QIIME2 Pipeline

qiime2/2022.2

# Raw multiplexed data

**Location**: /proj/carrlab/HTSF/221213\_UNC22\_1006\_000000000\_KRCWW

Undetermined\_S0\_L001\_R1\_001.fastq.gz

Undetermined\_S0\_L001\_R2\_001.fastq.gz

Undetermined\_S0\_L001\_I1\_001.fastq.gz

\* Could not upload files to UNCC cluster

# Rename files for QIIME2 pipeline

**Location**: /nas/longleaf/home/aasorgen/AnorexiaNervosa\_Stress/emp-paired-end-sequences

Undetermined\_S0\_L001\_R1\_001.fastq.gz -> forward.fastq.gz

Undetermined\_S0\_L001\_R2\_001.fastq.gz -> reverse.fastq.gz

Undetermined\_S0\_L001\_I1\_001.fastq.gz -> barcodes.fastq.gz

# Import sequences into QIIME2

qiime tools import \

--type EMPPairedEndSequences \

--input-path ${inputPath}/emp-paired-end-sequences \

--output-path ${seqs}/emp-paired-end-sequences.qza

# Demultiplex sequences

qiime demux emp-paired \

--m-barcodes-file ${inputPath}/ANIGMA\_MAPPING.txt \

--m-barcodes-column BarcodeSequence \

--p-rev-comp-mapping-barcodes \

--i-seqs ${seqs}/emp-paired-end-sequences.qza \

--o-per-sample-sequences ${seqs}/demux-full.qza \

--o-error-correction-details ${details}/demux-details.qza

# Remove adapters/primers

qiime cutadapt trim-paired \

--i-demultiplexed-sequences ${seqs}demux-full.qza \

--p-front-f GAGTGCCAGCMGCCGCGGTAA \

--p-front-r ACGGACTACDBGGGTWTCTAAT \

--p-error-rate 0 \

--o-trimmed-sequences ${seqs}/trimmed-seqs.qza \

--verbose

qiime demux summarize \

--i-data ${seqs}/trimmed-seqs.qza \

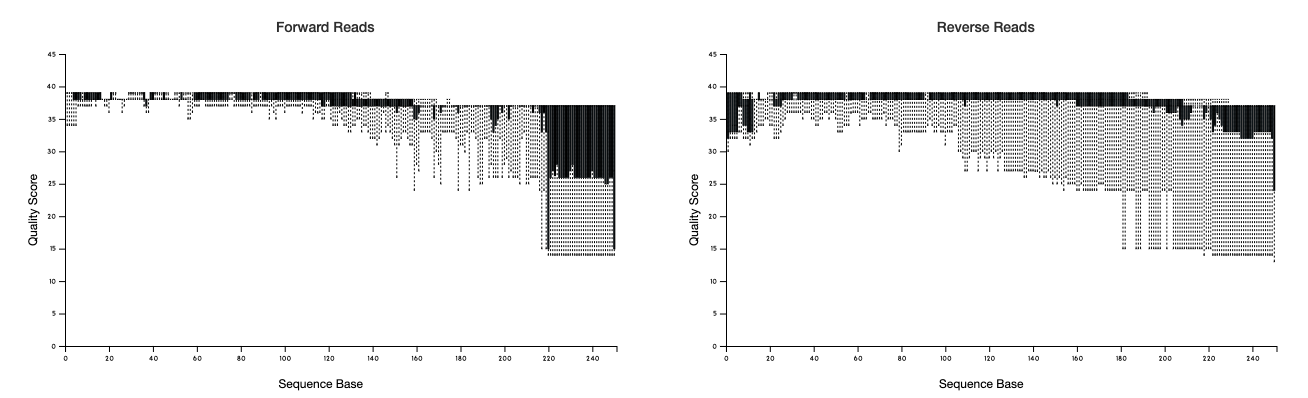
--o-visualization ${seqs}/trimmed-seqs.qzv

Total samples: 362

Total sequence counts: 12,690,965

Read length: 250 NT

# Denoise and quality trim sequences



trimmed-seqs.qzv

qiime dada2 denoise-paired \

--i-demultiplexed-seqs ${seqs}/trimmed-seqs.qza \

--p-trim-left-f 0 \

--p-trim-left-r 0 \

--p-trunc-len-f 200 \

--p-trunc-len-r 200 \

--o-table ${tables}/ASVtable.qza \

--o-representative-sequences ${seqs}/ASVseqs.qza \

--o-denoising-stats ${tables}/denoising-stats.qza \

--verbose \

--p-n-threads 16

# Filter out samples that aren’t Kylie’s

qiime feature-table filter-samples \

--i-table ${tables}/ASVtable.qza \

--m-metadata-file ${inputPath}/ANIGMA\_MAPPING.txt \

--p-where '[person]="Kylie"' \

--o-filtered-table Kylie-ASVtable.qza

# Remove ASVs with a frequency < 10

qiime feature-table filter-features \

--i-table ${tables}/ASVtable.qza \

--p-min-frequency 10 \

--o-filtered-table ${tables}/filtered-ASVtable.qza

# Remove sequences filtered out of ASV table

qiime feature-table filter-seqs \

--i-data ${seqs}/ASVseqs.qza \

--i-table ${tables}/filtered-ASVtable.qza \

--o-filtered-data ${seqs}/filtered-ASVseqs.qza

# Produce readable filtered ASV table file

qiime feature-table summarize \

--i-table ${tables}/filtered-ASVtable.qza \

--o-visualization ${visualizations}/filtered-ASVtable.qzv

Number of samples: 304

Number of features: 5,334

Total frequency: 6,287,928

# Remove ASVs with a total frequency < 0.01%

Cutoff = 628

qiime feature-table filter-features \

--i-table ${tables}/filtered-ASVtable.qza \

--p-min-frequency ${cutoff} \

--o-filtered-table ${tables}/filtered-ASVtable-abun.qza

# Produce readable filtered ASV table file

qiime feature-table summarize \

--i-table ${tables}/filtered-ASVtable-abun.qza \

--o-visualization ${visualizations}/filtered-ASVtable-abun.qzv

Number of samples: 303

Number of features: 1,656

Total frequency: 5,642,228

# Filter out super low reads

Cutoff = 1000

Max = 51,197 (Sample 240)

Min = 1 (Sample 48)

qiime feature-table filter-samples \

--i-table ${tables}/filtered-ASVtable-abun.qza \

--p-min-frequency ${cutoff} \

--o-filtered-table ${tables}/final-filtered-table.qza

qiime feature-table summarize \

--i-table ${tables}/final-filtered-table.qza \

--o-visualization ${visualizations}/final-filtered-table.qzv

31 samples removed

<1,000: 178, 158, 113, 116

<100: 105, 336, 315, 168, 132, 123, 70, 117, 64, 86

<10: 88, 80, 98, 309, 122, 126, 130, 190, 34, 79, 102, 110, 118, 121, 77, 90, 48

# Filter sequences

qiime feature-table filter-seqs \

--i-data ${seqs}/filtered-ASVseqs.qza \

--i-table ${tables}/final-filtered-table.qza \

--o-filtered-data ${seqs}/final-filtered-seqs.qza

qiime feature-table tabulate-seqs \

--i-data ${seqs}/final-filtered-seqs.qza \

--o-visualization ${visualizations}/final-filtered-seqs.qzv

# Assign taxonomy to ASVs using Silva 138

qiime feature-classifier classify-consensus-vsearch \

--i-query ${seqs}/final-filtered-seqs.qza \

--i-reference-reads ${inputPath}/silva-138-99-seqs.qza \

--i-reference-taxonomy ${inputPath}/silva-138-99-tax.qza \

--o-classification ${tables}/taxonomy.qza

qiime metadata tabulate \

--m-input-file ${tables}/taxonomy.qza \

--o-visualization ${visualizations}/taxonomy.qzv

# Remove all mitochondria/chloroplast identification from taxonomy

qiime taxa filter-table \

--i-table ${tables}/final-filtered-table.qza \

--i-taxonomy ${tables}/taxonomy.qza \

--p-exclude mitochondria,chloroplast \

--o-filtered-table ${tables}/taxa-table.qza

qiime feature-table summarize \

--i-table ${tables}/taxa-table.qza \

--o-visualization ${visualizations}/taxa-table.qzv

# Make taxonomy barplots

qiime taxa barplot \

--i-table ${tables}/taxa-table.qza \

--m-metadata-file ${metadata} \

--i-taxonomy ${tables}/taxonomy.qza \

--o-visualization ${visualizations}/taxonomy-barplot.qzv

# Generate count tables

Open taxonomy-barplot.qzv in QIIME2 View

Download csv for each level

Phylum – level-2

Class – level-3

Order – level-4

Family – level-5

Genus – level-6

Species – level-7