Germ-free-PAT 16S OTU

Author: Victoria Ruiz & Thomas Battaglia

Introduction

This Notebook is meant to contain the 16S_OTU data found within the respective manuscript. It contains only the code used to generate the figures found within the main text. This entire dataset is publically available in QIITA under the ID 10527. More details about the procedures used to generate the data can be found within the **Methods** section of the manuscript. The table found within the data folder has been processed to remove any OTU less than 0.01% relative abudance. This table does not include the germ-free samples.

Generate alpha diversity metrics for barplots (Figure 3b)

These commands will generate the alpha diversity values to re-create **Figure 3b**. Faiths Phylogenetic Diversity is first calculated and sumamrized into Treatment and Timepoints with Mean and SD. This data was placed into PRISM 7 to be graphed. The table used to make the PRISM plot is found in `analysis/alphadiv/pd_tree_summarized.txt`

[8]
```
# Make directory to store results
mkdir -p analysis/alphadiv/

# Generate Faiths PD values for all samples
alpha_diversity.py \
-i data/otu_table_m0001.biom \
-o analysis/alphadiv/pd_whole_tree.txt \
-m PD_whole_tree \
-t data/rep_set.tre
```

[9]
```
# Add alpha diversity to the mapping file
 add_alpha_to_mapping_file.py \
 -i analysis/alphadiv/pd_whole_tree.txt \
 -o data/gfPAT_mapping_file_walpha.txt \
 -m data/gfPAT_mapping_file.txt
```

```
# Summarize the data for use with PRISM
```

[10]
```
## Final table used to generate plot
Rscript scripts/alpha_for_prism.R \
-i data/gfPAT_mapping_file_walpha.txt \
-o analysis/alphadiv/pd_tree_summarized.txt \
--group Group \
--time Type_of_sample \
--alpha PD_whole_tree_alpha
```

Generate alpha diversity metrics for barplots (Figure 3c)

This command generates an HTML file that uses the program Emperor to view 3D PCoA data. The file `analysis/betadiv/unweighted_unifrac_emperor_pcoa_plot/index.html` is the output used to generate the figure within the manuscript.

[16]
```
# Run bdiv through plots
beta_diversity_through_plots.py \
-i data/otu_table_m0001.biom \
-o analysis/betadiv_pcoa/ \
-t data/rep_set.tre \
-m data/gfPAT_mapping_file.txt
```

[17]
```
# Make a folder to store results
mkdir -p analysis/betadiv_pcoa/Fecal_D21_stats

# Filter to keep only Fecal (D21) samples and compute ADONIS
filter_distance_matrix.py \
-i analysis/betadiv_pcoa/unweighted_unifrac_dm.txt \
-o analysis/betadiv_pcoa/Fecal_D21_stats/uw_unifrac_dm_Fecal.txt \
-m data/gfPAT_mapping_file.txt \
-s "Type_of_sample:Fecal"

# Run ADONIS test
compare_categories.py \
-i analysis/betadiv_pcoa/Fecal_D21_stats/uw_unifrac_dm_Fecal.txt \
-o analysis/betadiv_pcoa/Fecal_D21_stats/ \
-m data/gfPAT_mapping_file.txt \
-c Treatment_Sex \
--method adonis \
-n 999
```

[18]
```
# Make a folder to store results
mkdir -p analysis/betadiv_pcoa/Colon_stats

# Filter to keep only Colon samples and compute ADONIS
```

```
filter_distance_matrix.py \
-i analysis/betadiv_pcoa/unweighted_unifrac_dm.txt \
-o analysis/betadiv_pcoa/Colon_stats/uw_unifrac_dm_Colon.txt \
-m data/gfPAT_mapping_file.txt \
-s "Type_of_sample:Colon"

# Run ADONIS test
compare_categories.py \
-i analysis/betadiv_pcoa/Colon_stats/uw_unifrac_dm_Colon.txt \
-o analysis/betadiv_pcoa/Colon_stats/ \
-m data/gfPAT_mapping_file.txt \
-c Treatment_Sex \
--method adonis \
-n 999
```