

# Fecal Microbiota Transplant (FMT) Qiime 2 Analysis

Tess Keating

# Background

- In a study by Kang et al., children with autism and gastrointestinal disorders were treated with fecal microbiota transplants to reduce their symptoms
- For 18 weeks, their microbiomes and severity of their symptoms were monitored through fecal swab and stool samples
- An Illumina 16s rRNA library was prepared for sequencing the samples

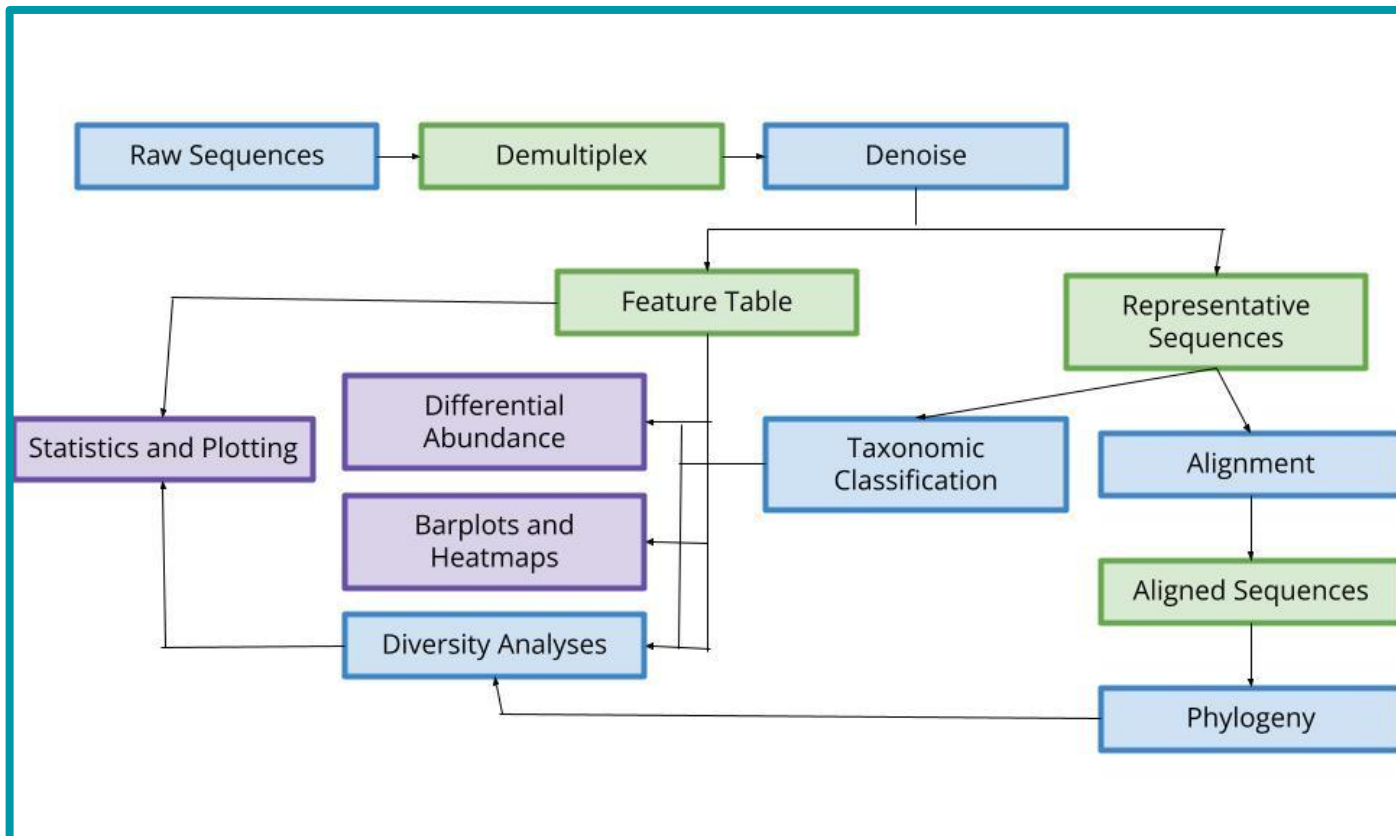
Kang, DW., Adams, J.B., Gregory, A.C. *et al.* Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome* 5, 10 (2017).

<https://doi.org/10.1186/s40168-016-0225-7>

# Goals

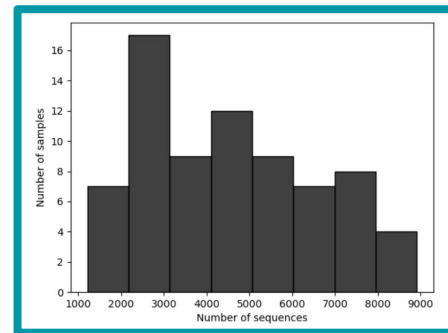
- Using sequenced data from the study by Kang et al., a bioinformatic pathway analysis was performed
- Raw reads were denoised and merged, sequences were aligned and classified, phylogenies were created, and diversity metrics were assessed using Qiime and VSCode
- Bioinformatic techniques were used to generate results in the form of representative figures and visualizations

# Pipeline



# Methods: Obtaining, Denoising, and Merging Data

- *wget*: obtains initial data files (sample metadata and 10% subsample data)
- *demux summarize*: plots sequence quality to assess reads before denoising
- *dada2 denoise-single*: denoises sequences (removes errors and increases accuracy)
- *metadata tabulate*: tabulates denoised stats (amount of filtered, denoised, and non-chimeric read inputs)
- *feature-table tabulate-seqs*: gives sequence lengths
- *feature-table merge*: combines two feature tables
- *feature-table merge-seqs*: gives sequence lengths
- *feature-table summarize*: tabulates and plots frequency stats
- *feature-table tabulate-seqs*: makes merged table with sequence lengths of each feature



Feature ID	Sequence Length	Sequence
002e78333d6cf2b11aa7a5ba03dd2c68	137	AGCGTTATCCGGATTATTGGGTTTAAAGGGCGCGTAGGCGGGACGTCAGTCAAGTCAGCGGTAAAGACTGCAGCTAAACTGTAGCACGCCGTTGAAACTGTCACCTGGAGACGAGACGAGGGAGGCGGAACAAGTGAA
0046913ae6f9e12dbd889671ed26c09d	137	AGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCGAGGCGGTTCACTCAGTCTGTCTTAAAGTGCGGGGCTTAACCCCGTAGGGGACGGAACGCTGAAGTATAGAGTATCGAGAGGAAAGCAGAATTCTCTAGT
009c3d1fd56cf5682c875f959d9fee33	137	AGCGTTATCCGGATTATTGGGTTTAAAGGGTGCCTAGACGGGAAATTAAGTTAGTTGTGAAATCCCTCGGCTTAAGTGAAGCACTGCAACTAAACTGGTTTTCTTGAGTGCTGAGAGGAAAGTGGAAATCTCTAG
00c7ff8350c1f1874014b473b1890bf3	137	GGCGTTATCCGGATTATTGGGTTTAAAGGGAGCGTAGGCCGTGGGTTAAGTGTGTTGTGAAATCCGGTTGCTCAACATCCGGTTTGCAGCGCATCTGTCTCACTTGAGTGCACACACGACGAGCGGAATTCGTCG
014ed0286a2651609cc4a145b7ea6788	137	GGCGTTATCCGGATTATTGGGTTTAAAGGGAGCGTAGGCCGCGAGGTAAGCGTGTGTGAAATGTACCGGCTCAACCGGTGAATTGCAGCGCAACTGTCTGGCTTGAGTGCACGGTAAGTAGGCGGAATTCATGG

# Methods: Aligning Sequences

- *alignment mafft*: aligns sequences of feature table
- *alignment mask*: removes ambiguous positions from alignment

# Methods: Classifying Sequences

- *wget*: obtains 16s rRNA human stool classifier from SILVA database
- *feature-classifier classify-sklearn*: assigns taxonomy to rep seqs
- *metadata tabulate*: tabulates taxa and confidence of each feature

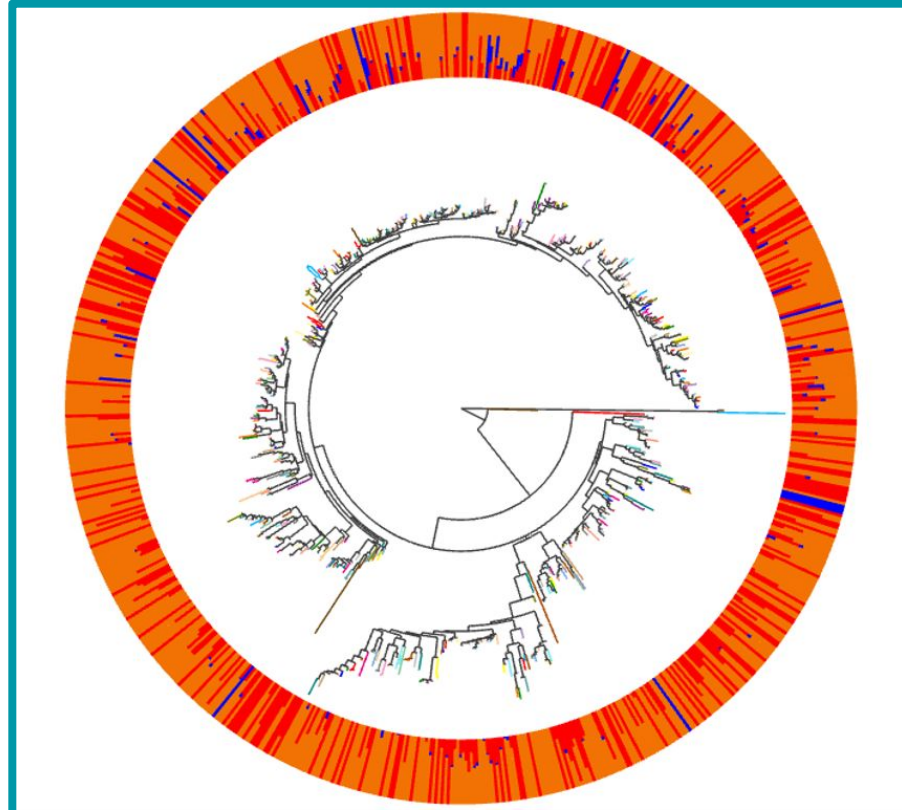
Feature ID #q2-types	Taxon categorical	Confidence categorical
002e78333d6cf2b11aa7a5ba03dd2c68	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Bacteroidales; f__Marinifilaceae; g__Butyrlicimonas; s__uncultured_bacterium	0.9443730139226642
0046913ae6f9e12dbd889671ed26c09d	d__Bacteria; p__Firmicutes; c__Negativicutes; o__Veillonellales-Selenomonadales; f__Veillonellaceae; g__Megasphaera; s__uncultured_bacterium	0.9972867127847357
009c3d1fd56cf5682c875f959d9fee33	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridia_UCG-014; f__Clostridia_UCG-014; g__Clostridia_UCG-014	0.9999999999982891
00c7ff8350c1f1874014b473b1890bf3	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Bacteroidales; f__Prevotellaceae; g__Prevotella; s__Prevotella_bergensis	0.9999999722448368
014ed0286a2651609cc4a145b7ea6788	d__Bacteria; p__Bacteroidota; c__Bacteroidia; o__Bacteroidales; f__Prevotellaceae; g__Prevotella; s__Prevotella_buccalis	0.9999876928118754

# Methods: Making Phylogenetic Tree

- *phylogeny fasttree*: makes tree from aligned sequences
- *phylogeny midpoint-root*: roots tree
- *empress tree-plot*: adds taxa to rooted tree
- *empress community-plot*: plots phylogenies and taxonomic community data



# Results: Phylogenetic Tree



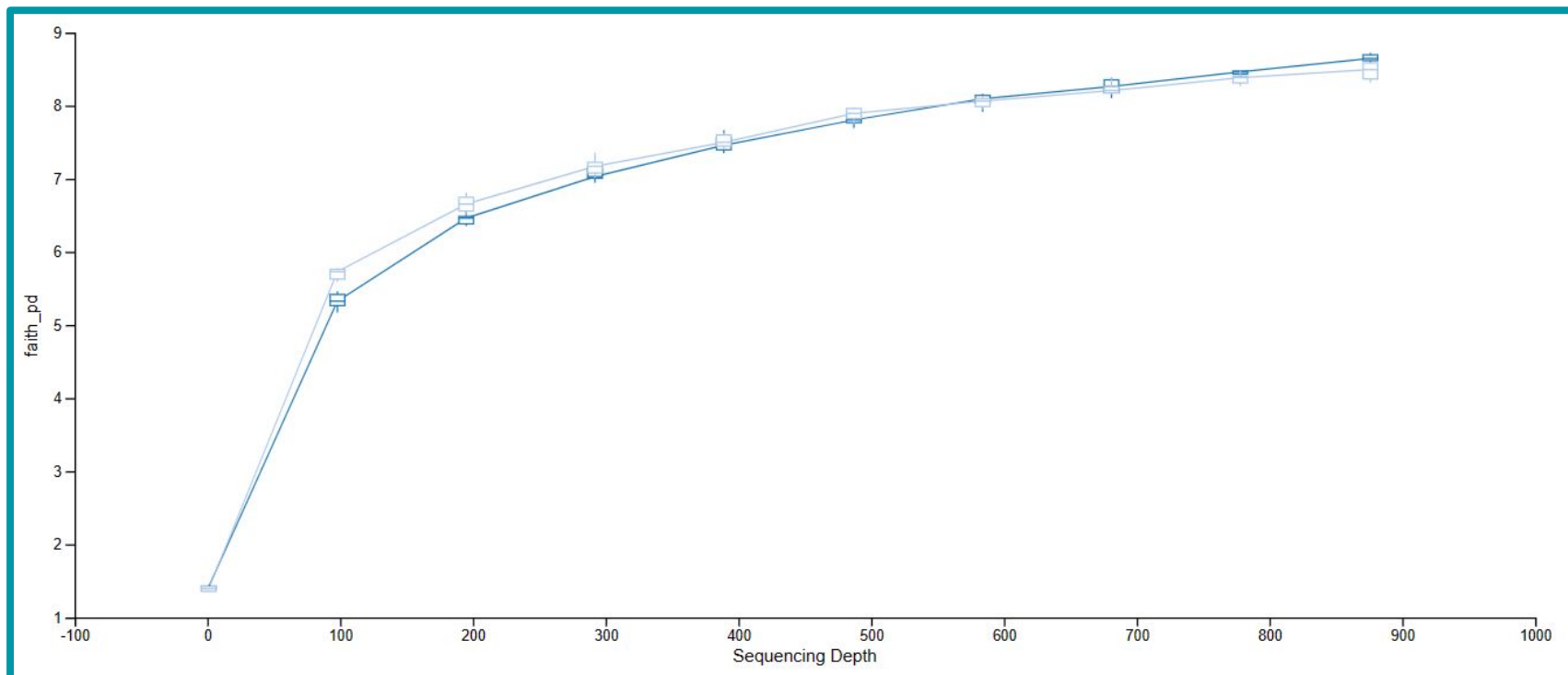
# Methods: Assessing Diversity Metrics

- *feature-table filter-samples*: filters samples to compare control and treatment groups
- *feature-table summarize*: makes filtered table to determine sequence depth
- *diversity core-metrics-phylogenetic*: makes and plots alpha and beta diversity metrics using sequence depth

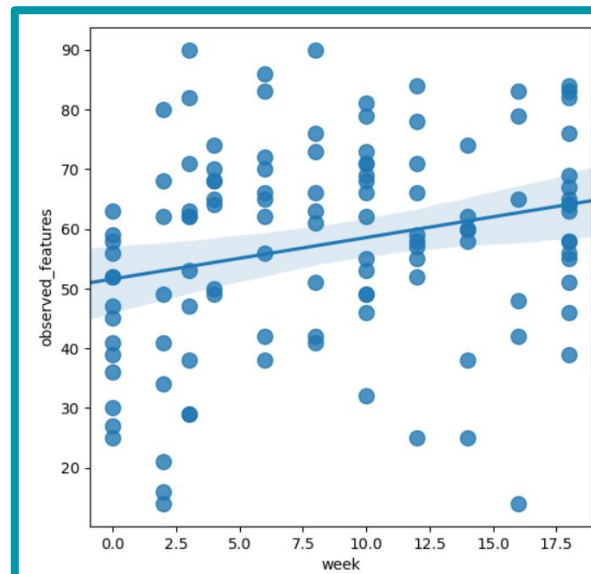
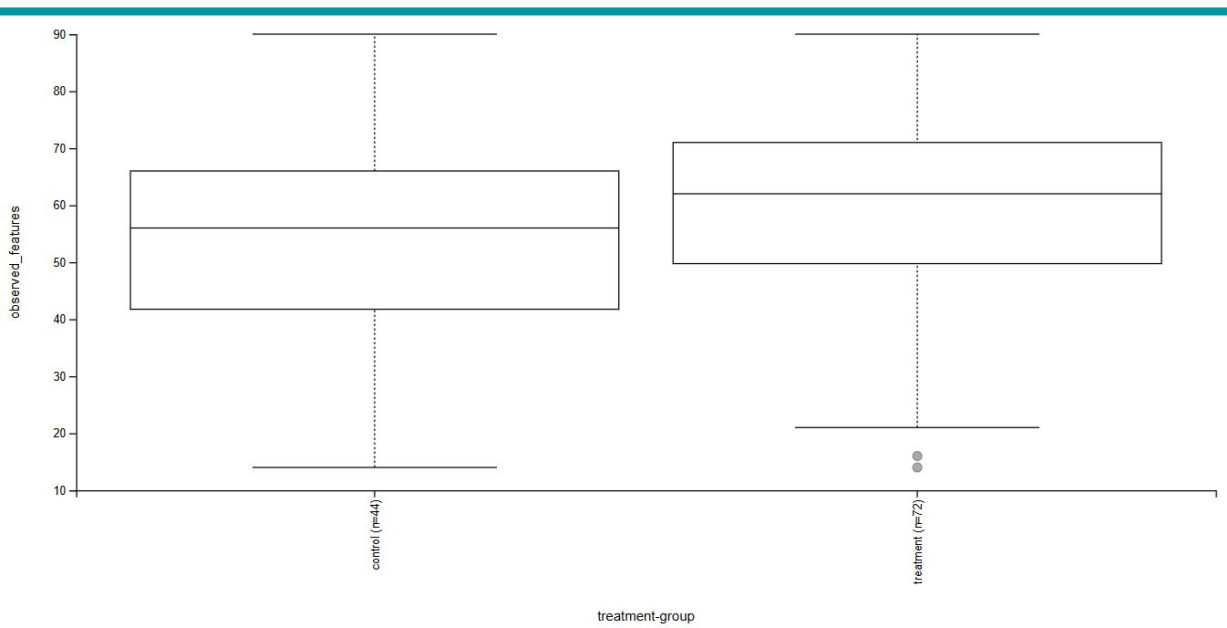
## Alpha Diversity:

- *diversity alpha-rarefaction*: verifies sequence depth and plots depth vs diversity
- *diversity alpha-group-significance*: plots alpha diversity vs observed features
- *longitudinal linear-mixed-effects*: makes alpha diversity linear plot with weekly treatment vs diversity

# Results: Alpha Rarefaction Plot



# Results: Alpha Diversity Plots

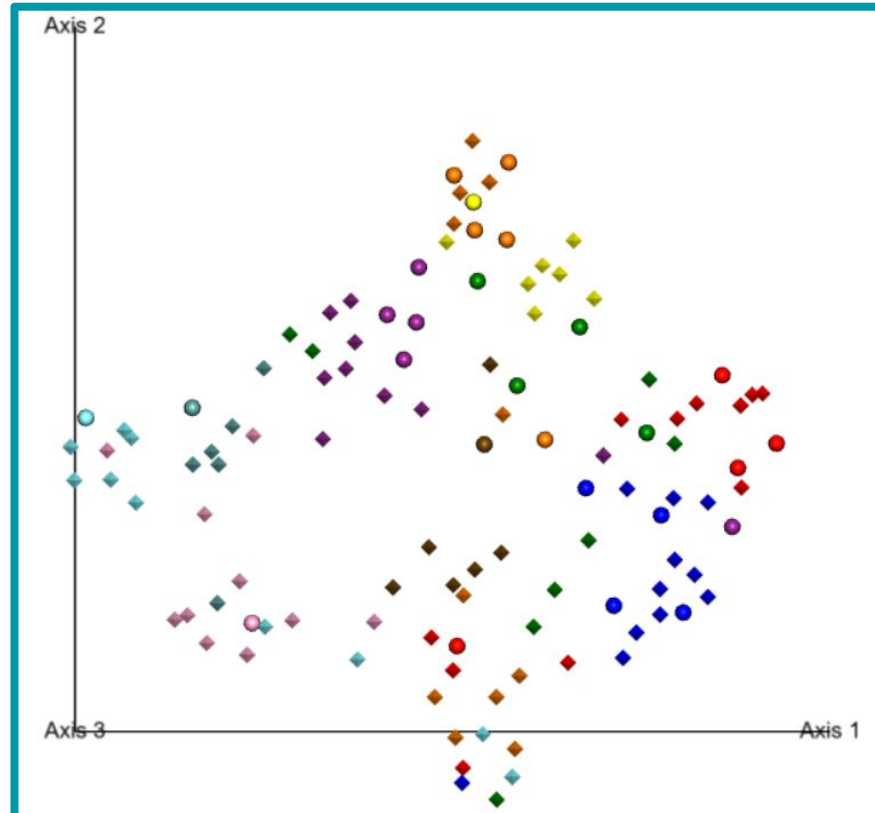


# Methods: Assessing Diversity Metrics

## Beta Diversity:

- *diversity umap*: reduces dimensions of beta diversity metrics using unweighted and weighted unifrac matrices
- *metadata tabulate*: tabulates unifrac matrices using diversity values (Faith's phylogenetic diversity, evenness, and Shannon diversity)
- *taxa barplot*: makes taxonomy barplot
- *emperor plot*: plots umap and pcoa data from beta diversity unifrac matrices

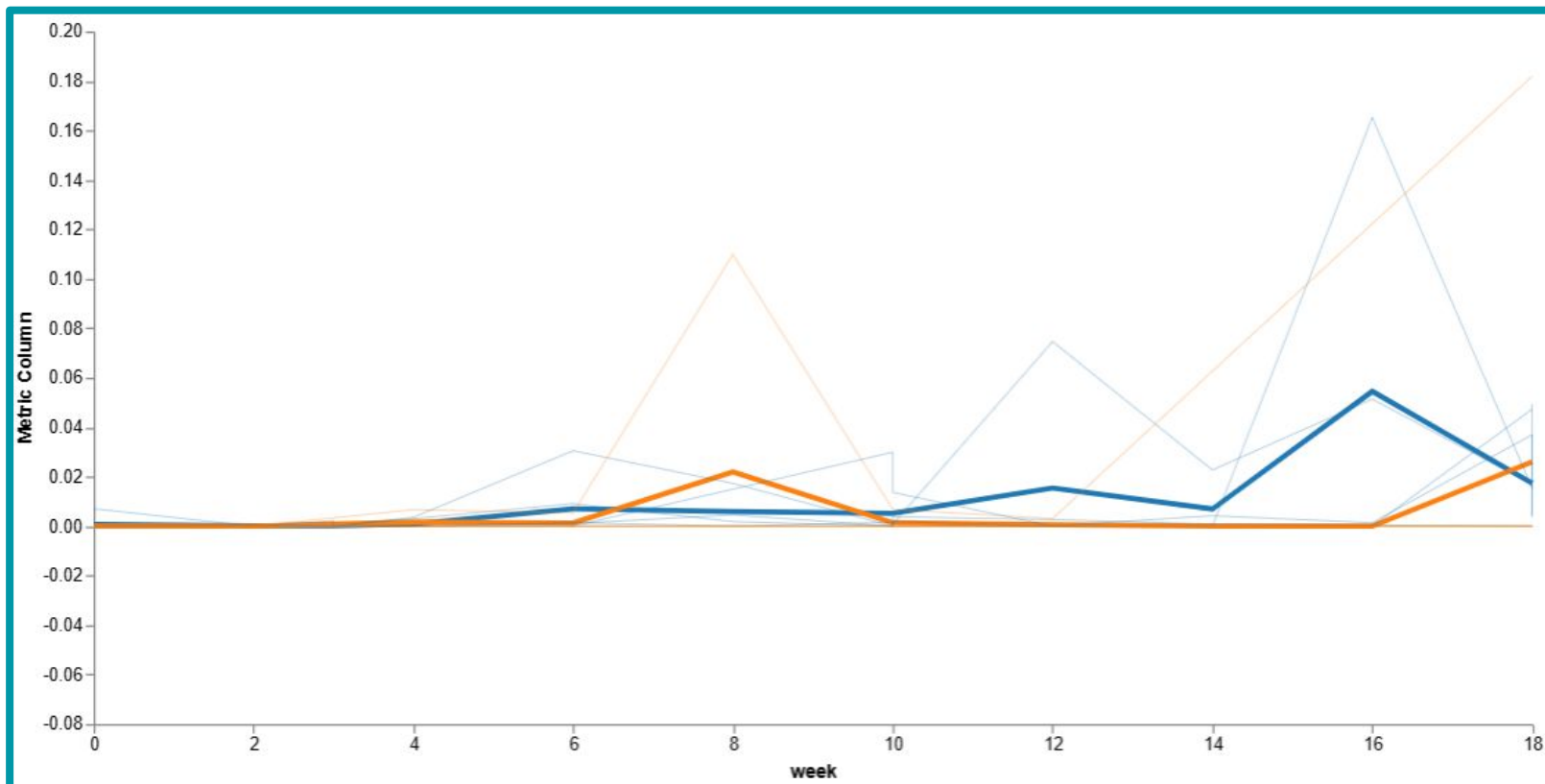
# Results: Beta Diversity Plot



# Methods: Assessing Diversity Metrics

- *taxa collapse*: adds taxa to feature table
- *feature-table filter-features-conditionally*: filters abundance of genera in feature table
- *feature-table relative-frequency*: converts counts in filtered feature table to relative frequencies
- *longitudinal volatility*: makes longitudinal volatility plot using metadata, diversity metrics, and taxa and relative frequencies from table
- *longitudinal feature-volatility*: makes volatility control plot to identify features that change over time

# Results: Volatility Control Plot





# Conclusions

- My full script contains all commands that I used for this project and can be found in my repo on GitHub
- I ran into issues getting the empress command and pushing from my local repo to my GitHub repo
- Qiime is a useful program for performing bioinformatic analyses

# Sources

Kang, DW., Adams, J.B., Gregory, A.C. *et al.* Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome* 5, 10 (2017).

<https://doi.org/10.1186/s40168-016-0225-7>

<https://docs.qiime2.org/2024.2/tutorials/fmt/>

<https://docs.qiime2.org/2022.2/tutorials/phylogeny/#sequence-alignment>

<https://zenodo.org/records/6395539#.ZGE7pHbMJhE>

<https://docs.qiime2.org/2024.2/tutorials/feature-classifier/>

<https://docs.qiime2.org/2024.2/tutorials/phylogeny/#fasttree>

<https://library.qiime2.org/plugins/empress/32/>

<https://github.com/biocore/empress#tutorial-using-empress-in-qiime-2>

<https://docs.qiime2.org/2024.2/tutorials/filtering/>

<https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/040-even-sampling.html>

<https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/050-core-metrics.html>

<https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/060-alpha-diversity.html>

<https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/070-beta-diversity.html>

<https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/080-longitudinal.html>