dunfield_lab_qiime2_pipeline_e_pdf

May 28, 2019

1 qiime2 v.2019.4 dunfield lab tutorial

- Relative paths are used
- All commands imply terminal is in the current working directory
- Knowledge of linux terminal is highly incouraged

1.1 Activate qiime2 environment

[ ]:
```
conda activate qiime2-2019.4
```

1.2 Optionally enable qiime-specific autocompletion

[ ]:
```
source tab-qiime
```

1.3 Navigate to your working directory

- qiime2_lab_tutorial already folder contains raw data
- lab_pipeline folder contains trained NaiveBayes Classifier
- mapping file

Next step expects that the folder exists (example)

[ ]:
```
cd ~/Desktop/qiime2_lab_tutorial/lab_pipeline
```

1.4 Let's view the directory content

[57]:
```
ls -l
```

```
total 216584
-rwxrwxrwx 1 root root     50718 May 27 23:39
dunfield_lab_qiime2_pipeline_e.ipynb
-rwxrwxrwx 1 root root     70275 May 27 23:35 dunfield_lab_qiime2_pipeline.ipynb
-rwxrwxrwx 1 root root      2580 Mar 25 12:39 mappingfile_upd4.csv
-rwxrwxrwx 1 root root 221649684 May 18 09:59
v3v4_silva132_classifier_wps2_2groups.qza
```

1.5 Trim primers

- Various primer trimming tools exist
  - cutadapt
  - bbmap
  - qiime2 native tools
  - trimmomatic
  - manual removal...

It is critical to remove non-biological sequences from the data. We will remove our 16S V3-V4 region (Bacteria-specific primer set) primers sequences using cutadapt * f-primer CCTACGGGNGGCWGCAG * r-primer GACTACHVGGGTATCTAATCC

[ ]:

1.5.1 Making directories

[58]:
```
mkdir primer_trimmed_fastqs; mkdir cutadapt_logs
```

1.5.2 Primer trimming w cutadapt with a help of a little script :)

! Expects to contain our data in the raw_data folder in a parent directory

- Run the following script in a command line:

```
for file1 in ../raw_data/*_R1_*.fastq.gz; do
    file2="${file1%_R1_001.fastq.gz}_R2_001.fastq.gz"

    fname1=`basename $file1`
    fname2=`basename $file2`
    `cutadapt --pair-filter any -j 4 --no-indels --discard-untrimmed \
    -g CCTACGGGNGGCWGCAG -G GACTACHVGGGTATCTAATCC \
    -o primer_trimmed_fastqs/$fname1 -p primer_trimmed_fastqs/$fname2 \
    $file1 $file2 \
    > cutadapt_logs/${fname1}_cutadapt_log.txt`
done
```

1.6 Import trimmed FASTQs as a QIIME2 artifact

To keep the directory clean you can put the artifact files in a new directory

[ ]:
```
mkdir paired_reads_qza
```

1.6.1 Casava 1.8 single-end demultiplexed fastq

Format description

In the Casava 1.8 demultiplexed (single-end) format, there is one fastq.gz file for each sample in the study which contains the single-end reads for that sample. The file name includes the sample identifier and should look like L2S357_15_L001_R1_001.fastq.gz. The underscore-separated fields in this file name are:

```
the sample identifier,
the barcode sequence or a barcode identifier,
the lane number,
the direction of the read (i.e. only R1, because these are single-end reads), and
the set number.
```

Obtaining example data

1.6.2 Importing...

[61]:
```
qiime tools import --type SampleData[PairedEndSequencesWithQuality] \
                   --input-path primer_trimmed_fastqs \
                   --output-path paired_reads_qza/reads_trimmed.qza \
                   --input-format CasavaOneEightSingleLanePerSampleDirFmt
```

```
Imported primer_trimmed_fastqs as CasavaOneEightSingleLanePerSampleDirFmt
to paired_reads_qza/reads_trimmed.qza
```

- Our reads are now ready to be used by qiime2

1.7 Quality control w/ deblur:

Currently deblur doesn't support paired-end reads ### Using VSEARCH for joining:

[62]:
```
qiime vsearch join-pairs \
--i-demultiplexed-seqs paired_reads_qza/reads_trimmed.qza \
--o-joined-sequences paired_reads_qza/reads_trimmed_joined.qza
```

```
Saved SampleData[JoinedSequencesWithQuality] to:
paired_reads_qza/reads_trimmed_joined.qza
```

1.7.1 Filter out low-quality reads.

This command will filter out low-quality reads based on the default options. (this step may take a while)

[63]:
```
qiime quality-filter q-score-joined \
--i-demux paired_reads_qza/reads_trimmed_joined.qza \
--o-filter-stats filt_stats.qza \
--o-filtered-sequences paired_reads_qza/reads_trimmed_joined_filt.qza
```

```
Saved SampleData[JoinedSequencesWithQuality] to:
paired_reads_qza/reads_trimmed_joined_filt.qza
Saved QualityFilterStats to: filt_stats.qza
```

1.7.2 Deblur Workflow

This workflow is 16S sequences, for other amplicon regions, you can use the denoise-other option in the command and specify a reference database.

Note that you will need to trim all sequences to the same length with the –p-trim-length option. In order to determine the correct length to trim down to, run the following QC:

1.7.3 To find appropriate deblur parameters we need to summarize our joined reads

[64]:
```
qiime demux summarize \
--i-data paired_reads_qza/reads_trimmed_joined_filt.qza \
--o-visualization reads_trimmed_joined_filt_summary.qzv
```

```
Saved Visualization to: reads_trimmed_joined_filt_summary.qzv
```

1.7.4 View the obtained visualization

[65]:
```
qiime tools view reads_trimmed_joined_filt_summary.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[20069:20069:0527/234619.288113:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[20029:20050:0527/234619.290330:ERROR:browser_process_sub_thread.cc(209)] Waited
4 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.7.5 Qiime help on importing/exporting/viewing artefacts

[66]:
```
qiime tools --help
```

```
Usage: qiime tools [OPTIONS] COMMAND [ARGS]...

  Tools for working with QIIME 2 files.

Options:
  --help      Show this message and exit.

Commands:
  citations         Print citations for a QIIME 2 result.
  export            Export data from a QIIME 2 Artifact or a
Visualization
  extract           Extract a QIIME 2 Artifact or Visualization
archive.
  import            Import data into a new QIIME 2 Artifact.
  inspect-metadata  Inspect columns available in metadata.
  peek              Take a peek at a QIIME 2 Artifact or
Visualization.
  validate          Validate data in a QIIME 2 Artifact.
  view              View a QIIME 2 Visualization.
```

1.7.6 Explore provenance w/ https://view.qiime2.org

Showing on denoise-16S

[67]:
```
qiime deblur denoise-16S --help
```

```
Usage: qiime deblur denoise-16S [OPTIONS]

  Perform sequence quality control for Illumina data using the Deblur
  workflow with a 16S reference as a positive filter. Only forward reads are
  supported at this time. The specific reference used is the 88% OTUs from
  Greengenes 13_8. This mode of operation should only be used when data were
  generated from a 16S amplicon protocol on an Illumina platform. The
  reference is only used to assess whether each sequence is likely to be 16S
  by a local alignment using SortMeRNA with a permissive e-value; the
  reference is not used to characterize the sequences.

Inputs:
  --i-demultiplexed-seqs ARTIFACT
SampleData[SequencesWithQuality |
    PairedEndSequencesWithQuality | JoinedSequencesWithQuality]
                         The demultiplexed sequences to be denoised.
[required]
Parameters:
  --p-trim-length INTEGER
                         Sequence trim length, specify -1 to disable
                         trimming.
[required]
  --p-left-trim-len INTEGER
    Range(0, None)       Sequence trimming from the 5' end. A value of
0 will
                         disable this trim.                      [default:
0]
  --p-sample-stats / --p-no-sample-stats
                         If true, gather stats per sample.    [default:
False]
  --p-mean-error NUMBER  The mean per nucleotide error, used for
original
                         sequence estimate.                  [default:
0.005]
  --p-indel-prob NUMBER  Insertion/deletion (indel) probability
(same for N
                         indels).                             [default:
0.01]
  --p-indel-max INTEGER  Maximum number of insertion/deletions.
[default: 3]
  --p-min-reads INTEGER  Retain only features appearing at least
min-reads
```

```
                          times across all samples in the resulting feature
                          table.                                    [default:
10]
  --p-min-size INTEGER    In each sample, discard all features with
an
                          abundance less than min-size.
[default: 2]
  --p-jobs-to-start INTEGER
                          Number of jobs to start (if to run in parallel).
                                                                    [default:
1]
  --p-hashed-feature-ids / --p-no-hashed-feature-ids
                          If true, hash the feature IDs.          [default:
True]
Outputs:
  --o-table ARTIFACT FeatureTable[Frequency]
                          The resulting denoised feature table.
[required]
  --o-representative-sequences ARTIFACT
FeatureData[Sequence]
                          The resulting feature sequences.
[required]
  --o-stats ARTIFACT      Per-sample stats if requested.
    DeblurStats
[required]
Miscellaneous:
  --output-dir PATH       Output unspecified results to a directory
  --verbose / --quiet     Display verbose output to
stdout and/or stderr
                          during execution of this action. Or silence output if
                          execution is successful (silence is golden).
  --citations             Show citations and exit.
  --help                  Show this message and exit.
```

1.7.7 Denoising w/ deblur

- Here I'm using the default behaviour of –p-min-reads = 10
- Reads are trimmed to 402nt which retains is at least 98% of the reads (this step may take a while depending on the size of your data . . .)

[106]:
```
qiime deblur denoise-16S \
--i-demultiplexed-seqs paired_reads_qza/reads_trimmed_joined_filt.qza \
--p-trim-length 402 \
--p-sample-stats \
--p-jobs-to-start 8 \
--p-min-reads 10 \
--output-dir deblur_output
```

```
Saved FeatureTable[Frequency] to: deblur_output/table.qza
Saved FeatureData[Sequence] to:
deblur_output/representative_sequences.qza
Saved DeblurStats to: deblur_output/stats.qza
```

1.7.8 Output is saved in the deblur_output folder

let's summarise our deblur output

[107]:
```
qiime deblur visualize-stats \
  --i-deblur-stats deblur_output/stats.qza \
  --o-visualization deblur_output/deblur-stats.qzv
```

```
Saved Visualization to: deblur_output/deblur-stats.qzv
```

[108]:
```
qiime tools view deblur_output/deblur-stats.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after quitting.[29817:29838:0527/235316.922441:ERRO
R:browser_process_sub_thread.cc(209)] Waited 5 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

[109]:
```
qiime feature-table summarize \
--i-table deblur_output/table.qza \
--o-visualization deblur_output/deblur_table_summary.qzv
```

```
Saved Visualization to: deblur_output/deblur_table_summary.qzv
```

[110]:
```
qiime tools view deblur_output/deblur_table_summary.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after quitting.[30210:30231:0527/235353.890697:ERRO
R:browser_process_sub_thread.cc(209)] Waited 5 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.7.9 Tabulate representative sequences

[111]:
```
qiime feature-table tabulate-seqs \
  --i-data deblur_output/representative_sequences.qza \
  --o-visualization representative_sequences.qzv
```

```
Saved Visualization to: representative_sequences.qzv
```

[112]:
```
qiime tools view representative_sequences.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[30607:30607:0527/235411.374129:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[30568:30589:0527/235411.374893:ERROR:browser_process_sub_thread.cc(209)] Waited
3 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.8 Building phylogeny with FastTree

1.8.1 Making multiple-sequence alignment

We'll need to make a multiple-sequence alignment of the ASVs before running FastTree.

[113]:
```
mkdir fast_tree_out
```

```
mkdir: cannot create directory fast_tree_out: File exists
```

[114]:
```
qiime alignment mafft \
--i-sequences deblur_output/representative_sequences.qza \
--p-n-threads 8 \
--o-alignment fast_tree_out/rep_seqs_mafft.qza
```

```
Saved FeatureData[AlignedSequence] to: fast_tree_out/rep_seqs_mafft.qza
```

1.8.2 Filtering multiple-sequence alignment

Variable positions in the alignment need to be masked before FastTree is run, which can be done with this command:

[115]:
```
qiime alignment mask --i-alignment fast_tree_out/rep_seqs_mafft.qza \
  --o-masked-alignment fast_tree_out/rep_seqs_mafft_masked.qza
```

```
Saved FeatureData[AlignedSequence] to:
fast_tree_out/rep_seqs_mafft_masked.qza
```

1.8.3 Running FastTree

Finally FastTree can be run on this masked multiple-sequence alignment:

[116]:
```
qiime phylogeny fasttree \
--i-alignment fast_tree_out/rep_seqs_mafft_masked.qza \
--p-n-threads 4 \
--o-tree fast_tree_out/rep_seqs_aligned_masked_tree
```

```
Saved Phylogeny[Unrooted] to:
fast_tree_out/rep_seqs_aligned_masked_tree.qza
```

1.8.4 Make a rooted tree

Use midpoint root

[117]:
```
qiime phylogeny midpoint-root \
--i-tree fast_tree_out/rep_seqs_aligned_masked_tree.qza \
--o-rooted-tree fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza
```

```
Saved Phylogeny[Rooted] to:
fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza
```

1.8.5 Generate rarefaction curves

- Useful QC step
- Determine maximum depth for the rarefaction using following (I'm using 8000):

[118]:
```
qiime tools view deblur_output/deblur_table_summary.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[31267:31267:0527/235529.491841:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[31226:31247:0527/235529.502609:ERROR:browser_process_sub_thread.cc(209)] Waited
5 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

[119]:
```
qiime diversity alpha-rarefaction --help
```

```
Usage: qiime diversity alpha-rarefaction [OPTIONS]

  Generate interactive alpha rarefaction curves by computing rarefactions
  between `min_depth` and `max_depth`. The number of intermediate depths to
  compute is controlled by the `steps` parameter, with n `iterations` being
  computed at each rarefaction depth. If sample metadata is provided,
  samples may be grouped based on distinct values within a metadata column.

Inputs:
  --i-table ARTIFACT FeatureTable[Frequency]
                           Feature table to compute rarefaction curves from.
[required]
  --i-phylogeny ARTIFACT  Optional phylogeny for phylogenetic
metrics.
    Phylogeny[Rooted]
[optional]
Parameters:
  --p-max-depth INTEGER   The maximum rarefaction depth. Must be
```

```
greater than
    Range(1, None)          min-depth.
[required]
  --p-metrics TEXT... Choices('margalef', 'simpson',
    'michaelis_menten_fit', 'simpson_e', 'gini_index', 'faith_pd',
    'mcintosh_d', 'mcintosh_e', 'enspie', 'doubles', 'brillouin_d',
    'goods_coverage', 'lladser_pe', 'ace', 'fisher_alpha', 'shannon',
    'dominance', 'heip_e', 'pielou_e', 'chao1', 'robbins', 'menhinick',
    'singles', 'berger_parker_d', 'observed_otus')
                            The metrics to be measured. By default computes
                            observed_otus, shannon, and if phylogeny is
                            provided, faith_pd.
[optional]
  --m-metadata-file METADATA...
    (multiple arguments     The sample metadata.
     will be merged)
[optional]
  --p-min-depth INTEGER     The minimum rarefaction depth.
    Range(1, None)
[default: 1]
  --p-steps INTEGER         The number of rarefaction depths to
include between
    Range(2, None)          min-depth and max-depth.
[default: 10]
  --p-iterations INTEGER    The number of rarefied feature tables to
compute at
    Range(1, None)          each step.
[default: 10]
Outputs:
  --o-visualization VISUALIZATION
[required]
Miscellaneous:
  --output-dir PATH         Output unspecified results to a
directory
  --verbose / --quiet       Display verbose output to
stdout and/or stderr
                            during execution of this action. Or silence output
                            if execution is successful (silence is golden).
  --citations               Show citations and exit.
  --help                    Show this message and exit.
```

[ ]:

[ ]:

[ ]:

[120]:
```
qiime diversity alpha-rarefaction \
--i-table deblur_output/table.qza \
--p-max-depth 8000 \
--p-metrics simpson \
--p-metrics faith_pd \
--p-metrics dominance \
--p-metrics chao1 \
--p-metrics observed_otus \
--p-metrics shannon \
--p-steps 20 \
--i-phylogeny fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza \
--o-visualization rarefaction_curves.qzv
```

```
Saved Visualization to: rarefaction_curves.qzv
```

[121]:
```
qiime tools view rarefaction_curves.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after quitting.[31652:31673:0527/235612.774071:ERRO
R:browser_process_sub_thread.cc(209)] Waited 3 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.8.6 Using metadata w rarefaction

[122]:
```
qiime diversity alpha-rarefaction \
--i-table deblur_output/table.qza \
--p-max-depth 8000 \
--p-steps 20 \
--i-phylogeny fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza \
--m-metadata-file mappingfile_upd4.csv \
--o-visualization rarefaction_metadata_curves.qzv
```

```
Plugin error from diversity:
  The following IDs are not present in the metadata: 'nm1-9a', 'o1', 'o29'
Debug info has been saved to /tmp/qiime2-q2cli-err-6vb4gl4x.log
```

1.8.7 Ups! We seem to have an error...

- We need to remove those samples from FeatureTable
- Also, we have 2 samples that contain no data, let's remove them as well

[123]:
```
echo SampleID > samples-to-exclude.tsv
```

[ ]:
```
echo nm1-9a >> samples-to-exclude.tsv
```

[ ]:
```
echo o1 >> samples-to-exclude.tsv
```

[ ]:
```
echo o29 >> samples-to-exclude.tsv
```

[ ]:
```
echo o20 >> samples-to-exclude.tsv
```

[ ]:
```
echo o7 >> samples-to-exclude.tsv
```

1.8.8 Filtering out samples

[124]:
```
qiime feature-table filter-samples \
  --p-exclude-ids \
  --i-table deblur_output/table.qza \
  --m-metadata-file samples-to-exclude.tsv \
  --o-filtered-table id-filtered-deblur-table.qza
```

```
Saved FeatureTable[Frequency] to: id-filtered-deblur-table.qza
```

1.8.9 Let's run it again

- Pay attention that we are supplying updated FeatureTable as its –i-table argument

[125]:
```
qiime diversity alpha-rarefaction \
--i-table id-filtered-deblur-table.qza \
--p-max-depth 8000 \
--p-steps 20 \
--i-phylogeny fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza \
--m-metadata-file mappingfile_upd4.csv \
--o-visualization rarefaction_metadata_curves.qzv
```

```
Saved Visualization to: rarefaction_metadata_curves.qzv
```

[126]:
```
qiime tools view rarefaction_metadata_curves.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[32156:32156:0527/235720.460100:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[32116:32137:0527/235720.466302:ERROR:browser_process_sub_thread.cc(209)] Waited
5 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.8.10 Optional step to re-summarize our filtered FeatureTable

[127]:
```
qiime feature-table summarize \
  --i-table id-filtered-deblur-table.qza \
  --o-visualization id-filtered-deblur-table.qzv \
  --m-sample-metadata-file mappingfile_upd4.csv
```

```
Saved Visualization to: id-filtered-deblur-table.qzv
```

[128]:
```
qiime tools view id-filtered-deblur-table.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[32516:32516:0527/235731.148830:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[32474:32495:0527/235731.158822:ERROR:browser_process_sub_thread.cc(209)] Waited
3 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.9 Assign taxonomy

- Could be assigned to ASVs using a Naive-Bayes classifier
- This classifier was trained using SILVA 132 database and is specific for v3v4 region
- Contains edits for WPS-2 (Rubrimentifilales and AS-11)
- Could be trained de novo, but RAM intensive
- Qiime version sensitive

(this step may take a long time to complete . . .)

[129]:
```
qiime feature-classifier classify-sklearn \
--i-reads deblur_output/representative_sequences.qza \
--i-classifier v3v4_silva132_classifier_wps2_2groups.qza \
--output-dir taxonomy
```

```
Saved FeatureData[Taxonomy] to: taxonomy/classification.qza
```

1.9.1 Our taxonomy folder now contains classification.qza file

let's explore the results..

Following command export the classification as a tsv-file

[130]:
```
qiime tools export \
--input-path taxonomy/classification.qza \
--output-path taxonomy
```

```
Exported taxonomy/classification.qza as TSVTaxonomyDirectoryFormat to
directory taxonomy
```

1.9.2 At last..., Our Beloved Bar-Chart :)

[131]:
```
qiime taxa barplot \
--i-table id-filtered-deblur-table.qza \
--i-taxonomy taxonomy/classification.qza \
--m-metadata-file mappingfile_upd4.csv \
--o-visualization taxonomy/taxa_barplot.qzv
```

```
Saved Visualization to: taxonomy/taxa_barplot.qzv
```

[132]:
```
qiime tools view taxonomy/taxa_barplot.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[6326:6375:0528/003825.571767:ERROR:browser_process_sub_thread.cc(209)]
Waited 3 ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.10 Finally, let's calculate core diversity metrics

- For this step we need to select a reasonable rarefaction value
- Let's have a look at our FeatureTable again

[133]:
```
qiime tools view id-filtered-deblur-table.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[6800:6800:0528/003836.431192:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[6759:6780:0528/003836.436214:ERROR:browser_process_sub_thread.cc(209)] Waited 5
ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

400 seems to be a good number in this case, we are loosing only 2 samples

[134]:
```
qiime diversity core-metrics-phylogenetic \
--i-phylogeny fast_tree_out/rep_seqs_mafft_masked_tree_rooted.qza \
--i-table id-filtered-deblur-table.qza \
--p-sampling-depth 400 \
--m-metadata-file mappingfile_upd4.csv \
--output-dir core-metrics
```

```
Saved FeatureTable[Frequency] to: core-metrics/rarefied_table.qza
Saved SampleData[AlphaDiversity] % Properties('phylogenetic') to: core-
metrics/faith_pd_vector.qza
Saved SampleData[AlphaDiversity] to: core-
metrics/observed_otus_vector.qza
Saved SampleData[AlphaDiversity] to: core-metrics/shannon_vector.qza
Saved SampleData[AlphaDiversity] to: core-metrics/evenness_vector.qza
Saved DistanceMatrix % Properties('phylogenetic') to: core-
metrics/unweighted_unifrac_distance_matrix.qza
Saved DistanceMatrix % Properties('phylogenetic') to: core-
metrics/weighted_unifrac_distance_matrix.qza
Saved DistanceMatrix to: core-metrics/jaccard_distance_matrix.qza
Saved DistanceMatrix to: core-metrics/bray_curtis_distance_matrix.qza
Saved PCoAResults to: core-metrics/unweighted_unifrac_pcoa_results.qza
Saved PCoAResults to: core-metrics/weighted_unifrac_pcoa_results.qza
Saved PCoAResults to: core-metrics/jaccard_pcoa_results.qza
Saved PCoAResults to: core-metrics/bray_curtis_pcoa_results.qza
Saved Visualization to: core-metrics/unweighted_unifrac_emperor.qzv
Saved Visualization to: core-metrics/weighted_unifrac_emperor.qzv
Saved Visualization to: core-metrics/jaccard_emperor.qzv
Saved Visualization to: core-metrics/bray_curtis_emperor.qzv
```

let's view an ordination plot

[135]:
```
qiime tools view core-metrics/weighted_unifrac_emperor.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[7169:7226:0528/003856.615757:ERROR:browser_process_sub_thread.cc(209)]
Waited 3 ms for network service
[7278:7278:0528/003856.616153:ERROR:sandbox_linux.cc(364)] InitializeSandbox()
called with multiple threads in process gpu-process.
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.10.1 Alpha diversity group significance test

- An example of just one metric

[136]:
```
qiime diversity alpha-group-significance \
--i-alpha-diversity core-metrics/faith_pd_vector.qza \
--m-metadata-file mappingfile_upd4.csv \
--o-visualization core-metrics/faith-pd-group-significance.qzv
```

```
Saved Visualization to: core-metrics/faith-pd-group-significance.qzv
```

[137]:
```
qiime tools view core-metrics/faith-pd-group-significance.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[7683:7683:0528/003911.628849:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[7643:7664:0528/003911.631947:ERROR:browser_process_sub_thread.cc(209)] Waited 5
ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.10.2 Beta diversity group significance test

- lets test weighted unifrac

[138]:
```
qiime diversity beta-group-significance \
--i-distance-matrix core-metrics/weighted_unifrac_distance_matrix.qza \
--m-metadata-file mappingfile_upd4.csv \
--m-metadata-column Location \
--p-pairwise \
--o-visualization core-metrics/unweighted-unifrac-bodysite-significance.qzv
```

```
Saved Visualization to: core-metrics/unweighted-unifrac-bodysite-
significance.qzv
```

[139]:
```
qiime tools view core-metrics/unweighted-unifrac-bodysite-significance.qzv
```

```
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
accessible or work correctly after
quitting.[8054:8054:0528/003925.363841:ERROR:sandbox_linux.cc(364)]
InitializeSandbox() called with multiple threads in process gpu-process.
[8012:8033:0528/003925.379743:ERROR:browser_process_sub_thread.cc(209)] Waited 5
ms for network service
Opening in existing browser session.

Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be
```

1.11 Bonus part: Exporting FeatureTables (biom files)

- qiime2 keeps taxonomy separately
- therefore exporting biom files with taxonomy needs some additional steps

Here we are slightly modifying the header of the taxonomy file

[140]:
```
sed -i -e '1 s/Feature/#OTUID/' -e '1 s/Taxon/taxonomy/' taxonomy/taxonomy.tsv
```

1.11.1 Let's export the biom file into the folder and convert to tsv there

[141]:
```
qiime tools export \
--input-path id-filtered-deblur-table.qza \
--output-path id-filtered-deblur-table-exported
```

```
Exported id-filtered-deblur-table.qza as BIOMV210DirFmt to directory id-
filtered-deblur-table-exported
```

[142]:
```
biom add-metadata \
-i id-filtered-deblur-table-exported/feature-table.biom \
-o id-filtered-deblur-table-exported/feature-table_w_tax.biom \
--observation-metadata-fp taxonomy/taxonomy.tsv \
--sc-separated taxonomy
```

1.11.2 And finally a familiar biom-convert :)

[143]:
```
biom convert \
-i id-filtered-deblur-table-exported/feature-table_w_tax.biom \
-o id-filtered-deblur-table-exported/feature-table.tsv \
--to-tsv --header-key taxonomy
```

1.12 Good-bye!