Phân tích hệ vi sinh vật bằng giải trình tự thế hệ mới

Báo cáo viên: ThS. Lương Bắc An

Nội dung:
- Thu nhận mẫu và tách chiết DNA
- Chuẩn bị thư viện giải trình tự
- Phân tích dữ liệu giải trình tự bằng BaseSpace (Illumina).
- Phân tích mở rộng với công cụ QIIME2.

Thu nhận mẫu và tách chiết DNA

ZYMO RESEARCH

Microbiomics
Made Simple™

HostZERO™ Microbial DNA Kit
Streamlined host depletion and bacterial DNA isolation from mammalian samples

Step 1
Depletion of eukaryotic host DNA from sample.

Add Host Depletion Solution directly to sample.

Step 2
Unbiased lysis of remaining microbial cells.

Lyse sample with ZR BashingBead™ Lysis Tube (0.1 mm & 0.5 mm).

Step 3
Isolation of microbial DNA.

Bind, wash, and elute DNA with Zymo-Spin™ IC-Z.

Ultra-Pure Microbial DNA

- Lấy mẫu đại diện (mẫu, chứng buffer bảo quản)
- Bảo quản mẫu trong dung dịch bảo quản phù hợp.
- Tách chiết bằng bộ kít phù hợp (có thể tách gram -/+)
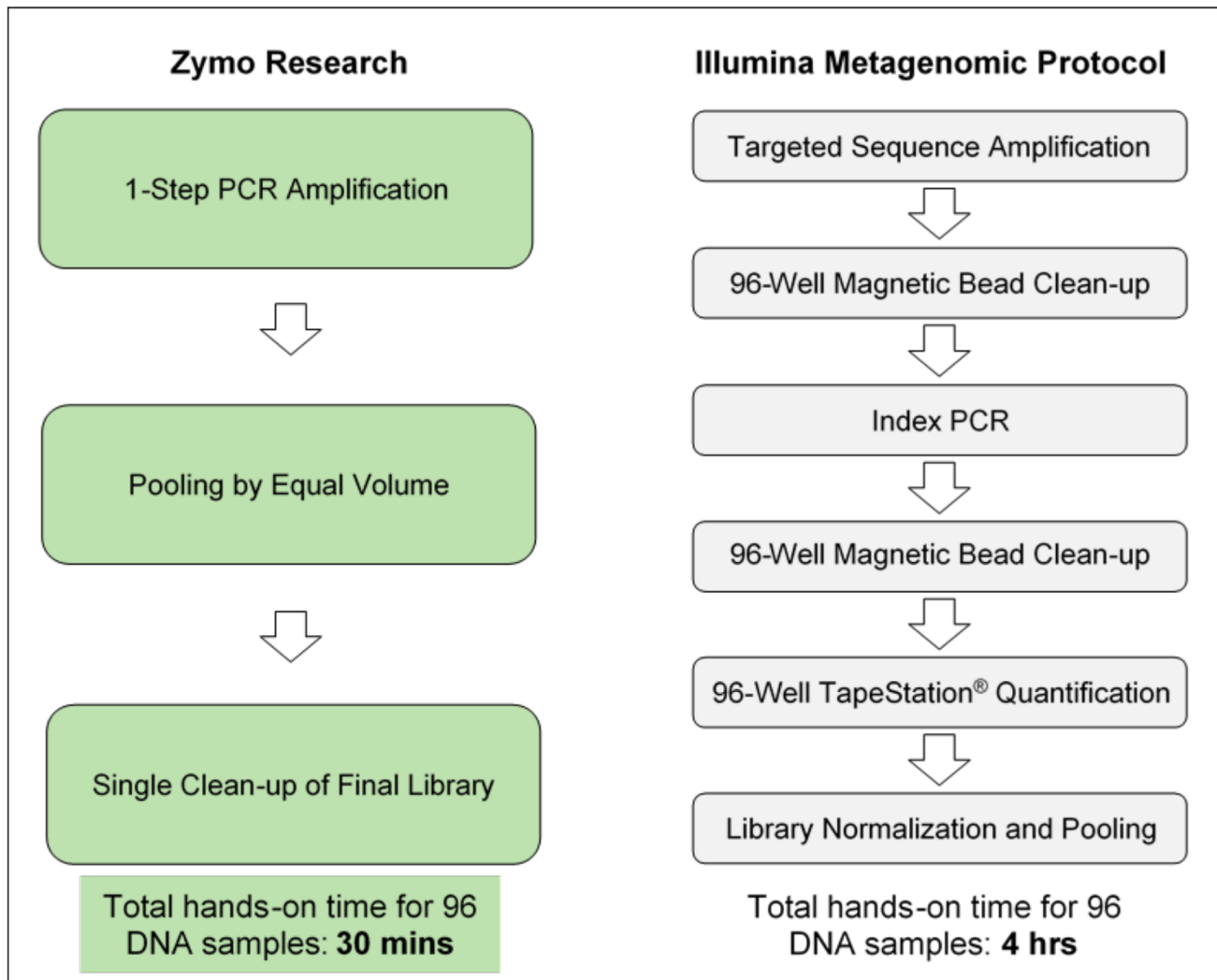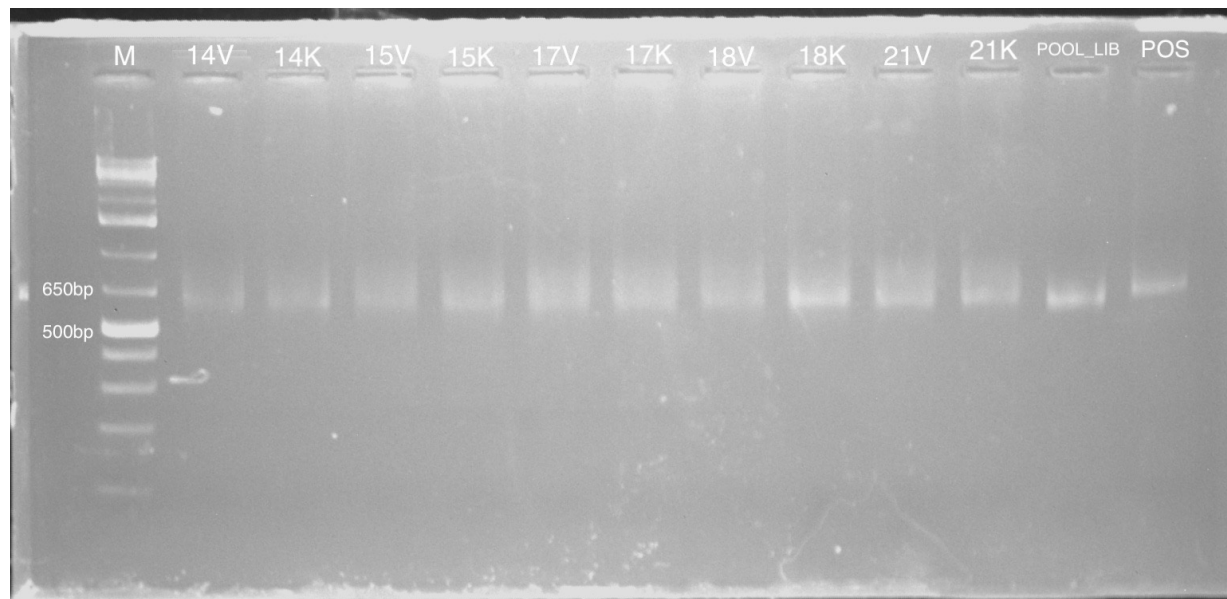- Sử dụng bộ kít có khả năng deplete host DNA.
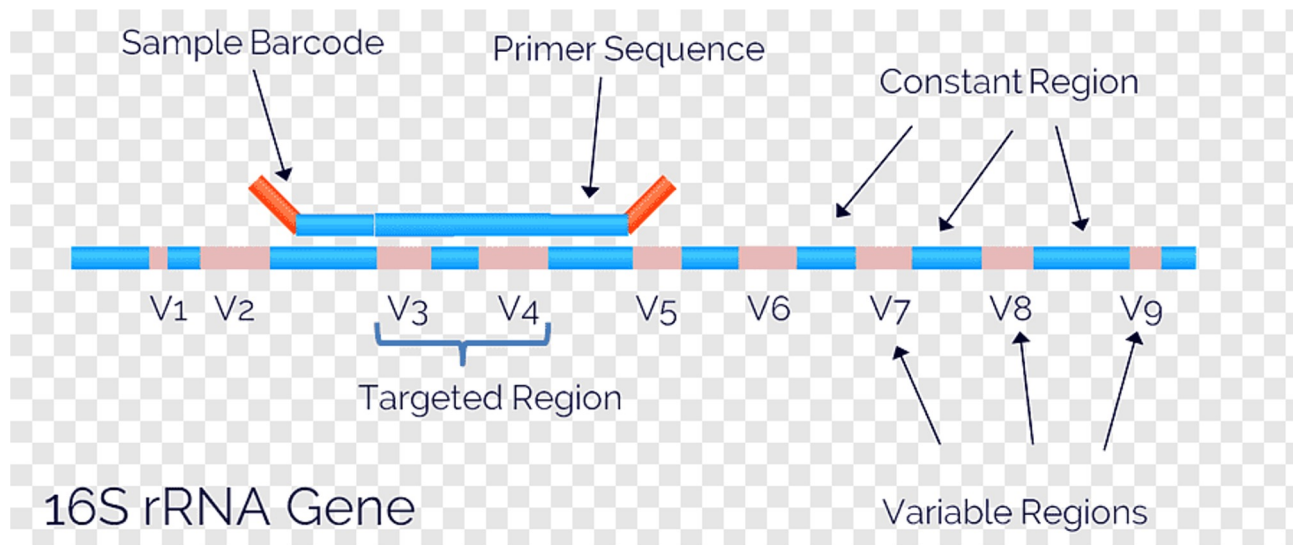
Chuẩn bị thư viện giải trình tự

Quick-16S™ Plus NGS Library Prep Kit (V3-V4)
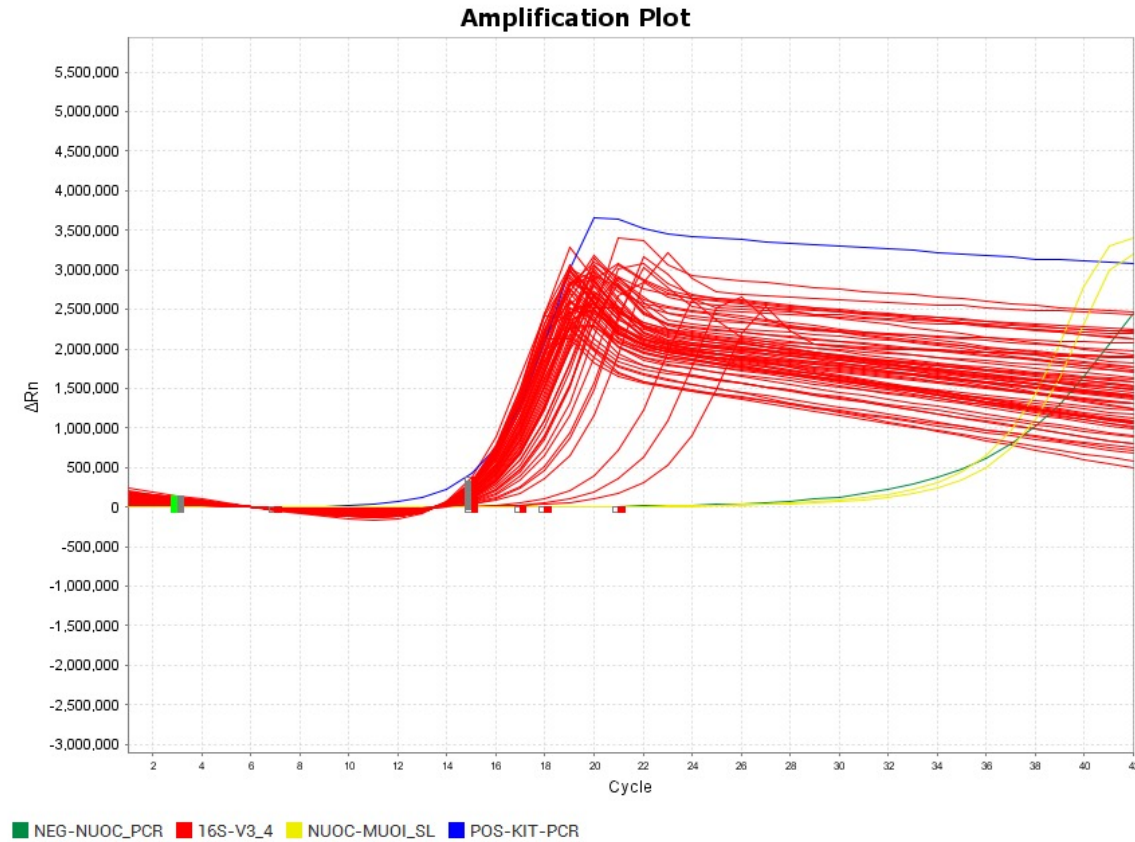Fastest, normalization-free 16S library prep

16S rRNA Gene

Pooling library and run MiSeq

Set positive sample Ct equal 15

Ct: Volume (µl) = 0.000598x^2 - 0.0637x + 3.178; x = Ct.

Amplification Plot

Legend: NEG-NUOC_PCR, 16S-V3_4, NUOC-MUOI_SL, POS-KIT-PCR

Illumina MiSeq® Setup:

The MiSeq® Reagent Kit v3 (600-cycle) and a final library loading concentration of 10-12pM^2 with 15% PhiX spike-in is recommended.

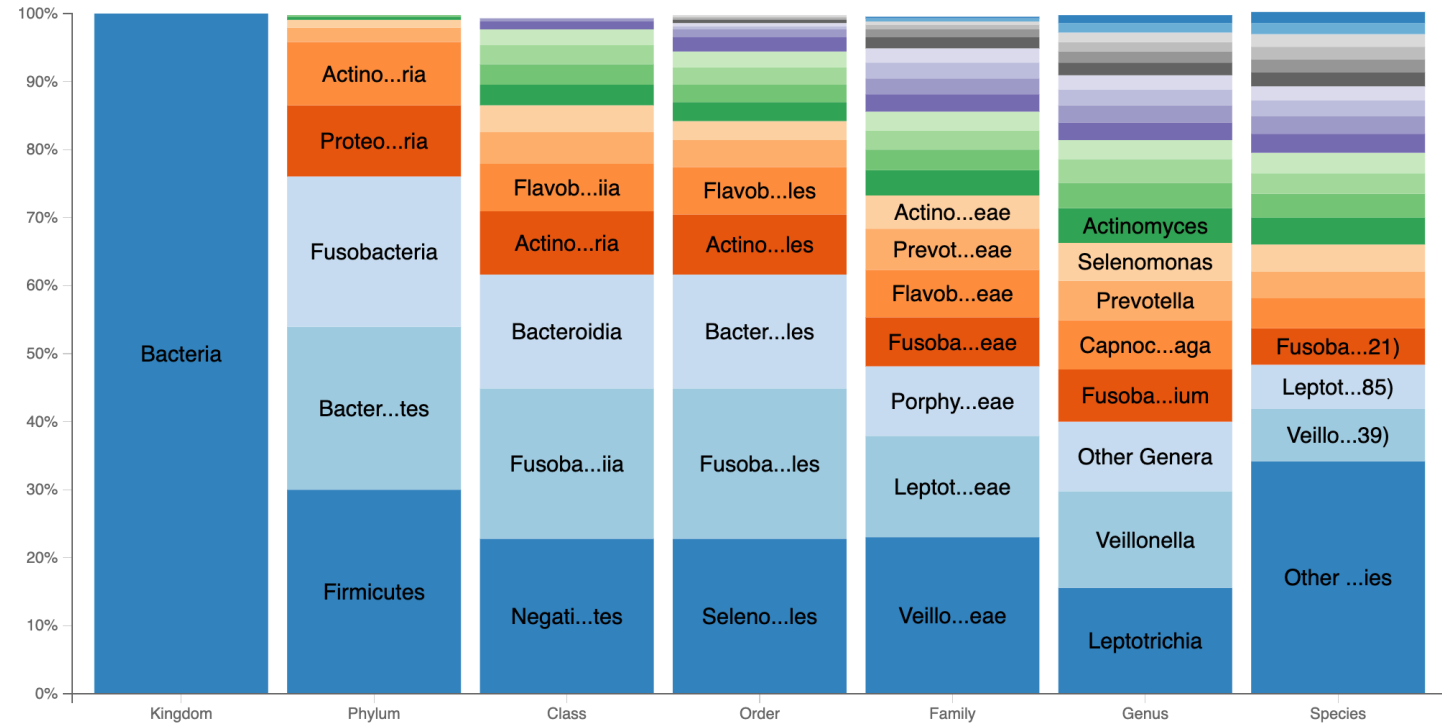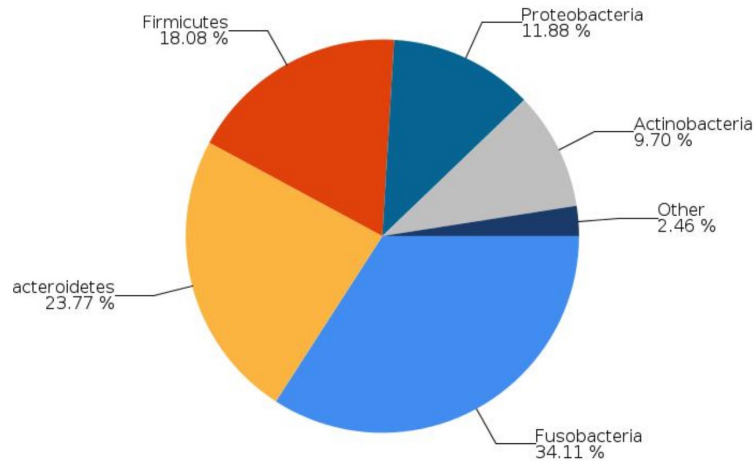# Basespace Analysis

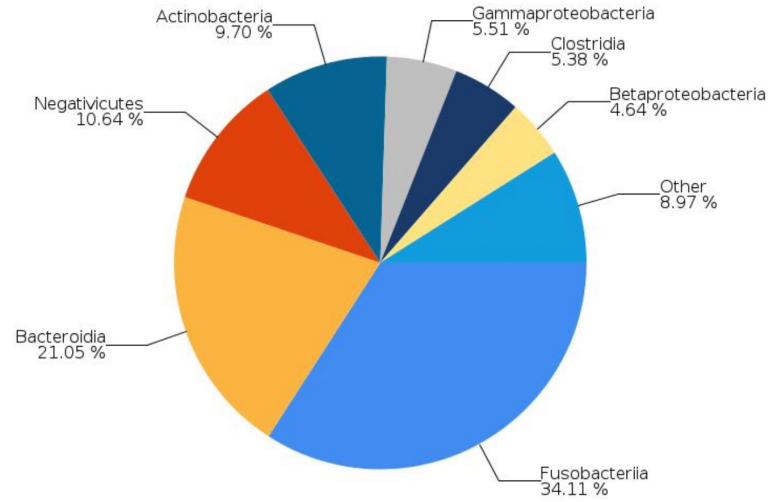**TOP 20 CLASSIFICATION RESULTS BY TAXONOMIC LEVEL**

This column chart shows the relative abundance of the top 20 classification results within each taxonomic level. Mouse over any category to see its description and abundance.
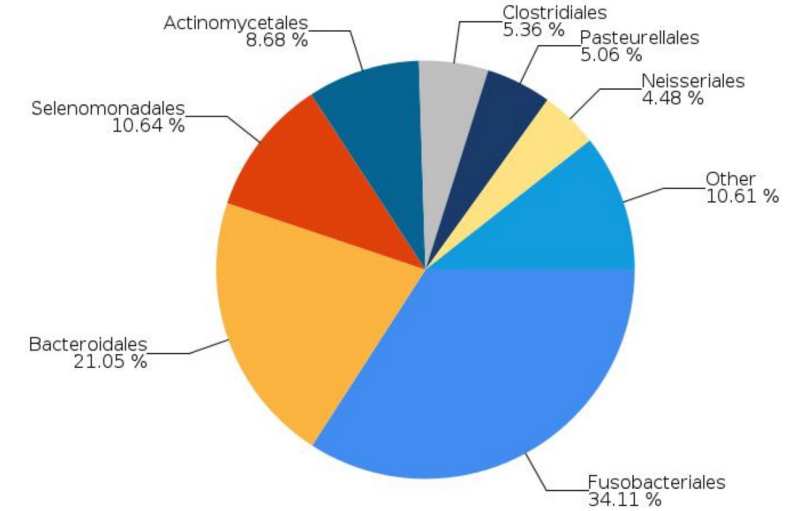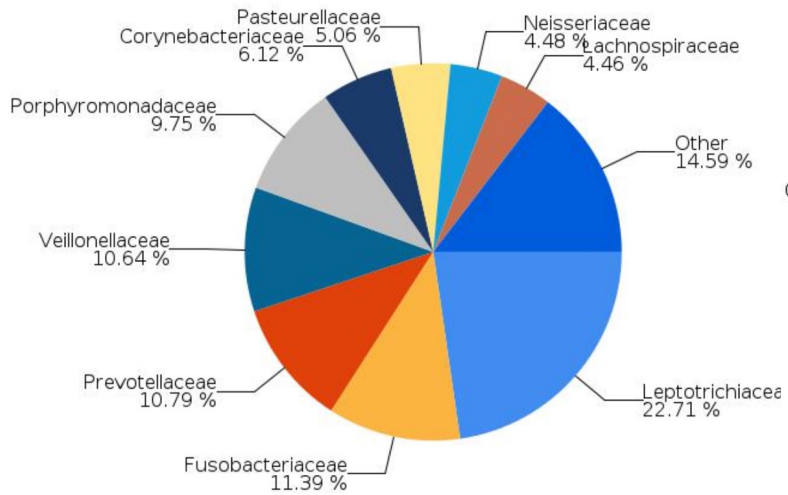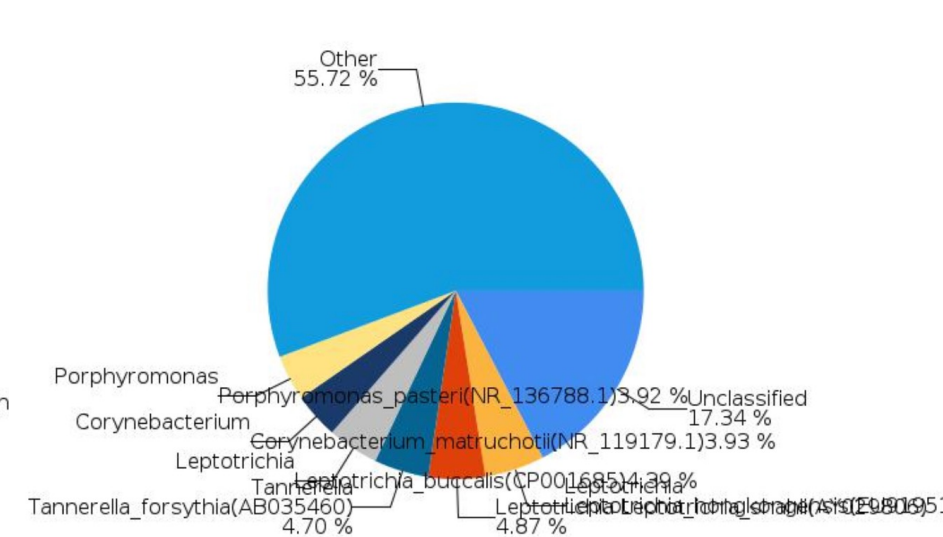
Top Phylum Classification Results

- Proteobacteria 11.88 %
- Firmicutes 18.08 %
- Actinobacteria 9.70 %
- Other 2.46 %
- acteroidetes 23.77 %
- Fusobacteria 34.11 %

Top Class Classification Results

- Actinobacteria 9.70 %
- Gammaproteobacteria 5.51 %
- Clostridia 5.38 %
- Betaproteobacteria 4.64 %
- Negativicutes 10.64 %
- Other 8.97 %
- Bacteroidia 21.05 %
- Fusobacteriia 34.11 %

Top Order Classification Results

- Actinomycetales 8.68 %
- Clostridiales 5.36 %
- Pasteurellales 5.06 %
- Selenomonadales 10.64 %
- Neisseriales 4.48 %
- Other 10.61 %
- Bacteroidales 21.05 %
- Fusobacteriales 34.11 %

Top Family Classification Results

- Pasteurellaceae 5.06 %
- Corynebacteriaceae 6.12 %
- Neisseriaceae 4.48 %
- Lachnospiraceae 4.46 %
- Porphyromonadaceae 9.75 %
- Other 14.59 %
- Veillonellaceae 10.64 %
- Prevotellaceae 10.79 %
- Leptotrichiaceae 22.71 %
- Fusobacteriaceae 11.39 %

- Neisseria 4.33 %
- Porphyromonas 4.59 %
- Other 29.75 %
- Tannerella 4.70 %
- Corynebacterium 6.12 %
- Selenomonas 6.48 %
- Leptotrich 22.68 %
- Prevotella 9.97 %
- Fusobacterium 11.39 %

Top Species Classification Results

- Other 55.72 %
- Porphyromonas
- Porphyromonas_pasteri(NR_136788.1) 3.92 %
- Unclassified 17.34 %
- Corynebacterium
- Corynebacterium_matruchotii(NR_119179.1) 3.93 %
- Leptotrichia
- Leptotrichia_buccalis(CP001685) 4.39 %
- Tannerella
- Tannerella_forsythia(AB035460)
- Leptotrichia
- 4.70 %
- 4.87 %

https://docs.qiime2.org/2024.2/tutorials/overview/
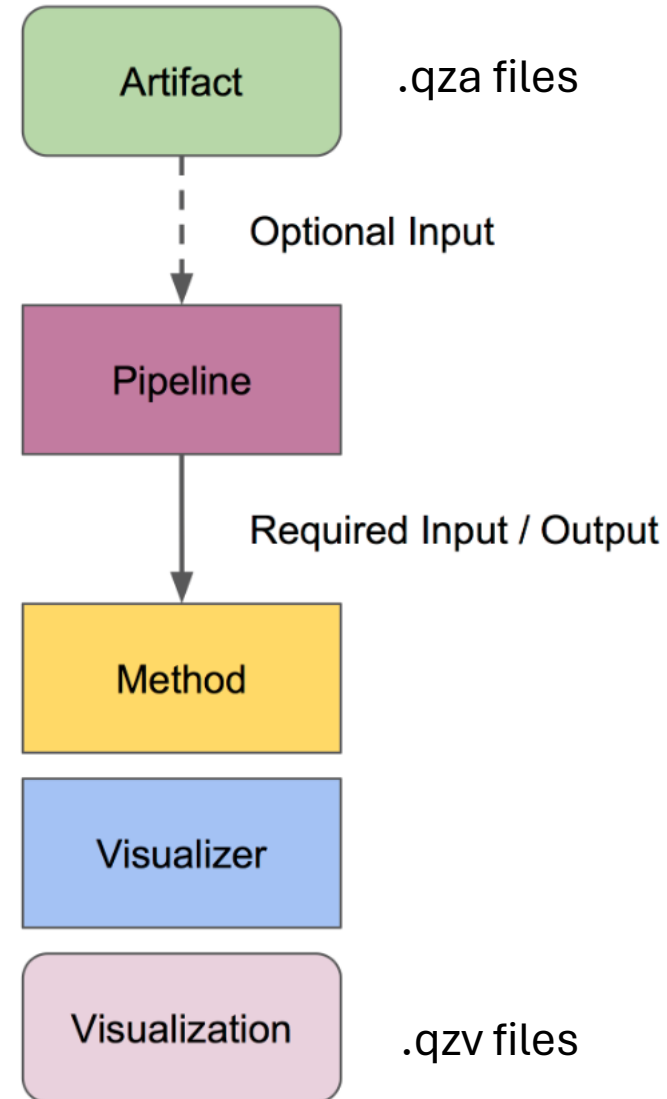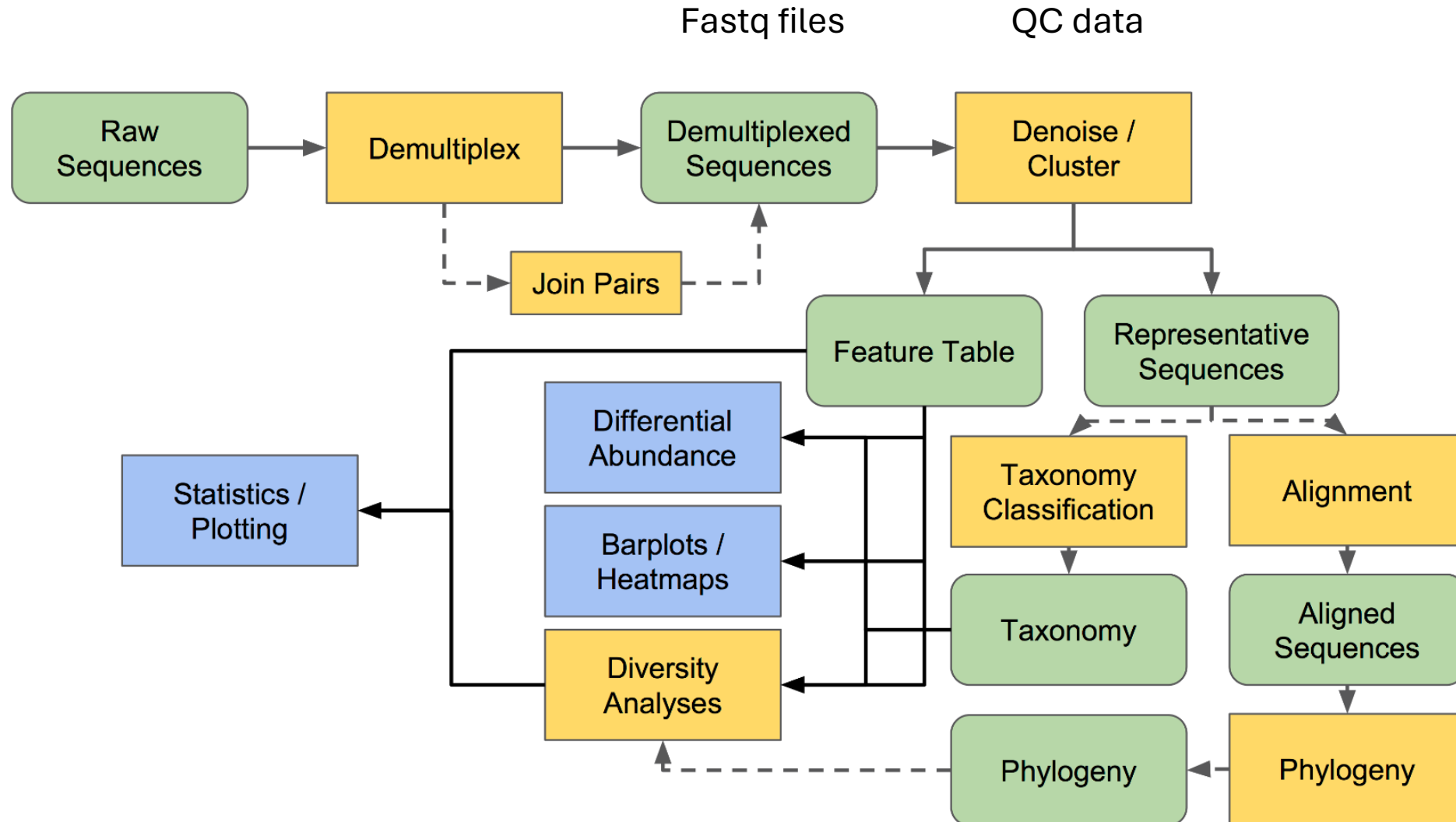
- Run in conda virtual env.
- Command line program.
- Linux, Unix, Window.
- Variable plugins.

General workflow

Visualization

.qzv files

Import sequencing data

```
qiime tools import \
  --type "SampleData[SequencesWithQuality]" \
  --input-format SingleEndFastqManifestPhred33V2 \
  --input-path ./manifest.tsv \
  --output-path ./demux_seqs.qza

qiime demux summarize \
  --i-data ./demux_seqs.qza \
  --o-visualization ./ngoc_demux.qzv

qiime tools view ngoc_demux.qzv
```

Sequencing quality control and feature table construction (Denoised) DADA2

qiime dada2 denoise-paired \
  --i-demultiplexed-seqs demux.qza \
  --p-trim-left-f 7 \
  --p-trim-left-r 6 \
  --p-trunc-len-f 290 \
  --p-trunc-len-r 220 \
  --o-table table.qza \
  --o-representative-sequences rep-seqs.qza \
  --o-denoising-stats denoising-stats.qza


```
qiime tools view ngoc_denosing_stats.qzv
```

Core-metrics-phylogenetic

```
qiime diversity core-metrics-phylogenetic \
 --i-phylogeny phylogeny-align-to-tree-mafft-fasttree/rooted_tree.qza \
  --i-table table.qza \
  --p-sampling-depth 6000 \
  --m-metadata-file metadata.tsv \
  --output-dir diversity-core-metrics-phylogenetic
```

Rarefaction alpha

```
qiime diversity alpha-rarefaction \
  --i-table table.qza \
  --i-phylogeny phylogeny-align-to-tree-mafft-fasttree/rooted_tree.qza \
  --p-max-depth 9000 \
  --m-metadata-file sample-metadata.tsv \
  --o-visualization alpha-rarefaction.qzv
```

Alpha diversity

- Shannon's diversity index (a quantitative measure of community richness)

- Observed Features (a qualitative measure of community richness).

`qiime tools view alpha_rarefaction.qzv`

- Faith's Phylogenetic Diversity (a qualitative measure of community richness that incorporates phylogenetic relationships between the features).

`qiime tools view faith_pd_group_significance.qzv`

Beta diversity

- Jaccard distance (a qualitative measure of community dissimilarity)

- Bray-Curtis distance (a quantitative measure of community dissimilarity)

`qiime tools view bray_curtis_emperor.qzv`

Taxonomic analysis

```
qiime feature-classifier classify-sklearn \
  --i-classifier gg-13-8-99-515-806-nb-classifier.qza \
  --i-reads rep-seqs.qza \
  --o-classification taxonomy.qza
qiime metadata tabulate \
  --m-input-file taxonomy.qza \
  --o-visualization taxonomy.qzv
```
qiime tools view taxonomy.qzv

```
qiime taxa barplot \
  --i-table table.qza \
  --i-taxonomy taxonomy.qza \
  --m-metadata-file sample-metadata.tsv \
  --o-visualization taxa-bar-plots.qzv
```

qiime tools view taxa-bar-plot.qzv

Differential abundance testing with ANCOM-BC

Volcano plot