|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Supplemental Table S1** Indicator taxa of each diet type | | | | | | | | | | |
| **Diet Type** | **Kingdom** | **Phylum** | **Class** | **Order** | **Family** | **A**a | **B**b | **Stat**c | ***p*-value** | **Significance**d |
| Carnivores | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Paraprevotellaceae | 0.52182 | 0.93421 | 0.698 | 0.005 | \*\* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Bacteroidaceae | 0.47046 | 0.94737 | 0.668 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | NA | 0.45071 | 0.97368 | 0.662 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Ruminococcaceae | 0.42695 | 0.98684 | 0.649 | 0.01 | \*\* |
|  | Archaea | Euryarchaeota | Methanobacteria | Methanobacteriales | Methanobacteriaceae | 0.48695 | 0.81579 | 0.63 | 0.015 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Pseudomonadales | Moraxellaceae | 0.67693 | 0.57895 | 0.626 | 0.03 | \* |
|  | Bacteria | Proteobacteria | Deltaproteobacteria | Desulfovibrionales | Desulfovibrionaceae | 0.44659 | 0.82895 | 0.608 | 0.02 | \* |
|  | Bacteria | Firmicutes | Bacilli | Bacillales | Bacillaceae | 0.85734 | 0.42105 | 0.601 | 0.01 | \*\* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Rikenellaceae | 0.57925 | 0.61842 | 0.599 | 0.01 | \*\* |
|  | Bacteria | Verrucomicrobia | Verruco-5 | WCHB1-41 | RFP12 | 0.6023 | 0.59211 | 0.597 | 0.02 | \* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | RF16 | 0.61911 | 0.56579 | 0.592 | 0.005 | \*\* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | NA | 0.46875 | 0.68421 | 0.566 | 0.01 | \*\* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | BS11 | 0.65889 | 0.47368 | 0.559 | 0.005 | \*\* |
|  | Bacteria | Verrucomicrobia | Verrucomicrobiae | Verrucomicrobiales | Verrucomicrobiaceae | 0.52635 | 0.59211 | 0.558 | 0.03 | \* |
|  | Bacteria | Firmicutes | Bacilli | Bacillales | NA | 0.60782 | 0.47368 | 0.537 | 0.005 | \*\* |
|  | Bacteria | Lentisphaerae | Lentisphaeria | Victivallales | Victivallaceae | 0.51616 | 0.53947 | 0.528 | 0.025 | \* |
|  | Bacteria | Tenericutes | Mollicutes | Mycoplasmatales | Mycoplasmataceae | 0.71649 | 0.38158 | 0.523 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Bacilli | Bacillales | Planococcaceae | 0.50867 | 0.52632 | 0.517 | 0.01 | \*\* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Xanthomonadales | Xanthomonadaceae | 0.80679 | 0.30263 | 0.494 | 0.04 | \* |
|  | Bacteria | Proteobacteria | NA | NA | NA | 0.49476 | 0.48684 | 0.491 | 0.045 | \* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Dehalobacteriaceae | 0.43401 | 0.53947 | 0.484 | 0.035 | \* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Tissierellaceae | 0.86307 | 0.26316 | 0.477 | 0.045 | \* |
|  | Archaea | Euryarchaeota | Methanomicrobia | Methanomicrobiales | Methanocorpusculaceae | 0.73809 | 0.30263 | 0.473 | 0.005 | \*\* |
|  | Bacteria | Synergistetes | Synergistia | Synergistales | Synergistaceae | 0.51148 | 0.42105 | 0.464 | 0.02 | \* |
|  | Bacteria | Actinobacteria | Actinobacteria | Actinomycetales | Corynebacteriaceae | 0.53968 | 0.36842 | 0.446 | 0.025 | \* |
|  | Bacteria | Firmicutes | Bacilli | Lactobacillales | Aerococcaceae | 0.8495 | 0.22368 | 0.436 | 0.005 | \*\* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | NA | NA | 0.96019 | 0.17105 | 0.405 | 0.015 | \* |
|  | Bacteria | Firmicutes | Bacilli | Lactobacillales | Carnobacteriaceae | 0.84526 | 0.17105 | 0.38 | 0.005 | \*\* |
|  | Archaea | Euryarchaeota | Methanomicrobia | Methanosarcinales | Methanosarcinaceae | 0.6848 | 0.17105 | 0.342 | 0.005 | \*\* |
|  | Bacteria | Deferribacteres | Deferribacteres | Deferribacterales | Deferribacteraceae | 0.6431 | 0.15789 | 0.319 | 0.03 | \* |
|  | Bacteria | Tenericutes | NA | NA | NA | 0.85183 | 0.10526 | 0.299 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | EtOH8 | 0.72117 | 0.11842 | 0.292 | 0.025 | \* |
|  | Bacteria | Fibrobacteres | Fibrobacteria | Fibrobacterales | NA | 0.65154 | 0.11842 | 0.278 | 0.015 | \* |
|  | Bacteria | Actinobacteria | NA | NA | NA | 0.94088 | 0.07895 | 0.273 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Bacilli | NA | NA | 0.58384 | 0.11842 | 0.263 | 0.05 | \* |
|  | Bacteria | Actinobacteria | Actinobacteria | Actinomycetales | Bogoriellaceae | 0.91918 | 0.06579 | 0.246 | 0.015 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Alteromonadales | Chromatiaceae | 0.83872 | 0.06579 | 0.235 | 0.01 | \*\* |
|  | Bacteria | Spirochaetes | Leptospirae | Leptospirales | Sediment-4 | 1 | 0.05263 | 0.229 | 0.005 | \*\* |
|  | Bacteria | Actinobacteria | Actinobacteria | Actinomycetales | Nocardiopsaceae | 0.95619 | 0.05263 | 0.224 | 0.04 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Chromatiales | Ectothiorhodospiraceae | 1 | 0.03947 | 0.199 | 0.01 | \*\* |
|  | Bacteria | TM7 | TM7-3 | I025 | NA | 1 | 0.03947 | 0.199 | 0.01 | \*\* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Thiotrichales | Piscirickettsiaceae | 0.81038 | 0.03947 | 0.179 | 0.045 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Oceanospirillaceae | 1 | 0.02632 | 0.162 | 0.045 | \* |
| Herbivores | Bacteria | Bacteroidetes | Flavobacteriia | Flavobacteriales | Flavobacteriaceae | 0.72466 | 0.2931 | 0.461 | 0.01 | \*\* |
|  | Bacteria | Actinobacteria | Actinobacteria | Actinomycetales | Microbacteriaceae | 0.68402 | 0.26724 | 0.428 | 0.025 | \* |
|  | Bacteria | Firmicutes | Bacilli | Bacillales | Alicyclobacillaceae | 0.76756 | 0.18103 | 0.373 | 0.005 | \*\* |
|  | Bacteria | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Rhodospirillaceae | 0.69929 | 0.19828 | 0.372 | 0.015 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Xanthomonadales | Sinobacteraceae | 0.61505 | 0.14655 | 0.3 | 0.045 | \* |
|  | Bacteria | Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | 0.82609 | 0.06034 | 0.223 | 0.045 | \* |
|  | Bacteria | Actinobacteria | Acidimicrobiia | Acidimicrobiales | Microthrixaceae | 1 | 0.0431 | 0.208 | 0.035 | \* |
|  | Bacteria | Proteobacteria | Deltaproteobacteria | Myxococcales | Cystobacteraceae | 0.77075 | 0.0431 | 0.182 | 0.045 | \* |
| Omnivores | Bacteria | NA | NA | NA | NA | 0.7079 | 0.9615 | 0.825 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Erysipelotrichi | Erysipelotrichales | Erysipelotrichaceae | 0.4911 | 0.9519 | 0.684 | 0.015 | \* |
|  | Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Prevotellaceae | 0.5025 | 0.9038 | 0.674 | 0.02 | \* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Lachnospiraceae | 0.4173 | 0.9904 | 0.643 | 0.03 | \* |
|  | Bacteria | Verrucomicrobia | Opitutae | Cerasicoccales | Cerasicoccaceae | 0.8264 | 0.5 | 0.643 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Bacilli | Lactobacillales | Lactobacillaceae | 0.8889 | 0.4615 | 0.641 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Christensenellaceae | 0.4781 | 0.7885 | 0.614 | 0.015 | \* |
|  | Bacteria | Spirochaetes | Spirochaetes | Spirochaetales | Spirochaetaceae | 0.4857 | 0.7692 | 0.611 | 0.025 | \* |
|  | Archaea | Euryarchaeota | Thermoplasmata | E2 | Methanomassiliicoccaceae | 0.4831 | 0.7212 | 0.59 | 0.005 | \*\* |
|  | Bacteria | Elusimicrobia | Elusimicrobia | Elusimicrobiales | Elusimicrobiaceae | 0.6105 | 0.4808 | 0.542 | 0.01 | \*\* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Aeromonadales | Succinivibrionaceae | 0.5165 | 0.5481 | 0.532 | 0.05 | \* |
|  | Bacteria | Proteobacteria | Gammaproteobacteria | Vibrionales | Vibrionaceae | 0.696 | 0.2596 | 0.425 | 0.005 | \*\* |
|  | Bacteria | Spirochaetes | Brevinematae | Brevinematales | Brevinemataceae | 0.6709 | 0.2019 | 0.368 | 0.005 | \*\* |
|  | Bacteria | Verrucomicrobia | Verruco-5 | WCHB1-41 | WCHB1-25 | 0.7964 | 0.1635 | 0.361 | 0.005 | \*\* |
|  | Bacteria | Firmicutes | Clostridia | Clostridiales | Syntrophomonadaceae | 0.5982 | 0.1827 | 0.331 | 0.03 | \* |

a specificity (the fraction of the taxa found in the indicated diet type)

b fidelity (the fraction of samples of the indicated diet type that contain at least 1 sequence of the taxa)

c indicator taxa value

d significance level: \* = 0.01 ; \*\* = 0.001

NA indicates the taxanomic rank was not annotated

**Supplemental Script S1**

QIIME2 script

####Import the sequences and zoo\_manifest.txt file into our home directory

#Go to the data directory (/data) and make a directory “project 2”.

cd /data && mkdir project2

#Make 2 directories (1 for each research question) in the directory “project2”

#project\_2\_trophic\_level directory will contain files for diet type analysis

#project2\_omnivore directory will contain files for the omnivores' plant-based diet analysis

cd project2

mkdir project2\_trophic\_level

mkdir project2\_omnivore

####Import sequence data of the captivity paper and generate a “demux.qza file” in the “project2” directory

#note: the sequences are already demultiplexed

qiime tools import \

--type 'SampleData[SequencesWithQuality]' \

--input-path /mnt/datasets/project\_2/zoo/zoo\_manifest.txt \

--output-path demux.qza \

--input-format SingleEndFastqManifestPhred33V2

####Generate a “demux.qzv” file to view the summary of the sequence (e.g. counts of sequence per sample and sequence qualities at each position).

qiime demux summarize \

--i-data demux.qza \

--o-visualization demux.qzv

#Secure copy demux.qzv from the server to the desktop and visualize it on the QIIME2 web-based visualization tool https://view.qiime2.org/

#run this code on a local machine

cd /mnt/c/Users/lenax/Desktop/MICB447\_Project2

scp root@10.19.139.147:/data/project2/demux.qzv .

####Perform quality control on the metadata using dada2.

#Exploring the “demux.qzv” file, we will decide on the value for m and n depending on the sequence quality at each position. m= # of bases to trim off at the beginning of the sequence; n= at which position we would truncate each sequence.

#Create new session

screen -S project2

qiime dada2 denoise-single \

--i-demultiplexed-seqs demux.qza \

--p-trim-left 0 \

--p-trunc-len 150 \

--o-representative-sequences rep-seqs.qza \

--o-table table.qza \

--o-denoising-stats stats.qza

#Generate stats.qzv file

qiime metadata tabulate \

--m-input-file stats.qza \

--o-visualization stats.qzv

#Secure copy stats.qzv from the server to the desktop and visualize it on the QIIME2 web-based visualization tool https://view.qiime2.org/

#run this code on a local machine

cd /mnt/c/Users/lenax/Desktop/MICB447\_Project2

scp root@10.19.139.147:/data/project2/stats.qzv .

################################################################################################################

#####The following section would be specifically for the diet type analysis

###Pick a sampling depth (n) for diet type analysis by exploring at “table.qzv” and “alpha-rarefaction.qzv”

#Move into directory for diet type analysis

cd project2\_trophic\_level

#Generate “table.qzv”

qiime feature-table summarize \

--i-table /data/project2/table.qza \

--o-visualization table.qzv \

--m-sample-metadata-file /mnt/datasets/project\_2/zoo/zoo\_metadata.txt

#Secure copy table.qzv from the server to the desktop and visualize it on the QIIME2 web-based visualization tool https://view.qiime2.org/

#run this code on a local machine

cd /mnt/c/Users/lenax/Desktop/MICB447\_Project2

scp root@10.19.139.147:/data/project2/table.qzv .

#Generate a tree for phylogenetic diversity analyses

screen -S tree

qiime phylogeny align-to-tree-mafft-fasttree \

--i-sequences /data/project2/rep-seqs.qza \

--o-alignment aligned-rep-seqs.qza \

--o-masked-alignment masked-aligned-rep-seqs.qza \

--o-tree unrooted-tree.qza \

--o-rooted-tree rooted-tree.qza

###Generate a alpha rarefaction plot (for diet type) “Alpha-rarefaction.qzv”

#We made a new metadata file “zoo\_metadata\_new\_columns.txt” (adding column “plant-based food”, “animal-based food” and “Diet\_Type”, replacing / in the column with \_, replacing 1,2,3 in the “Tropic\_level” column to one, two, three)

#On our local machine, secure copy “zoo\_metadata\_new\_columns.txt” and “zoo\_metadata\_new\_columns.xlsx.tsv” to the group server

scp file.txt root@10.19.139.147:/data/project2/

screen -S rarefaction

qiime diversity alpha-rarefaction \

--i-table /data/project2/table.qza \

--i-phylogeny rooted-tree.qza \

--p-max-depth 12018 \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--o-visualization alpha-rarefaction12018.qzv

#pick a rarefaction depth n by exploring "table.qzv" file and the alpha rarefaction plot "alpha-rarefaction12018.qzv" file

#Move table.qzv and table.qza (for trophic level question) into project2\_trophic\_level directory.

mv /data/project2/table.qza /data/project2/project2\_trophic\_level/table.qza

mv /data/project2/table.qzv /data/project2/project2\_trophic\_level/table.qzv

###Generate core metrics for subsequent diversity analyses

screen -S core-metrics-phylogenetic

qiime diversity core-metrics-phylogenetic \

--i-phylogeny rooted-tree.qza \

--i-table table.qza \

--p-sampling-depth 12018 \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--output-dir core-metrics-results-trophic

####run alpha diversity analysis for diet type in QIIME2

#generate Kruskal-Wallis test values for alpha diversity analysis for diet type

qiime diversity alpha-group-significance \

--i-alpha-diversity core-metrics-results-trophic/observed\_features\_vector.qza \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--o-visualization core-metrics-results-trophic/observed\_features-group-significance.qzv

####run beta diversity analysis for diet type in QIIME2

#generate PERMANOVA statistical test values for beta diversity analysis (based on weighted UniFrac) for diet type

qiime diversity beta-group-significance \

--i-distance-matrix core-metrics-results-trophic/weighted\_unifrac\_distance\_matrix.qza \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--m-metadata-column Diet\_type \

--o-visualization core-metrics-results-trophic/weighted-unifrac-diet-type-significance.qzv \

--p-pairwise

##############################################################################

#####The following section would be for the omnivores' plant-based diet analysis

####Filter the metadata table to include omnivores only and name it as “omnivore-diet-filtered-table.qza”

cd project2\_omnivore

qiime feature-table filter-samples \

--i-table /data/project2/table.qza \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--p-where "[Diet\_Type]='O’" \

--o-filtered-table omnivore-diet-filtered-table.qza

#Generate “omnivore-diet-filtered-tabled.qzv” file

qiime feature-table summarize \

--i-table omnivore-diet-filtered-table.qza \

--o-visualization omnivore-diet-filtered-table.qzv \

--m-sample-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt

#Secure copy “omnivore-diet-filter-table.qzv” from server into local desktop

to visualize through QIIME2 VIEW.

#Run this on the local Terminal window.

scp root@10.19.139.147:/data/project2/project2\_omnivore/omnivore-diet-filtered-table.qzv .

####Generate a alpha rarefaction plot “alpha-rarefaction-omnivore-diet.qzv”

qiime diversity alpha-rarefaction \

--i-table omnivore-diet-filtered-table.qza \

--i-phylogeny /data/project2/project2\_trophic\_level/rooted-tree.qza \

--p-max-depth 218310 \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--o-visualization alpha-rarefaction-omnivore-diet218310.qzv

qiime diversity alpha-rarefaction \

--i-table omnivore-diet-filtered-table.qza \

--i-phylogeny /data/project2/project2\_trophic\_level/rooted-tree.qza \

--p-max-depth 30000 \

--m-metadata-file /data/project2/zoo\_metadata\_new\_columns.txt \

--o-visualization alpha-rarefaction-omnivore-diet30000.qzv

#Secure copy “alpha-rarefaction-omnivore-diet30000.qzv” and “alpha-rarefaction-omnivore-diet218310.qzv” from server into local desktop to visualize through QIIME2 VIEW.

#Run these steps on the local computer.

scp root@10.19.139.147:/data/project2/project2\_omnivore/alpha-rarefaction-omnivore-diet30000.qzv .

#Pick a rarefaction depth n by exploring “omnivore-diet-filtered-table.qzv” file and the alpha rarefaction plot “alpha-rarefaction-omnivore-diet30000.qzv” and “alpha-rarefaction-omnivore-diet218310.qzv”

#copy the rooted-tree.qza from the project2\_trophic\_level directory to the project2\_omnivore directory

cp /data/project2/project2\_trophic\_level/rooted-tree.qza /data/project2/project2\_omnivore/rooted-tree.qza

####Generate core metrics for subsequent diversity analyses

qiime diversity core-metrics-phylogenetic \

--i-phylogeny rooted-tree.qza \

--i-table omnivore-diet-filtered-table.qza \

--p-sampling-depth 17292 \

--m-metadata-file /data/project2/zoo\_metadata\_categorical\_omnivore.txt \

--output-dir core-metrics-results-omnivore

####run alpha diversity analysis for omnivores' plant-based diet in QIIME2

#generate Kruskal-Wallis test values for alpha diversity analysis for omnivores' plant-based diet

qiime diversity alpha-group-significance \

--i-alpha-diversity core-metrics-results-omnivore/observed\_features\_vector.qza \

--m-metadata-file /data/project2/zoo\_metadata\_categorical\_omnivore.txt \

--o-visualization core-metrics-results-omnivore/observed-features-group-significance.qzv

####run beta diversity analysis for omnivores' plant-based diet in QIIME2

#generate PERMANOVA statistical test values for beta diversity analysis (based on weighted UniFrac) for omnivores' plant-based diet

qiime diversity beta-group-significance \

--i-distance-matrix core-metrics-results-omnivore/weighted\_unifrac\_distance\_matrix.qza \

--m-metadata-file /data/project2/zoo\_metadata\_categorical\_omnivore.txt \

--m-metadata-column plant-based-food \

--o-visualization core-metrics-results-omnivore/weighted-unifrac-omnivore-box-significance.qzv \

--p-pairwise

########################################################################################################################################

#diet type indicator taxa

#generate indicator taxa on different taxonomic ranks for diet types

####Assign taxonomic information

cd /data/project2

# Download and unzip the chosen database “Greengenes”

wget "ftp://greengenes.microbio.me/greengenes\_release/gg\_13\_5/gg\_13\_8\_otus.tar.gz"

gunzip gg\_13\_8\_otus.tar.gz

tar -xvf gg\_13\_8\_otus.tar

# Greengenes 99% database used # Extract files and convert them to .qza

qiime tools import \

--type 'FeatureData[Sequence]' \

--input-path gg\_13\_8\_otus/rep\_set/99\_otus.fasta \

--output-path ref-otus.qza

qiime tools import \

--type 'FeatureData[Taxonomy]' \

--input-format HeaderlessTSVTaxonomyFormat \

--input-path gg\_13\_8\_otus/taxonomy/99\_otu\_taxonomy.txt \

--output-path ref-taxonomy.qza

# Primer sequence used in the captivity paper- from the link Ilan sent: Original sequences: 515F (Caporaso)–806R (Caporaso), reverse-barcoded:

FWD:GTGCCAGCMGCCGCGGTAA; REV:GGACTACHVGGGTWTCTAAT https://earthmicrobiome.org/protocols-and-standards/16s/

# Dada2 truncation length: 150

# Extract study study-specific reference read

screen -S taxonomy2

qiime feature-classifier extract-reads \

--i-sequences ref-otus.qza \

--p-f-primer GTGCCAGCMGCCGCGGTAA \

--p-r-primer GGACTACHVGGGTWTCTAAT\

--p-trunc-len 150 \

--o-reads ref-seqs2.qza

# Train a classifier using the reference reads

qiime feature-classifier fit-classifier-naive-bayes \

--i-reference-reads ref-seqs2.qza \

--i-reference-taxonomy ref-taxonomy.qza \

--o-classifier classifier.qza

# Assign taxonomy to sequences

qiime feature-classifier classify-sklearn \

--i-classifier classifier.qza \

--i-reads rep-seqs.qza \

--o-classification taxonomy.qza

####Export and generate BIOM data for table.qza, taxonomy.qza, and rooted-tree.qza

cd /data/project2/

qiime tools export \

--input-path taxonomy.qza \

--output-path exported

cd /data/project2/project2\_trophic\_level

qiime tools export \

--input-path table.qza \

--output-path exported

qiime tools export \

--input-path rooted-tree.qza \

--output-path exported

####Editing column names for the taxonomy.tsv

#Open file in a text editor on server

cd /data/project2/

nano exported/taxonomy.tsv

#Edit the column names and change

`Feature ID` to `#OTUID`

`Taxon` to `taxonomy`

`Confidence` to `confidence`

####Combining taxonomy with BIOM data

cd /data/project2/project2\_trophic\_level

biom add-metadata \

-i exported/feature-table.biom \

-o exported/table-with-taxonomy.biom \

--observation-metadata-fp /data/project2/exported/taxonomy.tsv \

--sc-separated taxonomy

#Secure copy “table-with-taxonomy.biom” from the server to the desktop#run this code on a local machine

scp root@10.19.139.147:/data/project2/project2\_trophic\_level/exported/table-with-taxonomy.biom .

**Supplemental Script S2**

R Script

###########################################################################################

###### R Script for Diet Type

####generate box-and-whiskers plot for alpha diversity analysis (based on observed features) for diet type in Rstudio

#load packages

library(dplyr)

library(phyloseq)

library(ggplot2)

### Alpha diversity function

# richness function

richness = function(x){

return(sum(x>0))

}

### Load data for diet type

biom = import\_biom("diet\_type/table-with-taxonomy.biom")

taxa\_table = otu\_table(biom)

taxonomy = tax\_table(biom)

metadata = read.table("diet\_type/zoo\_metadata\_new\_columns.txt",sep="\t",header=T,row.names = 1, comment.char = "")

# select only samples with metadata

microbial\_samples = colnames(taxa\_table)

metadata\_samples = rownames(metadata)

which\_metadata = c()

for (i in 1:dim(taxa\_table)[2]){

which\_metadata = c(which\_metadata,which(metadata\_samples == microbial\_samples[i]))

}

metadata = metadata[which\_metadata,]

### Calculate alpha diversity

metadata$richness = apply(taxa\_table,2,richness)

#richness plot for diet\_type question

ggplot(metadata,aes(x=Diet\_Type,y=richness))+

geom\_boxplot(outlier.shape=NA)+

geom\_point(position=position\_jitter(width=0.1))+

ylab("Observed Features")+

xlab("Diet Type")+

labs(title="Alpha Diversity")+theme(plot.title = element\_text(hjust = 0.5))+

scale\_x\_discrete(labels = c("Carnivores (n=68)","Herbivores (n=116)","Omnivores (n=104)"))

#find mean, max, min, median, and sd of diversity for each group

metadata%>%

group\_by(Diet\_Type)%>%

summarise(Mean=mean(richness),

Max=max(richness),

Min=min(richness),

Median=median(richness),

Std=sd(richness))

#end

####generate PCoA plot for beta diversity analysis (based on weighted UniFrac) for diet type in Rstudio

#Loading packages

library(tidyverse)

library(vegan)

library(phyloseq)

library(ape)

#Load Data

biom\_file <- import\_biom("table-with-taxonomy.biom")

metadata <-import\_qiime\_sample\_data("zoo\_metadata\_new\_columns.txt")

tree <-read\_tree\_greengenes("tree.nwk")

tree <- multi2di(tree)

#Combine all objects into phyloseq object

physeq <-merge\_phyloseq(biom\_file, metadata, tree)

#Overview of phyloseq object

physeq

#Set a set of random numbers.

set.seed(711)

#Visualize beta diversity through graphing PCoA plot for diet type question

#Diversity requires rarefied taxa tables. We decided on a rarefaction depth of 12018 for diet type.

physeq\_rar <-rarefy\_even\_depth(physeq, sample.size = 12018)

#Convert to RA (relative abundance)

physeq\_rar\_RA <- transform\_sample\_counts(physeq\_rar, function(x) x/sum(x))

#Weighted Unifrac

#We define the type of analysis with ordinate() and setting the method to PCoA, setting distance to type of beta diversity analysis.

ord <- ordinate(physeq\_rar\_RA, method = "PCoA",

distance = "wunifrac")

#Plot data

plot\_ordination(physeq\_rar\_RA, ord,

type = "sample",

color = "Diet\_Type",

#shape = (shapes possible)

title = "PCoA(Weighted Unifrac)") +

#Manually adjust colours for point

scale\_colour\_manual(values = c("blue", "red", "orange"),

labels = c("Carnivores", "Herbivores",

"Omnivores")) +

stat\_ellipse(type = "norm", size =1) +

guides(colour = guide\_legend("Diet Type")) +

theme\_bw(base\_size = 14)

########################################################################################################################################################

######R Script for Omnivores' plant-based diet

####generate correlation analysis for alpha diversity analysis (based on observed features) for omnivores' plant-based diet in Rstudio

#load package

library(dplyr)

library(phyloseq)

#library(labdsv)

library(ggplot2)

### Alpha diversity function

# richness function

richness = function(x){

return(sum(x>0))

}

### Load data

biom = import\_biom("table-with-taxonomy.biom")

taxa\_table = otu\_table(biom)

taxonomy = tax\_table(biom)

metadata = read.table("zoo\_metadata\_new\_columns.txt",sep="\t",header=T,row.names = 1, comment.char = "")

# select only samples with metadata

microbial\_samples = colnames(taxa\_table)

metadata\_samples = rownames(metadata)

which\_metadata = c()

for (i in 1:dim(taxa\_table)[2]){

which\_metadata = c(which\_metadata,which(metadata\_samples == microbial\_samples[i]))

}

metadata = metadata[which\_metadata,]

### Calculate alpha diversity

metadata$richness = apply(taxa\_table,2,richness)

metadata$shannons = apply(taxa\_table,2,shannons)

metadata$evenness = apply(taxa\_table,2,evenness)

### Continuous variable

ggplot(metadata,aes(x=plant.based.food,y=richness)) +

ylab("Observed Features")+

scale\_x\_continuous(name="Plant-based Composition of Diet (%)")+

labs(title="Alpha Diversity")+theme(plot.title = element\_text(hjust = 0.5))+

geom\_point() +

geom\_smooth(method="lm",formula=y~x)

# Correlation

cor.test(metadata$plant.based.food,metadata$richness)

####generate PCoA plot for beta diversity analysis (based on weighted UniFrac) for omnivores' plant-based diet in Rstudio

#Loading packages

library(tidyverse)

library(vegan)

library(phyloseq)

library(ape)

#Load Data

biom\_file <- import\_biom("table-with-taxonomy.biom")

metadata <-import\_qiime\_sample\_data("zoo\_metadata\_new\_columns.txt")

tree <-read\_tree\_greengenes("tree.nwk")

tree <- multi2di(tree)

#Change animal.based.diet and plant.based.diet into categorical variable

#With or without, remove hash tag to execute command

#metadata$animal.base.food.categorical = as.factor(metadata$animal.based.food)

metadata$plant.based.food.categorical = as.factor(metadata$plant.based.food)

#Combine all objects into phyloseq object

physeq <-merge\_phyloseq(biom\_file, metadata, tree)

#Overview of phyloseq object

physeq

#Set set of random numbers.

set.seed(711)

#Beta diversity PCoA plot for trophic level/diet type question

#Diversity requires rarefied taxa tables

#Earlier in the project, we decided on a rarefaction depth of 12018 for diet type.

physeq\_rar <-rarefy\_even\_depth(physeq, sample.size = 17292)

#Convert to RA (relative abundance)

physeq\_rar\_RA <- transform\_sample\_counts(physeq\_rar, function(x) x/sum(x))

#Subset out omnivorous data only

physeq\_rar\_RA\_omni = subset\_samples(physeq,Diet\_Type=="O")

#Weighted Unifrac

#We define the type of analysis with ordinate() and setting the

#method to PCoA, setting distance to type of beta diversity analysis.

ord <- ordinate(physeq\_rar\_RA\_omni, method = "PCoA",

distance = "wunifrac")

#Plot data (plant based diet weighted unifrac)

plot\_ordination(physeq\_rar\_RA\_omni, ord,

type = "sample",

color = "plant.based.food.categorical",

#shape = (shapes possible)

title = "PCoA(Weighted Unifrac)") +

#Manually adjust colours for point

scale\_colour\_manual(values = c("red", "orange", "green", "blue", "purple")) +

#labels = c()) +

stat\_ellipse(type = "norm", size =1) +

# scale\_color\_gradient(low = "Blue", high = "Red")+

guides(colour = guide\_legend("Plant-Based Diet (%)")) +

theme\_bw(base\_size = 14)