

Qiime 2

- importing data in Manifest file

```
qiime tools import \
```

```
> --input-path manifest.tsv \
```

```
> --output-path reads.qza \
```

```
> --type 'SampleData[PairedEndSequencesWithQuality]' \
```

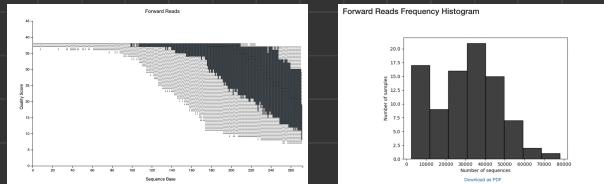
```
> --input-format PairedEndFastqManifestPhred33V2 \
```

- Summarizing the data, visualizing it, and getting quality scores

```
qiime demux summarize \
```

```
> --i-data reads.qza \
```

```
> --o-visualization reads.qzv
```



```
qiime tools view reads.qzv
```

↳ quality scores ranged between 20 & 30 are good, lower not so good

↳ based on the quality scores, find at which bases to cut off both the forward and reverse reads

- Quality control and denoise the data

```
qiime dada2 denoise-paired \
```

```
> --i-demultiplexed-seqs reads.qza \
```

```
> --p-trim-left-f # \
```

```
> --p-trim-left-r # \
```

```
> --p-trunc-len-f # \
```

```
> --p-trunc-len-r # \
```

```
> --o-table table.qza \
```

```
> --o-representative-sequences rep_seqs.qza \
```

```
> --o-denoising-stats denoising-stats.qza
```

sample-id	input	filtered	denoised	merged	non-chimeric	percentage of input passed filter	percentage of input merged	percentage of input non-chimeric
	numerical	numerical	numerical	numerical	numerical			
sample-1	62596	36711	58.65	35727	32239	51.5	28580	45.66
sample-10	49936	28984	58.04	28440	26847	53.76	24290	48.64
sample-11	42635	26336	61.77	26124	25535	59.89	25434	59.66
sample-12	34085	19760	57.97	19337	18191	53.37	17494	51.32
sample-13	37209	18336	49.28	18136	17902	48.11	14840	39.88
sample-14	4636	2632	56.77	2453	2155	46.48	2155	46.48
sample-15	8542	4916	57.55	4778	4690	54.91	4690	54.91
sample-16	2696	1389	51.52	1269	1053	39.06	1053	39.06
sample-17	5529	2597	46.97	2360	1823	32.97	1823	32.97
sample-18	8827	4772	54.06	4466	3849	43.6	3849	43.6

```
qiime metadata tabulate \
```

```
> --m-input-file denoising-stats.qza \
```

```
> --o-visualization denoising-stats
```

```
qiime tools view denoising-stats.qzv
```

- Build a phylogenetic tree
 - perform multiple sequence alignment with mafft
 - qiime alignment mafft \


```
> -- i-sequences rep-seqs.qza \
> -- o-alignment aligned-rep-seqs
```
 - mask or filter the alignment
 - qiime alignment mask \

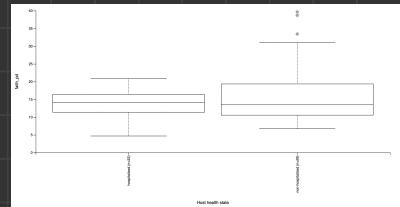

```
> -- i-alignment aligned-rep-seqs.qza \
> -- o-masked-alignment masked-aligned-rep-seqs.qza
```
 - Build an unrooted tree using fasttree
 - qiime phylogeny fasttree \


```
> --i-alignment masked-aligned-rep-seqs.qza \
> --o-tree unrooted-tree
```
 - Root the tree
 - qiime phylogeny midpoint-root \


```
> --i-tree unrooted-tree.qza \
> --o-rooted-tree rooted-tree
```
- Diversity analysis
 - run the default core diversity metrics
 - qiime diversity core-metrics-phylogenetic \


```
> --i-phylogeny rooted-tree.qza \
> --i-table table.qza \
> --p-sampling-depth # \ → the base sample size you want
> --m-metadata-file manifest.tsv \
> --output-dir core-metrics-results
```
 - alpha diversity metrics
 - qiime diversity alpha-group-significance \


```
> --i-alpha-diversity core-metrics-results/faith-pd-vector.qza \
> --m-metadata-file manifest.tsv \
> --o-visualization core-metrics-results/faith-pd-group-significance
```



qiime tools view core-metrics-results/faith-pd-group-significance.qzv

- visualize beta diversity Emperor plots

qiime tools view core-metrics-results/unweighted-unifrac-emperor.qzv

- Alpha rarefaction plots

qiime diversity alpha-rarefaction\

>-- i-table core-metrics-results/rarefied_table.qza\

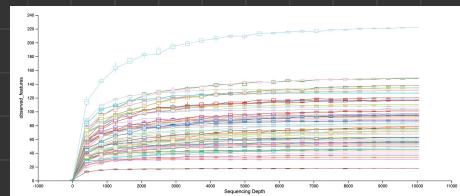
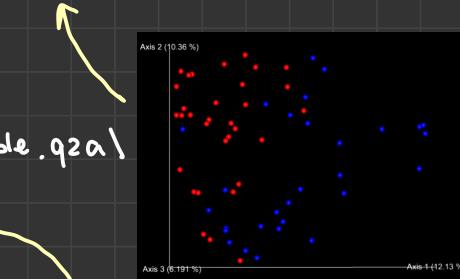
>-- p-max-depth # \

>-- m-metadata-file manifest.tsv\

>-- p-steps 25 \ → default is 10

>-- o-visualization alpha-rarefaction.qzv

qiime tools view alpha-rarefaction.qzv



Training the Classifier

- Extract reference reads

qiime feature-classifier extract-reads\

>-- i-sequences silva132-99.qza \ → Could use another database

>-- p-f-primer insert_f_primer\ like 85_0TU

>-- p-r-primer insert_r_primer\

>-- p-trunc-len # \

>-- o-reads ref_seqs

- Train the classifier

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

>-- i-reference-reads ref_seqs.qza\

>-- i-reference-taxonomy silva132-99_ref_taxonomy.qza\

>-- o-classifier classifier.qza

Assign Taxonomy

qiime feature-classifier classify-sklearn\

>-- i-classifier classifier.qza\

>-- i-reads rep_seqs.qza\

>-- o-classification taxonomy.qza

★ - classifier is
unite-ver8-99-classifier
-09.02.2020.qza

Visualize taxonomy

qiime metadata tabulate \

```
> -- i-input-file taxonomy.qza \n> -- o-visualization taxonomy
```

qiime tools view taxonomy.qzv

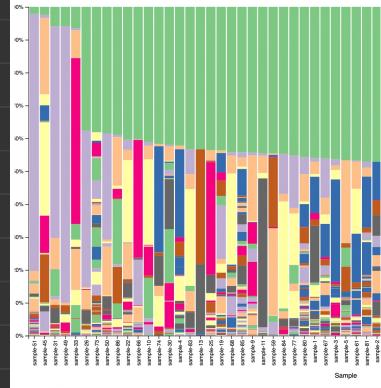
Feature ID Filepath	Taxon category
00008d8aefc7302721d437a429d8a97f	D_0_Bacteria
0025cbef9aace4594205de011954eb2	D_0_Bacteria
0024cd5790008b19559a2986519832a2	D_0_Bacteria
00761909cc420215d35542b4838354	D_0_Bacteria
00903c29393cc7d841bad5b5d17afe	D_0_Bacteria
009fb47982490050ff622e90e7a0535	D_0_Bacteria
00a01ca07b78ed50300a83729370cfc	D_0_Bacteria
00a1a1dc7a23ee338974e1987a91945	D_0_Bacteria,D_1_Firmicutes,D_2_Bacilli,D_3_Lactobacillales,D_4_Lactobacillaceae,D_5
00e54032890742090032747292cb4b53	D_0_Bacteria,D_1_Cyanobacteria,D_2_Oxypotobacteria,D_3_Nostocales,D_4_Nostoc
00b34ef4b777a53bb65a1e3cadfa3a	D_0_Bacteria,D_1_Firmicutes,D_2_Bacilli,D_3_Lactobacillales,D_4_Lactobacillaceae,D_5
00d54239f97d3be2823d3e3099e3030	D_0_Bacteria
00ec68758f30191aa772a129a7d5297	D_0_Bacteria

Make barplots of taxonomy data

qiime taxa barplot \

```
> -- i-table core_metrics_results/rarefied_table.qza \n> -- i-taxonomy taxonomy.qza \n> -- m-metadata-file manifest.tsv \n> -- o-visualization taxa-bar-plots.qzv
```

qiime tools view taxa-bar-plots.qzv



When importing single end Fastq files:

qiime tools import \

```
> -- type 'SampleData[SequencesWithQuality]' \n> -- input-path manifest.tsv \n> -- output-path reads.qza \n> -- input-format SingleEndFastqManifestPhred33V2
```

QIIME Picrust2 plugin

To access: /users/SydneySanchez/Desktop/q2-picrust2-2021.11_0/picrust2

Full pipeline:

```
$ qiime picrust2 full-pipeline \
> --i-table biom.qza \
> --i-seq seqs.qza \
> --output-dir q2-picrust2-output \
> --p-threads 1 \
> --p-hsp-method pic \
> --p-max-nsti 2 \
> --verbose
```

The required inputs i-table and i-seq, need to correspond to QIIME artifacts of types FeatureTable[Frequency] and FeatureData[Sequence]. The FeatureTable needs to contain the abundances of ASVs (a BIOM table) and the sequence file needs to be a FASTA file containing the sequences for each ASV

For a description of the command:

```
$ qiime picrust2 full-pipeline --help
```

The output files:

- ec-metagenome.qza - EC metagenome predictions
- ko-metagenome.qza - KO metagenome predictions
- pathway-abundance.qza - MetaCyc pathway abundance predictions
 - ↳ the artifacts are all of type FeatureTable[Frequency], so they can be used with QIIME plugins

To visualize the summary information:

```
$ qiime feature-table summarize \
> --i-table pathway-abundance.qza \
> --o-visualization pathway-abundance.qzv
$ qiime tools view pathway-abundance.qzv
```

For core diversity metrics:

```
$ qiime diversity core-metrics \
> --i-table pathway-abundance.qza \
> --p-sampling-depth # → # = minimum sample pathway abundance found \
> --m-metadata-file metadata.tsv \
cont'd ...
```

```
> --output-dir pathabun_core_metrics_out \
> --p-n-jobs 1
```

To use the tables outside of qiime, convert to BIOM format:

```
# qiime tools export \
> --input-path pathway-abundance.qza \
> --output-path pathabun-exported
biom convert \
- i pathabun-exported/feature-table.biom \
- o pathabun-exported/feature-table.biom.tsv \
--to-tsv
```

} converts a BIOM file to plain-text