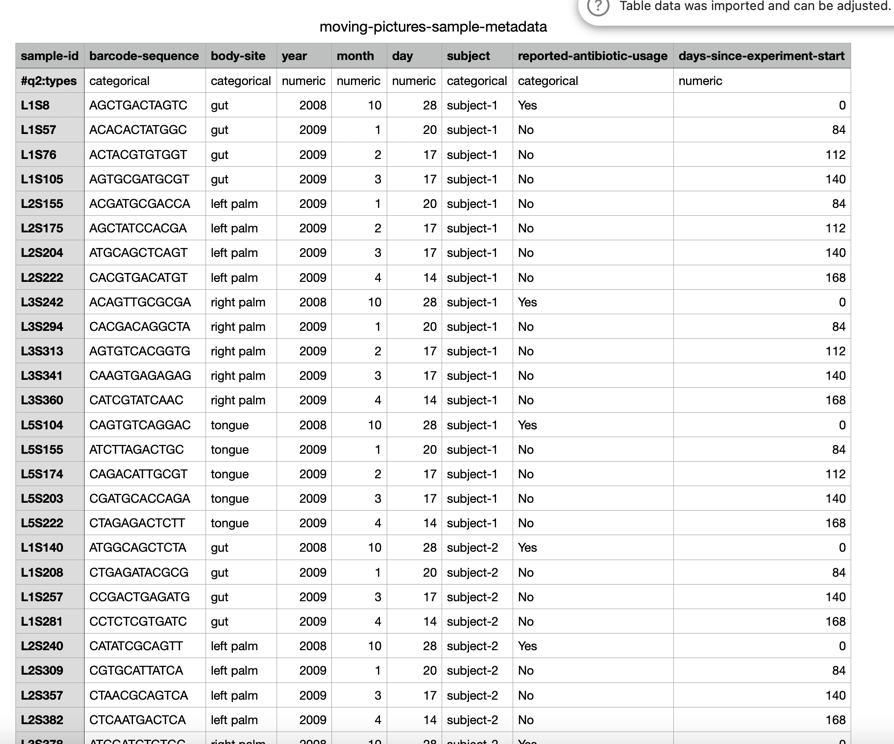
Nhi Nguyen

QIIME2

1. Metadata file:

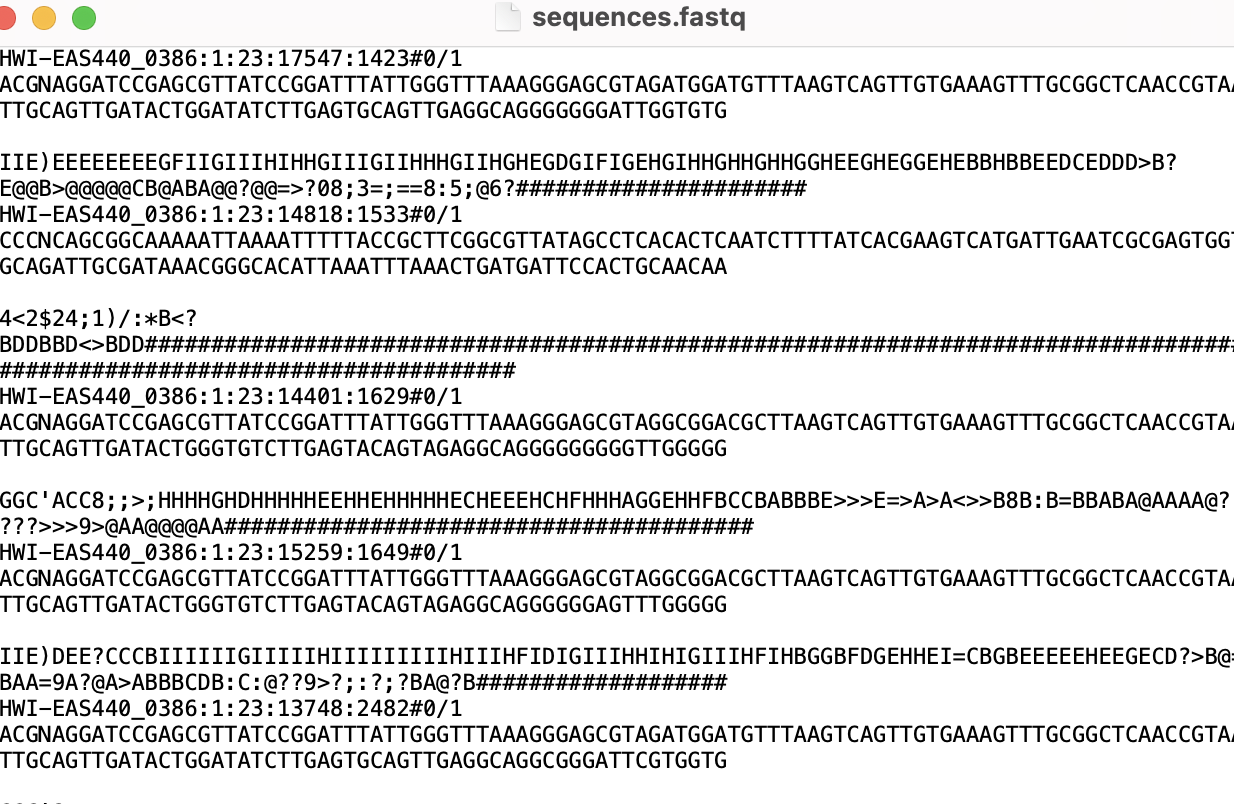


1. Barcode File:

A screenshot of a computer

Description automatically generated

Barcode.fastq



Sequence.fastq

1. A .qza file including citations and provenance

A screenshot of a computer

Description automatically generated

Req-seqs.qza

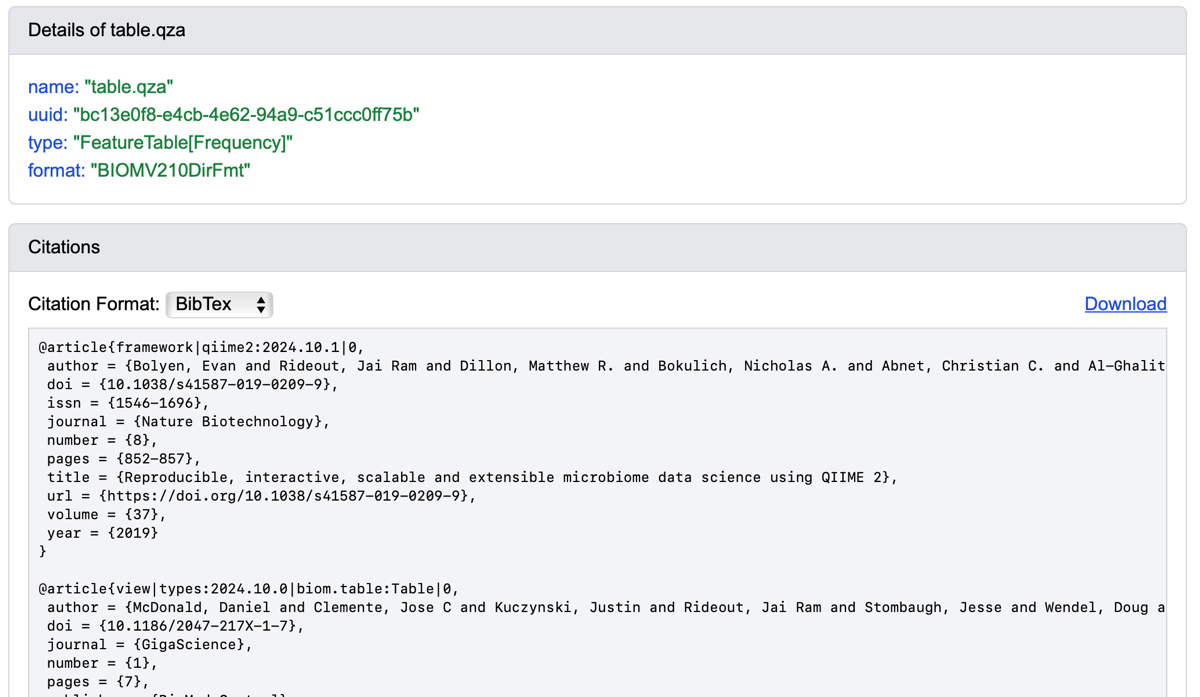


Table.qza

A diagram of a diagram

Description automatically generated

Table.qza (Prevenance)

A screenshot of a computer

Description automatically generated

Taxanomy.qza

A diagram of a diagram

Description automatically generated

Taxanomy.qza (Provenance)

1. The req-seq file

A screenshot of a computer

Description automatically generated

1. Demultiplexing

* demux.qza – demultiplexed sequences
* demux-details.qza – error correction info
* demux.qzv – visualization of sequence quality

A screenshot of a cell phone

Description automatically generated

A graph with numbers and a number of samples

Description automatically generated

A screenshot of a screenshot of a number

Description automatically generated

A graph of a quality score

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

Demux.qzv shows the quality of the sequence after separating by sample.

Helps you to decide where to trim before cleaning the data.

So in the figure, the quality drop after base 90 , so need to trim there before cleaning the data.

1. **Stat DADA2 Denoising:**

* table.qza – feature table
* rep-seqs.qza – representative sequences
* stats-dada2.qza – denoising stats
* stats-dada2.qzv – stats visualization

DADA2 is to helps to removes bad sequence, fix error , find clean sequence .

A screenshot of a data

Description automatically generated

This is DADA2 denoising summary, so for each sample, it show how many reads were kept after filtering, denoising ,and removing chimeras. Most sample kept over 50% of reads, which consider good quality data.

**Feature table summary**

Table.qzv

A screenshot of a computer

Description automatically generated

A graph of a bar

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

**Representative sequence:**

Req-seqs.qzv

A screenshot of a computer

Description automatically generated

**Phylogeny:**

* aligned-rep-seqs.qza, masked-aligned-rep-seqs.qza
* unrooted-tree.qza, rooted-tree.qza

**Core diversity Analysis**:

* core-metrics-results/ :
* faith\_pd\_vector.qza, evenness\_vector.qza,
* unweighted\_unifrac\_distance\_matrix.qza,
* Emperor plots:
* unweighted\_unifrac\_emperor.qzv
* bray\_curtis\_emperor.qzv, etc.

A screenshot of a computer

Description automatically generated

Bray- curtis emperor days since experiment start

A screenshot of a computer

Description automatically generated

Bray curtis emperor

A screenshot of a video game

Description automatically generated

Jaccard emperor

A screenshot of a computer

Description automatically generated

unweighted\_unifrac\_emperor.qzv (beta diversity colored by site)

A screenshot of a computer

Description automatically generated

Weighted unifrac emperor

A screenshot of a computer

Description automatically generated

Beta-diversity plot colored by reported-antibiotic-usage

A screenshot of a diagram

Description automatically generated

Alpha diversity boxplot (evenness-group-significance) (body site)

A screenshot of a graph

Description automatically generated

Alpha diversity boxplot (evenness-group-significance) (reported antibiotic usage)

A graph with a diagram

Description automatically generated with medium confidence

Faith phylogenetic diversity (body site)

A screenshot of a computer screen

Description automatically generated

Reported antibiotic usage (faith pd group significance)

**Group significance:**

unweighted-unifrac-body-site-significance.qzv

A diagram of a graph

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

A screenshot of a diagram

Description automatically generated

A screenshot of a graph

Description automatically generated

Alpha Rarefaction:

A graph of different colored lines

Description automatically generated

alpha-rarefaction.qzv

**Taxanomy:**

A screenshot of a computer

Description automatically generated

**Taxa bar plots**

A screenshot of a computer

Description automatically generated

**Differential Abundance (gut only):**

A screenshot of a graph

Description automatically generated

**Results:**

**Beta-Diversity Plots:**

4 distance metrics were used : Bray Curtis, Jaccard, Unweighted Unifrac, Weighted unifrac

Each plot was rotated to show all 3 principal coordinate (PC1, PC2, PC3) and colored by body site.

For Bray Curtis, the plot show distinct clustering by location , means that each site has different ,microbial communities . With axis 1, explains the most variations (32.42%)

Bray curtis plot colored by antibiotic usage show that people who took anitibiotic have different microbial communities.

Weighted UniFrac has the highest value on axis 1, with 69%, means that it explains the most difference between sampled based on both species’ presence and abundance.

**Alpha-Diversity Plots:**

Evenness show how balanced the microbial community is

Faith’s Phylogenetic diversity show how many different types of microbes there are

By body site:

Gut and tongue have lower diversity and lower evenness. Gut has the lowest phylogenetic diversity and evenness, meaning it has fewer and more dominant microbes.

By antibiotic:

People who used antibiotics have lower diversity in both metrics

They have less rich and less balanced in microbial communities

From the bar plot, we can see that gut samples had more Bacteroides. Oral samples had more Streptococcus and Veilonella. Palm samples had more Corynebacterium and Propionibacterium.

These bar plots helps to explain the diversity result, that different body part support different microbes.