**Tutorial: Sequence paired end data without barcode**

QIIME2 website: https://docs.qiime2.org/2023.2/tutorials/importing/

**Step1: #Load QIIME2 in bigpurple**

module add condaenvs/cpu/5.3.1/qiime2-2020.8

module load miniconda2/4.5.4

**Step2: #Covert BAM file to fastq paired end file**

module add samtools/1.9

samtools bam2fq input.bam > output.fq

#Spilt the Fastq file to paired end fastq files (forward.fq and backward.fq)

module add python/cpu/3.8.11

python3 split\_fq\_to\_paired\_end.py <absolute path to input fq> <prefix of output fq> <path to the directory of output>

**Step3: Process with QIIME2**

#Before importing fastq files, we need to create a text or csv documentation and load the existing fastq file names, file path and directions.

# exmaple template:

sample-id,absolute-filepath,direction

Sample1,/path/sample1\_R1.fq,forward

Sample1,/path/sample1\_R1.fq,backward

#Data import

qiime tools import --type 'SampleData[PairedEndSequencesWithQuality]' --input-path . --output-path demux.qza --input-format PairedEndFastqManifestPhred33

# Visualization for demux.qza

qiime demux summarize --i-data demux.qza --o-visualization demux.qzv

#Denoise with default parameters

## This step is preferred using sbatch commend due to large dataset

qiime dada2 denoise-paired --i-demultiplexed-seqs demux.qza --p-trim-left-f 13 --p-trim-left-r 13 --p-trunc-len-f 150 --p-trunc-len-r 150 --o-table table.qza --o-representative-sequences rep-seqs.qza --o-denoising-stats denoising-stats.qza

#Feature table (OTU table)

qiime feature-table summarize --i-table table.qza --o-visualization table.qzv --m-sample-metadata-file metadata.txt

#Output phylogenetic tree structure information

qiime phylogeny align-to-tree-mafft-fasttree --i-sequences 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

#Taxonomy table and its visualization

##Classifier: "2022.10.backbone.full-length.nb" http://ftp.microbio.me/greengenes\_release/2022.10/2022.10.backbone.full-length.nb.qza

##This step is preferred using sbatch commend due to large dataset (need more memory)

qiime feature-classifier classify-sklearn --i-classifier 2022.10.backbone.full-length.nb.qza --i-reads rep-seqs.qza --o-classification taxonomy.qza

qiime metadata tabulate --m-input-file taxonomy.qza --o-visualization taxonomy.qzv

**Step4: #Convert qiime 2 output (.qza) to phyloseq version using R**

Tutorial: https://forum.qiime2.org/t/tutorial-integrating-qiime2-and-r-for-data-visualization-and-analysis-using-qiime2r/4121

library(qiime2R)

library(phyloseq)

library(MicrobeR) # Data visualization

library(microbiome) # Data analysis and visualization

library(plyr)

pilot\_physeq <- qza\_to\_phyloseq(

features = "table.qza",

tree = "rooted-tree.qza",

taxonomy = "taxonomy.qza",

metadata = "metadata.txt"

)

pilot\_physeq

Downstream analysis references:

https://joey711.github.io/phyloseq/index.html