PulsePAT- 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.

Filter the table to only show fecal samples

```
[48]   # Keep only fecal samples for analyses
       filter_samples_from_otu_table.py \
       -i data/pulsePAT_m0001.biom \
       -o data/pulsePAT_m0001_fecal.biom \
       -m data/pulsePAT_mapping.txt \
       --output_mapping_fp data/pulsePAT_mapping_fecal.txt \
       -s "Type_of_sample:Fecal"
```

Generate Alpha Diveristy Metrics (Figure 1b)

These commands will generate the alpha diversity values to re-create **Figure 1b**. Faiths Phylogenetic Diversity is first calculated and sumamrized into Treatment and Timepoints with Mean and SEM. 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`

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

       # Generate alpha diversity values for all samples
       alpha_diversity.py \
```

```
    -i data/pulsePAT_m0001_fecal.biom \
    -o analysis/alphadiv/pd_whole_tree.txt \
    -m PD_whole_tree \
    -t data/rep_set.tre
```

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

```
[53]   # Summarize the data for use with PRISM
       ## Final table used to generate plot
       Rscript scripts/alpha_for_prism.R \
       -i data/pulsePAT_mapping_fecal_walpha.txt \
       -o analysis/alphadiv/pd_tree_summarized.txt \
       --group Treatment \
       --time Day_of_life \
       --alpha PD_whole_tree_alpha
```

Generate PCoA of Beta-diveristy metrics (Figure 1c)

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.

```
[54]   # Run bdiv through plots
       beta_diversity_through_plots.py \
       -i data/pulsePAT_m0001_fecal.biom \
       -o analysis/betadiv_pcoa/ \
       -t data/rep_set.tre \
       -m data/pulsePAT_mapping_fecal.txt
```

Generate Beta-diversity comparisons (Figure 1d)

```
[55]    # Generate pairwise values for Unweighted-UniFrac distances
       make_distance_boxplots.py \
       -d analysis/betadiv_pcoa/unweighted_unifrac_dm.txt \
       -o analysis/betadiv_barplot/ \
       -m data/pulsePAT_mapping_fecal.txt \
```

```
-f "Treatment" \
--save_raw_data
```

Make taxanomic abundance plots (Figure 1e)

These commands will generate the graphs found within the multi-panel figure 3e. The command must be run on each of the 6 permutations. The final PDF is composed of multiple plots across the different Treatment groups and genders.

[56]
```
# Split the table by Treatment and Gender
split_otu_table.py \
-i data/pulsePAT_m0001.biom \
-o data/per_study_otu_tables \
-m data/pulsePAT_mapping.txt \
-f Sex,Treatment
```

[57]
```
# Plot barplot with only Male-Control
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Male_Treatment_Control__.b
iom \
-o analysis/taxa_plots/male_control \
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Male_Treatment_Control__
.txt \
-c Day_of_life_tissue \
--sort

# Plot barplot with only Male-PAT1
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Male_Treatment_PAT1__.biom
\
-o analysis/taxa_plots/male_pat1 \
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Male_Treatment_PAT1__.tx
t \
-c Day_of_life_tissue \
--sort

# Plot barplot with only Male-PAT3
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Male_Treatment_PAT3__.biom
\
-o analysis/taxa_plots/male_pat3 \
```

```
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Male_Treatment_PAT3__.tx
t \
-c Day_of_life_tissue \
--sort

# - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - #

# Plot barplot with only Female-Control
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Female_Treatment_Control__
.biom \
-o analysis/taxa_plots/female_control \
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Female_Treatment_Control
__.txt \
-c Day_of_life_tissue \
--sort

# Plot barplot with only Female-PAT1
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Female_Treatment_PAT1__.bi
om \
-o analysis/taxa_plots/female_pat1 \
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Female_Treatment_PAT1__.
txt \
-c Day_of_life_tissue \
--sort

# Plot barplot with only Female-PAT3
summarize_taxa_through_plots.py \
-i
data/per_study_otu_tables/pulsePAT_m0001__Sex_Female_Treatment_PAT3__.bi
om \
-o analysis/taxa_plots/female_pat3 \
-m
data/per_study_otu_tables/pulsePAT_mapping__Sex_Female_Treatment_PAT3__.
txt \
-c Day_of_life_tissue \
--sort
```