

Microbiota amplicOn  
CHaracterization Implement

**MOCHI**

**User Guide**

# TABLE OF CONTENTS

## Analysis

|                                    |    |
|------------------------------------|----|
| Sequence preprocessing .....       | 3  |
| (A) Sequence summary .....         | 3  |
| (B) Sequence denoising .....       | 6  |
| (C) Taxonomic classification.....  | 8  |
| Taxanomy analysis .....            | 11 |
| (A) Upload files .....             | 11 |
| (B) Inspect result.....            | 13 |
| 1. Taxonomic table .....           | 13 |
| 2. Taxonomic barplot.....          | 14 |
| 3. Taxonomic heatmap .....         | 15 |
| 4. Krona .....                     | 16 |
| 5. Alpha diversity .....           | 17 |
| 6. Beta diversity.....             | 20 |
| 7. Phylogenetic diversity.....     | 23 |
| 8. ANCOM .....                     | 30 |
| Function analysis .....            | 33 |
| (A) Upload files.....              | 33 |
| (B) Inspect result.....            | 34 |
| 1. Function annotation table ..... | 34 |
| 2. Function barplot.....           | 35 |

## ANALYSIS

# Sequence preprocessing

### (A) Sequence summary

1. Select “**Sequence Preprocessing**” on the top bar, choose “**Step 1. Sequence summary**”.

The screenshot shows the MOCHI web interface. At the top, there is a navigation bar with links for Home, Sequence Preprocessing (which is currently selected and highlighted in blue), Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar, a blue header bar says "Welcome to MOCHI". Underneath the header, there are three steps listed: Step 1. Sequence summary, Step 2. Sequence denoising, and Step 3. Taxonomy classification. A red arrow points to the "Step 1. Sequence summary" link.

2. Press the “**Select the directory**” button to open selection window. Alternatively, you could press the “Demo” button to download the example files.

This screenshot shows the "Sequence files" section of the MOCHI interface. It has a blue background and contains a text input field labeled "Please choose the directory containing sequence files (\*.fastq.gz)". Below the input field is a red-bordered button labeled "Select the directory". Further down is a "Demo" button. To the right of this section, there is a "Start!" button and a "1. Sequences summary (for single end)" section.

#### 1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

(2) Inspect the result.

**! Important:** The sequence files should be placed under “seqs\_folder” where MOCHI was installed.

3. Select a directory containing sequence files.

This screenshot shows a file selection dialog box. At the top, it says "Please select a directory" and has a close button "X". Below that are buttons for "Create new folder", "Sort content", and a pencil icon. The main area is divided into "Directories" and "Content". The "Directories" section shows a tree view with "raw\_data" expanded, revealing a folder named "single\_seqs\_demo" which is highlighted with a red box and a yellow arrow pointing to it, labeled "Select folder". The "Content" section lists numerous fastq.gz files with their sizes. At the bottom right of the dialog are two buttons: "Confirm selection" (highlighted with a yellow arrow) and "Cancel".

| L1S105_9_L001_R1_001.fastq.gz  | 889.1 kB |  |
|--------------------------------|----------|--|
| L1S140_6_L001_R1_001.fastq.gz  | 675.3 kB |  |
| L1S208_10_L001_R1_001.fastq.gz | 861.9 kB |  |
| L1S257_11_L001_R1_001.fastq.gz | 633.7 kB |  |
| L1S281_5_L001_R1_001.fastq.gz  | 685.7 kB |  |
| L1S57_13_L001_R1_001.fastq.gz  | 896.4 kB |  |
| L1S76_12_L001_R1_001.fastq.gz  | 767.0 kB |  |
| L1S8_8_L001_R1_001.fastq.gz    | 887.4 kB |  |
| L2S155_25_L001_R1_001.fastq.gz | 770.0 kB |  |
| L2S175_27_L001_R1_001.fastq.gz | 873.6 kB |  |
| L2S204_1_L001_R1_001.fastq.gz  | 615.1 kB |  |
| L2S222_23_L001_R1_001.fastq.gz | 736.1 kB |  |
| L2S240_7_L001_R1_001.fastq.gz  | 929.2 kB |  |
| L2S309_33_L001_R1_001.fastq.gz | 314.7 kB |  |
| L2S357_15_L001_R1_001.fastq.gz | 502.5 kB |  |
| L2S382_34_L001_R1_001.fastq.gz | 792.2 kB |  |
| L3S242_19_L001_R1_001.fastq.gz | 149.3 kB |  |

4. **Sequence type** (single-end or paired-end) is now automatically detected. If not correct, you can choose manually.

**Sequence files**  
Please choose the directory containing sequence files (\*.fastq.gz)

Select the directory  
 Demo

**Sequence type**  
Choose the sequence type  
 Single end

## 1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

Start!

(2) Inspect the result.

View!  log file

On the **View** webpage, you can download the result by right click **Save as ...**

Example output for sequence summary

Example result for single end

5. Choose the **primers** or check the box when using primer-trimmed reads.

**Primer sequences**

Check this if your sequences are primer-trimmed reads

Choose the forward primer sequence  
 515F

Choose the reverse primer sequence  
 806R

Show primer sequences

6. Set the number of threads to run the analysis. If zero is provided, all available cores will be used. If you do not know the number to enter, leave it to the default number (all threads - 2).

**Computing setting**

Number of threads MOCHI can use  
 0

Show the number of threads on system

7. Click on the “**Start!**” button. (If you wish to preview what the result will look like, press the “Example result for single/pair end” button to open a demo result.)

**Sequence files**  
Please choose the directory containing sequence files (\*.fastq.gz)

Select the directory

## 1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

Start!

8. Please wait while running. When complete, a popup window will display.

The screenshot shows the MQCHI web application. On the left, there's a sidebar with sections for "Primer sequences" and "Sequence files". Under "Primer sequences", there's a checkbox for "Check this if your sequences are primer-trimmed reads" and a text input field for "Choose the forward primer sequence". On the right, there's a main content area with tabs for "Sequence Preprocessing", "Taxonomy Analysis", "Function Analysis", and "Tutorial". A central message box displays "Successful!" and a note: "This analysis took 16.96 secs. You can click the button View to inspect the result." Below this, a red box highlights a "Please wait..." message with a circular loading icon. At the bottom, there's a section titled "(2) Inspect the result." with two buttons: "View!" (highlighted with a red box) and "log file".

9. Click on the “View!” button to view the result.

This screenshot shows the same MQCHI interface as above, but with a red box highlighting the "View!" button under the "(2) Inspect the result." section. The rest of the interface, including the "Sequence type" section and the "View!" button itself, is visible.

## (B) Sequence denoising

1. Select “**Sequence Preprocessing**” on the top bar, choose “**Step 2. Sequence denoising**”.

The screenshot shows the MOCHI web interface. At the top, there is a navigation bar with links for Home, Sequence Preprocessing (which is currently selected), Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar, a blue header bar says "Welcome to MOCHI". On the right side of the header, there is some descriptive text about MOCHI being built on QIIME2 and using the R package Shiny. A red arrow points to the "Sequence denoising" link in the "Sequence Preprocessing" dropdown menu.

2. Set the **position** and the **quality score** to trim the sequences.

### \* Starting and ending position:

Base pairs below starting position and above ending position will be filtered out. For instance, setting the starting position to 5 and the ending position to 120 will get sequences from 5 to 120 bp with 115 bp long. In addition, reads shorter than the ending position will be discarded. In this case, sequences less than 120 bp are removed. If the ending position is set to 0, no truncation or length filtering will be performed.

### \* Quality score:

Nucleotides with quality score less than or equal to specified value are truncated. Resulting reads shorter than ending position are discarded.

This screenshot shows the "Sequence trimming" section of the MOCHI interface. It includes two input fields: "The start position" (set to 0) and "The end position" (set to 0). Below these fields are two buttons: "learn more" and "Demo". A red arrow points to the "The start position" field. Another red arrow points to the "The end position" field.

This screenshot shows the "2. Sequence denoising (DADA2) for Single end" section of the MOCHI interface. It features a "Start!" button and a "Inspect the sequence denoising result" section with several options: "Show summary table", "Show seqs info", "Show filter info", "Show alpha rarefaction", and "log file". Below this is an "Example output for sequence denoising" section with tabs for "Summary table", "Seqs info", "Filter info", and "Alpha rarefaction". A red arrow points to the "Quality score threshold" input field in the "Sequence trimming" section of the adjacent screenshot.

3. Set the parameter of **chimera**, **training error model**, **threads** and upload the **metadata**.

### \* Chimeric reads filter:

A chimeric read is a sequence originate from multiple or parent sequences. Chimeric reads are generally considered contaminant. Whereas a chimeric read can be interpreted as a novel sequence, it is in fact an artifact. The higher this value is, the more chimeric reads will be used in analysis. For most cases, 1 is the default value.

### \* Training error model:

Input the number of reads for training the error model. Smaller numbers will result in a shorter run time but a less reliable error model. The default value is 1,000,000.

### \* Threads:

Specify numbers of thread to speed up the analytical process. Increasing the number of threads will decrease the running time. When zero is provided, all available cores will be used. If you do not know the number to enter, leave it to the default number.

### \* Integrating the metadata (Optional):

If the metadata is provided, the results will have metadata information.

The screenshot shows the MOCHI interface with several configuration sections:

- Chimeric reads filtering**: A section with a text input field containing "1" and a "learn more" button.
- Error model training**: A section with a text input field containing "1000000" and a "learn more" button.
- Metadata Integrating (optional)**: A section with a "Browse..." button, a "No file selected" message, and a "learn more" button.
- Computing setting**: A section with a text input field containing "0" and a "learn more" button.

- Click on the “**Start to denoise!**” button. (If you wish to preview what the result will look like, press the buttons in the “Example output for sequence denoising” section.)

The screenshot shows the Sequence trimming section with a text input field containing "5".

## 2. Sequence denoising (DADA2) for Single end

(1) Start to denoise.

**Start!**

- Please wait while running. When complete, a popup window will display.

The screenshot shows the Quality score filtering section with a text input field containing "0".



- Click on the buttons in the “result section” to view the results.

The screenshot shows the MOCHI interface with the following results:

- MOCCHI** navigation bar: Home, Sequence Preprocessing, Taxonomy Analysis, Function Analysis, Tutorial.
- Denoising successfully!** message in a modal window.
- Sequence trimming** section: "The start position" set to "0".
- Message in the main area: "This analysis took 48.32 secs. You can inspect the results!"
- Start to denoise!** button.

## (C) Taxonomic classification

1. Select “**Sequence Preprocessing**” on the top bar, choose “**Step 3. Taxonomic classification**”.

The screenshot shows the MOCHI web interface. At the top, there is a navigation bar with links for Home, Sequence Preprocessing (which is currently selected), Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar, a blue header bar says "Welcome to MOCHI". On the left, there is a sidebar with a "Database" section containing a dropdown menu set to "Silva (Not detected)", an "Auto download database" button, and a "Demo" button. The main content area has a title "3. Taxonomy classification" and two sections: "(1) Classify taxonomy" with a "Start!" button, and "(2) Inspect the taxonomy classification result" with "View!" and "log file" buttons. A red arrow points from the "Step 3. Taxonomy classification" link in the top navigation to the corresponding section in the main content area.

2. Download and select **database** (Silva, Greengene, or PR2) to predict taxa.

This screenshot shows the "Database" selection page. It features a title "Database" and a sub-instruction "Select the reference database for taxonomy classification." Below this is a "Choose the database" section with a dropdown menu currently set to "Silva (Not detected)" and an "Auto download database" button. A red arrow points from the "Auto download database" button to the "Database" section. At the bottom is a "Demo" button.

### 3. Taxonomy classification

#### (1) Classify taxonomy

Start!

#### (2) Inspect the taxonomy classification result.

View!  log file

#### 2.1. Automatically download database:

- \* Select a database from the drop-down menu “**Choose the database**”. Press the “**Auto download database**” button. The latest database will be pulled from the server. The downloading process may take a while depending on the file size and network speed.

#### 2.2. Manually download database:

- \* Silva: Follow the [link](#). Choose a version and download. Decompressing the downloaded file. Copy two folders, “**rep\_set**” and “**taxonomy**”, to the folder “**taxa\_database/silva**”.
- \* Greengenes: Follow the [link](#). Choose a version and download the corresponding “**otus.tar.gz**”. Decompressing the downloaded files. Copy two folders, “**rep\_set**” and “**taxonomy**”, to the folder “**taxa\_database/greengenes**”.
- \* PR2: Follow the [link](#). Choose a version and download the corresponding “**pr2\_version\_X.XX.X\_16S\_mothur.fasta.gz**” and “**pr2\_version\_X.XX.X\_16S\_mothur.tax.gz**”. Decompressing the downloaded files. Copy file “**pr2\_version\_X.XX.X\_16S\_mothur.fasta**” to the folder “**taxa\_database/PR2/18S/seqs**” and file “**pr2\_version\_X.XX.X\_16S\_mothur.tax**” to the folder “**taxa\_database/PR2/18S/taxonomy**”.

3. Check if your **primers** are correct.

### Reference sequence filtering

**1. Check primers**  
If incorrect, go to 'Sequence summary' to select the correct primer.

Your forward primer is **8F** now.

Your reverse primer is **519R** now.

[The taxonomic table](#) [The ASVs table](#) [The seqs data](#)

**Example output for taxonomy classification**

[taxonomy result](#)

4. Set the minimum and maximum length to filter reference sequence.

#### \* Minimum and maximum length:

Reference sequences shorter than the specified values will be discarded. The default values are the minimum and the maximum length of denoised sequences. To disable length filtering, set the values to zero.

**2. Filter the reference sequence by length**

**Minimum length**

**Maximum length**

[learn more](#)

5. Set the number of threads to run the analysis. If zero is provided, all available cores will be used. If you do not know the number to enter, leave it to the default number (all threads - 2).

**Computing setting**

Number of threads MOCHI can use

6. Click on the "**Start!**" button. (If you wish to preview what the result will look like, press the "Example output for taxonomy classification" button.)

**Database**  
Select the reference database for taxonomy classification.

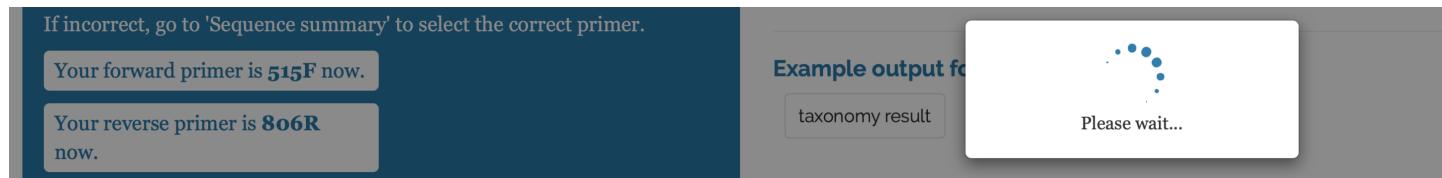
**Choose the database**

### 3. Taxonomy classification

(1) Classify taxonomy

**Start!**

7. Please wait while running. When complete, a popup window will display.



MOCHI Home Sequence Preprocessing Analysis Taxonomy OTU Table Database  
Database Select the reference database for taxonomy classification.  
Taxonomy classification has been finished!  
This analysis took 55.92 secs. You can inspect the results!

8. Click on the “**View!**” button to view the result. Download the files before proceeding to taxonomy analysis.

Auto download database  
Demo  
**Reference sequence filtering**  
**1. Check primers**  
If incorrect, go to 'Sequence summary' to select the correct primer.  
Your forward primer is **8F** now.

(2) Inspect the taxonomy classification result.  
**View!** log file

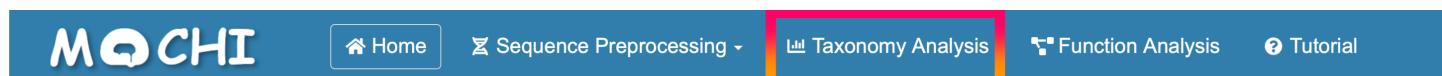
(3) Download the files for the next step.  
The taxonomic table The ASVs table The seqs data

Example output for taxonomy classification  
taxonomy result

# Taxanomy analysis

## (A) Upload files

1. Select “**Taxanomy analysis**” on the top bar.



### Welcome to MOCHI! (Microbiota amplicOn CHaracterization Implement)

MOCHI is a 16S or 18S microbiota amplicon rRNA analytical tool for microbiota based primarily on QIIME2 with a friendly web interface powered by the R package of Shiny. MOCHI may also be downloaded and operated locally.

2. In the left panel, press the “Browse” buttons to upload **metadata**, **taxonomic table** and **ASVs table** files. These files can be downloaded from “Sequence Preprocessing - Taxonomic classification” section. Please see [Sequence preprocessing / Taxonomic classification / step 8](#). Alternatively, you could press the “Demo” button to download the example files first and then upload. If sequences are 18S rRNA, please check the “18S rRNA” box.

#### \* Metadata (.tsv):

The first column name must be **#SampleID**.

#### \* Taxonomic table file (.qza):

You could upload self-derived taxonomic table file from QIIME2.

#### \* ASVs table (.qza):

An “amplicon sequence variant” table is a higher-resolution analogue of the traditional OTU table.

The image shows the left side of the MOCHI web interface, which contains three separate upload sections. Each section has a label at the top, a "Browse..." button, and a "No file selected" message. The first section is labeled "Upload the metadata file" and has a red border around it. The second section is labeled "Upload the taxonomic table file". The third section is labeled "Upload the ASVs table file". At the bottom of the left panel, there is a checkbox for "18S rRNA" and a "Demo" button. On the right side of the interface, there are several navigation links: "Taxonomic table", "Taxonomic barplot", "Taxonomic heatmap", and "Krona".

### 3. The results will be displayed on the right panel once the files are uploaded.

**Upload the metadata file**

Browse... Metadata\_example.tsv  
Upload complete

**Upload the taxonomic table file**

Browse... Taxonomic\_table\_example.qza  
Upload complete

**Upload the ASVs table file**

Browse... ASVs\_table\_example.qza  
Upload complete

**■ 18S rRNA**

Demo

Taxonomic table   Taxonomic barplot   Taxonomic heatmap   Krona   Alpha diversity   Beta diversity   Phylogenetic diversity   ANCOM

**Choose the group**

SampleID    barcode.sequence    body.site    year    month    day    subject    reported.antibiotic.usage    days.since.experiment.start

Show **10** entries

| Kingdom (K=3) | Phylum (K=21)  | Class (K=42)         | Order (K=69)     | Family (K=134)   | Genus (K=281)           | Species (K=356) | L1S8 | L1S57 | L1S76 | L1S105 | L2S155 | L2S175 | L2 |
|---------------|----------------|----------------------|------------------|------------------|-------------------------|-----------------|------|-------|-------|--------|--------|--------|----|
| Unassigned    | Unassigned     | Unassigned           | Unassigned       | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      | 0  |
| Archaea       | Crenarchaeota  | Thaumarchaeota       | Nitrosphaerales  | Nitrosphaeraceae | Candidatus Nitrosphaera | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      | 0  |
| Archaea       | Crenarchaeota  | Thaumarchaeota       | Nitrosphaerales  | Nitrosphaeraceae | Candidatus Nitrosphaera | SCA1145         | 0    | 0     | 0     | 0      | 0      | 0      | 0  |
| Bacteria      | Unassigned     | Unassigned           | Unassigned       | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 25     | 13     |    |
| Bacteria      | Acidobacteria  | [Chloracidobacteria] | RB41             | Ellin6075        | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |    |
| Bacteria      | Actinobacteria | Acidimicrobia        | Acidimicrobiales | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |    |
| Bacteria      | Actinobacteria | Actinobacteria       | Actinomycetales  | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 7      | 0      |    |
| Bacteria      | Actinobacteria | Actinobacteria       | Actinomycetaceae | Actinomyces      | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |    |
| Bacteria      | Actinobacteria | Actinobacteria       | Actinomycetales  | Actinomycetaceae | Actinomyces             | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |    |

Showing 1 to 10 of 385 entries

[Previous](#) 1 [2](#) [3](#) [4](#) [5](#) ... [39](#) [Next](#)

[Download Taxonomic table](#)

## (B) Inspect result

- MOCHI displays the results in eight approaches: (1) Taxonomic table, (2) Taxonomic barplot, (3) Taxonomic heatmap, (4) Krona, (5) Alpha diversity, (6) Beta diversity, (7) Phylogenetic diversity, and (8) ANCOM:

### 1. Taxonomic table

A table describes taxonomy information and read count.

Choose the group

SampleID barcode.sequence body.site year month day subject reported.antibiotic.usage days.since.experiment.start

Show 10 entries

Search:

| Kingdom (K=3) | Phylum (K=21)  | Class (K=42)              | Order (K=69)     | Family (K=134)   | Genus (K=281)           | Species (K=356) | L1S8 | L1S57 | L1S76 | L1S105 | L2S155 | L2S175 |
|---------------|----------------|---------------------------|------------------|------------------|-------------------------|-----------------|------|-------|-------|--------|--------|--------|
| Unassigned    | Unassigned     | Unassigned                | Unassigned       | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Archaea       | Crenarchaeota  | Thaumarchaeota            | Nitrosphaerales  | Nitrosphaeraceae | Candidatus Nitrosphaera | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Archaea       | Crenarchaeota  | Thaumarchaeota            | Nitrosphaerales  | Nitrosphaeraceae | Candidatus Nitrosphaera | SCA1145         | 0    | 0     | 0     | 0      | 0      | 0      |
| Bacteria      | Unassigned     | Unassigned                | Unassigned       | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 25     | 13     |
| Bacteria      | Acidobacteria  | [Chloracidobacteria] KB41 | Ellin6075        | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Bacteria      | Actinobacteria | Acidimicrobia             | Acidimicrobiales | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Bacteria      | Actinobacteria | Actinobacteria            | Actinomycetales  | Unassigned       | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 7      | 0      |
| Bacteria      | Actinobacteria | Actinobacteria            | Actinomycetales  | Actinomycetaceae | Unassigned              | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Bacteria      | Actinobacteria | Actinobacteria            | Actinomycetales  | Actinomycetaceae | Actinomyces             | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |
| Bacteria      | Actinobacteria | Actinobacteria            | Actinomycetales  | Actinomycetaceae | Actinomyces             | Unassigned      | 0    | 0     | 0     | 0      | 0      | 0      |

Showing 1 to 10 of 385 entries

Previous 1 2 3 4 5 ... 39 Next

Download Taxonomic table

Download button

#### \* Choose the group:

Select a group provided in the metadata to categorize read count (see below).

#### \* Taxonomy information:

The left part of the table (the first 7 columns) represents taxonomy information. The column name indicates taxonomy levels. K denotes the number of taxa at a given level.

#### \* Read count:

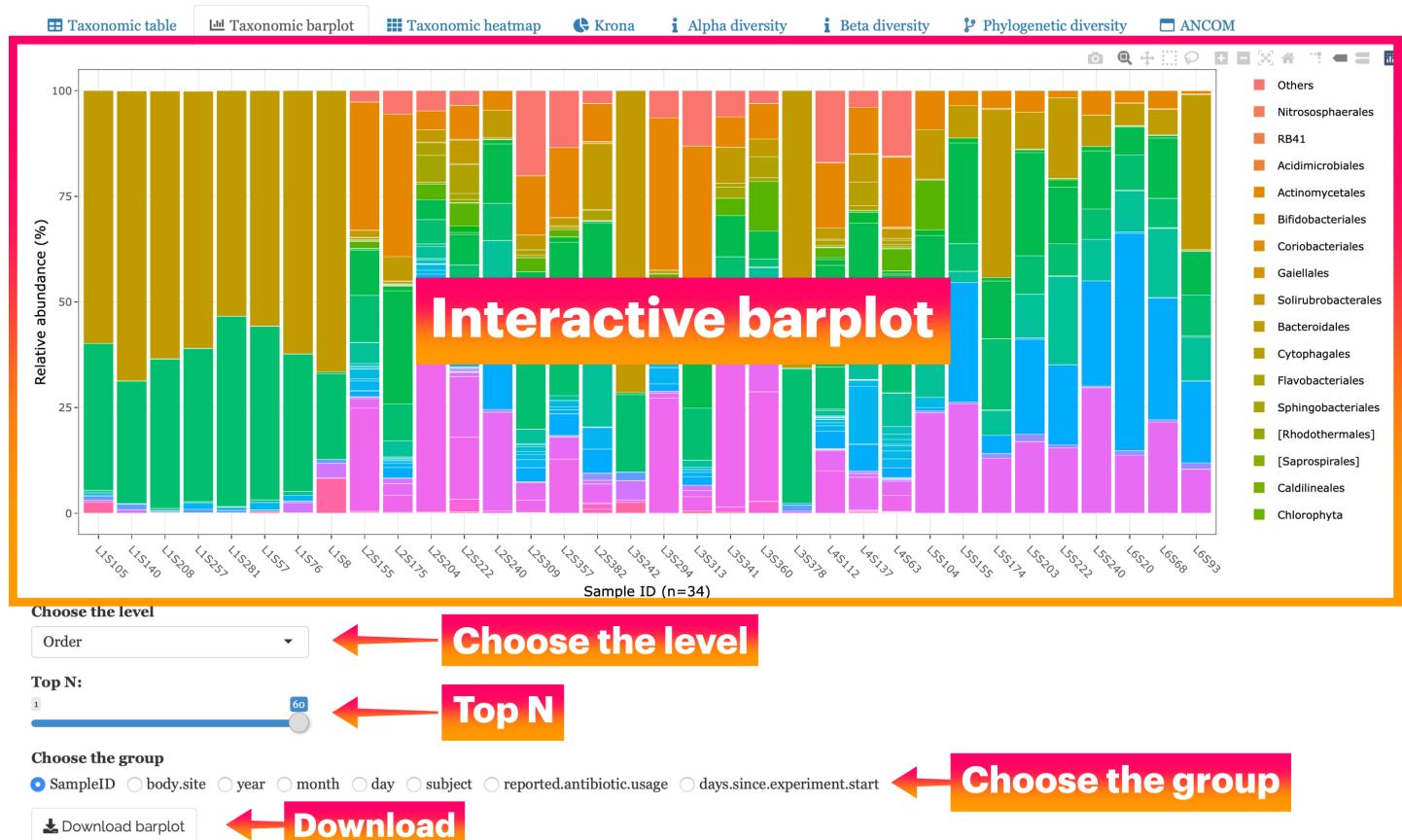
The right part of the table represents read count. The values of read count are categorized by the selected group. The column name indicates the variables of the selected group.

#### \* Download taxonomic table:

Click on the "Download Taxonomic table" button to download the displayed table.

## 2. Taxonomic barplot

An interactive barplot showing the percentage of taxa in all sample. Each taxon is represented by a sub-bar with different colors.



### \* Interactive barplot:

When a cursor hovers over the bar region, the information of species will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back.

### \* Choose the level:

The taxa in the plot will be presented at the selected taxonomic level.

### \* Top N:

Control the numbers of taxa displayed on the plot. When you select value N, the plot will show the union of the top N relatively abundant taxa in each sample. For example, if N = 2 is selected and the top 2 abundant taxa in Sample A and Sample B are “taxa\_1 and taxa\_2” and “taxa\_1 and taxa\_3”, respectively, the plot will show the relative abundance of taxa\_1, taxa\_2 and taxa\_3.

### \* Choose the group:

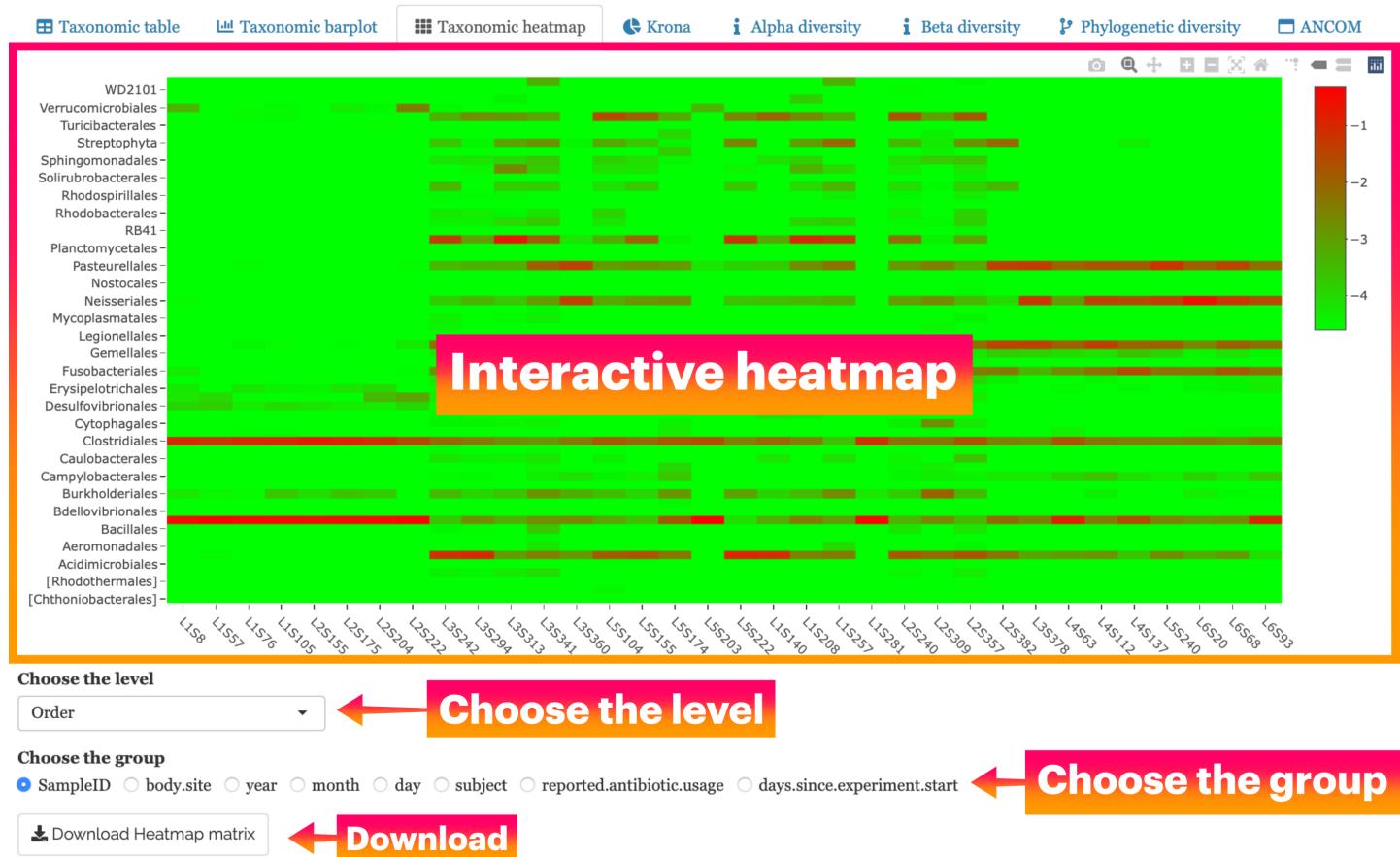
The barplot will be grouped based on the selected metadata.

### \* Download barplot:

Click on the “Download barplot” button to download the barplot. Alternatively, click on the camera icon on the top-right region of the barplot.

### 3. Taxonomic heatmap

An interactive heatmap showing the log10-transformation percentage of taxa in all sample. To prevent taking logarithm of zero, a small value of 0.01 is added to all percentage values before transformation. The transformed values are shown in color gradient.



#### \* Interactive heatmap:

When a cursor hovers over the heatmap, the information of transformed value will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back. Click on the camera icon on the top-right region of the heatmap to download the plot.

#### \* Choose the level:

The taxa in the plot will be presented at the selected taxonomic level.

#### \* Choose the group:

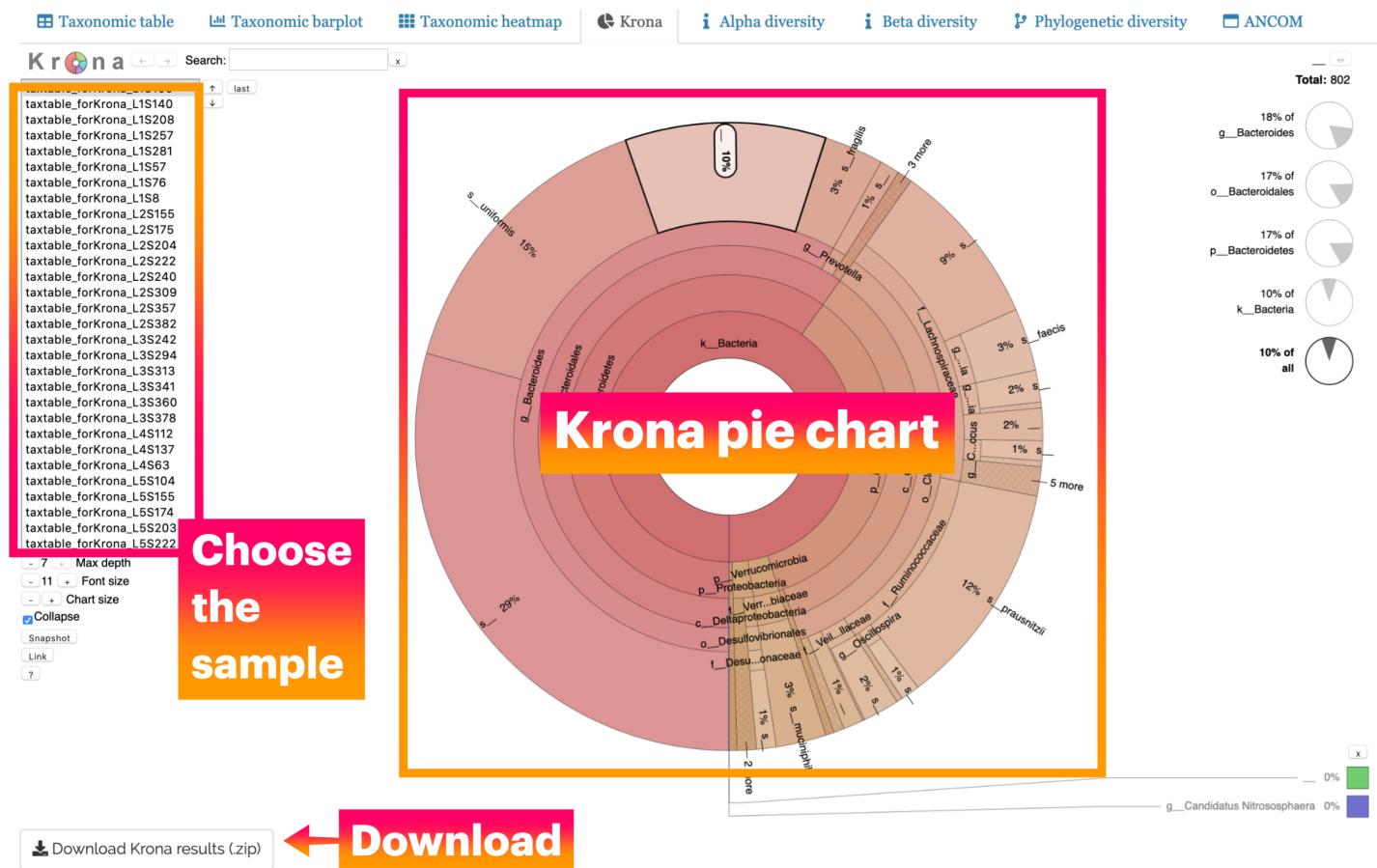
The heatmap will be grouped based on the selected metadata.

#### \* Download heatmap matrix:

Click on the "Download Heatmap matrix" button to download the heatmap matrix data.

## 4. Krona

A visualization tool allowing hierarchical data to be explored with zooming, multi-layered pie charts. [Get more information](#).



### \* Krona pie chart:

An interactive pie plot. Single click a taxon will show the ratio of the selected taxon over different taxonomy level. Double click a taxon will zoom in the selected taxonomy level. To zoom back, click the backspace button on the top-left region.

### \* Choose the sample:

Select the sample to switch to the corresponding pie plot.

### \* Download Krona results:

Click on the “Download Krona results (.zip)” button to download the interactive pie plot (html files with Javascript).

## 5. Alpha diversity

Evaluation of species diversity within samples. In MOCHI, we adapt 8 indexes (richness, Chao1, ACE, Shannon diversity, InvSimpson diversity, Shannon evenness, Simpson evenness, and Goods coverage).

### 5.1. Table

Table

Show 10 entries

Search:

|    | Sample | Richness | Chao1 | ACE | Shannon_diversty | Simspon_diversity | InvSimpson_diversity | Shannon_evenness | Simpson_evenness |
|----|--------|----------|-------|-----|------------------|-------------------|----------------------|------------------|------------------|
| 1  | L1S105 | 63       | 63    | 63  | 2.6808           | 0.8705            | 7.7201               | 0.4033           | 0.0015           |
| 2  | L1S140 | 65       | 65    | 65  | 2.6609           | 0.8519            | 6.7499               | 0.4004           | 0.0015           |
| 3  | L1S208 | 85       | 85    | 85  | 3.1189           | 0.8995            | 9.955                | 0.4693           | 0.0014           |
| 4  | L1S257 | 81       | 81    | 81  | 3.259            | 0.9256            | 13.4455              | 0.4903           | 0.0014           |
| 5  | L1S281 | 72       | 72    | 72  |                  |                   | 26                   | 0.4792           | 0.0014           |
| 6  | L1S57  | 70       | 70    | 70  |                  |                   | 67                   | 0.4368           | 0.0015           |
| 7  | L1S76  | 61       | 61    | 61  | 2.4883           | 0.7959            | 4.8999               | 0.3744           | 0.0016           |
| 8  | L1S8   | 44       | 44    | 44  | 2.2026           | 0.7939            | 4.851                | 0.3314           | 0.0016           |
| 9  | L2S155 | 109      | 109   | 109 | 3.5545           | 0.9388            | 16.3338              | 0.5348           | 0.0014           |
| 10 | L2S175 | 104      | 104   | 104 | 3.4387           | 0.9221            | 12.8439              | 0.5174           | 0.0014           |

Showing 1 to 10 of 34 entries

Previous 1 2 3 4 Next

[!\[\]\(848edf3a971f9d4a6acd664a9b2a684c\_img.jpg\) Download Alpha Diversity Table](#)

[Download](#)

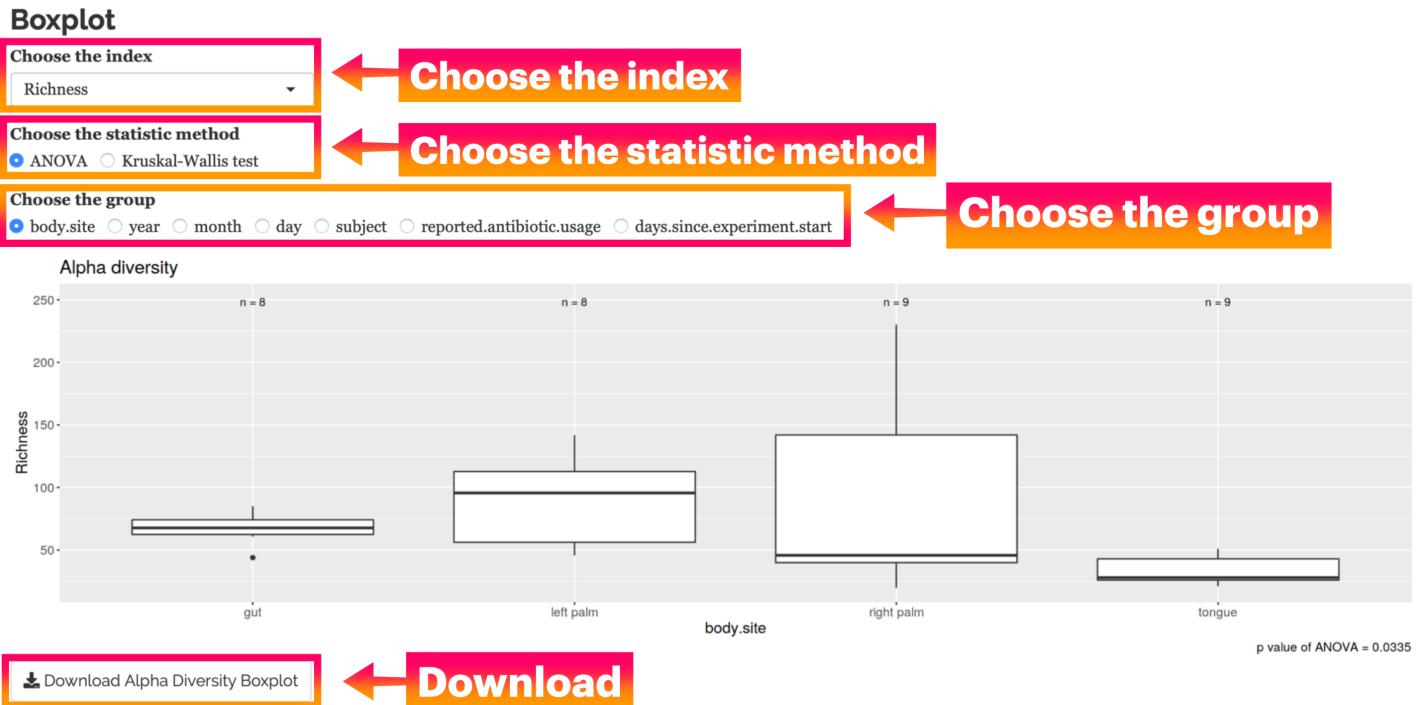
#### \* Alpha diversity table:

This table shows the values of 8 alpha diversity indexes.

#### \* Download alpha diversity table:

Click on the “Download Alpha Diversity Table” button to download the table.

## 5.2. Boxplot



\* Choose the index:

A boxplot will be presented with the selected index.

\* Choose the statistic method:

Select ANOVA (parametric method) or Kruskal-Wallis (nonparametric method) to test whether the distribution of the index is significantly different among the groups.

\* Choose the group:

The values of the index in the boxplot will be grouped based on the selected metadata.

\* Download Alpha diversity boxplot:

Click on the "Download Alpha Diversity Boxplot" button to download the boxplot.

### 5.3. Post-hoc analysis

## Post hoc analysis

### Tukey test

| Group A    | Group B    | Diff   | P value |
|------------|------------|--------|---------|
| tongue     | right palm | -52.78 | 0.06    |
| tongue     | left palm  | -57.46 | 0.04    |
| tongue     | gut        | -34.96 | 0.34    |
| right palm | left palm  | -4.68  | 1.00    |
| right palm | gut        | 17.82  | 0.82    |
| left palm  | gut        | 22.50  | 0.72    |

 Download Alpha Diversity statistical result

## Post hoc analysis

### Dunn test

| Group A    | Group B    | Z     | P value |
|------------|------------|-------|---------|
| gut        | left palm  | -0.58 | 0.28    |
| gut        | right palm | 0.76  | 0.22    |
| gut        | tongue     | 2.89  | 0.00    |
| left palm  | right palm | 1.35  | 0.09    |
| left palm  | tongue     | 3.49  | 0.00    |
| right palm | tongue     | 2.20  | 0.01    |

 Download Alpha Diversity statistical result

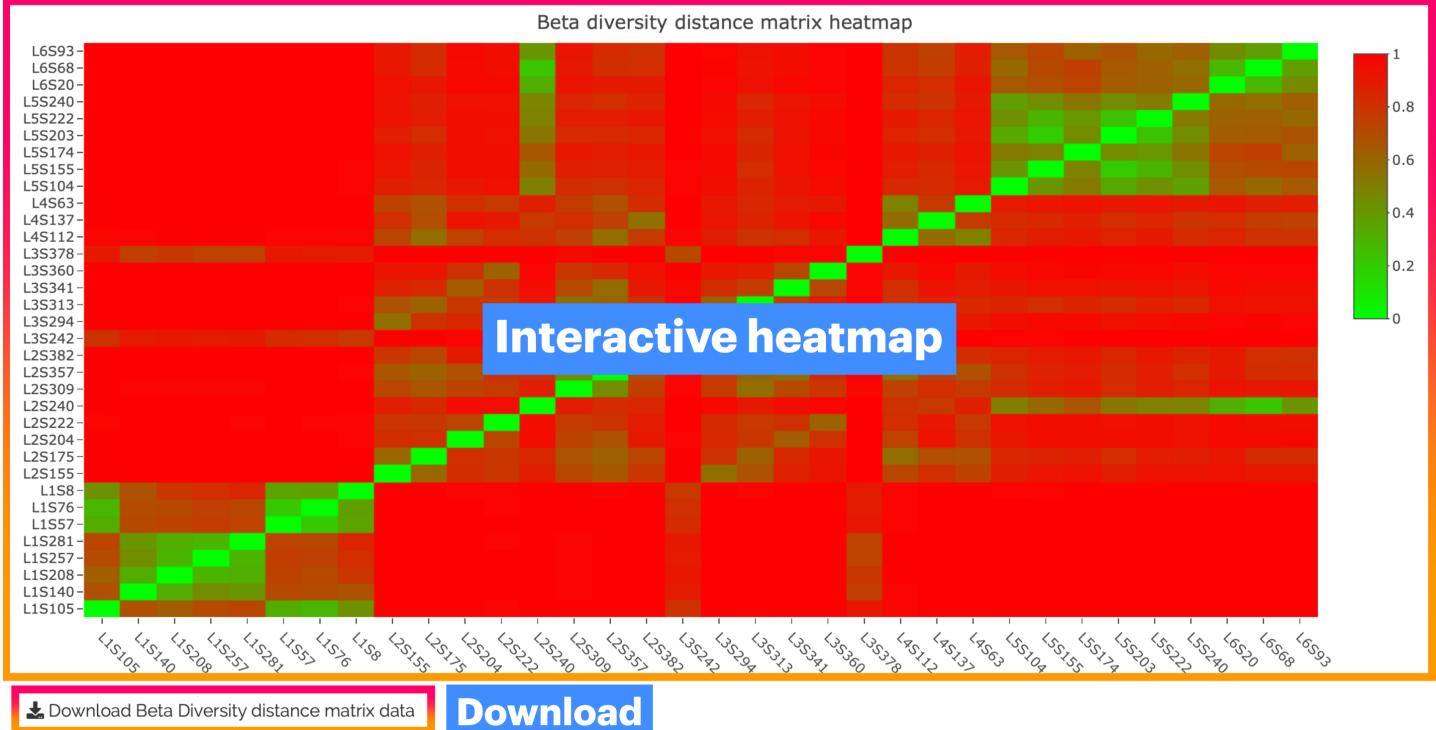
- \* If ANOVA is selected when creating the boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.
- \* Download Alpha diversity post-hoc test result:  
Click on the “Download Alpha Diversity statistical result” button to download the post-hoc test result.

## 6. Beta diversity

Evaluation of species diversity between samples. In MOCHI, we use the Bray-Curtis index.

### 6.1. Distance matrix

Beta diversity table (Bray-Curtis)



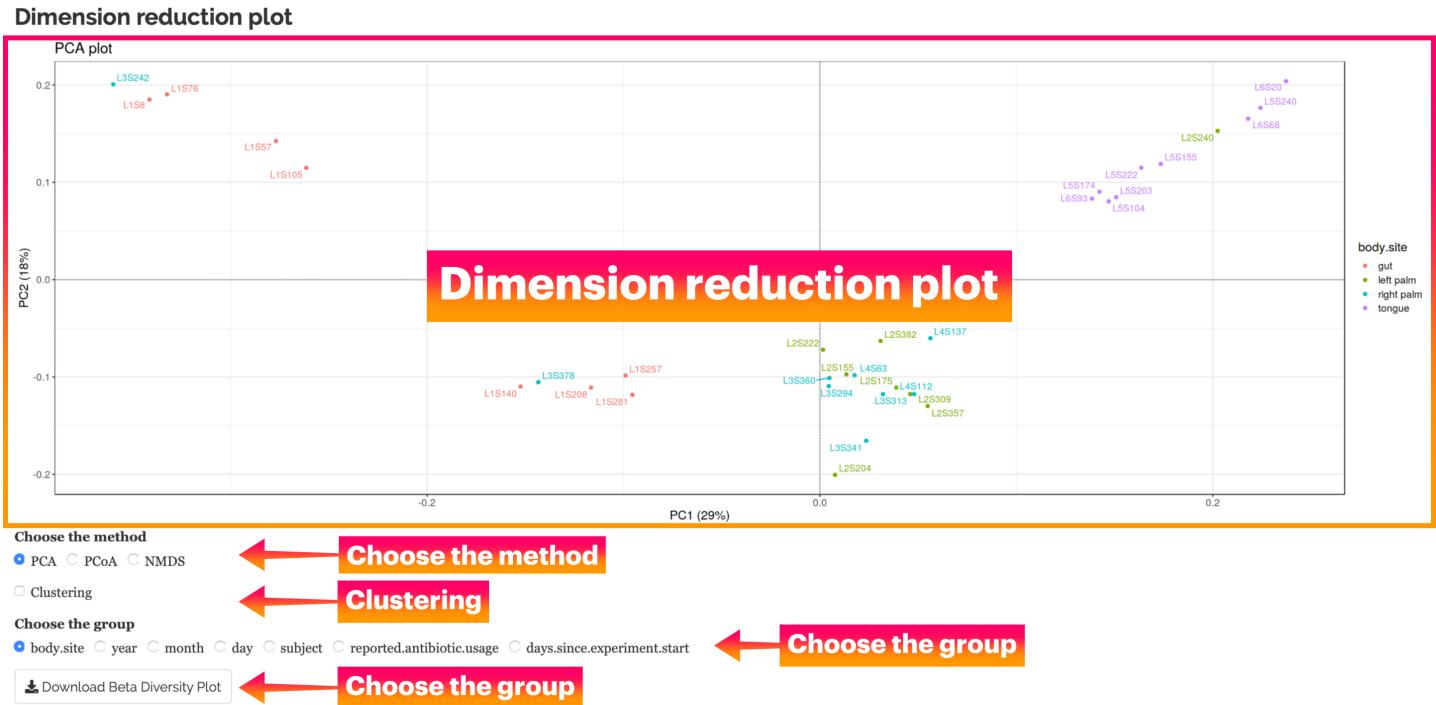
\* Interactive heatmap:

When a cursor hovers over the heatmap, the distance between species will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back. Click on the camera icon on the top-right region of the heatmap to download the plot.

\* Download distance matrix:

Click on the button "Download Beta Diversity distance matrix data" to download the matrix data.

## 6.2. Dimension-reduction plot



- \* We provide three dimension-reduction methods, including PCA (Principal Component Analysis), PCoA (Principal Co-ordinates Analysis) and NMDS (Non-metric Multidimensional Scaling) for visualization of beta diversity.
- \* Choose the method:  
Select a dimension reduction method.
- \* Clustering:  
When checking the box, samples in the same group will be surrounded by a circle.
- \* Choose the group:  
Samples will be labeled in colors based on the selected metadata.
- \* Download beta diversity plot:  
Click on the “Download Beta Diversity Plot” button to download the plot.

## 6.3. Statistical analysis

### Statistical analysis

#### PERMANOVA

| R <sup>2</sup> | P value |
|----------------|---------|
| 0.3999         | 0.001   |

[Download permanova table](#)

#### ANOSIM

| R      | P value |
|--------|---------|
| 0.6855 | 0.001   |

[Download ANOSIM table](#)

#### MRPP

| A      | Observe.delta | Expect.delta | P value |
|--------|---------------|--------------|---------|
| 0.2085 | 0.6886        | 0.8699       | 0.001   |

[Download MRPP table](#)

#### PERMANOVA pair

| Comparisons            | R <sup>2</sup> | P value |
|------------------------|----------------|---------|
| gut - left palm        | 0.3983         | 0.001   |
| gut - right palm       | 0.2834         | 0.001   |
| gut - tongue           | 0.5474         | 0.001   |
| left palm - right palm | 0.0585         | 0.541   |
| left palm - tongue     | 0.2985         | 0.001   |
| right palm - tongue    | 0.276          | 0.001   |

[Download permanova pair table](#)

#### ANOSIM pair

| Comparisons            | R       | P value |
|------------------------|---------|---------|
| gut - left palm        | 1       | 0.001   |
| gut - right palm       | 0.6686  | 0.001   |
| gut - tongue           | 1       | 0.001   |
| left palm - right palm | -0.0538 | 0.766   |
| left palm - tongue     | 0.6953  | 0.001   |
| right palm - tongue    | 0.5343  | 0.001   |

[Download ANOSIM pair table](#)

#### MRPP pair

| Comparisons            | A       | delta  | E.delta | P value |
|------------------------|---------|--------|---------|---------|
| gut - left palm        | 0.2055  | 0.6694 | 0.8426  | 0.001   |
| gut - right palm       | 0.1456  | 0.7342 | 0.8593  | 0.001   |
| gut - tongue           | 0.3046  | 0.5454 | 0.7842  | 0.002   |
| left palm - right palm | -0.0018 | 0.8318 | 0.8303  | 0.565   |
| left palm - tongue     | 0.1373  | 0.643  | 0.7453  | 0.002   |
| right palm - tongue    | 0.1412  | 0.7056 | 0.8216  | 0.002   |

[Download MRPP pair table](#)

- \* We provide three statistical methods, including PerMANOVA (Permutational Multivariate Analysis of Variance), ANOSIM (Analysis of Similarities) and MRPP (Multiple Response Permutation Procedure), to test whether beta diversity is significantly different among groups or between pairs of groups.

- \* Download statistical result table:

Click on the button below the table to download the statistical results.

## 7. Phylogenetic diversity

A species diversity considers the genetic distance. In MOCHI, we use Faith PD (a kind of alpha diversity which considers the genetic distance) and Unifrac distance (a kind of beta diversity which considers the genetic distance).

■ Taxonomic table ■ Taxonomic barplot ■ Taxonomic heatmap ■ Krona ■ Alpha diversity ■ Beta diversity

❖ Phylogenetic diversity ■ ANCOM

**Phylogenetic diversity is a species diversity considering genetic distance.**

**Upload the sequences data**

Upload

Browse... Seqs\_forPhylo\_example.qza  
Upload complete

**Input the sampling depth**

Sampling depth

898  
learn more

**Input the number of threads**

Number of threads

0  
The number of threads to use for multithreaded process. The default value is all threads-2.

Start!

Demo

\* **Upload the files:**

Upload the sequence file (.qza). If you have already finished the “Sequence Preprocessing” steps, download the file from “Sequence Preprocessing - Taxonomic classification” section and upload. Please see [Sequence preprocessing / Taxonomic classification / step 8](#).

\* **Sampling depth:**

Samples with total count smaller than set value will be dropped from the diversity analysis. The default value is the smallest total count among samples where no sample will be dropped.

\* **Number of threads:**

The number of threads to use for multithreaded process. The default value is all threads minus two.

\* **Start:**

Click on the “Start!” button to execute the analysis after above files and parameters have been uploaded and set.

- 7.1. Faith PD table: Faith PD (Faith's Phylogenetic Diversity) is a commonly used phylogenetic index. PD is a species diversity that considers genetic distance among species.

### Faith PD table

Show **10** entries

Search:

|    | SampleID | FaithPD          |
|----|----------|------------------|
| 1  | L1S105   | 7.03504527906064 |
| 2  | L1S140   | 6.81348963332276 |
| 3  | L1S208   | 7.56734619259508 |
| 4  | L1S257   | 6.85786737645975 |
| 5  | L1S281   | 6.75666778936291 |
| 6  | L1S57    | 6.63365160132782 |
| 7  | L1S76    | 6.33300452179527 |
| 8  | L1S8     | 5.36655055450142 |
| 9  | L2S155   | 18.3846874539932 |
| 10 | L2S175   | 16.260657028738  |

Showing 1 to 10 of 34 entries

Previous

1

2

3

4

Next

 Download Faith PD table

**Download**

\* Faith PD table:

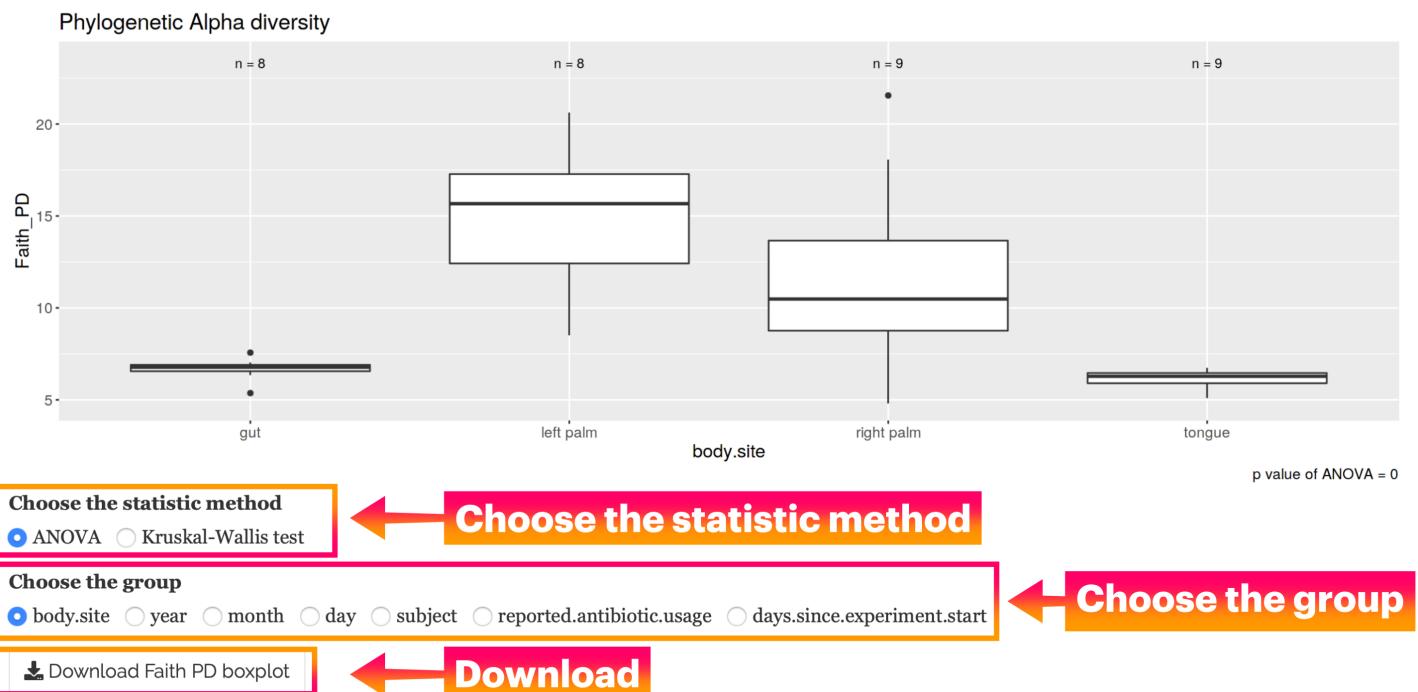
This table shows the Faith PD (phylogenetic diversity) of all samples.

\* Download Faith PD table:

Click on the “Download Faith PD table” button to download the table.

## 7.2. Faith PD boxplot: the distribution of Faith PD values using a boxplot.

### Faith PD boxplot



\* Choose the statistic method:

Select ANOVA (parametric method) or Kruskal-Wallis (nonparametric method) to test whether the distribution of Faith PD is significantly different among the groups.

\* Choose the group:

Faith PD will be grouped based on the selected metadata.

\* Download Faith PD boxplot:

Click on the “Download Faith PD Boxplot” button to download the boxplot.

### 7.3. Post-hoc analysis

#### Post hoc analysis

Tukey test

| Group A    | Group B    | Diff  | P value |
|------------|------------|-------|---------|
| tongue     | right palm | -5.60 | 0.01    |
| tongue     | left palm  | -8.65 | 0.00    |
| tongue     | gut        | -0.50 | 0.99    |
| right palm | left palm  | -3.05 | 0.30    |
| right palm | gut        | 5.10  | 0.03    |
| left palm  | gut        | 8.15  | 0.00    |

 Download Faith PD post hoc result

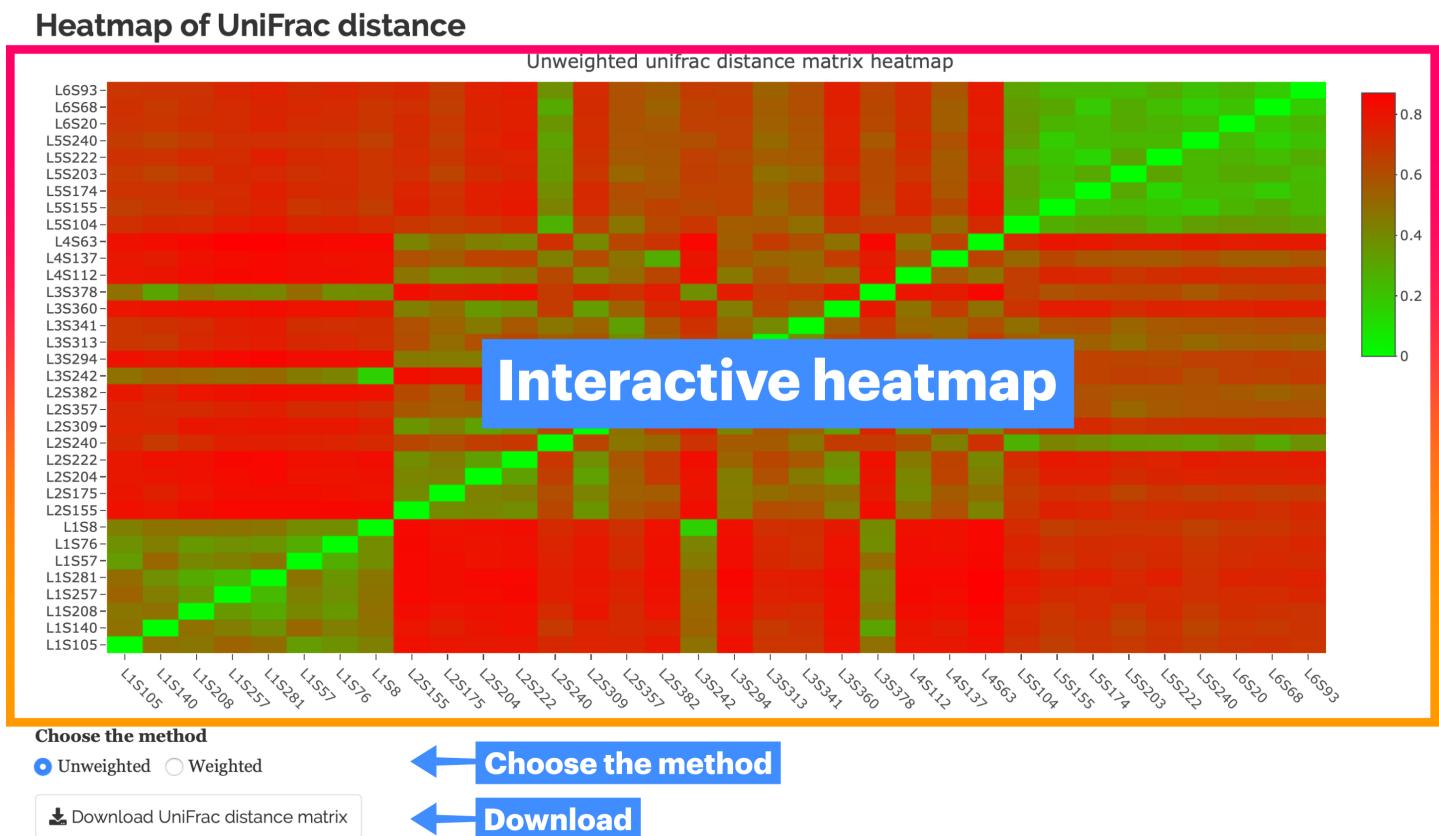
\* If ANOVA is selected when creating the Faith PD boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.

\* Download Faith PD post hoc result:

Click on the “Download Faith PD post hoc result” button to download the result.

---

## 7.4. Heatmap of UniFrac distance



\* **Interactive heatmap:**

When a cursor hovers over the heatmap, the information of species will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back. Click on the camera icon on the top-right region of the heatmap to download the plot.

\* **Choose the method:**

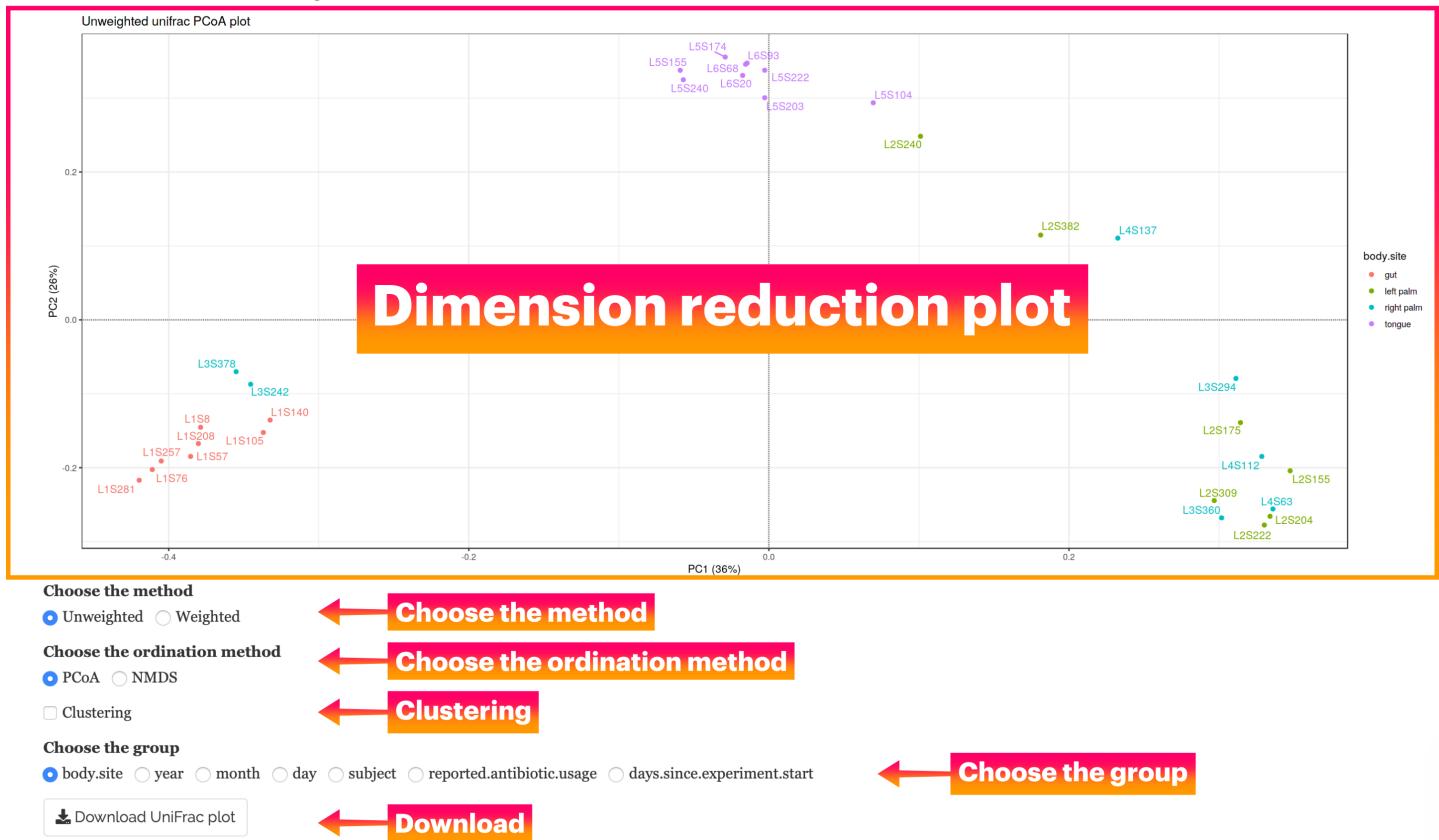
Select unweighted UniFrac (not consider the richness of taxa) or weighted UniFrac (consider the richness of taxa).

\* **Download heatmap matrix:**

Click on the “Download UniFrac distance matrix” button to download the matrix data.

## 7.5. Dimension-reduction plot of UniFrac distance

Dimension reduction plot of UniFrac distance



- \* We provide two dimension-reduction methods, including PCoA (Principal Co-ordinates Analysis) and NMDS (Non-metric Multidimensional Scaling) for visualizing UniFrac distance.
- \* Choose the method:  
Select unweighted UniFrac (not consider the richness of taxa) or weighted UniFrac (consider the richness of taxa).
- \* Choose the ordination method:  
Select a dimension reduction method.
- \* Clustering:  
When checking the box, samples in the same group will be surrounded by a circle.
- \* Choose the group:  
The samples in the plot will be labeled using colors based on the selected metadata.
- \* Download UniFrac plot:  
Click on the “Download UniFrac Plot” button to download the plot.

## 7.6. Statistical analysis

### Statistical analysis

| PerMANOVA      |         |
|----------------|---------|
| R <sup>2</sup> | P value |
| 0.1536         | 0.039   |

[Download PerMANOVA table](#)

| ANOSIM |         |
|--------|---------|
| R      | P value |
| 0.1171 | 0.041   |

[Download ANOSIM table](#)

| ANOSIM |               |              |         |
|--------|---------------|--------------|---------|
| A      | Observe.delta | Expect.delta | P value |
| 0.0458 | 0.5963        | 0.6249       | 0.028   |

[Download MRPP table](#)

| PerMANOVA pair         |                |         |
|------------------------|----------------|---------|
| Comparisons            | R <sup>2</sup> | P value |
| gut - left palm        | 0.5114         | 0.001   |
| gut - right palm       | 0.3407         | 0.002   |
| gut - tongue           | 0.6541         | 0.001   |
| left palm - right palm | 0.0598         | 0.488   |
| left palm - tongue     | 0.4782         | 0.001   |
| right palm - tongue    | 0.3359         | 0.002   |

[Download PerMANOVA pair table](#)

| ANOSIM pair            |         |         |
|------------------------|---------|---------|
| Comparisons            | R       | P value |
| gut - left palm        | 0.9989  | 0.002   |
| gut - right palm       | 0.5933  | 0.001   |
| gut - tongue           | 1       | 0.001   |
| left palm - right palm | -0.0395 | 0.664   |
| left palm - tongue     | 0.7956  | 0.001   |
| right palm - tongue    | 0.5442  | 0.001   |

[Download ANOSIM pair table](#)

| MRPP pair              |         |        |         |         |
|------------------------|---------|--------|---------|---------|
| Comparisons            | A       | delta  | E.delta | P value |
| gut - left palm        | 0.2681  | 0.4684 | 0.64    | 0.001   |
| gut - right palm       | 0.1624  | 0.5275 | 0.6298  | 0.003   |
| gut - tongue           | 0.3767  | 0.3298 | 0.5291  | 0.001   |
| left palm - right palm | -0.0024 | 0.5763 | 0.5749  | 0.527   |
| left palm - tongue     | 0.261   | 0.3785 | 0.5122  | 0.001   |
| right palm - tongue    | 0.2002  | 0.4393 | 0.5493  | 0.001   |

[Download MRPP pair table](#)

- \* We provide three statistical methods, including PerMANOVA (Permutational Multivariate Analysis of Variance), ANOSIM (Analysis of Similarities) and MRPP (Multiple Response Permutation Procedure), to test whether UniFrac distance is significantly different among groups or between pairs of groups.
- \* Download statistical result table:  
Click on the button below the table to download the statistical results.

## 8. ANCOM

Analyze composition of microbiomes. Used for comparing the composition of microbiomes in two or more populations. [Get more information](#).

The screenshot shows the QIIME 2 interface with the 'ANCOM' tab selected. At the top, there are several navigation links: Taxonomic table, Taxonomic barplot, Taxonomic heatmap, Krona, Alpha diversity, Beta diversity, Phylogenetic diversity, and ANCOM. Below these, a red box highlights the 'Choose the target comparison' section. Inside this section, the text 'Choose the target comparison (Should be categorical data)' is displayed, followed by a list of categories with radio buttons: body.site (selected), year, month, day, subject, reported.antibiotic.usage, and days.since.experiment.start. At the bottom left is a 'Start!' button, and at the bottom right is a larger 'Start' button.

\* Choose the target comparison:

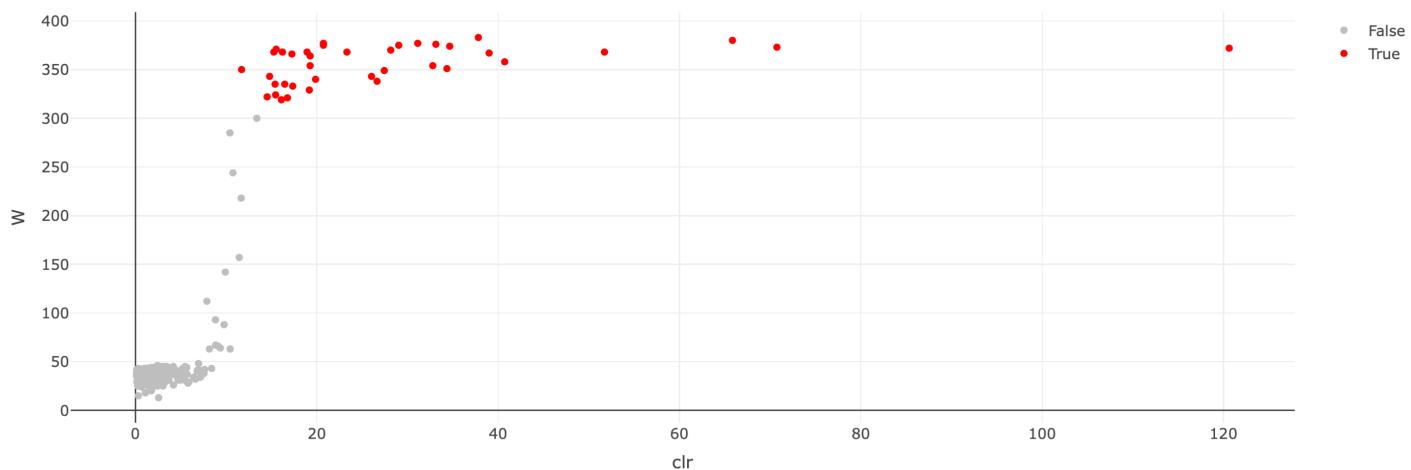
Select a group. ANCOM will then find significantly different abundant taxa among subgroups in that group.

\* Start:

Click on the “Start!” button to execute the analysis.

8.1. ANCOM volcano plot: An interactive plot shows the significantly different abundant taxa (red). When a cursor hovers over the dot, the taxa information will be presented.

## ANCOM Volcano Plot



The W value is essentially a count of the number of sub-hypotheses that have passed for a given species.

The clr represent log-fold change relative to the average microbe.

## 8.2. ANCOM statistical results: A table shows the W values for all taxa.

### ANCOM statistical results (Species with significant w value)

|    | Species  | W   |
|----|--|-----|
| 1  | k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;__                            | 383 |
| 2  | k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__uniformis                 | 380 |
| 3  | k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae | 377 |
| 4  | k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;__                           | 377 |
| 5  | k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__                          | 376 |
| 6  | k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__melaninogenica             | 375 |
| 7  | k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__            | 375 |
| 8  | k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium;s__prausnitzii             | 374 |
| 9  | k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira;s__                             | 373 |
| 10 | k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia;s__faecis                         | 372 |

Showing 1 to 10 of 38 entries

Previous 1 2 3 4 Next

[!\[\]\(6ef9aa63960241c7f0b6f0f9275edb17\_img.jpg\) Download the ANCOM result table \(Contain all species\)](#)

#### \* Download the ANCOM result table:

Click on the “Download the ANCOM result table” button to download the results. The table will contain the W values for all taxa.

# Function analysis

The database FAPROTAX is used to predict the function of microbiota.

## (A) Upload files

1. Select “**Function analysis**” on the top bar.



### Welcome to MOCHI! (Microbiota amplicOn CHaracterization Implement)

MOCHI is a 16S or 18S microbiota amplicon rRNA analytical tool for microbiota based primarily on QIIME2 with a friendly web interface powered by the R package of Shiny. MOCHI may also be downloaded and operated locally.

2. In the left panel, press the “**Browse**” buttons to upload metadata and taxonomic table. Alternatively, you could press the “Demo” button to download the example files first and then upload.

The image shows the MOCHI upload interface. On the left, there are two sections for uploading files: "Upload the metadata file" and "Upload the taxonomic table file". Each section has a "Browse..." button and a "No file selected" message. On the right, there are two buttons: "Function annotation table" and "Function barplot". At the bottom, there are three buttons: "Start!" (highlighted with a red box and a red arrow pointing to it), "Start" (with a red arrow pointing to it), and "Demo".

**FAPROTAX** is a database that maps prokaryotic clades (e.g. genera or species) to established metabolic or other ecologically relevant functions

Upload the metadata file ⓘ

Upload the taxonomic table file ⓘ

Function annotation table

Function barplot

Start!

Start

Demo

3. Click on the “**Start!**” button to conduct analysis.

## (B) Inspect result

### 1. Function annotation table

Display reads of the function types in every sample.

Function annotation table Function barplot

Summary

Summary

**Function annotation table**

Show 10  entries

| Type                                | L1S105 | L1S140 | L1S208 | L1S257 | L1S281 | L1S57 | L1S76 | L1S8 | L2S155 | L2S175 | L2S204 | L2S222 | L2S240 |
|-------------------------------------|--------|--------|--------|--------|--------|-------|-------|------|--------|--------|--------|--------|--------|
| 1 methanol_oxidation                | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 34     | 12     | 6      | 28     | 0      |
| 2 methylotrophy                     | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 34     | 12     | 6      | 28     | 0      |
| 3 aerobic_ammonia_oxidation         | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 0      | 0      | 0      | 0      | 0      |
| 4 nitrification                     | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 0      | 0      | 0      | 0      | 0      |
| 5 sulfate_respiration               | 79     | 40     |        |        |        |       |       |      |        |        | 0      | 2      | 3      |
| 6 sulfur_respiration                | 0      | 0      |        |        |        |       |       |      |        |        | 0      | 0      | 0      |
| 7 thiosulfate_respiration           | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 0      | 0      | 0      | 0      | 0      |
| 8 respiration_of_sulfur_compounds   | 79     | 40     | 21     | 27     | 24     | 35    | 29    | 62   | 0      | 0      | 2      | 3      | 0      |
| 9 arsenate_detoxification           | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 0      | 0      | 0      | 0      | 0      |
| 10 dissimilatory_arsenate_reduction | 0      | 0      | 0      | 0      | 0      | 0     | 0     | 0    | 0      | 0      | 0      | 0      | 0      |

Showing 1 to 10 of 54 entries

Previous 1 2 3 4 5 6 Next

[Download the function annotation table](#) [Download](#)

\* Summary: Essential information regarding the function prediction.

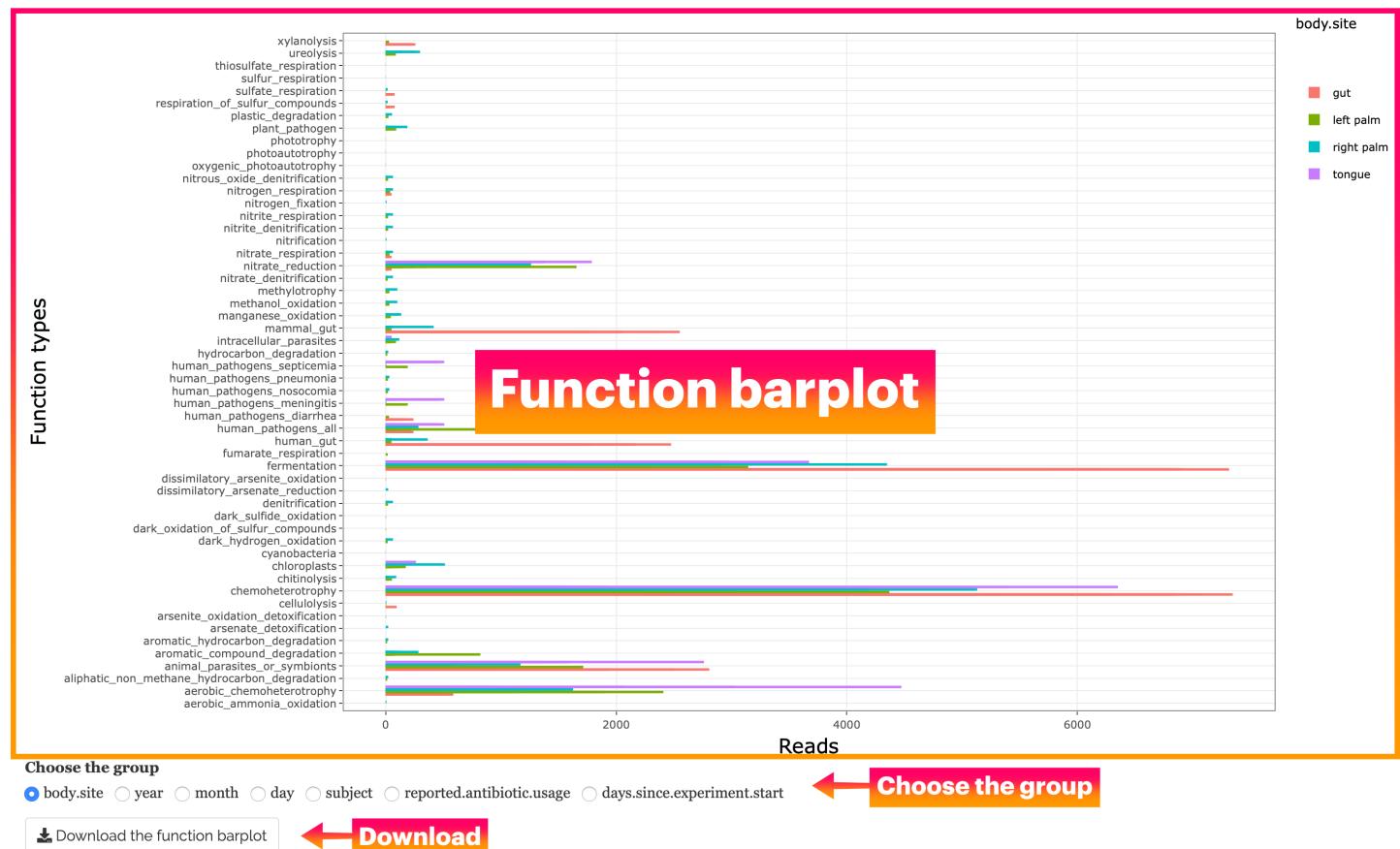
\* Download the function table:

Click on the “Download the function table” button to download the table.

## 2. Function barplot

The horizontal bars indicate reads of each function and are grouped based on the metadata.

### Function barplot



- \* Choose the group: The bars will be categorized based on the selected metadata.
- \* Download the function barplot:  
Click on the "Download the function barplot" button to download the barplot.