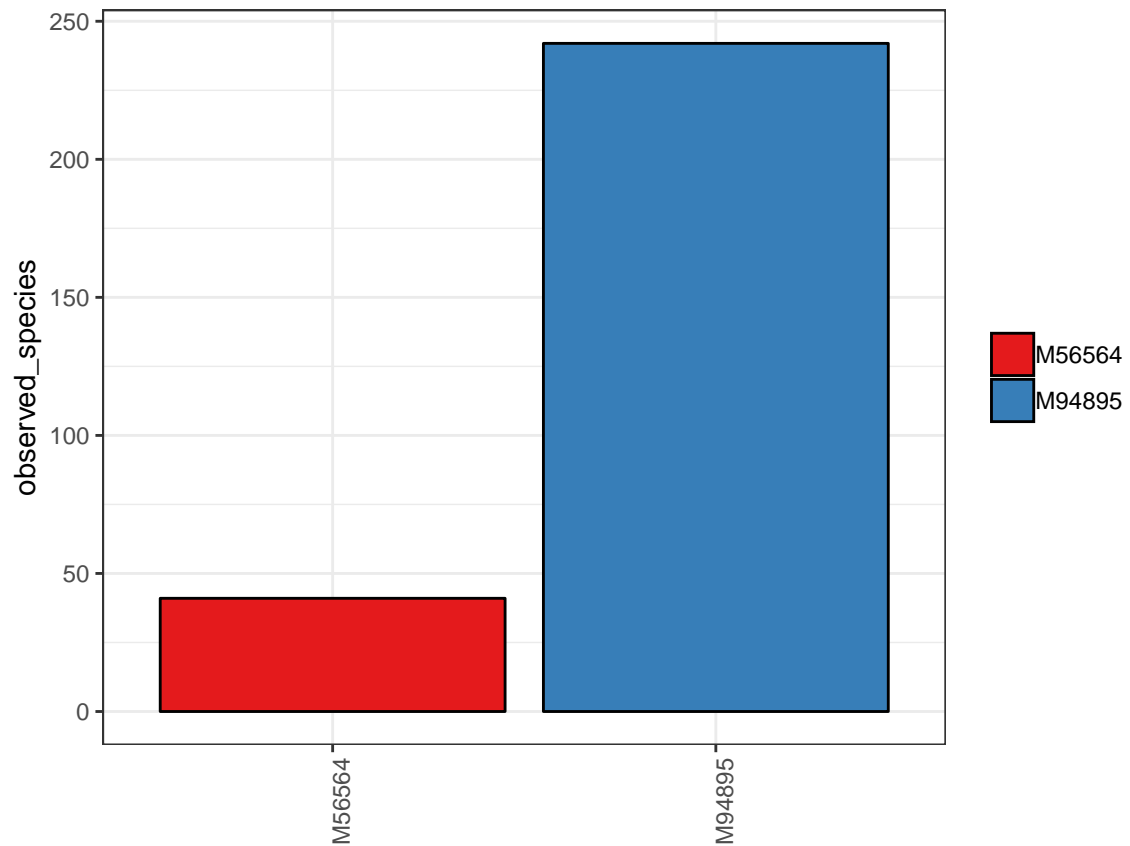Figure 5: Observed Species Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

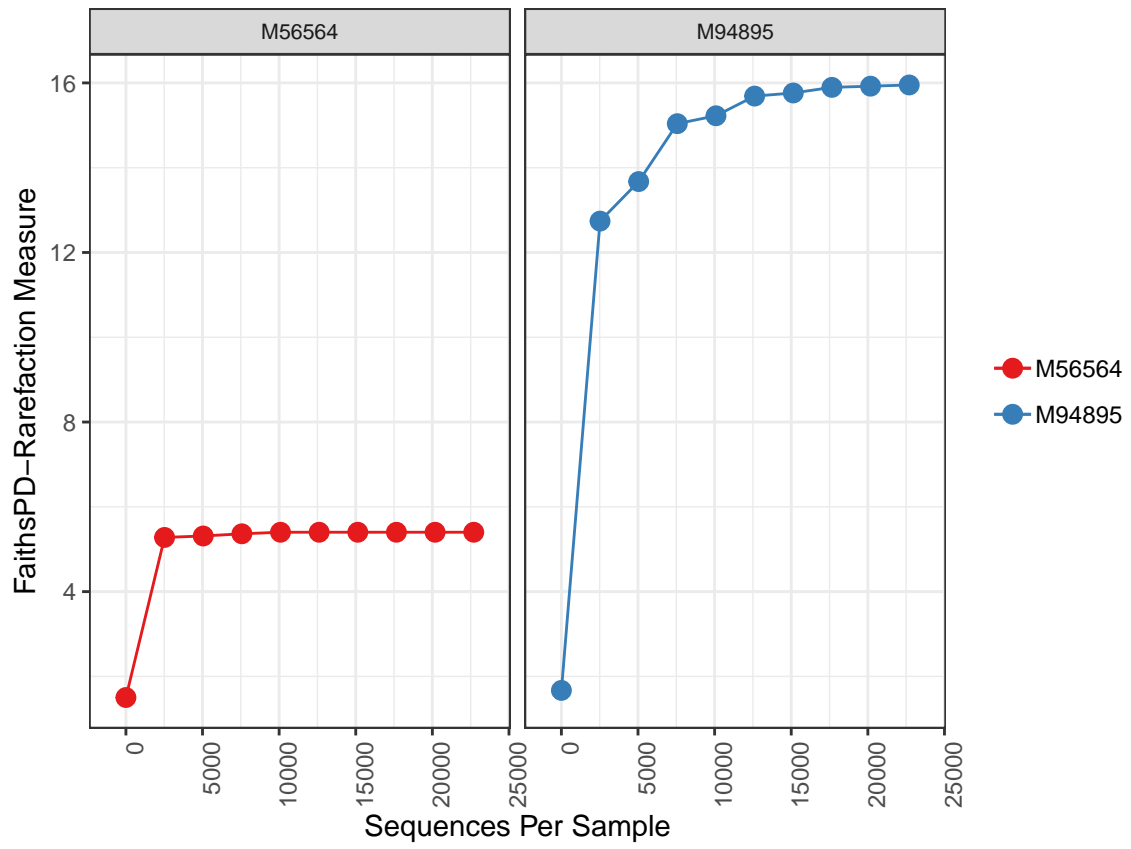Figure 6: Observed Species Measure of Alpha Diversity- Boxplot-Animal_ID.

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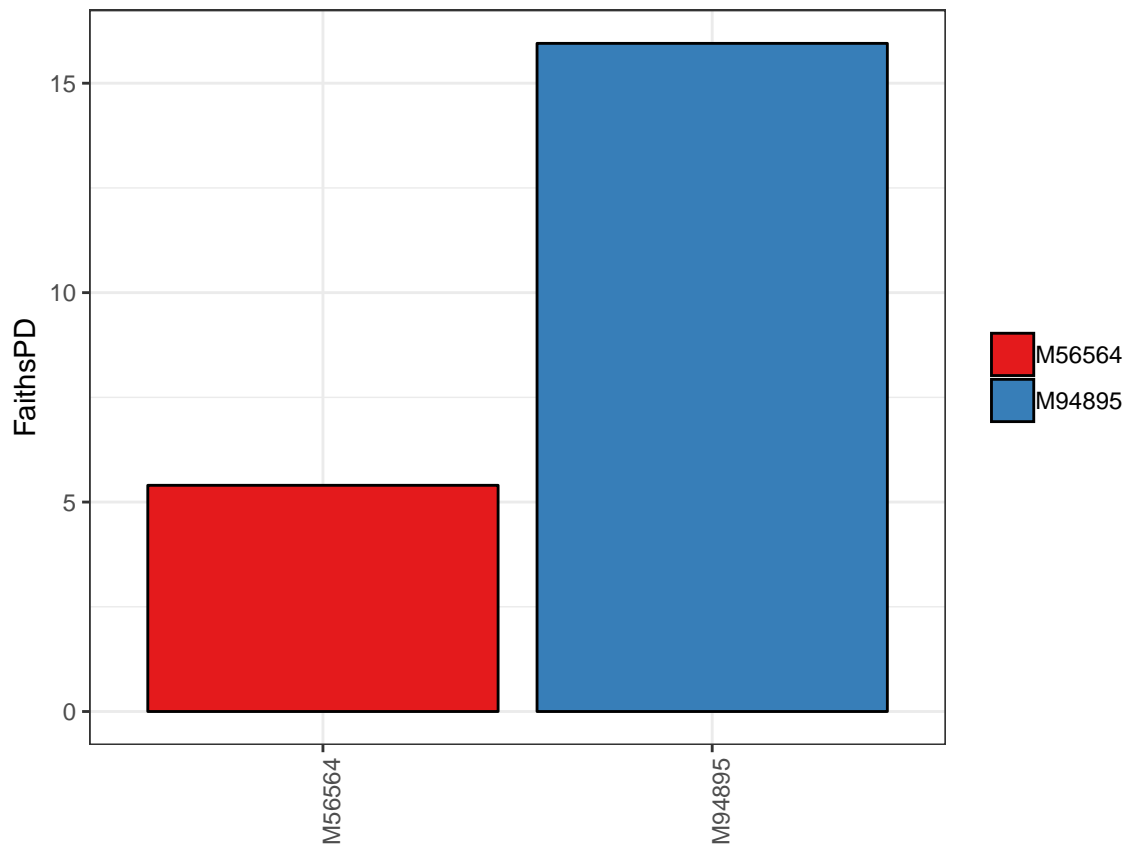

Figure 8: Faith's Phylogenetic Diversity Measure of Alpha Diversity- Boxplots-Animal_ID.
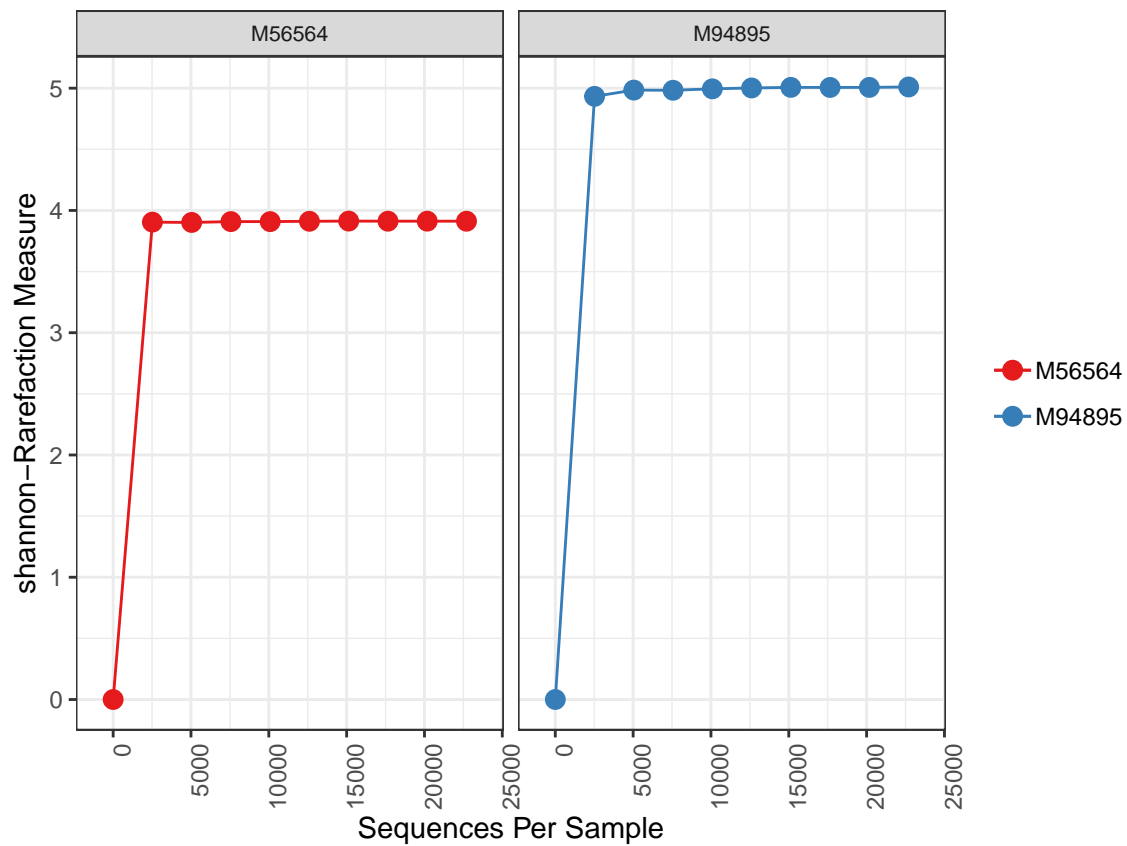
Figure 9: Shannon Measure of Alpha Diversity- Rarefaction curve-Animal_ID.
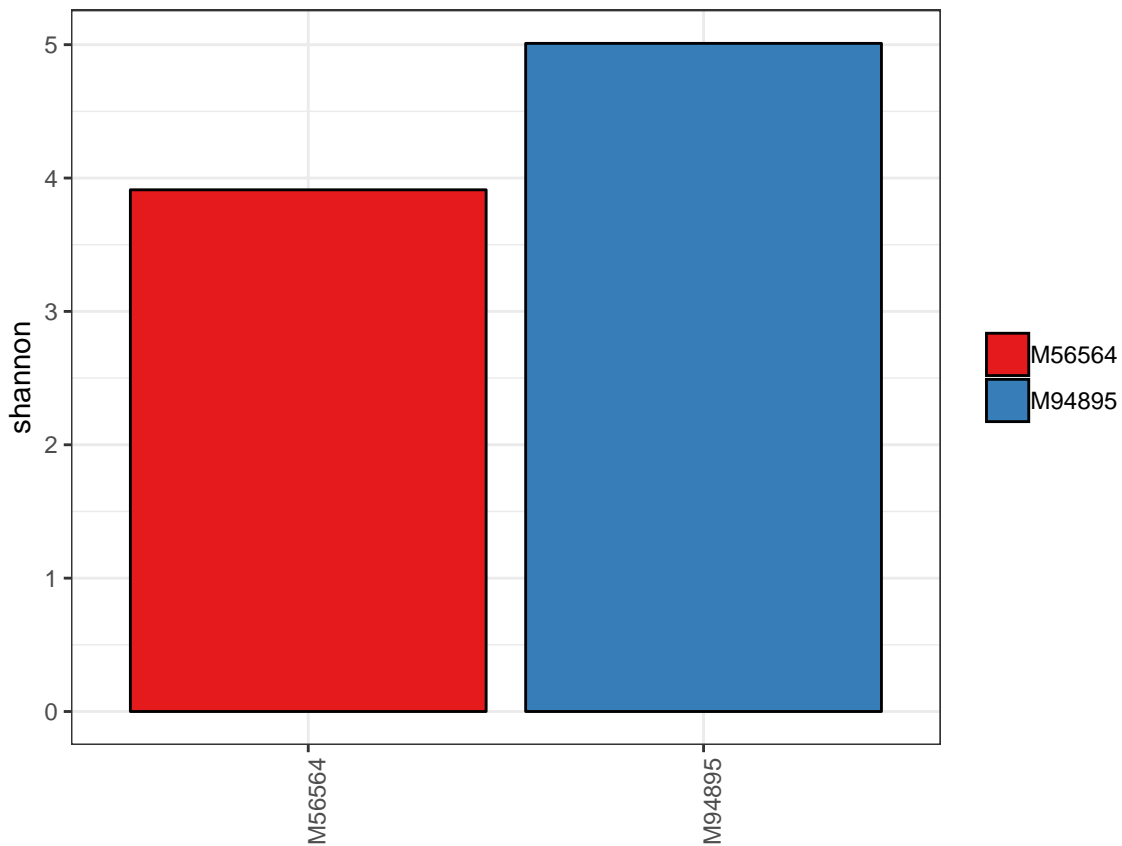

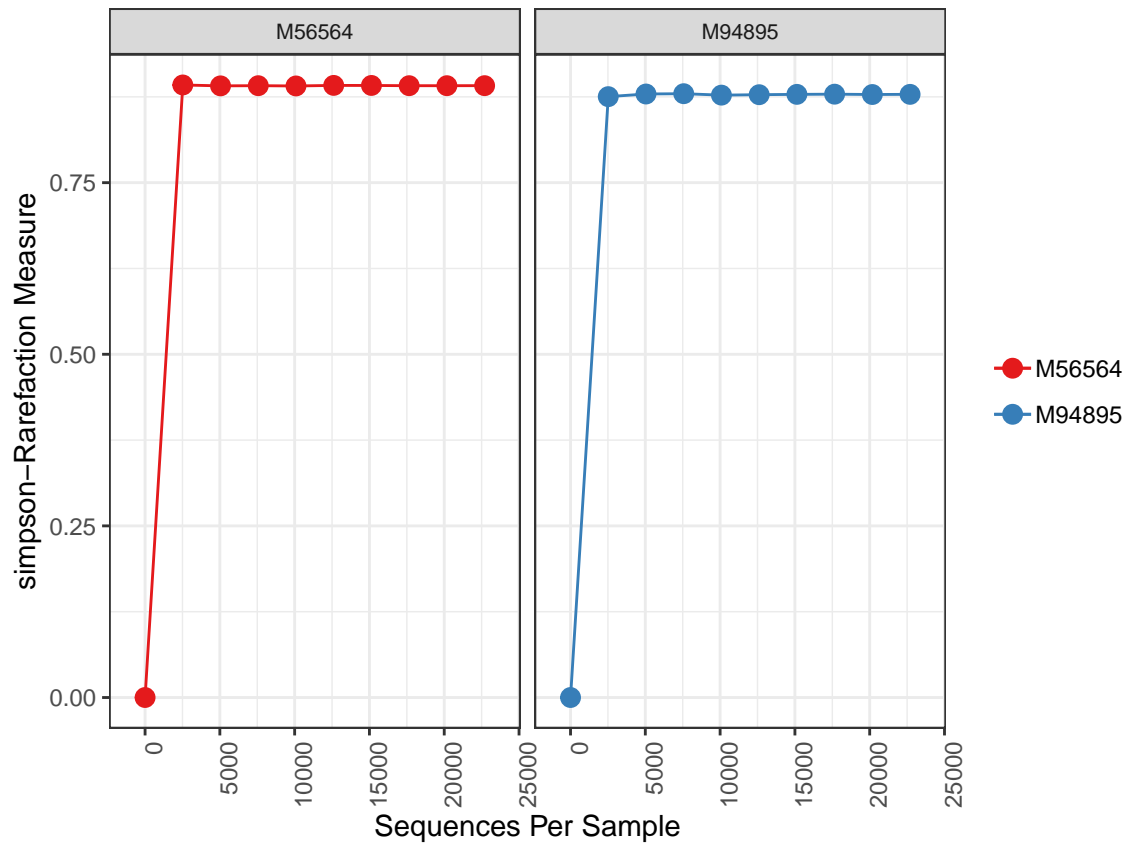Figure 10: Shannon Measure of Alpha Diversity- Boxplots-Animal_ID.

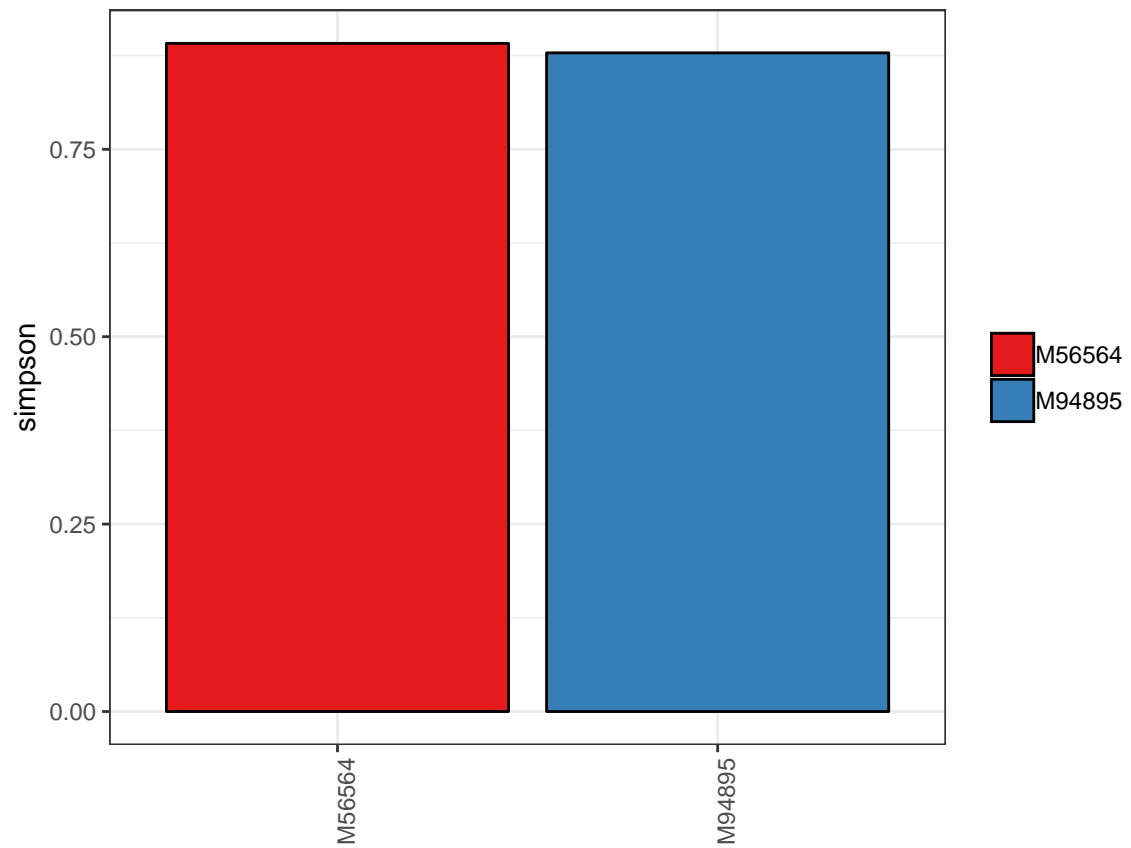Figure 11: Simpson Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

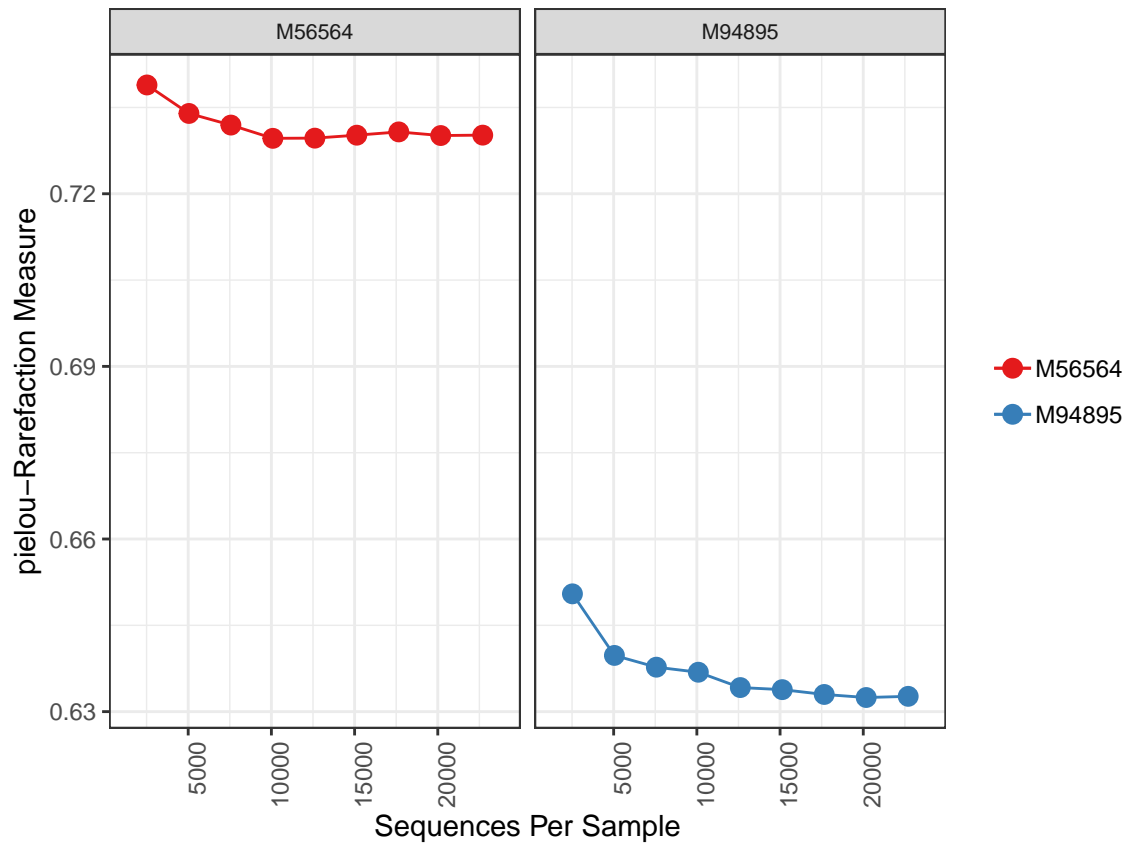Figure 12: Simpson Measure of Alpha Diversity- Boxplots-Animal_ID.

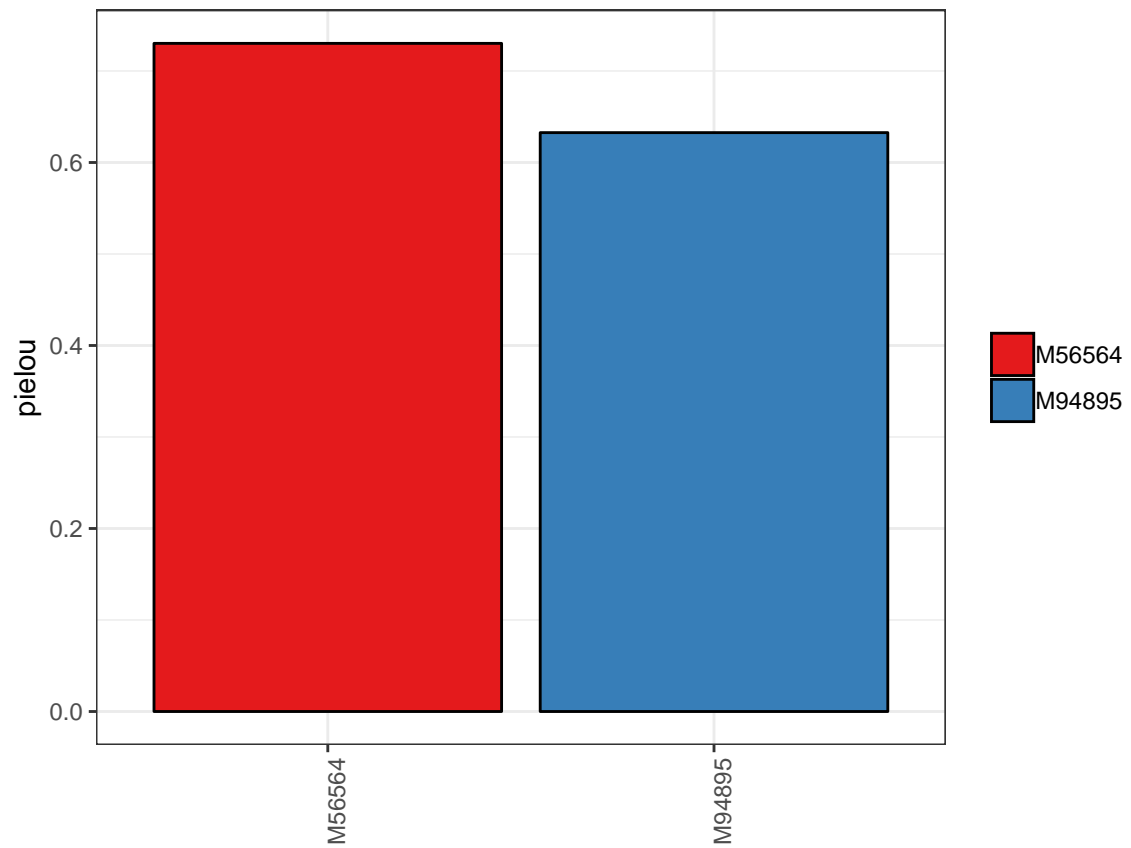Figure 13: Pielou's Evenness Measure of Alpha Diversity- Rarefaction curve-Animal_ID.

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).
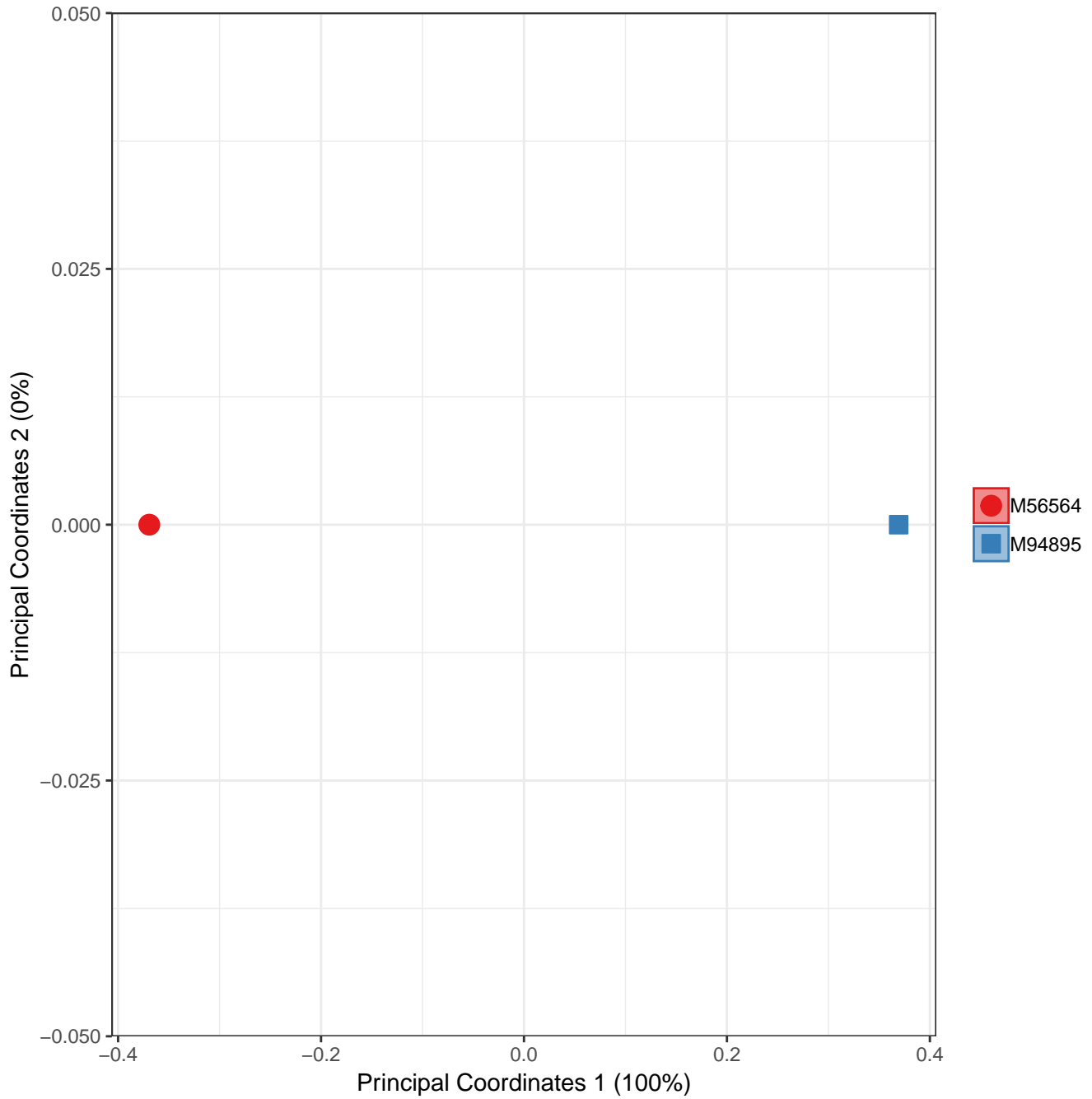

Figure 15: Unweighted Unifrac Measure of Beta Diversity-Animal_ID.
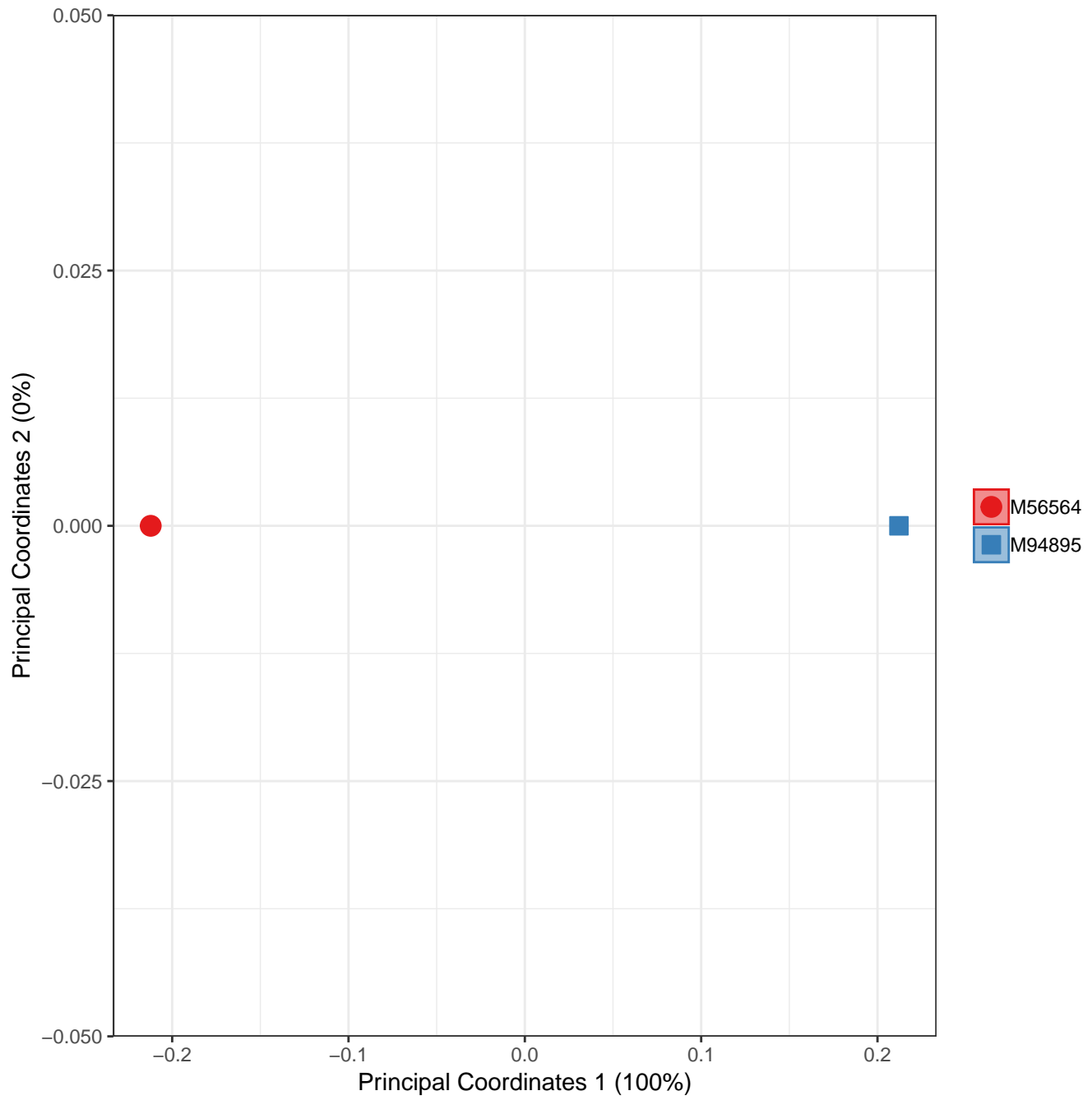
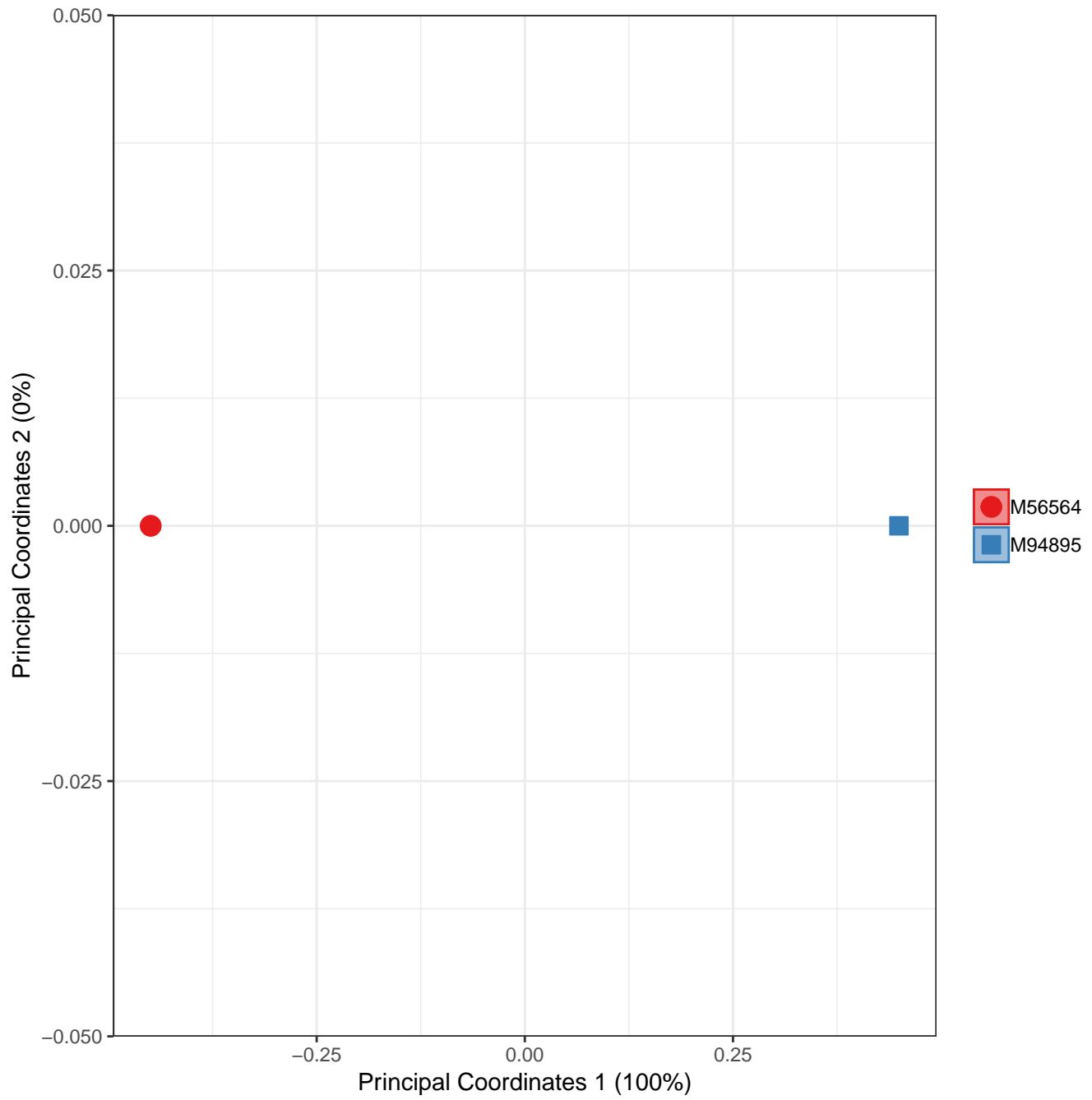Figure 16: Weighted Unifrac Measure of Beta Diversity-Animal_ID.

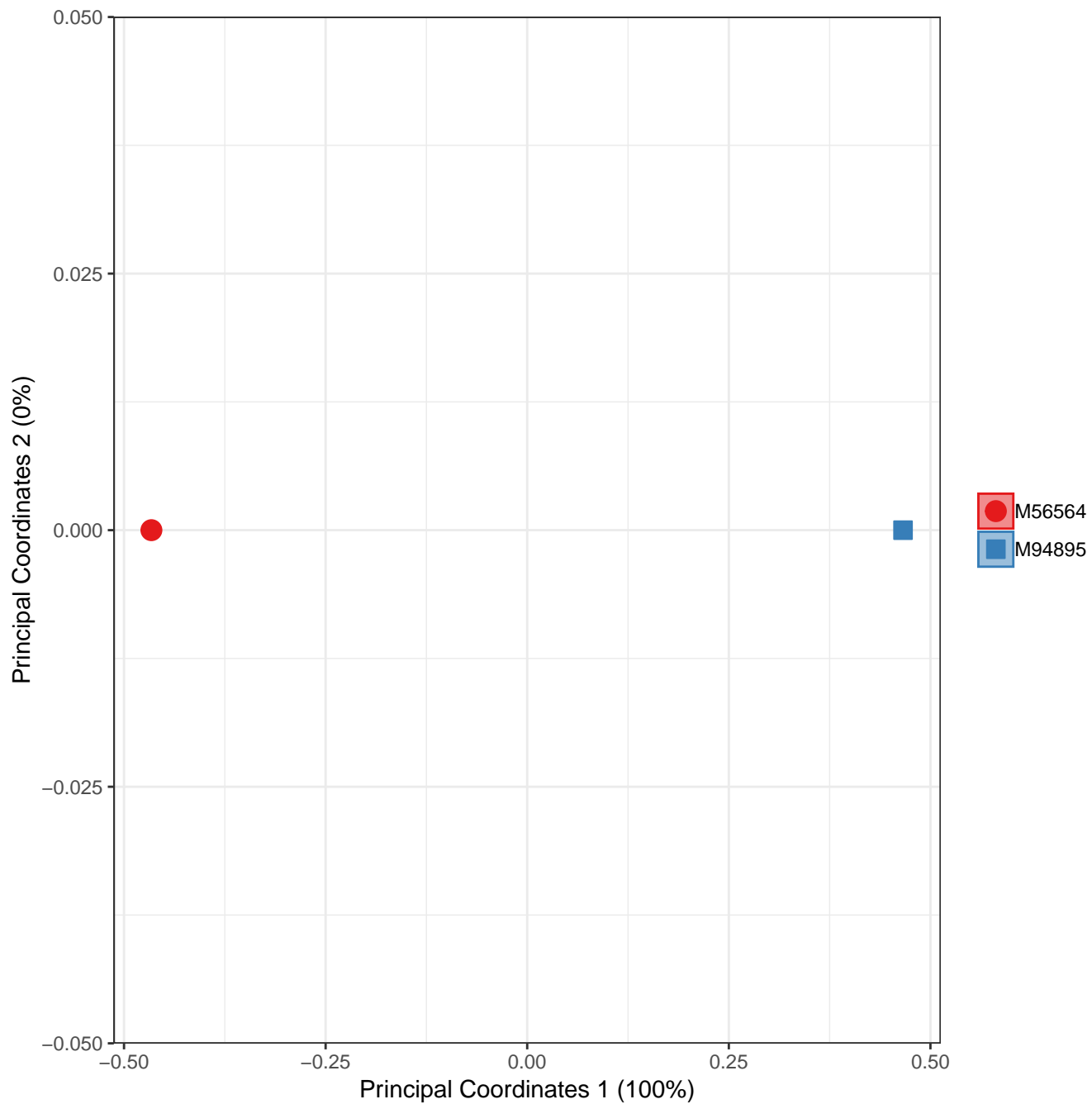Figure 17: Bray-Curtis Measure of Beta Diversity-Animal_ID.

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 abundace 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.
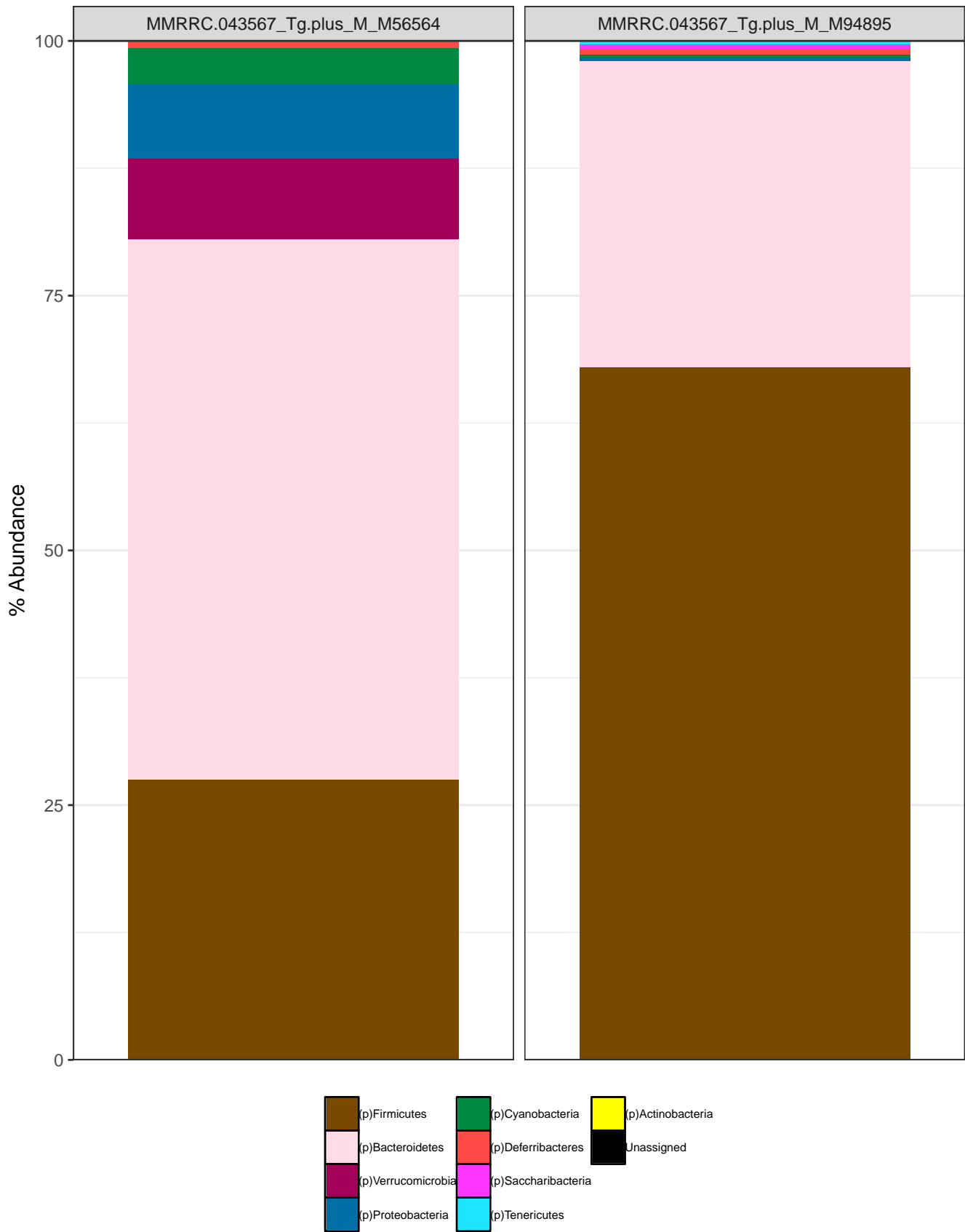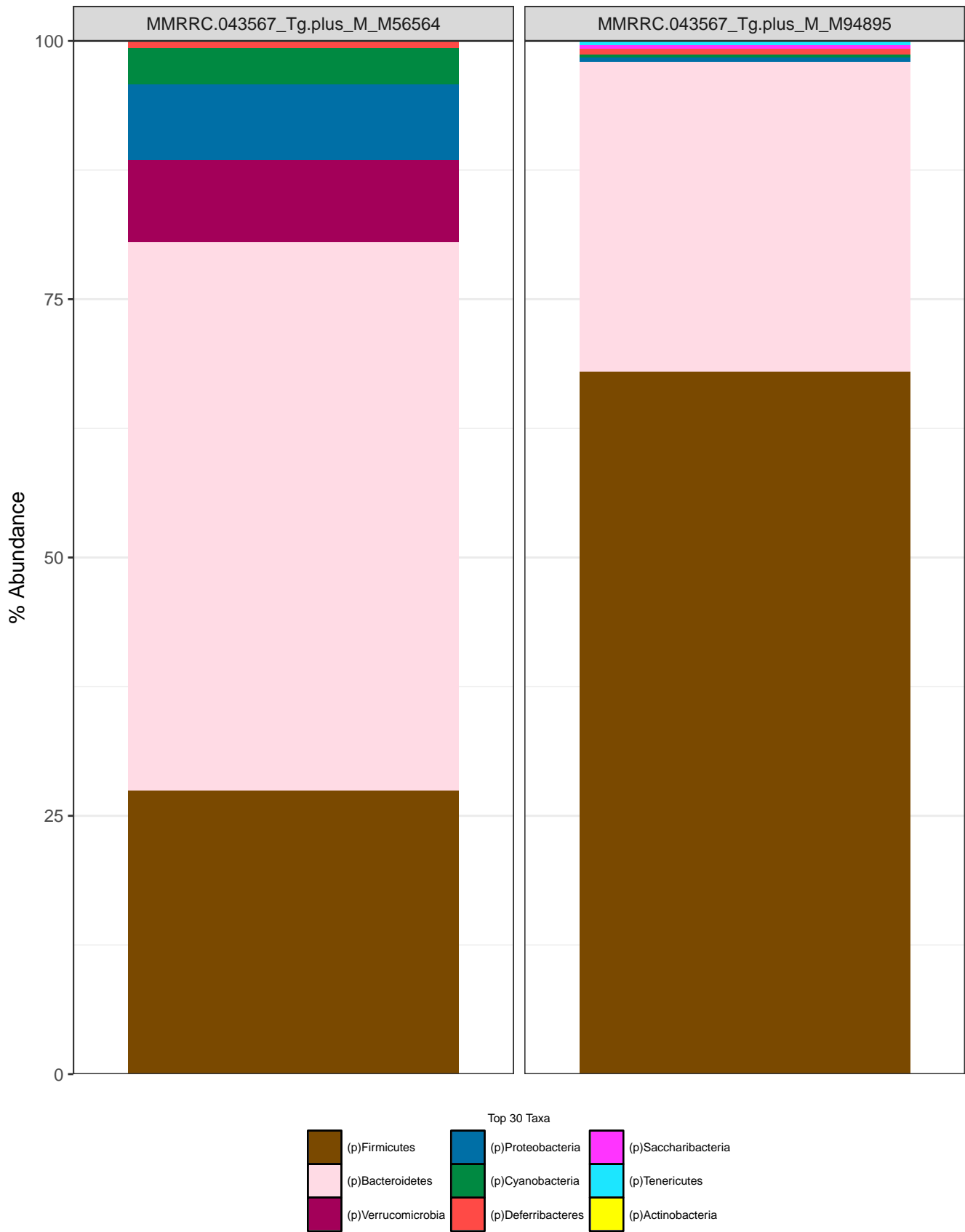
Figure 19: Phylum level abundances-UNFILTERED

Figure 20: Phylum level abundances-FILTERED: 1 phylum level taxa removed

Figure 21: Phylum level abundances-FILTERED & RESCALED

Figure 22: Class level abundances-UNFILTERED

Figure 23: Class level abundances-FILTERED: 2 class level taxa removed

Figure 24: Class level abundances-FILTERED & RESCALED
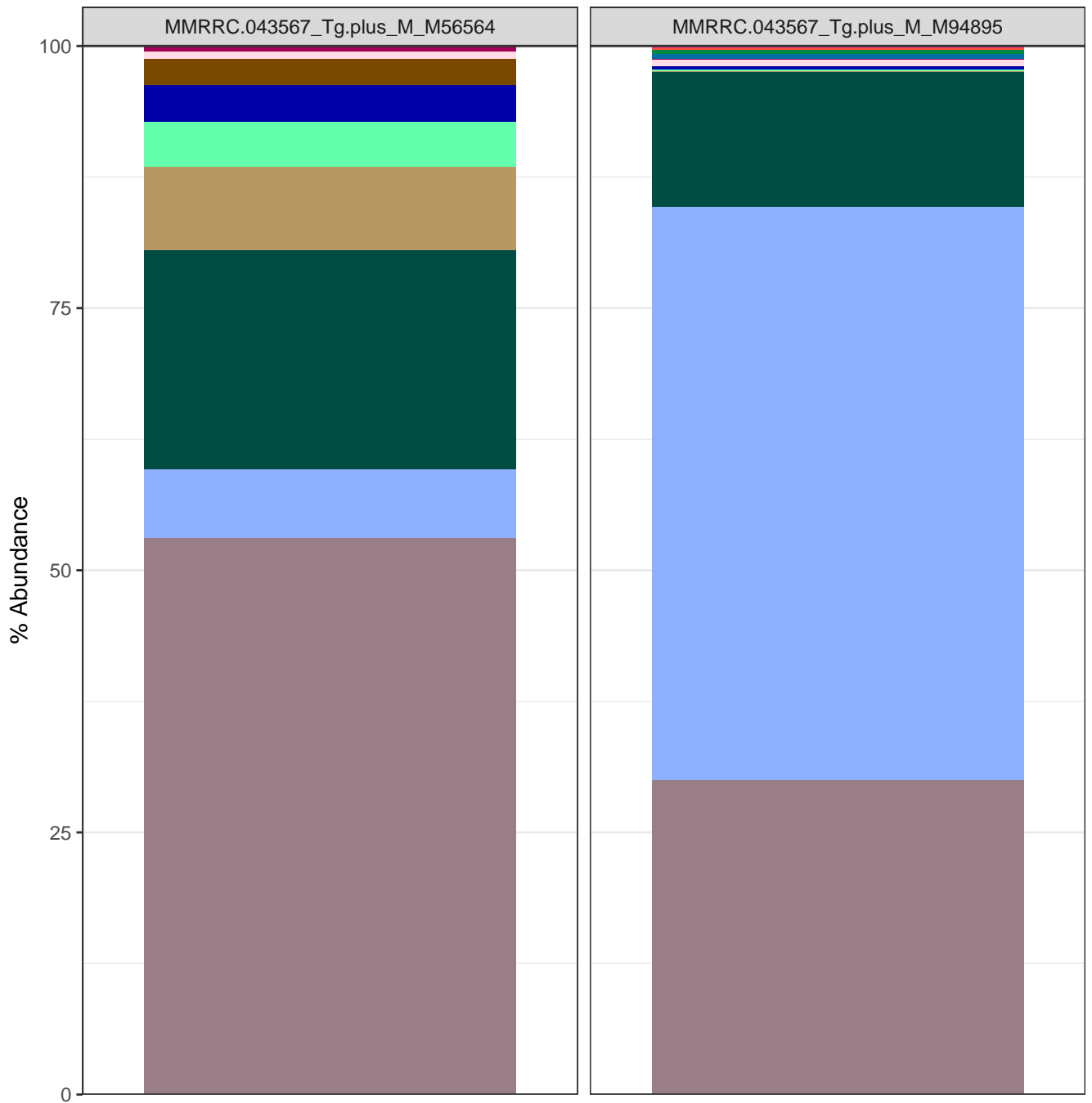
Figure 25: Order level abundances-UNFILTERED
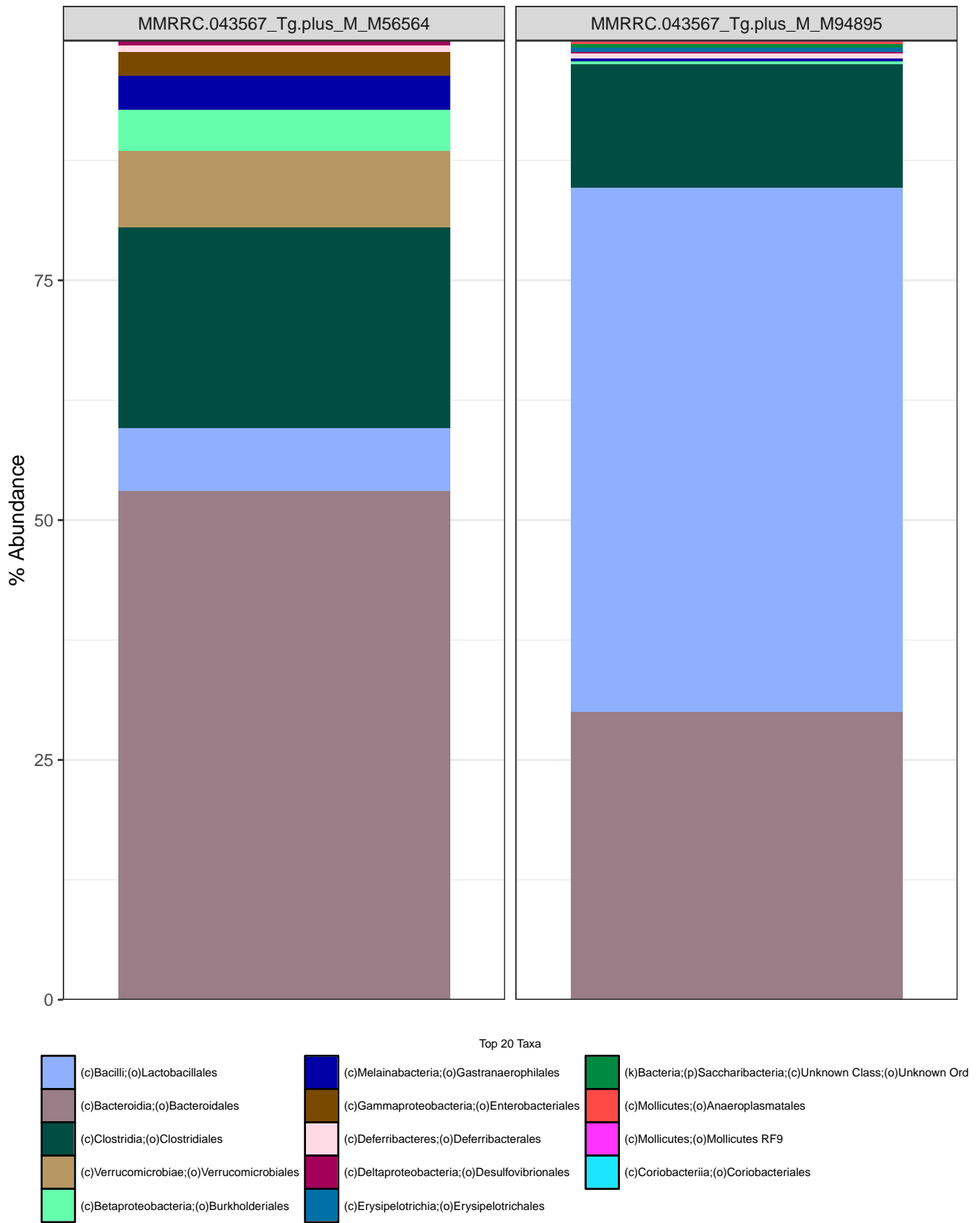
Figure 26: Order level abundances-FILTERED: 2 order level taxa removed
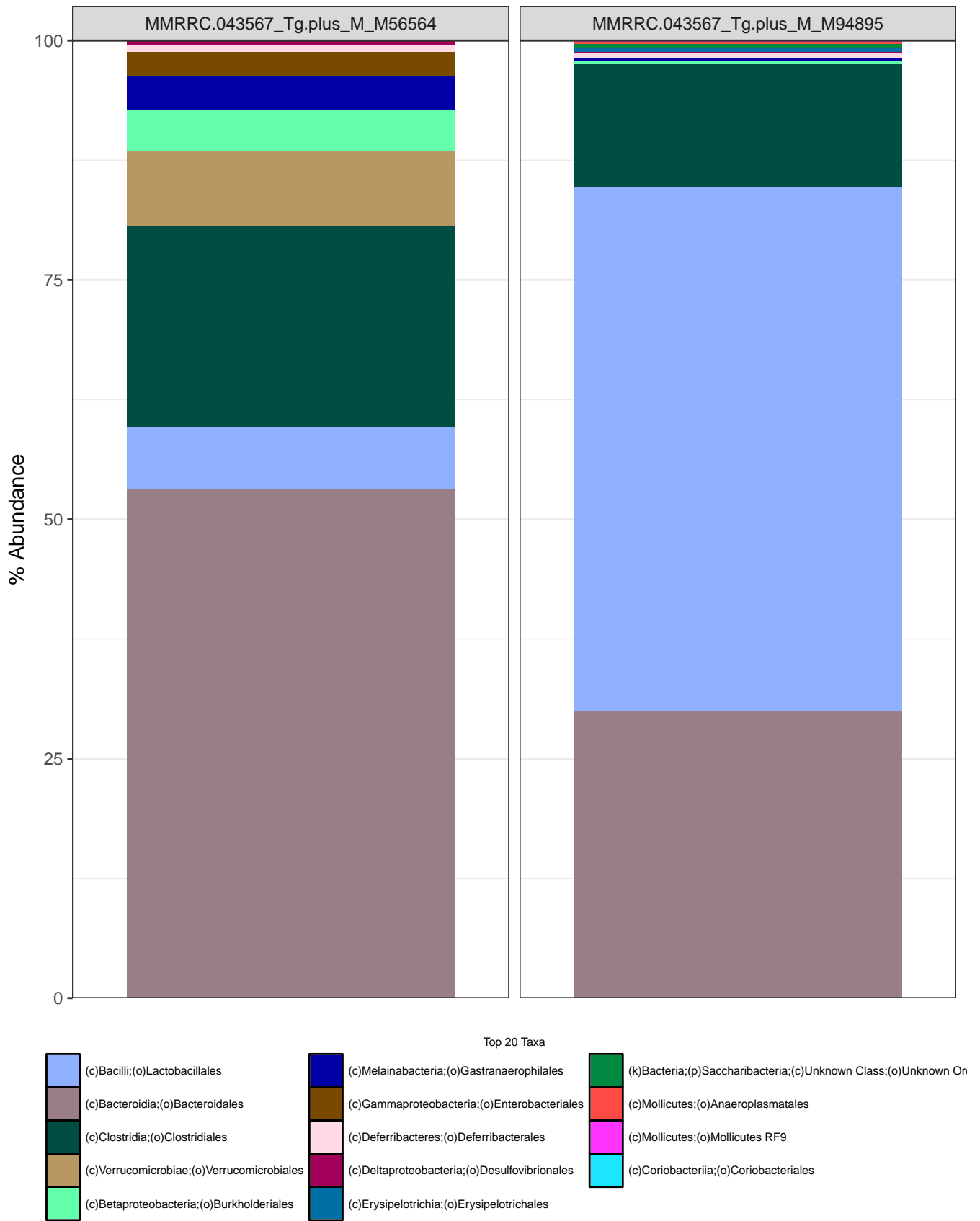
Figure 27: Order level abundances-FILTERED & RESCALED
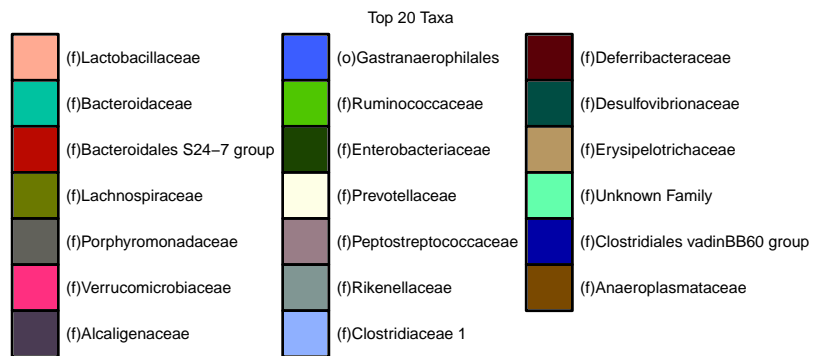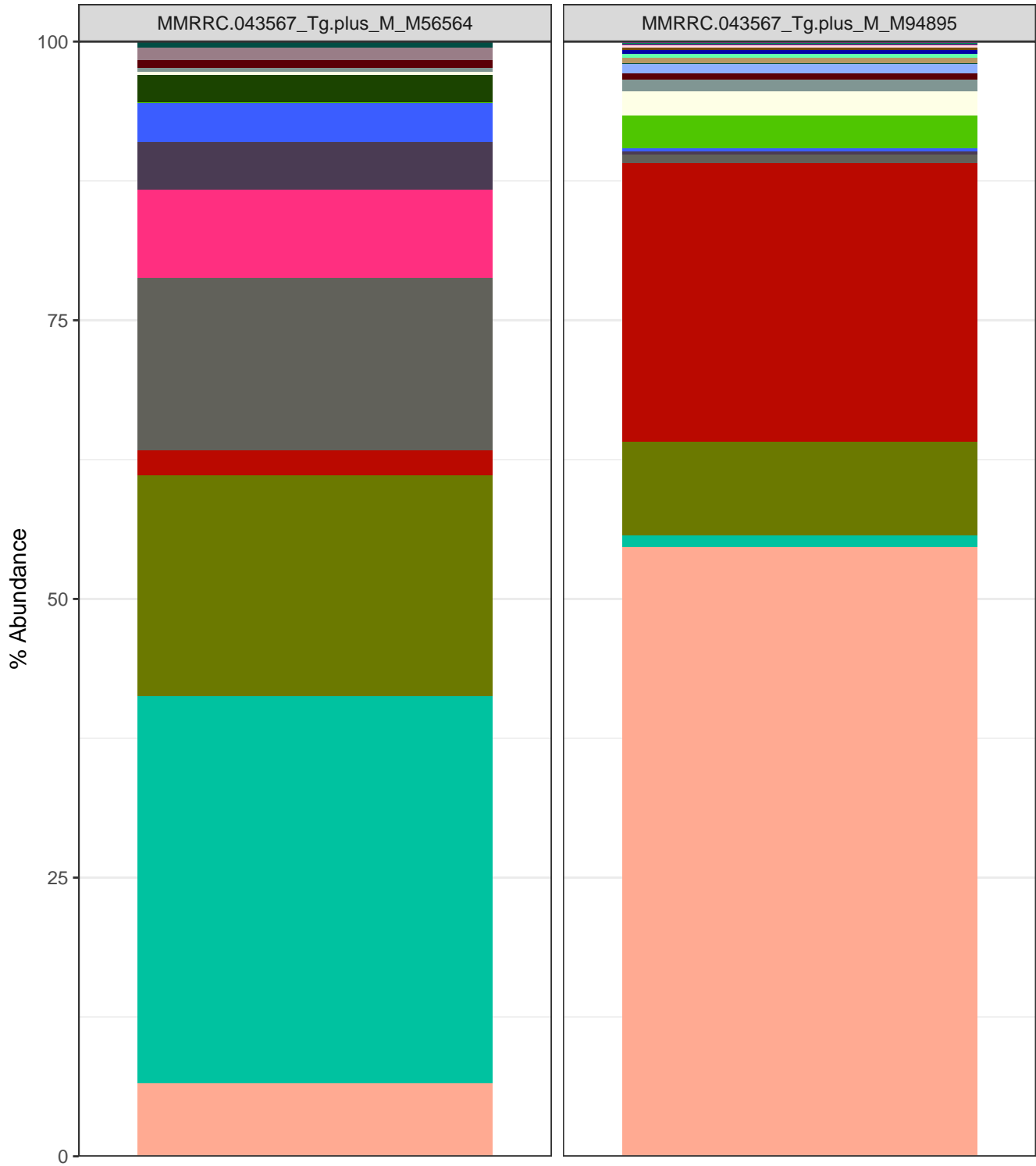
Figure 28: Family level abundances-UNFILTERED
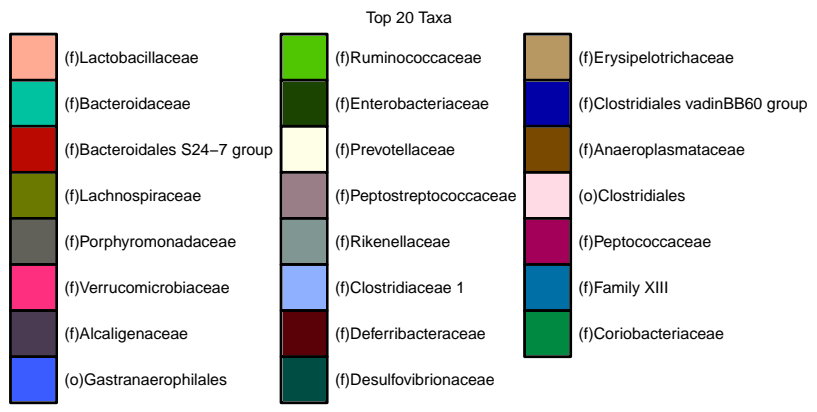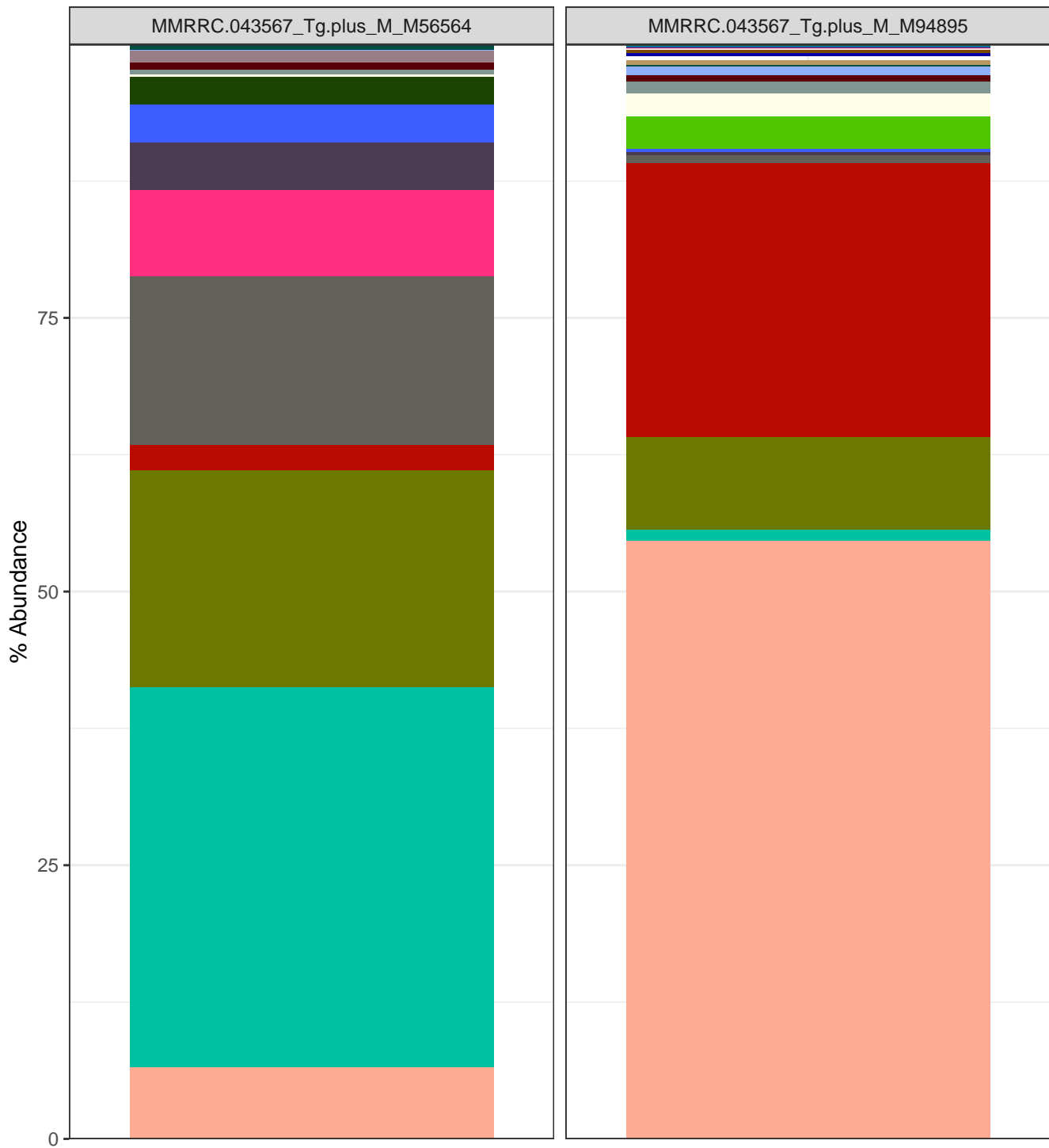
Figure 29: Family level abundances-FILTERED: 6 family level taxa removed
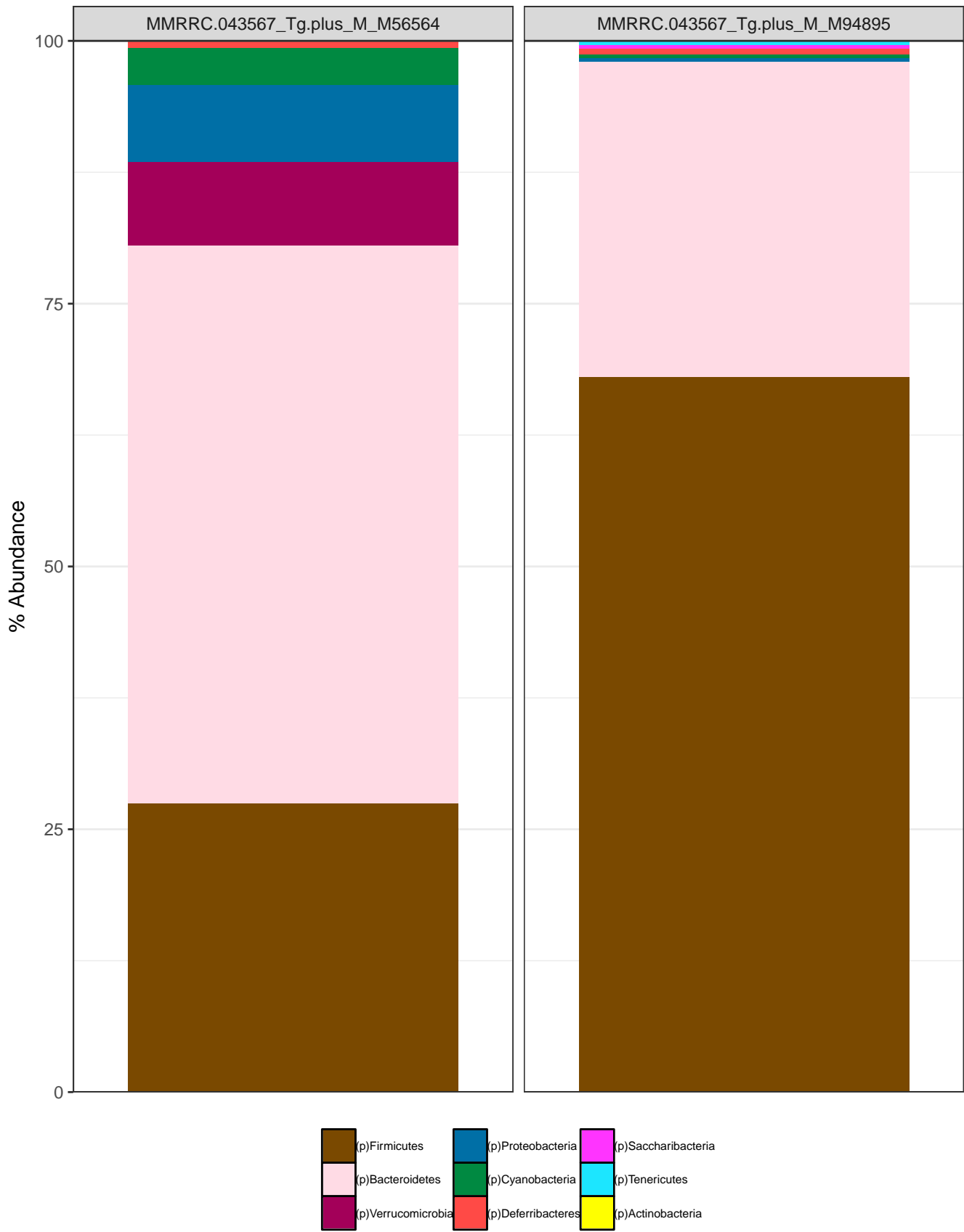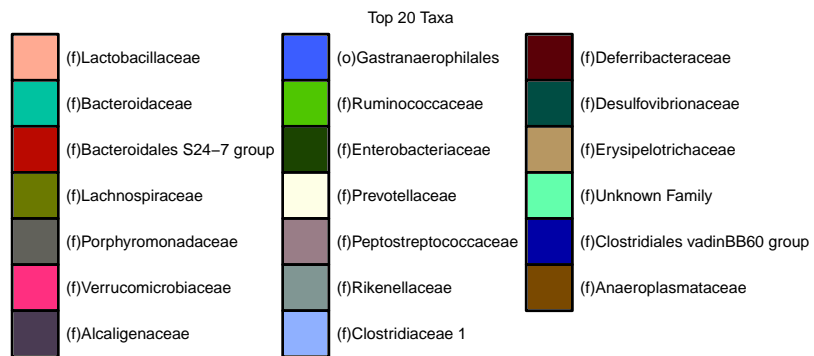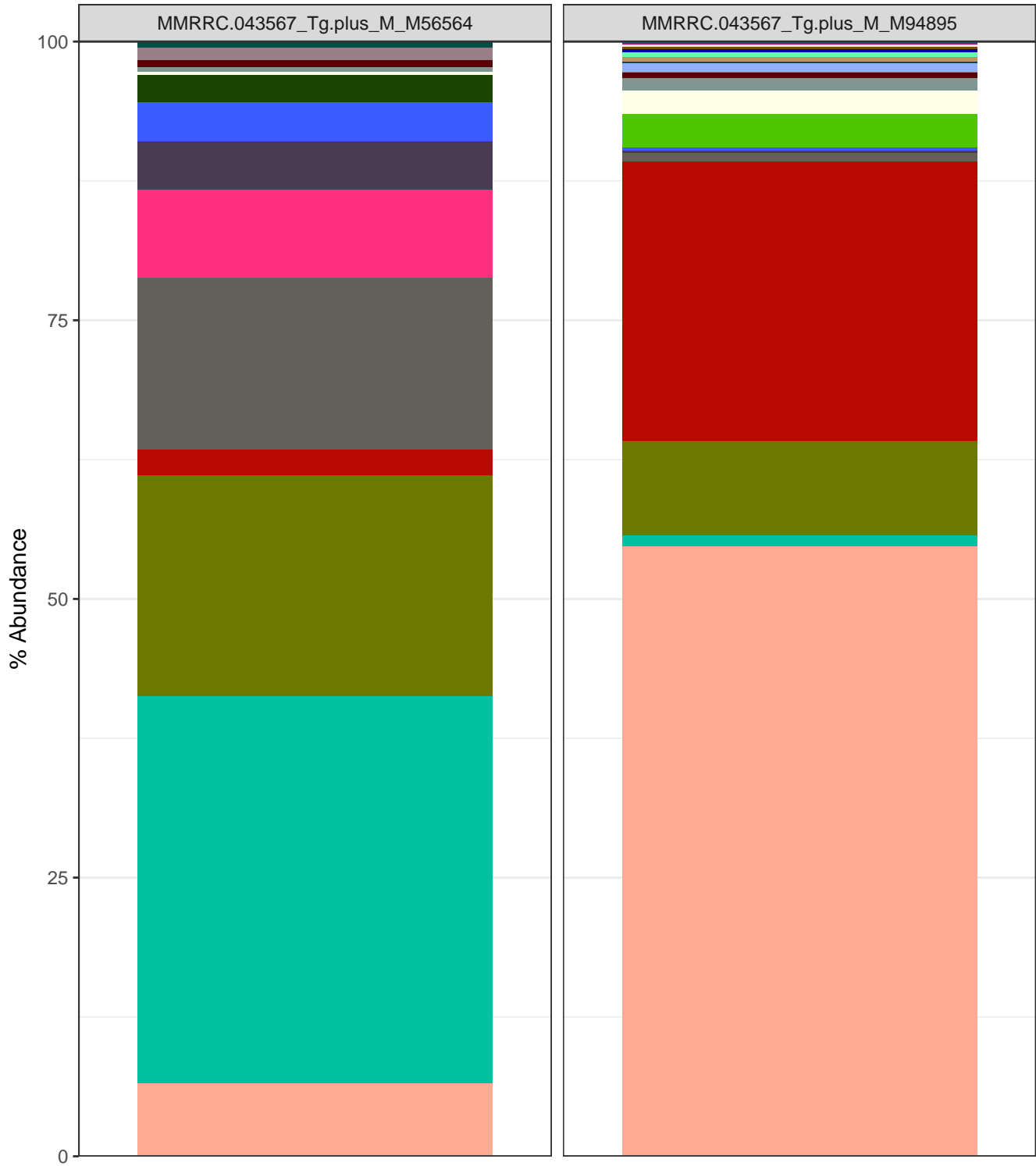
Figure 30: Family level abundances-FILTERED & RESCALED
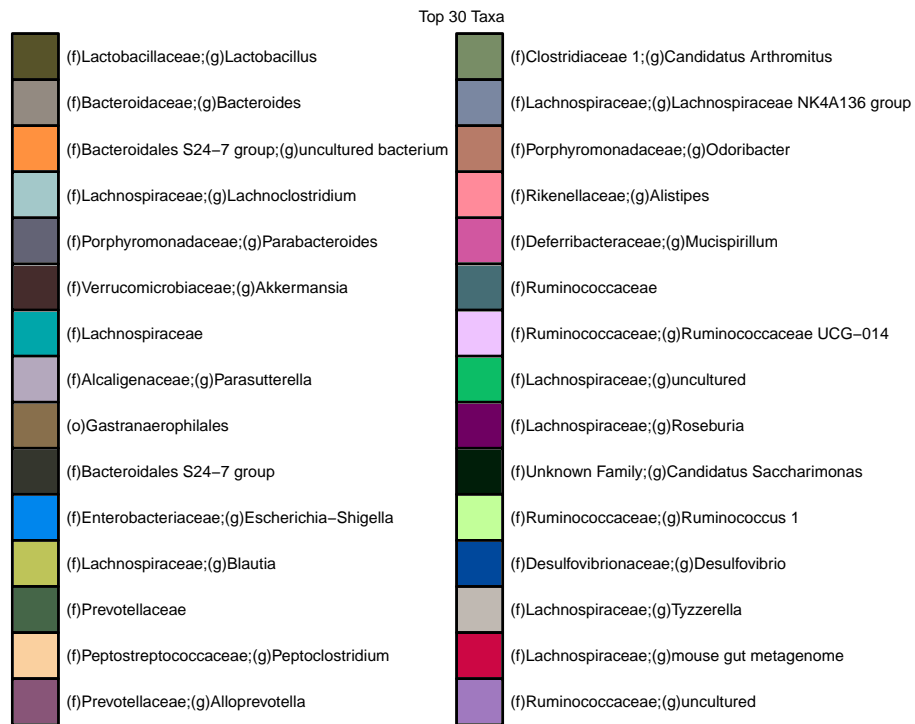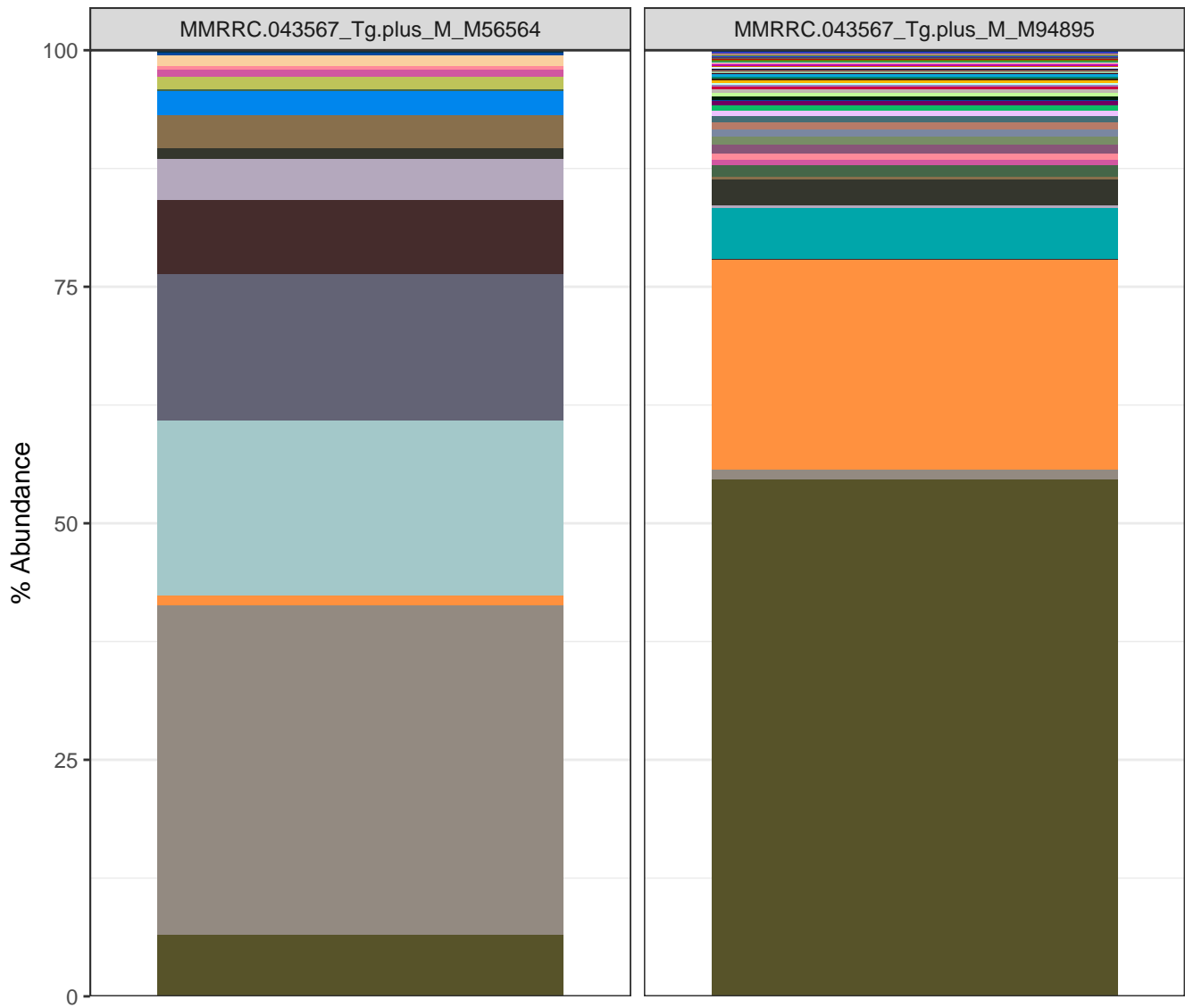
Figure 31: Genus level abundances-UNFILTERED
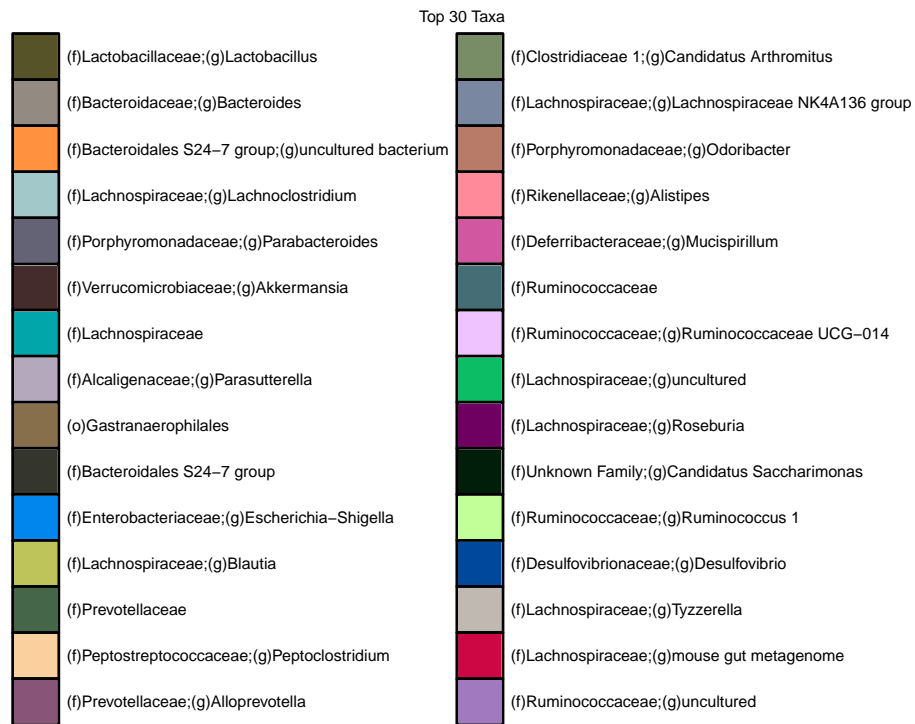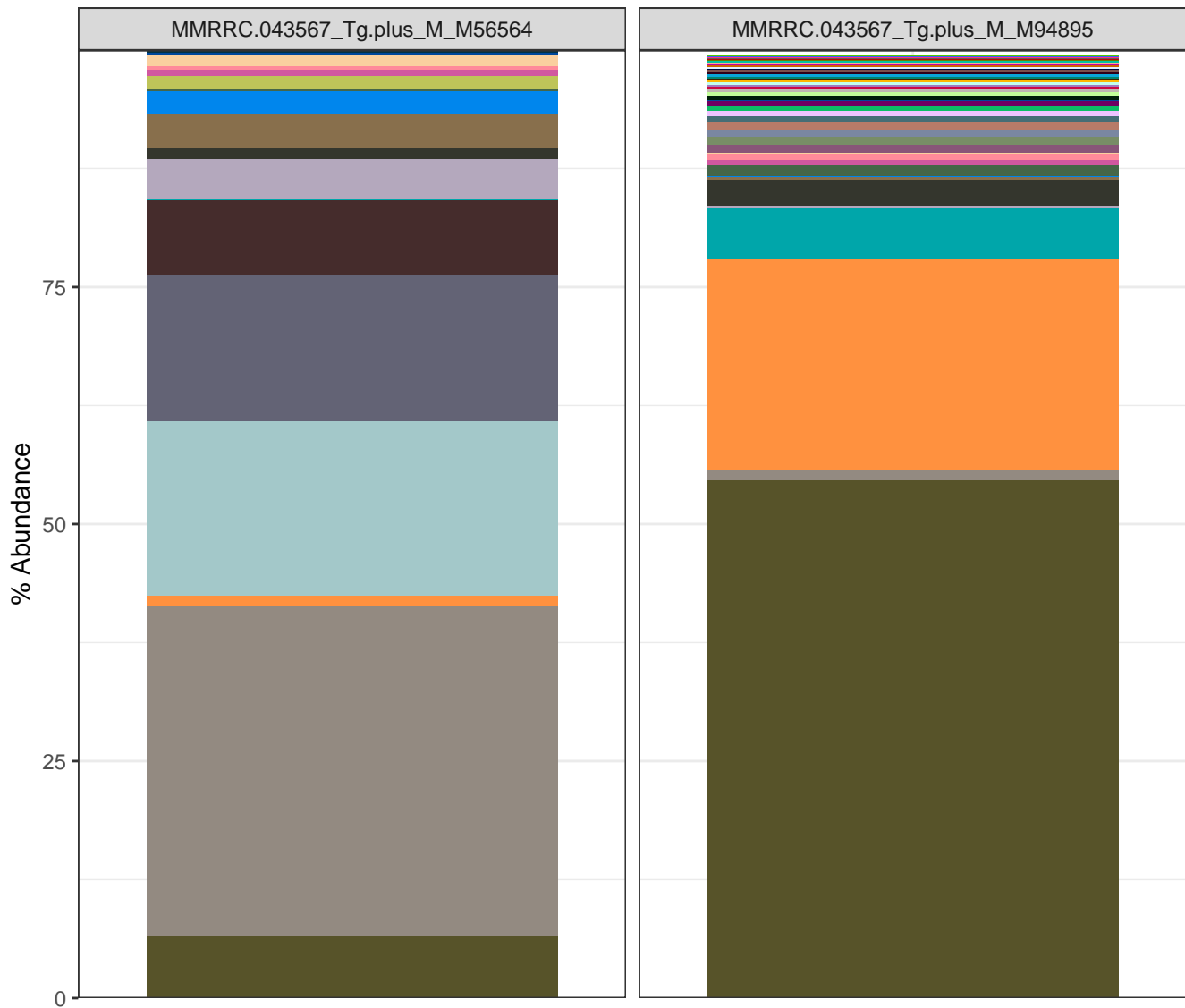
Figure 32: Genus level abundances-FILTERED: 18 genus level taxa removed
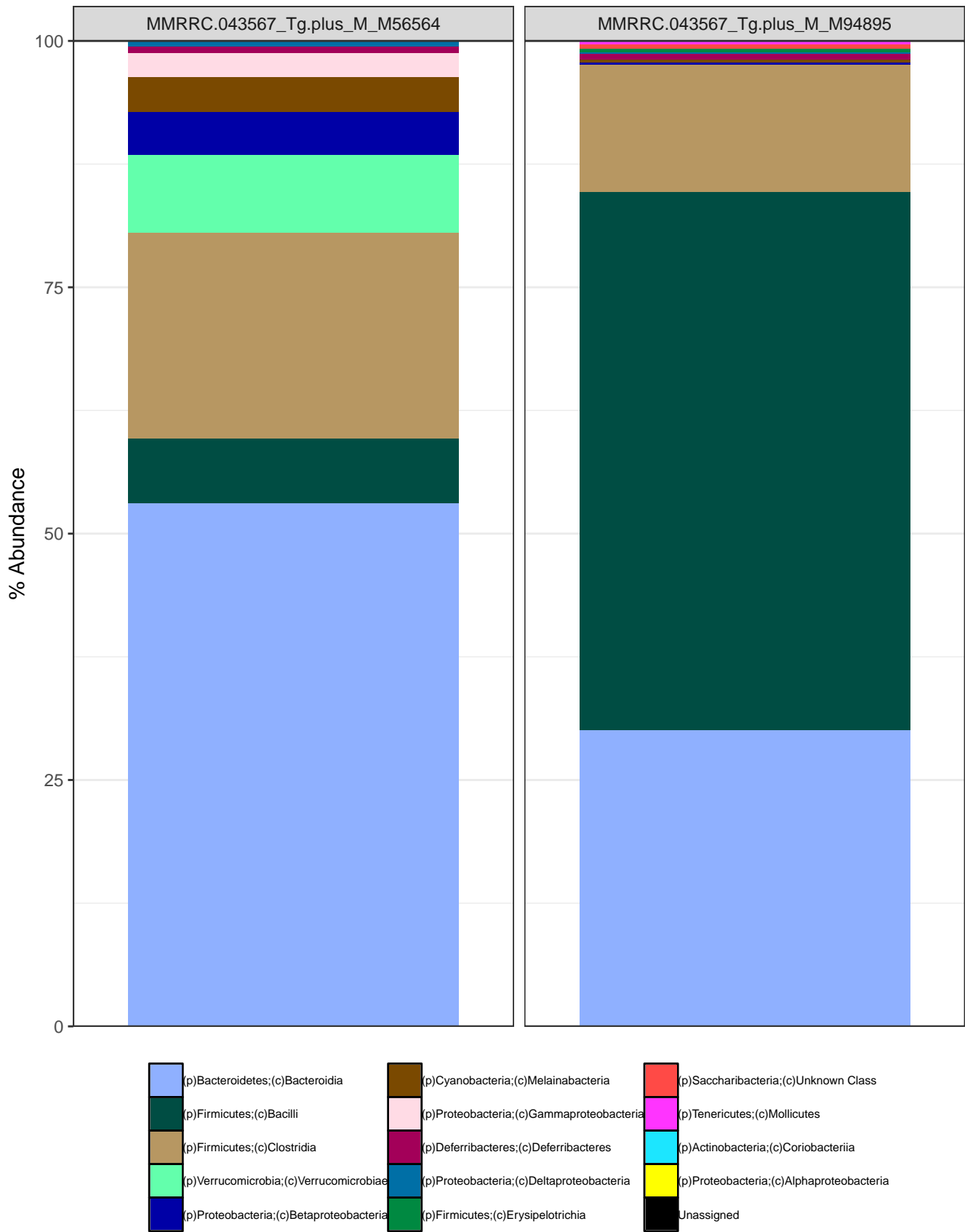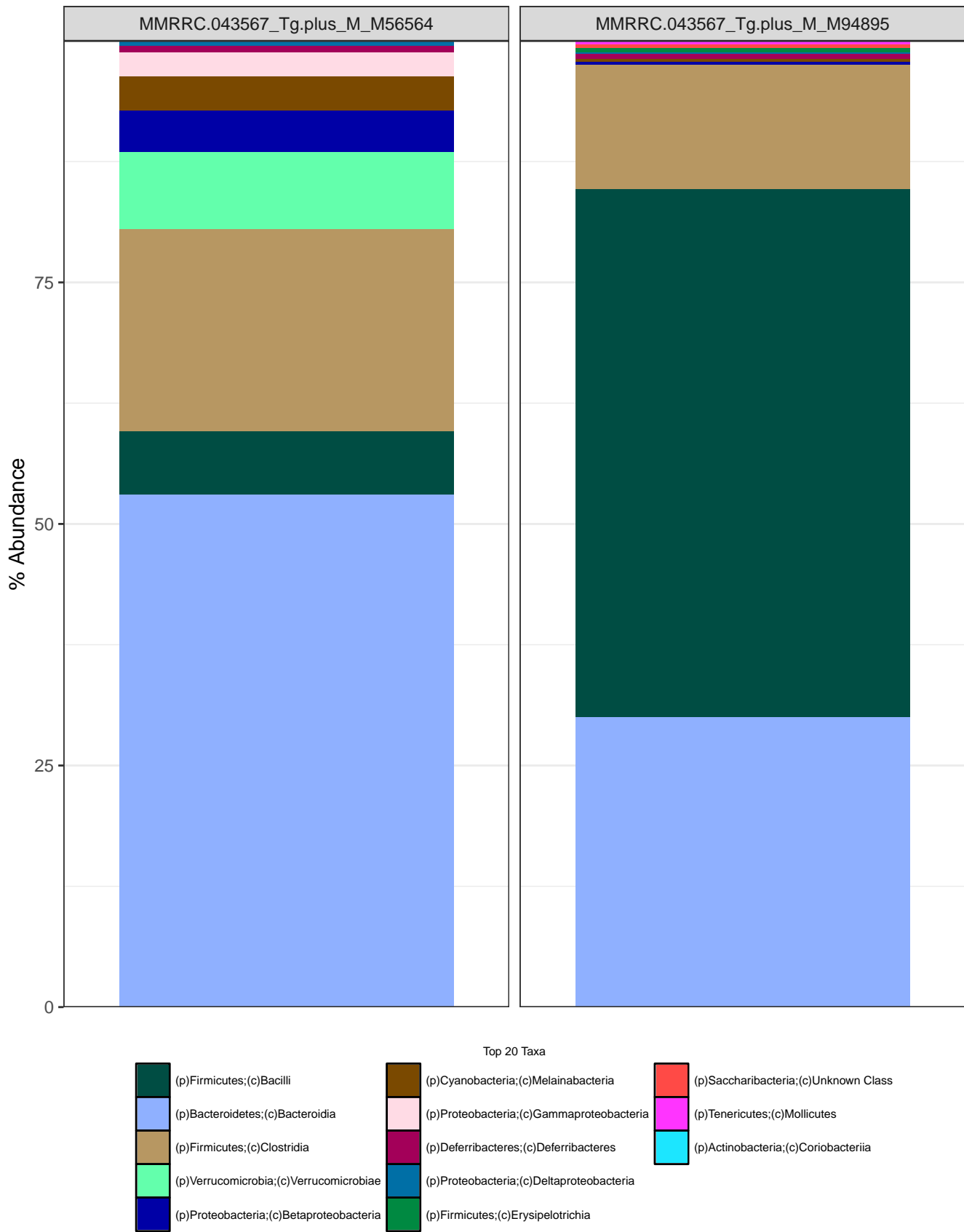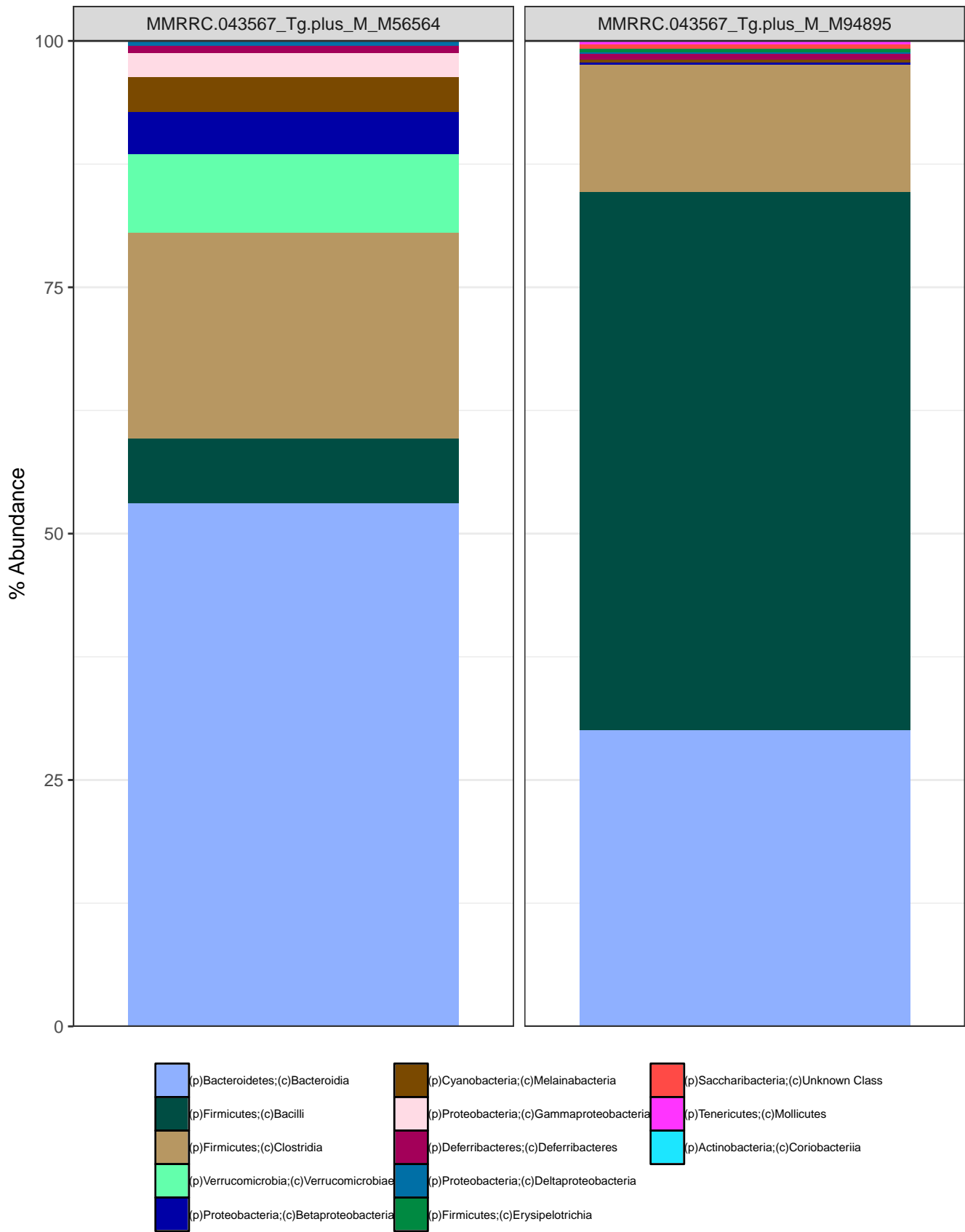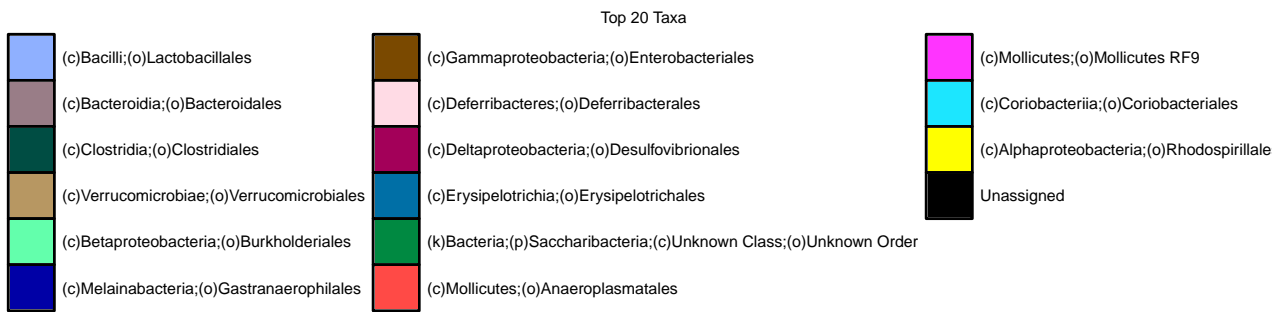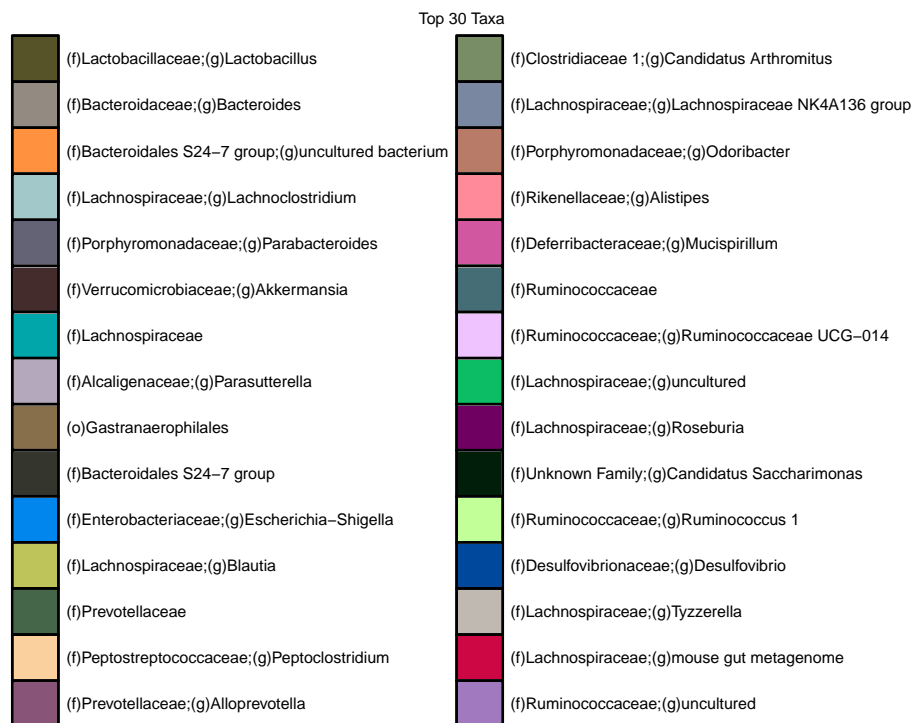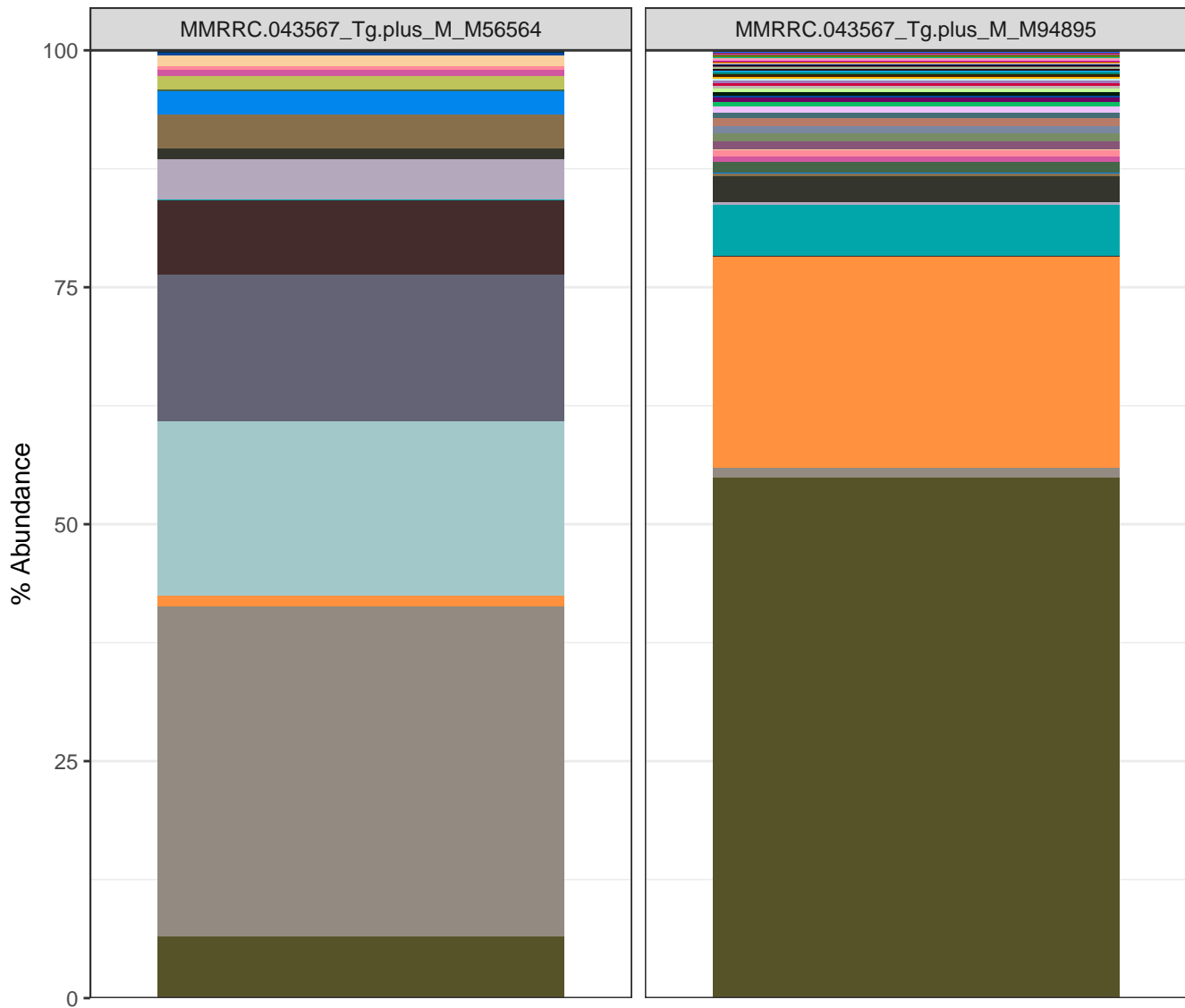
Figure 33: Genus level abundances-FILTERED & RESCALED

A   Appendix 1 (Taxa Abundance Tables)

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(p)Actinobacteria	0%	0.06%
(p)Bacteroidetes	53.09%	30.03%
(p)Cyanobacteria	3.52%	0.31%
(p)Deferribacteres	0.68%	0.55%
(p)Firmicutes	27.49%	67.99%
(p)Proteobacteria	7.32%	0.36%
(p)Saccharibacteria	0%	0.39%
(p)Tenericutes	0%	0.29%
(p)Verrucomicrobia	7.91%	0%
NA	**100**	**99.989**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(p)Actinobacteria;(c)Coriobacteriia	0%	0.06%
(p)Bacteroidetes;(c)Bacteroidia	53.09%	30.03%
(p)Cyanobacteria;(c)Melainabacteria	3.52%	0.31%
(p)Deferribacteres;(c)Deferribacteres	0.68%	0.55%
(p)Firmicutes;(c)Bacilli	6.54%	54.66%
(p)Firmicutes;(c)Clostridia	20.95%	12.9%
(p)Firmicutes;(c)Erysipelotrichia	0%	0.43%
(p)Proteobacteria;(c)Betaproteobacteria	4.31%	0.23%
(p)Proteobacteria;(c)Deltaproteobacteria	0.5%	0.13%
(p)Proteobacteria;(c)Gammaproteobacteria	2.48%	0%
(p)Saccharibacteria;(c)Unknown Class	0%	0.39%
(p)Tenericutes;(c)Mollicutes	0%	0.29%
(p)Verrucomicrobia;(c)Verrucomicrobiae	7.91%	0%
NA	**99.969**	**99.989**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(c)Bacilli;(o)Lactobacillales	6.54%	54.66%
(c)Bacteroidia;(o)Bacteroidales	53.09%	30.03%
(c)Betaproteobacteria;(o)Burkholderiales	4.31%	0.23%
(c)Clostridia;(o)Clostridiales	20.95%	12.9%
(c)Coriobacteriia;(o)Coriobacteriales	0%	0.06%
(c)Deferribacteres;(o)Deferribacterales	0.68%	0.55%
(c)Deltaproteobacteria;(o)Desulfovibrionales	0.5%	0.13%
(c)Erysipelotrichia;(o)Erysipelotrichales	0%	0.43%
(c)Gammaproteobacteria;(o)Enterobacteriales	2.48%	0%
(c)Melainabacteria;(o)Gastranaerophilales	3.52%	0.31%
(c)Mollicutes;(o)Anaeroplasmatales	0%	0.22%
(c)Mollicutes;(o)Mollicutes RF9	0%	0.06%
(c)Verrucomicrobiae;(o)Verrucomicrobiales	7.91%	0%
(k)Bacteria;(p)Saccharibacteria;(c)Unknown Class;(o)Unknown Order	0%	0.39%
NA	**99.969**	**99.989**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(f)Alcaligenaceae	4.31%	0.23%
(f)Anaeroplasmataceae	0%	0.22%
(f)Bacteroidaceae	34.77%	1.03%
(f)Bacteroidales S24-7 group	2.26%	25.03%
(f)Clostridiaceae 1	0%	0.8%
(f)Clostridiales vadinBB60 group	0%	0.32%
(f)Coriobacteriaceae	0%	0.06%
(f)Deferribacteraceae	0.68%	0.55%
(f)Desulfovibrionaceae	0.5%	0.13%
(f)Enterobacteriaceae	2.48%	0%
(f)Erysipelotrichaceae	0%	0.43%
(f)Family XIII	0%	0.08%
(f)Lachnospiraceae	19.83%	8.43%
(f)Lactobacillaceae	6.54%	54.66%
(f)Peptococcaceae	0%	0.1%
(f)Peptostreptococcaceae	1.12%	0%
(f)Porphyromonadaceae	15.42%	0.78%
(f)Prevotellaceae	0.23%	2.11%
(f)Rikenellaceae	0.41%	1.09%
(f)Ruminococcaceae	0%	2.98%
(f)Verrucomicrobiaceae	7.91%	0%
(k)Bacteria;(p)Saccharibacteria;(c)Unknown Class;(o)Unknown Order;(f)Unknown Family	0%	0.39%
(o)Clostridiales	0%	0.19%
(o)Gastranaerophilales	3.52%	0.29%
NA	**99.969**	**99.903**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(f)Alcaligenaceae;(g)Parasutterella	4.31%	0.23%
(f)Anaeroplasmataceae;(g)Anaeroplasma	0%	0.22%
(f)Bacteroidaceae;(g)Bacteroides	34.77%	1.03%
(f)Bacteroidales S24-7 group	1.15%	2.79%
(f)Bacteroidales S24-7 group;(g)uncultured bacterium	1.11%	22.24%
(f)Clostridiaceae 1;(g)Candidatus Arthromitus	0%	0.8%
(f)Clostridiales vadinBB60 group	0%	0.07%
(f)Clostridiales vadinBB60 group;(g)uncultured bacterium	0%	0.22%
(f)Coriobacteriaceae	0%	0.06%
(f)Deferribacteraceae;(g)Mucispirillum	0.68%	0.55%
(f)Desulfovibrionaceae;(g)Bilophila	0.2%	0%
(f)Desulfovibrionaceae;(g)Desulfovibrio	0.3%	0.13%
(f)Enterobacteriaceae;(g)Escherichia-Shigella	2.48%	0%
(f)Erysipelotrichaceae;(g)Holdemanella	0%	0.11%
(f)Erysipelotrichaceae;(g)Holdemania	0%	0.1%
(f)Erysipelotrichaceae;(g)uncultured	0%	0.17%
(f)Family XIII;(g)[Eubacterium] nodatum group	0%	0.07%
(f)Lachnospiraceae	0%	5.43%
(f)Lachnospiraceae;(g)[Eubacterium] eligens group	0%	0.14%
(f)Lachnospiraceae;(g)Blautia	1.36%	0%
(f)Lachnospiraceae;(g)Coprococcus 1	0%	0.08%
(f)Lachnospiraceae;(g)Lachnoclostridium	18.47%	0%
(f)Lachnospiraceae;(g)Lachnospiraceae NK4A136 group	0%	0.79%
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-001	0%	0.06%
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-002	0%	0.11%
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-008	0%	0.16%
(f)Lachnospiraceae;(g)mouse gut metagenome	0%	0.26%
(f)Lachnospiraceae;(g)Roseburia	0%	0.5%
(f)Lachnospiraceae;(g)Tyzzerella	0%	0.3%
(f)Lachnospiraceae;(g)uncultured	0%	0.51%
(f)Lactobacillaceae;(g)Lactobacillus	6.54%	54.66%
(f)Peptococcaceae;(g)uncultured	0%	0.06%
(f)Peptostreptococcaceae;(g)Peptoclostridium	1.12%	0%
(f)Porphyromonadaceae;(g)Odoribacter	0%	0.77%
(f)Porphyromonadaceae;(g)Parabacteroides	15.42%	0.01%
(f)Prevotellaceae	0.23%	1.18%
(f)Prevotellaceae;(g)Alloprevotella	0%	0.93%
(f)Rikenellaceae;(g)Alistipes	0.41%	0.71%
(f)Rikenellaceae;(g)Rikenella	0%	0.15%
(f)Rikenellaceae;(g)Rikenellaceae RC9 gut group	0%	0.22%
(f)Ruminococcaceae	0%	0.66%
(f)Ruminococcaceae;(g)[Eubacterium] coprostanoligenes group	0%	0.22%
(f)Ruminococcaceae;(g)Anaerotruncus	0%	0.13%
(f)Ruminococcaceae;(g)Butyricicoccus	0%	0.08%
(f)Ruminococcaceae;(g)Intestinimonas	0%	0.14%
(f)Ruminococcaceae;(g)Oscillibacter	0%	0.15%
(f)Ruminococcaceae;(g)Ruminiclostridium 5	0%	0.09%
(f)Ruminococcaceae;(g)Ruminiclostridium 9	0%	0.16%
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-014	0%	0.58%
(f)Ruminococcaceae;(g)Ruminococcus 1	0%	0.38%

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(f)Ruminococcaceae;(g)uncultured	0%	0.24%
(f)Unknown Family;(g)Candidatus Saccharimonas	0%	0.39%
(f)Verrucomicrobiaceae;(g)Akkermansia	7.91%	0%
(o)Clostridiales	0%	0.19%
(o)Gastranaerophilales	3.52%	0.29%
NA	**99.969**	**99.546**

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

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
Unassigned	0	0.0108
Total % Removed	**0**	**0.0108**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(p)Proteobacteria;(c)Alphaproteobacteria	0.0307	0.0000
Unassigned	0.0000	0.0108
Total % Removed	0.0307	0.0108

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

Taxa	MMRRC.043567__Tg.plus__M__M56564	MMRRC.043567__Tg.plus__M__M94895
(c)Alphaproteobacteria;(o)Rhodospirillales	0.0307	0.0000
Unassigned	0.0000	0.0108
Total % Removed	**0.0307**	**0.0108**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(f)Rhodospirillaceae	0.0307	0.0000
(f)uncultured Erysipelotrichia bacterium	0.0000	0.0259
(o)Mollicutes RF9	0.0000	0.0389
Unassigned	0.0000	0.0108
Total % Removed	**0.0307**	**0.0757**

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

Taxa	MMRRC.043567_Tg.plus_M_M56564	MMRRC.043567_Tg.plus_M_M94895
(f)Clostridiales vadinBB60 group;Ambiguous_taxa	0.0000	0.0216
(f)Erysipelotrichaceae;(g)Candidatus Stoquefichus	0.0000	0.0411
(f)Erysipelotrichaceae;(g)Faecalibaculum	0.0000	0.0130
(f)Family XIII;(g)Family XIII UCG-001	0.0000	0.0108
(f)Lachnospiraceae;(g)[Eubacterium] xylanophilum group	0.0000	0.0476
(f)Lachnospiraceae;(g)Lachnospiraceae UCG-004	0.0000	0.0324
(f)Peptococcaceae;(g)Dehalobacterium	0.0000	0.0238
(f)Peptococcaceae;(g)Peptococcus	0.0000	0.0108
(f)Rhodospirillaceae;(g)uncultured	0.0307	0.0000
(f)Rikenellaceae	0.0000	0.0108
(f)Ruminococcaceae;(g)Hydrogenoanaerobacterium	0.0000	0.0454
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-009	0.0000	0.0281
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-010	0.0000	0.0216
(f)Ruminococcaceae;(g)Ruminococcaceae UCG-013	0.0000	0.0497
(f)uncultured Erysipelotrichia bacterium;(g)uncultured Erysipelotrichia bacterium	0.0000	0.0259
(o)Mollicutes RF9	0.0000	0.0389
Unassigned	0.0000	0.0108
Total % Removed	**0.0307**	**0.4324**

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.