#Classifier need to be re-trained if diff lib prep method used

#Training classifier

#download gg\_13\_8\_otus.tar.gz file

#extract files

tar xvf ~/FL103/gg\_13\_8\_otus.tar.gz -C ~/FL103/gg\_files/

qiime tools import \

--type 'FeatureData[Sequence]' \

--input-path ~/FL103/gg\_files/99\_otus.fasta \

--output-path ~/FL103/gg\_files/99\_otus.qza

qiime tools import \

--type 'FeatureData[Taxonomy]' \

--input-format HeaderlessTSVTaxonomyFormat \

--input-path ~/FL103/gg\_files/99\_otu\_taxonomy.txt \

--output-path ~/FL103/gg\_files/ref-taxonomy.qza

qiime feature-classifier extract-reads \

--i-sequences ~/FL103/gg\_files/99\_otus.qza \

--p-f-primer ACTCCTACGGGAGGCAGCAG \

--p-r-primer GGACTACHVGGGTWTCTAAT \

--p-trunc-len 487 \

--p-min-length 300 \

--p-max-length 500 \

--o-reads ~/FL103/gg\_files/ref-seqs.qza

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

--i-reference-reads ~/FL103/gg\_files/ref-seqs.qza \

--i-reference-taxonomy ~/FL103/gg\_files/ref-taxonomy.qza \

--o-classifier ~/FL103/gg\_files/classifier.qza

#Using trained classifier

qiime feature-classifier classify-sklearn \

--i-classifier ~/FL103/gg\_files/classifier.qza \

--i-reads ~/FL103/DADA2\_350\_unfiltered/representative\_sequences.qza \

--o-classification ~/FL103/DADA2\_350\_unfiltered/trained\_tax/taxonomy.qza

#visualize/adding metadata to the taxonomy file

qiime metadata tabulate \

--m-input-file ~/FL103/DADA2\_350\_unfiltered/taxonomy.qza \

--o-visualization ~/FL103/DADA2\_350\_unfiltered/taxonomy.qzv

#filter out chloroplasts and mitochondria

qiime taxa filter-table \

--i-table ~/FL103/DADA2\_350\_unfiltered/table.qza \

--i-taxonomy ~/FL103/DADA2\_350\_unfiltered/taxonomy.qza \

--p-exclude mitochondria,chloroplast \

--o-filtered-table ~/FL103/DADA2\_350\_unfiltered/table-no-mitochondria-no-chloroplast.qza

srun --mem=64000 --partition=high --time=07:00:00 \

qiime feature-table filter-samples \

--i-table ~/FL103/DADA2\_350\_unfiltered/table-no-mitochondria-no-chloroplast.qza \

--m-metadata-file ~/FL103/FL103\_MapFile.txt \

--p-exclude-ids \

--p-where "SubID ='MockComm'" \

--o-filtered-table ~/FL103/DADA2\_350\_unfiltered/table\_filtered.qza

qiime feature-table summarize \

--i-table ~/FL103/DADA2\_350\_unfiltered/table-no-mitochondria-no-chloroplast.qza \

--o-visualization ~/FL103/DADA2\_350\_unfiltered/table-no-mitochondria-no-chloroplast.qzv \

--m-sample-metadata-file ~/FL103/FL103\_MapFile.txt

qiime feature-table summarize \

--i-table ~/FL103/DADA2\_350\_unfiltered/table\_filtered.qza \

--o-visualization ~/FL103/DADA2\_350\_unfiltered/table\_filtered.qzv \

--m-sample-metadata-file ~/FL103/FL103\_MapFile.txt

#generate interactive barplot of taxonomy

qiime taxa barplot \

--i-table ~/FL103/DADA2\_350\_unfiltered/diversity\_DADA2\_nomock/rarefied\_table.qza \

--i-taxonomy ~/FL103/DADA2\_350\_unfiltered/taxonomy.qza \

--m-metadata-file ~/FL103/FL103\_MapFile.txt \

--o-visualization ~/FL103/DADA2\_350\_unfiltered/taxa\_summary.qzv