

Beginner Microbiome Analysis 2025

Contents

I	Introduction	5
1	Welcome	7
2	Course Schedule	9
2.1	Meet Your Faculty	9
II	Modules	11
3	Module 1	13
4	Module 3: Microbiome statistics and visualizations	15
4.1	Table of Contents	15
4.2	Introduction	16
4.3	1. Build tree with SEPP QIIME 2 plugin	17
4.4	2. Generate rarefaction curves	18
4.5	3. Calculating diversity metrics and generating ordination plots .	18
4.6	4. Generate stacked bar chart of taxa relative abundances	20
4.7	5. Identifying differentially abundant features with ANCOM . . .	21
4.8	6. Identifying shared features across groups of samples	23
4.9	7. Heatmaps	23
4.10	8. Exporting data from QIIME 2 for use in other software	24
4.11	9. Considerations for 18S and ITS data	25
4.12	Answers	27

Part I

Introduction

Chapter 1

Welcome

Welcome to CBW's Beginner Microbiome Analysis Workshop!

Chapter 2

Course Schedule

Time (PST)	Day 1	Time (PST)	Day 2
8:30	Arrivals & Check-in	8:30	Arrivals
9:00	Welcome (Coordinator)	9:00	Module 3 Lecture
9:30	Module 1 Lecture		
10:30	Break (30min)	10:00	Break (30min)
11:00	Module 1 Lab	10:30	Module 3 Lab
13:00	Lunch (1h)	12:30	Class photo + Lunch (1h)
14:00	Module 2 Lecture	13:30	Module 4 Lecture
15:00	Break (30min)	14:30	Break (30min)
15:30	Module 2 Lab	15:00	Module 4 Lab
		17:00	Survey and Closing
			Remarks
17:30	Finished	17:30	Finished

2.1 Meet Your Faculty

Coming soon!

Part II

Modules

Chapter 3

Module 1

Coming soon!

Chapter 4

Module 3: Microbiome statistics and visualizations

Material by Juan Santana

This tutorial is part of the 2025 Canadian Bioinformatics Workshops Beginner Microbiome Analysis (Vancouver, BC, May 26-27). It is based on the Amplicon SOP v2 available on the Microbiome Helper repository and previous workshops designed by Diana Haider, Robert Beiko and Monica Alvaro Fuss.

4.1 Table of Contents

1. Build tree
2. Generate rarefaction curves
3. Calculating diversity metrics and generating ordination plots
4. Generate stacked bar chart of taxa relative abundances
5. Identifying differentially abundant features with ANCOM
6. Core microbiome analysis
7. Heatmaps
8. Exporting data from QIIME 2 for use in other software
9. Considerations for 18S and ITS data

4.2 Introduction

Modules 2 and 3 offer a step-by-step walkthrough of an end-to-end pipeline for analyzing high-throughput marker gene data using the command-line interface. Commonly used marker genes in microbiome research include 16S ribosomal RNA (rRNA) for prokaryotes, 18S rRNA for eukaryotes, and the internal transcribed spacer (ITS) region for fungi.

In this tutorial, we primarily focus on the same 16S rRNA dataset derived from wild blueberry (*Vaccinium angustifolium*) soil communities from natural and managed habitats we processed in Module 2. At the end of the tutorial, we also provide guidance on processing the 18S rRNA dataset from plastics incubated in a coastal marine environment (plastisphere), as well as the ITS dataset from the gut mycobiome (i.e., fungi) during pregnancy.

You can learn more about these studies in the following publications:

- Variation in Bacterial and Eukaryotic Communities Associated with Natural and Managed Wild Blueberry Habitats
- Metagenomic Functional Shifts to Plant Induced Environmental Changes -Microbial pioneers of plastic colonisation in coastal seawaters
- Landscape of the gut mycobiome dynamics during pregnancy and its relationship with host metabolism and pregnancy health

In Module 2, we covered the fundamentals of marker gene analysis, from processing raw reads to generating a filtered feature table. In Module 3, we will explore examples of downstream analyses used to draw biological insights from these data. The workflow described is integrated into the latest release of QIIME 2 (Quantitative Insights into Microbial Ecology, version 2025.4). As introduced in Module 2, this widely used microbiome bioinformatics platform is built around user-developed software packages called plugins, which operate on QIIME 2 artifact files (with the .qza extension). Documentation for these plugins—including tutorials and additional resources—can be found in the QIIME 2 user documentation. QIIME 2 also offers interpretable visualizations, which can be viewed by opening .qzv files in QIIME2 View

Let's set up our Module 3 directory inside **workspace** and create a symlink to the data we will be using (from Module 2).

```
cd ~/workspace
mkdir bmb_module3 bmb_module3/16S_analysis
cd bmb_module3/16S_analysis
ln -s ~/CourseData/MIC_data/bmb_module2/output/16S_analysis/deblur_output .
ln -s ~/CourseData/MIC_data/bmb_module2/output/16S_analysis/taxa .
ln -s ~/CourseData/MIC_data/bmb_module2/output/16S_analysis/Blueberry_16S_metadata.tsv
```


If you deactivated your QIIME2 environment, reactivate it with the command below.

```
conda activate qiime2-amplicon-2025.4
```

When you are finished this tutorial you can deactivate the conda environment using:

```
conda deactivate
```

Throughout this module, there are some questions aimed to help your understanding of some of the key concepts. You'll find the answers at the bottom of this page, but no one will be marking them.

4.3 1. Build tree with SEPP QIIME 2 plugin

SEPP (SATé-enabled Phylogenetic Placement) is a tool used to place short DNA sequences—such as 16S rRNA amplicon sequence variants (ASVs)—into an existing, high-quality reference phylogenetic tree. This is particularly helpful when you are working with microbiome data and want to infer evolutionary relationships more accurately. We will use QIIME 2's `q2-fragment-insertion` plugin to place ASVs derived from our **16S data** into a reference phylogenetic tree using the command below.

```
qiime fragment-insertion sepp
  --i-representative-sequences deblur_output/rep_seqs_final.qza \
  --i-reference-database ~/CourseData/MIC_data/bmb_module3/16S_analysis/sepp-refs-gg-13-8.qza \
  --o-tree asvs-tree.qza \
  --o-placements insertion-placements.qza \
  --p-threads 4
```

Due to memory constraints, you can instead copy the output into your folder with the following command:

```
ln -s ~/CourseData/MIC_data/bmb_module3/output/16S_analysis/asvs-tree.qza .
ln -s ~/CourseData/MIC_data/bmb_module3/output/16S_analysis/insertion-placements.qza .
```

High-quality reference phylogenetic trees can be downloaded from QIIME2's data resources. However, it is important to ensure that the reference tree

used for sequence placement matches the same reference database used for taxonomic classification (e.g., Greengenes or SILVA). In our workflow, we use `sepp-refs-gg-13-8.qza`, as our sequences were classified using the Greengenes database. Alternatively, custom reference files can be specified for placing other types of amplicons. However, for marker genes such as **18S** and **ITS**, the recommended approach is to construct a *de novo* phylogenetic tree, as outlined in section 9. Considerations for 18S and ITS data and further detailed in the Microbiome Helper repository.

4.4 2. Generate rarefaction curves

After inferring ASVs from our sequence reads, we can generate **rarefaction curves** to evaluate whether our sequencing depth was sufficient to capture most of the microbial diversity in each sample. The command below will generate these plots. The `--p-max-depth` parameter sets the maximum sequencing depth to sample across. Inspect the `deblur_table_final_summary.qzv` file (created in Module 2) using QIIME2 View to determine this number. For our 16S dataset, the sample with the highest number of reads contains 8,650 sequences, so we'll use that as the max depth.

```
qiime diversity alpha-rarefaction \
  --i-table deblur_output/deblur_table_final.qza \
  --p-max-depth 8650 \
  --p-steps 20 \
  --i-phylogeny asvs-tree.qza \
  --o-visualization rarefaction_curves.qzv
```

This will produce a `.qzv` file that can be opened in QIIME 2 View to explore the number of observed features (as well as two alpha diversity metrics: Shannon index and Faith's phylogenetic diversity) across rarefied sequencing depths for each sample.

Question 1: What is a good rarefaction depth for diversity analysis?

4.5 3. Calculating diversity metrics and generating ordination plots

QIIME 2 can calculate commonly used alpha and beta diversity metrics—including Faith's Phylogenetic Diversity, Shannon diversity, and UniFrac distances—using a single command. Ordination plots (e.g., PCoA plots based on weighted UniFrac distances) are also generated automatically. Before

4.5. 3. CALCULATING DIVERSITY METRICS AND GENERATING ORDINATION PLOTS19

calculating these metrics, QIIME 2 rarefies all samples to the same sequencing depth to ensure fair comparisons. The `--p-sampling-depth` parameter defines this cutoff. Any samples with fewer reads than this threshold will be excluded from the analysis. Based on the rarefaction curves we generated earlier, we determined that the lowest reasonable sequencing depth across all samples is **3,432** reads. Using this value ensures we retain all samples while still capturing most of the microbial diversity.

```
qiime diversity core-metrics-phylogenetic \  
  --i-table deblur_output/deblur_table_final.qza \  
  --i-phylogeny asvs-tree.qza \  
  --p-sampling-depth 3432 \  
  --m-metadata-file Blueberry_16S_metadata.tsv \  
  --p-n-jobs-or-threads 4 \  
  --output-dir diversity
```

This command will output a `diversity/` folder containing:

Alpha diversity vectors (e.g., Shannon, Faith's PD) Beta diversity distance matrices (e.g., UniFrac, Bray-Curtis) Rarefied feature table PCoA ordination plots (emperor.qzv files)

You can explore these ordination plots to visually assess whether microbial communities cluster based on sample groupings (e.g., natural vs. managed habitats). However, to statistically and visually compare alpha diversity across sample categories (e.g., habitat type), you have to generate boxplots and perform statistical tests using the following command:

```
qiime diversity alpha-group-significance \  
  --i-alpha-diversity diversity/shannon_vector.qza \  
  --m-metadata-file Blueberry_16S_metadata.tsv \  
  --o-visualization diversity/shannon_compare_groups.qzv
```

Hint: To analyze other alpha diversity metrics, simply replace `shannon_vector.qza` with another file from the `diversity/` directory. To view available metrics, run:

```
ls diversity/*_vector.qza
```

Note that you can also export this or any other diversity metric file (ending in `.qza`) using `qiime tools export` and analyze them with a different program (e.g., R, Python, Excel).

```
qiime tools export \
  --input-path diversity/shannon_vector.qza \
  --output-path exported_shannon
```

This will produce a file called `alpha-diversity.tsv` inside the `exported_shannon/` folder.

Question 2: are there any significant differences in alpha diversity between any of our metadata categories?

Question 3: which metadata category appears to provide more separation in the beta diversity PCoA plots?

If you want to run a PERMANOVA test to assess whether beta diversity differs significantly between groups, you can use the following command. Be sure to update the command with the specific beta diversity metric you're analyzing (e.g., Bray-Curtis, UniFrac) and replace category with the appropriate column name from your metadata file that defines your groups of interest.

```
qiime diversity beta-group-significance \
  --i-distance-matrix diversity/bray_curtis_distance_matrix.qza \
  --m-metadata-file Blueberry_16S_metadata.tsv \
  --m-metadata-column Category \
  --o-visualization diversity/bray_curtis_compare_groups_category.qzv
```

This will produce a `.qzv` file containing PERMANOVA results and boxplots for visualizing distances within and between groups.

Question 4: what do you mean I'm getting an error message? It should be working fine, it worked last time! I can't see what I'm doing wro... ooooooh there we go, I see now :)

4.6 4. Generate stacked bar chart of taxa relative abundances

A useful way to visualize the composition of microbial communities across samples is through interactive stacked bar charts of taxonomic relative abundances. This type of plot allows you to explore differences in microbial composition at various taxonomic levels (e.g., phylum, genus) across experimental groups. You can generate the visualization using the command below.

4.7. 5. IDENTIFYING DIFFERENTIALLY ABUNDANT FEATURES WITH ANCOM21

```
qiime taxa barplot \  
  --i-table deblur_output/deblur_table_final.qza \  
  --i-taxonomy taxa/classification.qza \  
  --m-metadata-file Blueberry_16S_metadata.tsv \  
  --o-visualization taxa_barplot.qzv
```

The interactive `.qzv` file generated with this command can be inspected with the QIIME2 viewer. In the plot you can explore dominant taxa across samples or groups, navigate between taxonomic levels (kingdom to species) and compare relative abundances between metadata-defined sample groups.

Hint: your metadata file includes any grouping columns you want to use to color or group the samples in the visualization (e.g., treatment, site, time point).

Question 5: can you identify any patterns between the metadata groups?

4.7 5. Identifying differentially abundant features with ANCOM

ANCOM (Analysis of Composition of Microbiomes) is a statistical method used to identify taxa (or features) that differ significantly in relative abundance between sample groups. ANCOM is robust to the compositional nature of microbiome data and does not assume a particular distribution for the features. However, ANCOM requires that all features in the table be non-zero, so a pseudocount (commonly set to 1) must first be added to avoid issues with zeros in the count matrix. The command below creates a new feature table where a count of 1 is added to all zero entries.

```
qiime composition add-pseudocount \  
  --i-table deblur_output/deblur_table_final.qza \  
  --p-pseudocount 1 \  
  --o-composition-table deblur_output/deblur_table_final_pseudocount.qza
```

We then run ANCOM using the command below. `--m-metadata-column` defines the groups you'd like to compare (e.g., Treatment, Site, Timepoint, etc.) according to the columns in your metadata file.

```
qiime composition ancom \  
  --i-table deblur_output/deblur_table_final_pseudocount.qza \  
  --m-metadata-file Blueberry_16S_metadata.tsv
```

```
--m-metadata-column category \
--output-dir ancom_output
```

The results from ANCOM will be saved in the `ancom_output/` directory. The key file is `ancom.qzv`, which can be viewed interactively at QIIME 2 View. The output highlights features (e.g., ASVs, OTUs, or taxa) that are differentially abundant between the specified groups based on their W statistic (number of significant pairwise comparisons).

Hint: ANCOM is sensitive to sparsity in the data. Consider filtering low-abundance or low-frequency features before running ANCOM to improve detection power and reduce false positives. Alternatively, we can use ANCOM-BC (see below)

ANCOM-BC addresses limitations in the original ANCOM method by incorporating bias correction and estimating absolute abundances from compositional data. It performs better under sparsity and unequal library sizes. Note that the `--p-formula` parameter lets you specify covariates or design models (e.g., Treatment + Timepoint), in our workflow we only test for significant abundances based on the “category” variable. In addition, we supply the original raw counts instead of the compositionally transformed table.

```
qiime composition ancombc \
  --i-table deblur_output/deblur_table_final.qza \
  --m-metadata-file Blueberry_16S_metadata.tsv \
  --p-formula category \
  --output-dir ancombc_output/
```

While `qiime composition ancombc` produces a `.qza` file of differentials (log-fold changes and statistics), QIIME 2 does not currently provide built-in commands to visualize this artifact as a `.qzv`. Therefore, we use another command to generate visualizations.

```
qiime composition da-barplot \
  --i-data ancombc_output/differentials.qza \
  --p-significance-threshold 0.05 \
  --p-effect-size-threshold 1.0 \
  --o-visualization ancombc_output/ancombc_results.qzv
```

The resultant `.qzv` barplot will show you features significantly enriched in one or another condition.

Question 6: Does ANCOM identify any differentially abundant taxa between any of the metadata groups? If so, which one(s)? Is the output from ANCOM-BC any different?

4.8 6. Identifying shared features across groups of samples

The core microbiome refers to the set of microbial taxa consistently found across a group of samples or within specific treatment groups. Identifying core taxa helps understand the stable and potentially important microbes in your system. In QIIME2 we can filter features by prevalence across samples using the command below. `-p-min-fraction 0.8` means features must be present in 80% of samples to be considered “core”.

```
qiime feature-table core-features \
  --i-table deblur_output/deblur_table_final.qza \
  --p-min-fraction 0.8 \
  --o-visualization core_features.qzv
```

This will output a feature table summary with the fraction of the ASVs shared across all samples. The core microbiome plugin `qiime feature-table core-features` doesn't have an option to group samples by treatment, so, if you want to identify the shared ASVs across samples from a group (e.g. managed) you need to subset your feature table first by treatment group using `qiime feature-table filter-samples`.

```
qiime feature-table filter-samples \
  --i-table deblur_output/deblur_table_final.qza \
  --m-metadata-file Blueberry_16S_metadata.tsv \
  --p-where "[category]='Managed'" \
  --o-filtered-table table_managed.qza
```

We can then run `qiime feature-table core-features` on the filtered table `table_managed.qza` to identify shared ASVs across “Managed” samples.

4.9 7. Heatmaps

Heatmaps visualize abundance patterns of taxa across samples or groups, allowing easy spotting of taxa that vary in abundance or co-occur across conditions. We can generate these heatmaps using `qiime feature-table heatmap`. However, raw feature tables at the ASV level often contain thousands of features, which can make heatmaps visually overwhelming and difficult to interpret. To create a clearer and more interpretable heatmap, you can collapse your feature table to a higher taxonomic rank, such as genus or family, before visualization, using the command below. `--p-level` sets the taxonomic rank, here we will collapse at the Class level.

```
qiime taxa collapse \
  --i-table deblur_output/deblur_table_final.qza \
  --i-taxonomy taxa/classification.qza \
  --p-level 3 \
  --o-collapsed-table deblur_output/table_class.qza
```

Now we generate the heatmap using the collapsed class-level table.

```
qiime feature-table heatmap \
  --i-table deblur_output/table_class.qza \
  --m-sample-metadata-file Blueberry_16S_metadata.tsv \
  --m-sample-metadata-column category \
  --p-metric braycurtis \
  --p-color-scheme RdYlBu \
  --o-visualization heatmap_class.qzv
```

Open the resulting `heatmap_class.qzv` in QIIME 2 View to interactively explore taxa abundance patterns across samples.

4.10 8. Exporting data from QIIME 2 for use in other software

While QIIME 2 offers a wide range of tools for microbial community analysis, you may want to perform additional custom analyses in software like R, Python, or MATLAB. To do so, you'll need to export your QIIME 2 artifacts into formats that are compatible with these platforms.

Representative sequences (i.e., ASVs) are stored in a `.qza` artifact that contains the DNA sequences used in downstream analysis. To export them into a standard FASTA format, use the following command:

```
qiime tools export \
  --input-path deblur_output/rep_seqs_final.qza \
  --output-path deblur_output_exported
```

Your sequences will be saved as `dna-sequences.fasta` inside the `deblur_output_16S_exported` folder. This file can be read by any downstream tool that accepts FASTA files.

BIOM (Biological Observation Matrix) is a standardized format for representing feature tables, typically containing: -Rows = features (e.g., ASVs, OTUs, taxa)

-Columns = samples -Cells = abundance values (counts, relative abundances, etc.) -Optional metadata (taxonomy, sample info)

BIOM files are widely used in microbiome analysis and supported by R packages like phyloseq, microbiome, and tools in Python such as scikit-bio and biom-format. To export a BIOM table (with taxonomy added as metadata) you can use the commands below.

```
#First we fix taxonomy header with sed (required for biom add-metadata)
sed -i -e '1 s/Feature/#Feature/' -e '1 s/Taxon/taxonomy/' taxa/taxonomy.tsv

#Second we export the raw feature table and create the biom table
qiime tools export \
  --input-path deblur_output/deblur_table_final.qza \
  --output-path deblur_output_exported

#Third we add taxonomy metadata to the BIOM file
biom add-metadata \
  -i deblur_output_exported/feature-table.biom \
  -o deblur_output_exported/feature-table_w_tax.biom \
  --observation-metadata-fp taxa/taxonomy.tsv \
  --sc-separated taxonomy

#Last we convert the BIOM file to TSV format (tab-separated values)
biom convert \
  -i deblur_output_exported/feature-table_w_tax.biom \
  -o deblur_output_exported/feature-table_w_tax.txt \
  --to-tsv \
  --header-key taxonomy
```

This will give you a plain-text feature table (feature-table_w_tax.txt) with taxonomy annotations in the header row, which is especially useful for tools like R (phyloseq), Excel, or even manual inspection.

To export the tree of your ASVs in a .nwk format, use the command below.

```
qiime tools export \
  --input-path asvs-tree.qza \
  --output-path deblur_output_exported
```

4.11 9. Considerations for 18S and ITS data

Most of the analyses described in this tutorial apply equally to 18S and ITS datasets, **with one key exception**: building a phylogenetic tree. Unlike 16S

data, 18S and ITS amplicons generally lack a universally accepted reference phylogeny. Therefore, we must generate *de novo* phylogenetic trees.

To start, create symbolic links to the relevant data folders for your 18S and ITS datasets:

```
cd ~/workspace/bmb_module3/
mkdir 18S_analysis
cd 18S_analysis
ln -s ~/CourseData/MIC_data/bmb_module2/output/18S_analysis/deblur_output .
ln -s ~/CourseData/MIC_data/bmb_module2/output/18S_analysis/taxa .
ln -s ~/CourseData/MIC_data/bmb_module2/output/18S_analysis/Plastisphere_18S_metadata.tsv .

cd ~/workspace/bmb_module3/
mkdir ITS_analysis
cd ITS_analysis
ln -s ~/CourseData/MIC_data/bmb_module2/output/ITS_analysis/deblur_output .
ln -s ~/CourseData/MIC_data/bmb_module2/output/ITS_analysis/taxa .
ln -s ~/CourseData/MIC_data/bmb_module2/output/ITS_analysis/pregnancy_ITS_metadata.tsv .
```

4.11.1 Building a *de novo* tree for 18S

First, we'll need to make a *de novo* multiple-sequence alignment of the ASVs using MAFFT.

```
cd ~/workspace/bmb_module3/18S_analysis
mkdir tree_out

qiime alignment mafft --i-sequences deblur_output/rep_seqs_final.qza \
                      --p-n-threads 4 \
                      --o-alignment tree_out/rep_seqs_final_aligned.qza
```

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

```
qiime alignment mask --i-alignment tree_out/rep_seqs_final_aligned.qza \
                    --o-masked-alignment tree_out/rep_seqs_final_aligned_masked.qza
```

We finally run FastTree on this masked multiple-sequence alignment to make our tree with the command below:

```
qiime phylogeny fasttree --i-alignment tree_out/rep_seqs_final_aligned_masked.qza \
                          --p-n-threads 4 \
                          --o-tree tree_out/rep_seqs_final_aligned_masked_tree
```

FastTree returns an unrooted tree. One basic way to add a root to a tree is to add it at the midpoint of the largest tip-to-tip distance in the tree, which is done with this command:

```
qiime phylogeny midpoint-root --i-tree tree_out/rep_seqs_final_aligned_masked_tree.qza \
                              --o-rooted-tree tree_out/rep_seqs_final_aligned_masked_tree_rooted.qza
```

Let's rename the tree and place it in our analysis folder.

```
cp tree_out/rep_seqs_final_aligned_masked_tree_rooted.qza asvs-tree.qza
```

Due to memory or time constraints, you may opt to use a tree that has already been computed. Use the following command to create symbolic links to precomputed trees

```
cd ~/workspace/bmb_module3/18S_analysis
ln -s ~/CourseData/MIC_data/bmb_module3/output/18S_analysis/asvs-tree.qza .

cd ~/workspace/bmb_module3/ITS_analysis
ln -s ~/CourseData/MIC_data/bmb_module3/output/ITS_analysis/asvs-tree.qza .
```

4.11.2 Building the ITS tree

Repeat the same commands shown above, but make sure you use the `deblur_output` from the ITS analysis. This will generate a *de novo* tree for your ITS dataset.

Once you have generated or linked the *de novo* trees for your 18S and ITS data, you can proceed with all downstream analyses — including diversity metrics, ordinations, and statistical testing — just as you would with 16S data.

4.12 Answers

Question 1: What is a good rarefaction depth for diversity analysis?

A cut-off of 4,000 reads will be sufficient: the curve plateaus around this depth and we won't exclude any samples.

Question 2: are there any significant differences in alpha diversity between any of our metadata categories?

There are significant differences in richness and phylogenetic diversity between forest and managed environment samples. There are no significant differences between bulk and rhizosphere soil.

Question 3: which metadata category appears to provide more separation in the beta diversity PCoA plots?

This is hard to say just by looking at the Emperor plots provided by QIIME2, but the forest/managed category appears to exhibit more distinct separation. The PERMANOVA test from the `beta-group-significance` command shows this as well.

Question 4: what do you mean I'm getting an error message? It should be working fine, it worked last time! I can't see what I'm doing wro... ooooooh there we go, I see now :)

There is a typo in `--m-metadata-column Category`: The "C" should be lowercase, matching the column header in your metadata file. ;)

Question 5: can you identify any patterns between the metadata groups?

Because stacked barcharts are limited in their analytical capabilities, it is hard to discern anything except very obvious patterns.

Question 6: Does ANCOM identify any differentially abundant taxa between forest and managed environments?

One ASVs is identified as differentially abundant between forest and managed environments using ANCOM. You can look up the identity of each ASV in the `taxonomy.tsv` file. With ANCOM-BC many more ASVs are identified as differentially abundant.