ANALYSIS OF GUT MICROBIOME COLONIZATION IN GERM-FREE MICE POST-FECAL MICROBIAL TRANSPLANTATION USING QIIME2

by Jude Aneke

The gut microbiome plays a crucial role in the health and disease resistance of its host. This metagenomic analysis aims to determine if transferred bacteria can colonize the recipient's gut by processing downsampled raw sequencing data using Qiime2. Specifically, we explore the differences in gut microbiome composition between wild and laboratory mice and assess the colonization success in germ-free mice following fecal microbial transplantation (FMT).

Names of groups:

W – wild mice,
L – laboratory mice,
WR – gnobiotic mice that get transplant from wild mice microbiome,
LR – gnobiotic mice that get transplant from laboratory mice microbiome.

Workflow

The steps for this analysis are outlined clearly and can be broken down into the following workflow:

1. Data Import

To import raw sequencing data into Qiime2:

```
qiime tools import \ --input-path
/home/masha_t/mice_microbiome/metadata_for_input.tsv \ --output-path
qza/gut_reads.qza \ --type 'SampleData[PairedEndSequencesWithQuality]' \ --
input-format PairedEndFastqManifestPhred33V2
```

2. Quality Control

To trim, filter, and correct sequencing errors using DADA2:

```
qiime dada2 denoise-paired \ --i-demultiplexed-seqs qza/gut_reads.qza \ --p-n-
threads 2 \ --p-trim-left-f 20 \ --p-trim-left-r 20 \ --p-trunc-len-f 190 \ --p-trunc-
len-r 170 \ --o-table qza/ASV_table.qza \ --o-representative-sequences
qza/rep_seq.qza \ --o-denoising-stats qza/reads.dada2.stats.qza
```

To visualize the quality control statistics:

```
qiime metadata tabulate \ --m-input-file qza/reads.dada2.stats.qza \ --o-
visualization qzv/reads.dada2.stats.qzv
```

Figure 1: Screenshot of Reads Statistics

sample-id (#q2:types)	input (numeric)	filtered (numeric)	percentage of input passed filter (numeric)	denoised (numeric)	merged (numeric)	percentage of input merged (numeric)	non-chimeric (numeric)	percentage of input non-chimeric (numeric)
W1	2138	1872	87.56	1792	1535	71.8	1495	69.93
L5	3311	2920	88.19	2818	2319	70.04	2271	68.59
W2	3714	3033	81.66	2918	2476	66.67	2392	64.4
W3	3211	2964	92.31	2863	2552	79.48	2450	76.3
L8	3498	3024	86.45	2923	2462	70.38	2458	70.27
W7	3501	2985	85.26	2933	2587	73.89	2490	71.12
W4	3619	3196	88.31	3079	2641	72.98	2589	71.54
L2	3669	3237	88.23	3127	2646	72.12	2646	72.12
L4	3836	3370	87.85	3255	2811	73.28	2730	71.17
W5	4301	3702	86.07	3519	2916	67.8	2771	64.43

Figure 1 shows the screenshot of reads statistics in ascending order of non-chimeric reads.

3. Taxonomic Classification

To classify the sequences taxonomically using a naive Bayes classifier:

```
qiime feature-classifier classify-sklearn \ --i-classifier /home/masha_t/mice_microbiome/classifier/gg_2022_10_backbone.v4.nb.qza \ --i-reads qza/rep_seq.qza \ --o-classification qza/taxonomy.qza
```

To create barplots of taxonomic composition:

```
qiime taxa barplot \ --i-table qza/ASV_table.qza \ --i-taxonomy qza/taxonomy.qza \ --m-metadata-file /home/masha_t/mice_microbiome/metadata.tsv \ --o-visualization qzv/taxonomy_barplot.qzv
```

Figure 2: Taxonomy Composition of Bacterial Phyla in Laboratory and Wild Mice

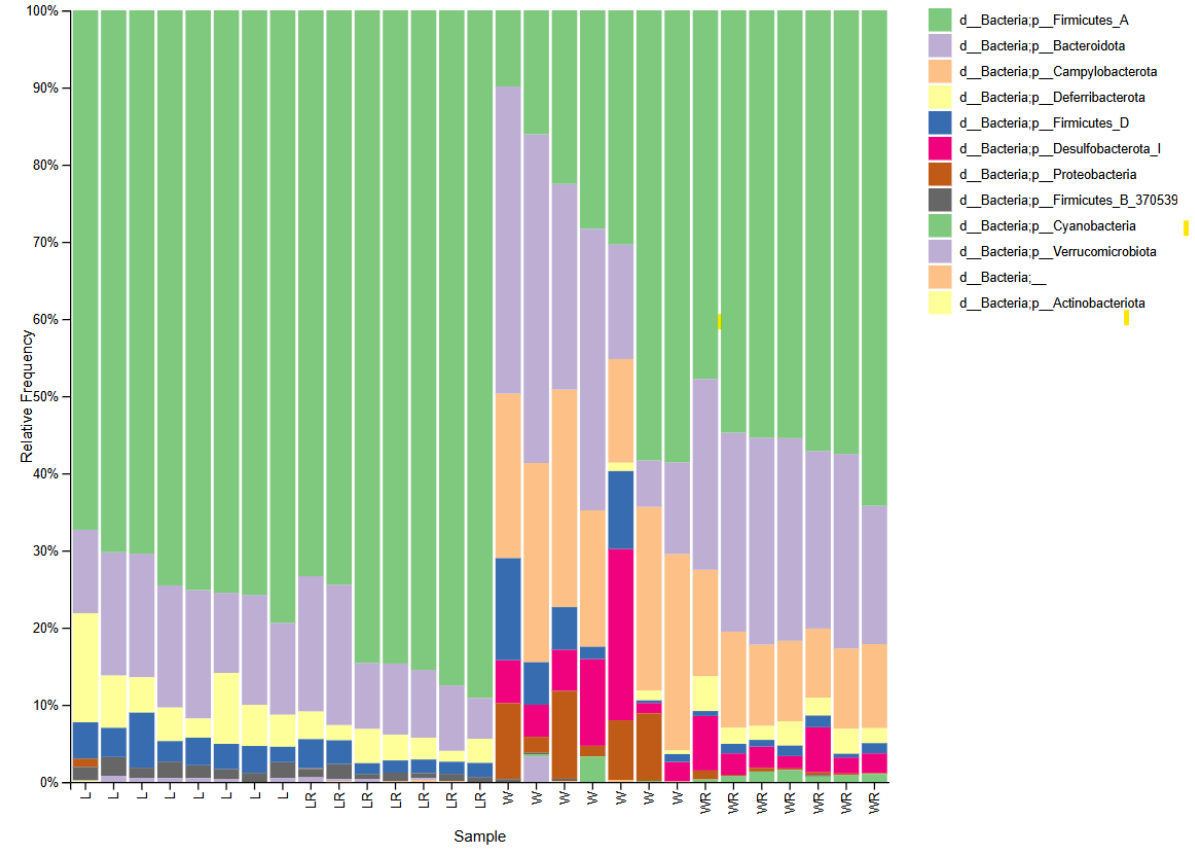
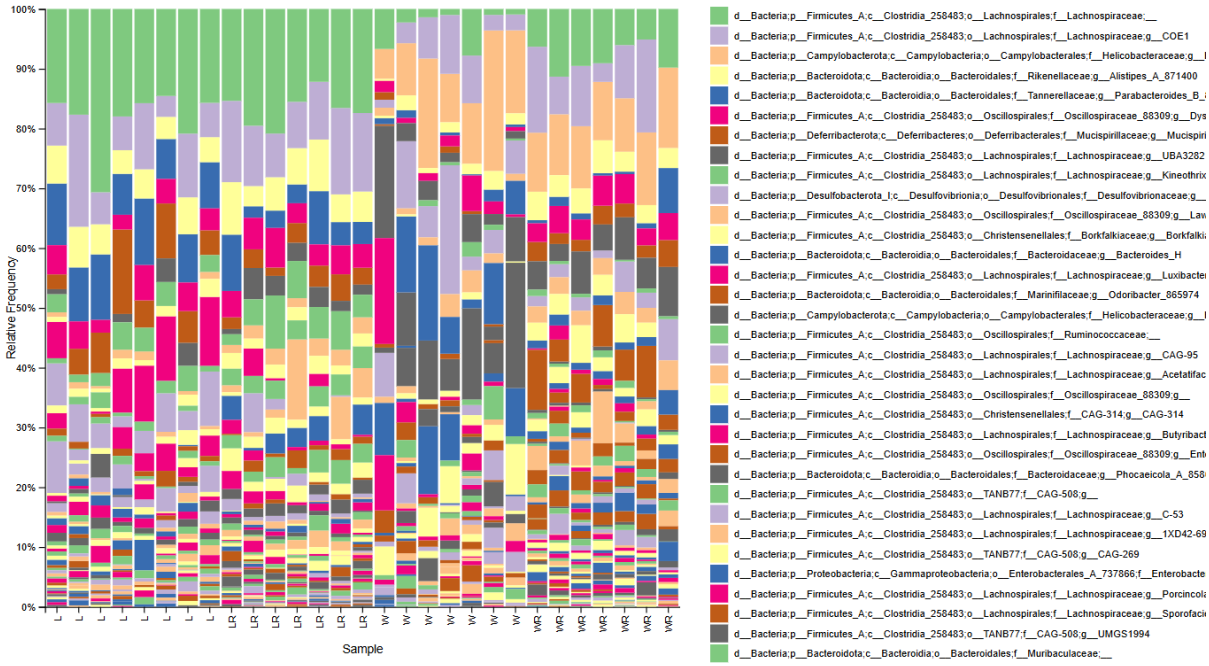


Figure 3: Taxonomy Composition of Bacterial Genera in Laboratory and Wild Mice



The microbiome from both wild and laboratory mice successfully colonized the guts of germ-free mice, as evidenced by the similarity in the microbiome composition between the donors

and recipients. The dominant bacterial phylum in both environments is Firmicutes_A (Figure 2). However, it was difficult to distinguish a single dominant bacterial genus. Generally, the laboratory mice predominantly harbor members of the Lachnospiraceae family, while the wild mice are primarily colonized by the genus Helicobacter_D (Figure 3). The barplot obviously shows high level of alpha diversity within all groups.

4. Diversity Analysis

To build a phylogenetic tree:

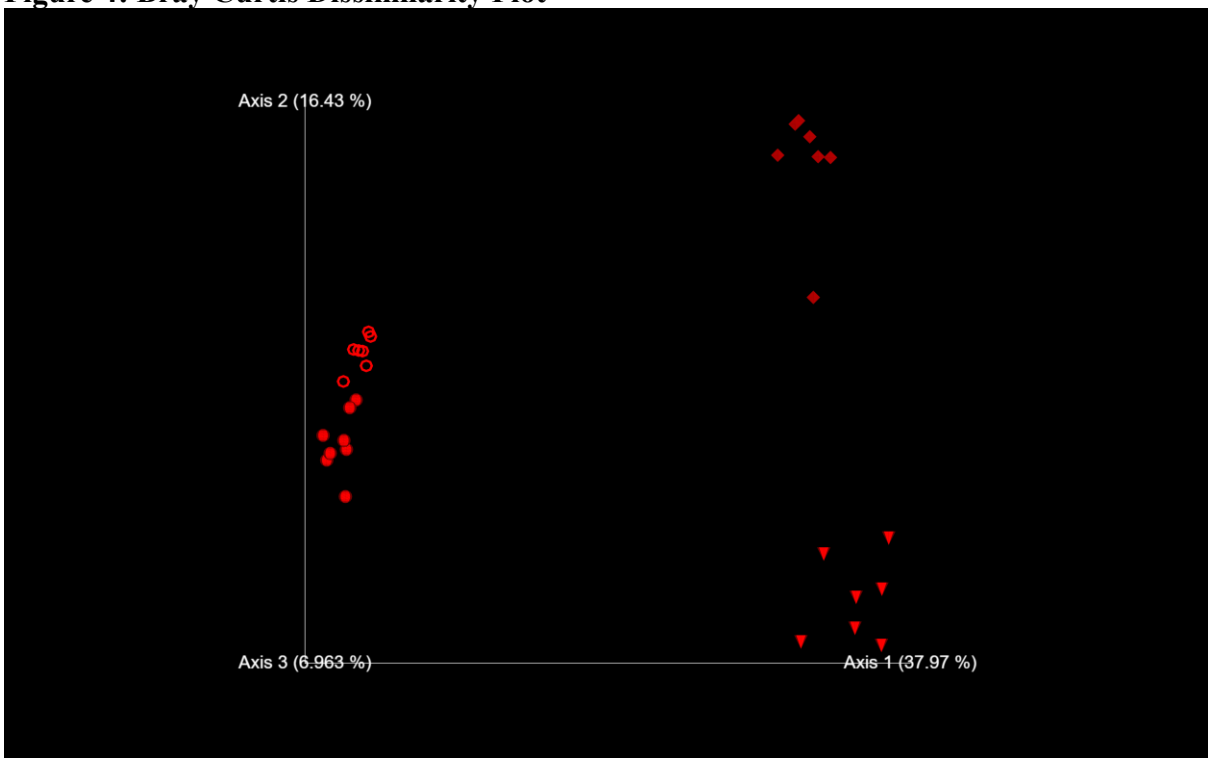
```
qiime phylogeny align-to-tree-mafft-fasttree \ --i-sequences qza/rep_seq.qza \ --o-alignment qza/aligned_rep_seq.qza \ --o-masked-alignment qza/masked_aligned_rep_seq.qza \ --o-tree qza/unrooted_tree.qza \ --o-rooted-tree qza/rooted_tree.qza
```

To estimate diversity metrics and generate visualizations:

```
qiime diversity core-metrics-phylogenetic \ --i-phylogeny qza/rooted_tree.qza \ --i-table qza/ASV_table.qza \ --p-sampling-depth 1495 \ --m-metadata-file /home/masha_t/mice_microbiome/metadata.tsv \ --output-dir mice_core_metrics_results
```

The reads have their smallest value at 1495, as shown in Figure 1. We downsampled to this value as a form of normalization to ensure that the diversity metrics and statistical comparisons are meaningful and not skewed by differences in sequencing depth. This allows for a fair comparison of microbial diversity and composition between groups, as the samples from each group contribute equally to the analysis.

Figure 4: Bray Curtis Dissimilarity Plot



Shapes:
L = Sphere
LW = Rings
W = Cone
WR = Diamond

In Figure 4, Axis 1 explains 37.97% of the total variation in the data, indicating significant differences in microbial communities for points separated along this axis. Axis 2 and Axis 3 explain 16.43% and 6.963% of the total variation respectively. L and LR have more similar microbial communities, as they are clustered together according to the principal component along each axis. W and WR also form distinct clusters, suggesting a similar pattern of microbial composition between wild mice and their recipients.

The clear separation along Axis 1 between the laboratory groups (L, LR) and the wild groups (W, WR) indicates differences in their microbiome compositions. Additionally, the separation along Axis 2 between W and WR indicates beta diversity in the microbiome composition between the wild donor mice and the wild recipient mice post-fecal microbial transplantation.

Figure 5: weighted UniFrac Dissimilarity Plot

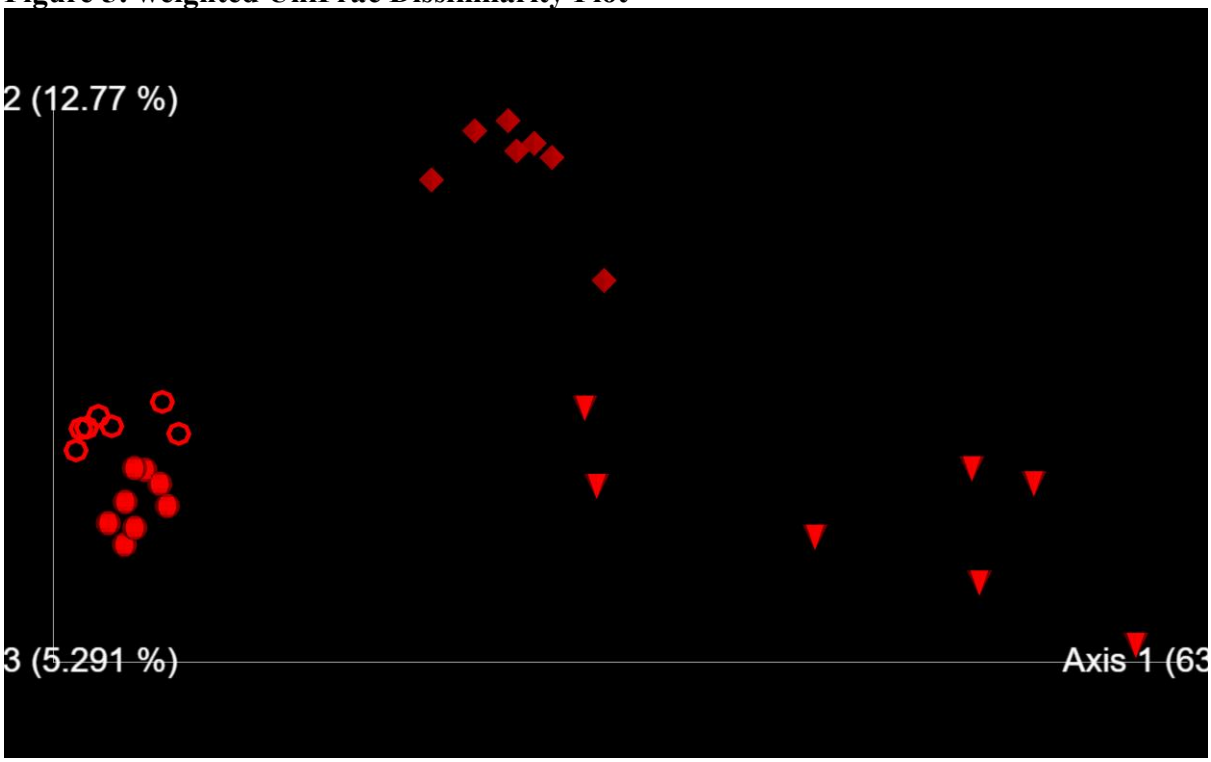


Figure 5, shows the weighted UniFrac principal coordinates analysis (PCoA) plot, which visualizes the beta diversity of microbial communities based on phylogenetic differences. Axis 1, Axis 2 and Axis 3 explain 63.56%, 12.77% and 5.29% of the total variation. The clear separation along Axis 1 between the laboratory groups (L, LR) and the wild groups (W, WR) indicates notable differences in their microbiome compositions. Additionally, the separation along Axis 2 between W (wild mice) and WR (wild recipient mice) further highlights the beta diversity in microbiome composition between the wild donor mice and the wild recipient mice post-fecal microbial transplantation.