

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
Taxonomy analysis .....	11
(A) Upload files .....	11
(B) Inspect results.....	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 highlighted), Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar, a blue header bar says "Welcome to MOCHI". Underneath the header, there is a sub-menu for "Sequence Preprocessing" with three options: Step 1. Sequence summary (highlighted with a red arrow), Step 2. Sequence denoising, and Step 3. Taxonomy classification. A small note below the sub-menu says "MOCHI is a 16S or 18S microbiota amplicon rRNA analysis pipeline based 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” button to upload sequence files in the “fastq.gz” format. Alternatively, you could press the “Demo” button to download the example files and upload. The size limit is 20 MB per file.

This screenshot shows the "Sequence files" section of the MOCHI interface. It includes a "Job ID: 8a7395918d" field, a "Not your job id?" link, a "Sequence files" heading, and a "Browse..." button with "No file selected" text. Below the browse button is a "Demo" button.

### 1. Sequence summary (for paired end)

(1) Summarize the paired-end sequences.

**Start!**

Example output for sequence summary

Example for paired end

- \* The Job ID is used to retrieve analyzed results, which will be stored on our server for two weeks. To recall the data, select “User Results” on the top bar, and enter the JobID.

This screenshot shows the "User Results" section of the MOCHI interface. It features a "Show the results" heading and a note that results will be deleted after two weeks. Below this is a "Job ID" input field with placeholder text "Input your job id" and a "Submit" button.

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

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

**Sequence type**  
Choose the sequence type

## 1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

(2) Inspect the result.

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

Example output for sequence summary

4. 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

Choose the reverse primer sequence

5. Set the number of threads to run the analysis. If zero, 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 result for single/pair end” button to open a demo result.)

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

## 1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

7. Please wait while it is running. When complete, a popup window will be displayed.

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 are checkboxes for "Check this if your sequences are primer-trimmed reads" and "Choose the forward primer sequence". The main content area has a header with tabs: Home, Sequence Preprocessing, Taxonomy Analysis, Function Analysis, and Tutorial. Below the tabs, a "Successful!" message is displayed with the text "This analysis took 16.96 secs. You can click the button View to inspect the result." A red box highlights a "Please wait..." message with a circular loading icon. At the bottom, there's a link to "Summarize the single-end sequences".

8. Click on the “View!” button to view the result. ([Example output](#))

The screenshot shows the MQCHI web application. On the left, there's a sidebar with a "Demo" button and a "Sequence type" section with the text "Choose the sequence type". The main content area has a header with tabs: Home, Sequence Preprocessing, Taxonomy Analysis, Function Analysis, and Tutorial. Below the tabs, a section titled "(2) Inspect the result." contains two buttons: "View!" and "log file". The "View!" button is highlighted with a red box. Below the buttons, the text "On the View webpage, you can download the result by right click Save as ..." 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 is a blue navigation bar with tabs for Home, Sequence Preprocessing (which is currently selected), Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar is a section titled "Welcome to MOCHI". On the right side of this section, there is a dropdown menu with three options: "Step 1. Sequence summary", "Step 2. Sequence denoising", and "Step 3. Taxonomy classification". A red arrow points from the text "Select ‘Sequence denoising’ on the top bar" to the "Step 2. Sequence denoising" option in the dropdown menu. Below the dropdown menu, there is some descriptive text about MOCHI and its underlying technology.

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 obtain 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 two sections of the MOCHI interface. On the left, under "Sequence trimming", there are fields for "The start position" (set to 0) and "The end position" (set to 0). Below these are buttons for "learn more" and "Demo". On the right, under "Quality score filtering", there is a field for "Quality score threshold" (set to 0) and a button for "learn more". Red arrows point from the text "Set the parameter of position and quality score to trim the sequences." to both the "The start position" field and the "Quality score threshold" field.

This screenshot shows the "Sequence denoising (DADA2) for Single end" section. It includes a "Start!" button and a "Inspect the sequence denoising result" section with links for "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 buttons for "Summary table", "Seqs info", "Filter info", and "Alpha rarefaction". Red arrows point from the text "Set the parameter of position and quality score to trim the sequences." to the "Start!" button and the "Inspect the sequence denoising result" section.

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

### \* Chimeric reads filter:

A chimeric read is a sequence originated from multiple 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 the analysis. For most cases, 1 is the default value.

### \* Training error model:

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

### \* Threads:

Specify the number of threads 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 configuration interface with four main sections highlighted by red arrows:

- Chimeric reads filtering**: A section for inputting the minimum fold-change value, currently set to 1. It includes a "learn more" link.
- Error model training**: A section for inputting the number of reads used for training the error model, currently set to 1000000. It includes a "learn more" link.
- Metadata Integrating (optional)**: A section for uploading metadata files. It includes a "Browse..." button, a file selection dropdown showing "No file selected", and a "learn more" link.
- Computing setting**: A section for specifying the number of threads MOCHI can use, currently set to 0. It includes a "learn more" link.

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

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

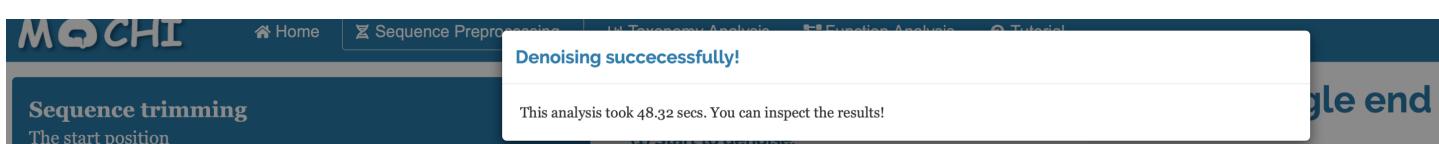
(1) Start to denoise.

**Start!**

5. Please wait while it is running. When complete, a popup window will be displayed.



6. Click on the button in the “result section” to view the results.  
(Example output: summary table, seqs info, filter info, alpha rarefaction)



## (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". Underneath the header, there is a sub-menu for "Sequence Preprocessing" with options: Step 1. Sequence summary, Step 2. Sequence denoising, Step 3. Taxonomy classification (which is highlighted with a red arrow), and a link to "Characterization Implement". A note at the bottom states: "MOCHI is a 16S or 18S microbiota amplicon rRNA analysis pipeline. MOCHI may also be downloaded and operated locally. Run QIIME2 with a friendly web interface powered by the R package of Shiny.".

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

The screenshot shows two pages from the MOCHI interface. On the left, the "Database" selection page has a title "Database" and a sub-instruction "Select the reference database for taxonomy classification." It features a dropdown menu labeled "Choose the database" with "Silva (Not detected)" selected, and a button "Auto download database" below it. Red arrows point from the text "Select a database from the drop-down menu" and "Follow the link" to these respective elements. On the right, the "3. Taxonomy classification" page has a title "3. Taxonomy classification" and a sub-instruction "(1) Classify taxonomy". It includes a "Start!" button and two links: "View!" and "log file". Below this, another section "(2) Inspect the taxonomy classification result." also has "View!" and "log file" links.

### 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 the network speed.

### 2.2. Manually download database:

- \* Silva: Follow the [link](#). Choose a version to 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 the reference sequence.

**\* Minimum and maximum length:**

Reference sequences not in range of 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**  
0

**Maximum length**  
0

[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  
0

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.

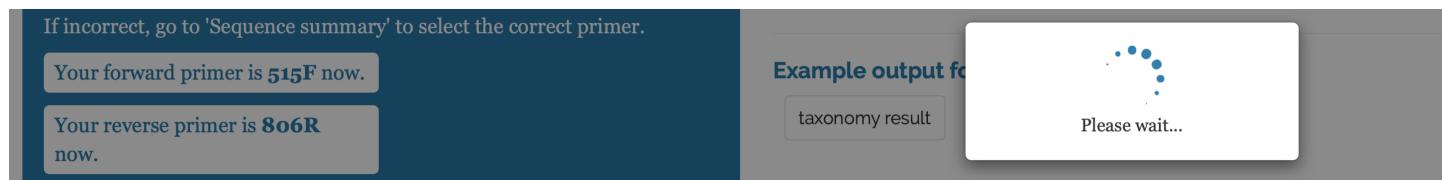
**Choose the database**  
Silva (Not detected) ▾

### 3. Taxonomy classification

(1) Classify taxonomy

**Start!**

7. Please wait while it is running. When complete, a popup window will be displayed.



**MQCHI** 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 the taxonomy analysis. ([Example output](#))

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](#)

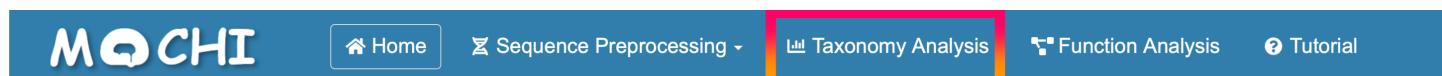
**(3) Download the files for the next step.**  
[The taxonomic table](#) [The ASVs table](#) [The seqs data](#)

**Example output for taxonomy classification**  
taxonomy result

# Taxonomy analysis

## (A) Upload files

1. Select “**Taxonomy 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” button to upload **metadata, taxonomic table** and **ASVs table** files. These files can be downloaded from the “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 them. 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.

This screenshot shows the 'Upload files' section of the MOCHI interface. It contains three main input fields: 'Upload the metadata file', 'Upload the taxonomic table file', and 'Upload the ASVs table file'. Each field has a 'Browse...' button and a 'No file selected' message. Below these is a checkbox for '18S rRNA' and a 'Demo' button. To the right of the input fields are four 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 results

- 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.

Taxonomy information							Read count					
Kingdom (K=3)	Phylum (K=21)	Class (K=42)	Order (K=69)	Family (K=134)	Genus (K=281)	Species (K=356)	LiS8	LiS57	LiS76	LiS105	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

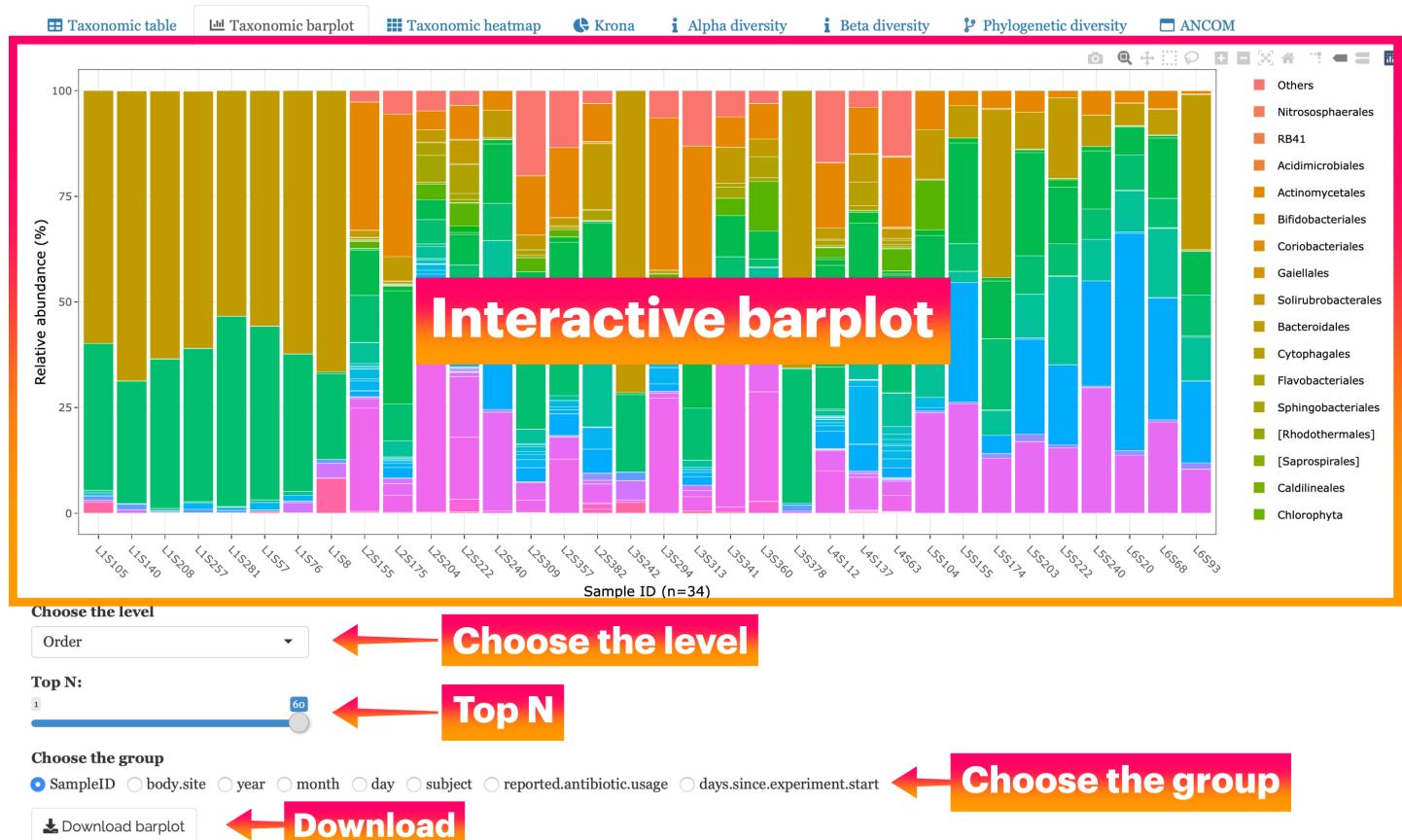
Showing 1 to 10 of 385 entries

Previous 1 2 3 4 5 ... 39 Next

- \* 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 samples. To prevent taking logarithm of zero, a small value of 0.01 is added to all percentage values before the 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:

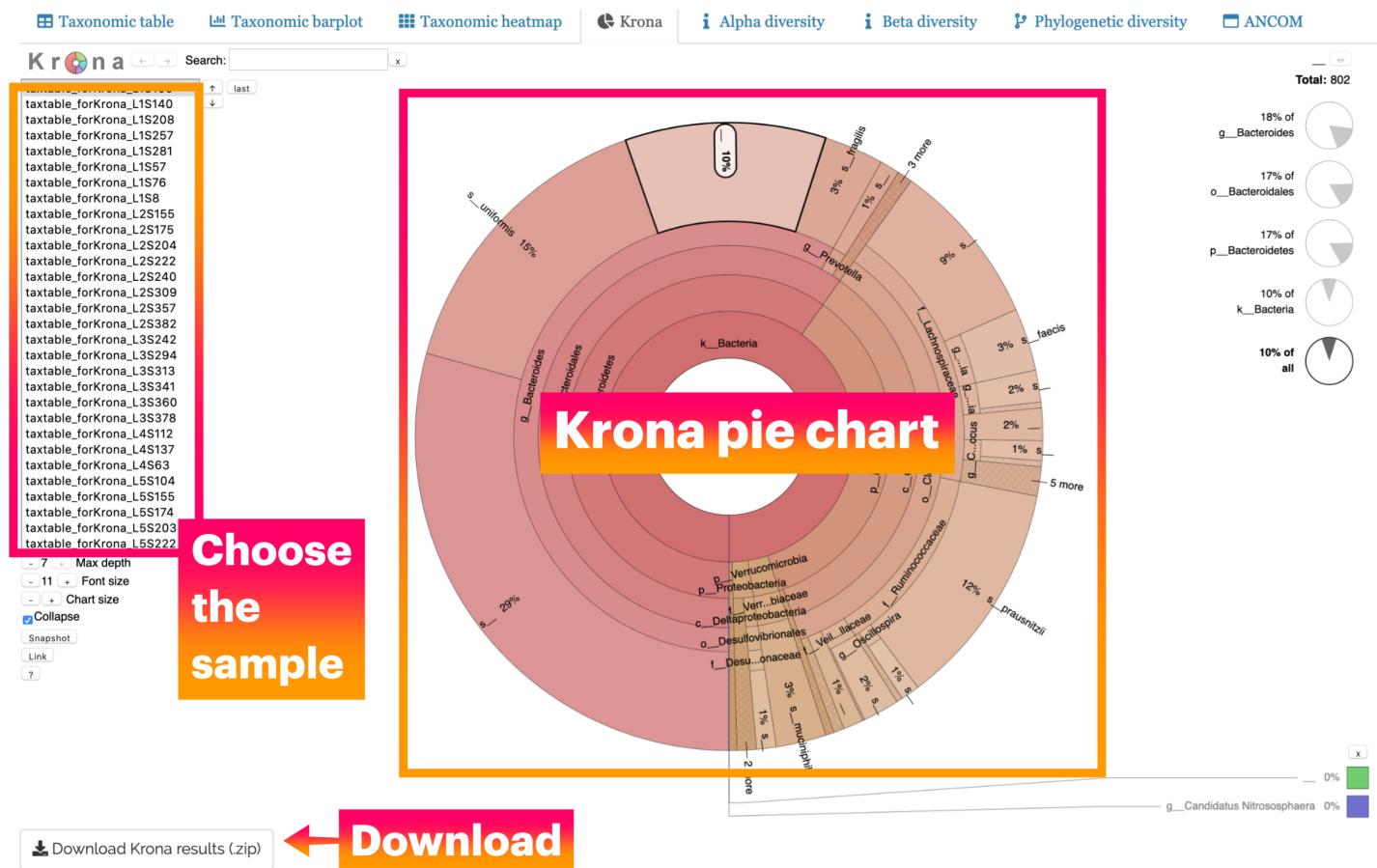
The heatmap will be grouped based on the selected metadata.

#### \* 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 Kona results:

Click on the “Download Kona 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

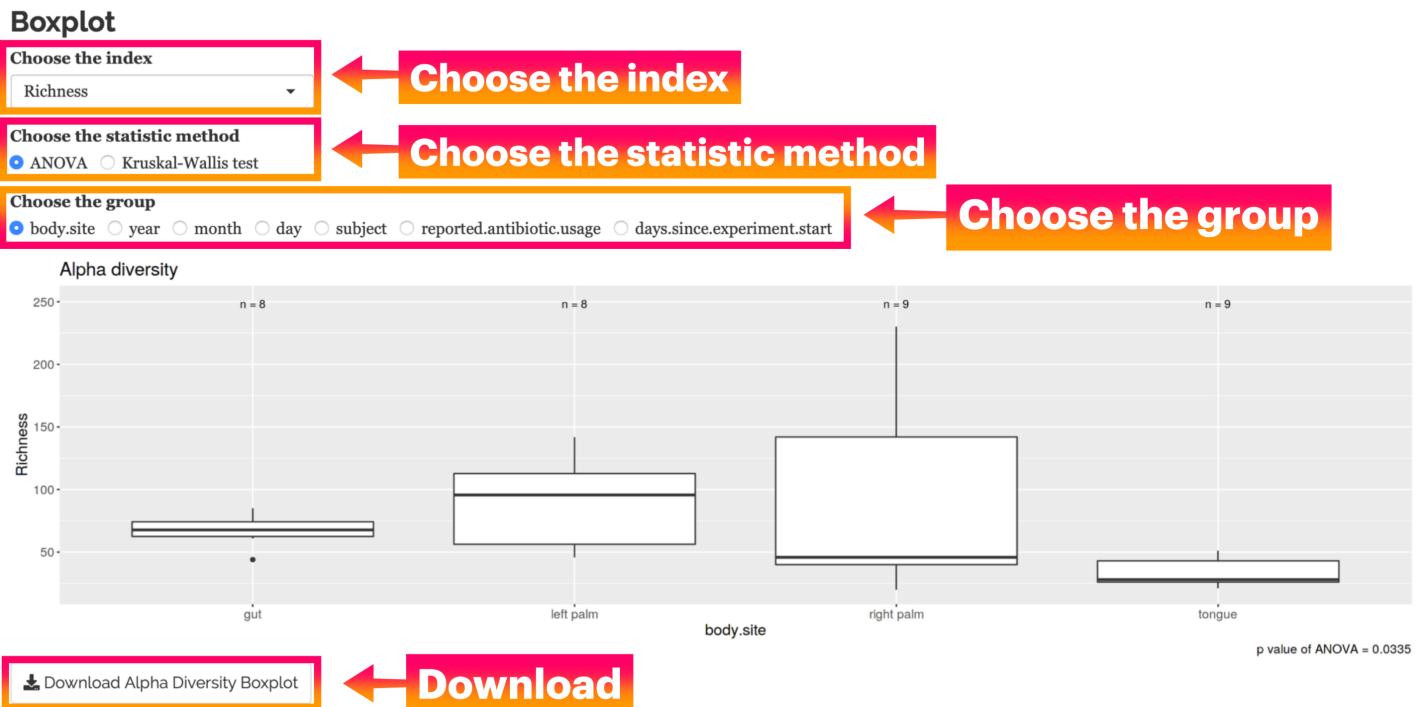
#### \* Alpha diversity table:

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 (a parametric method) or Kruskal-Wallis (a 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

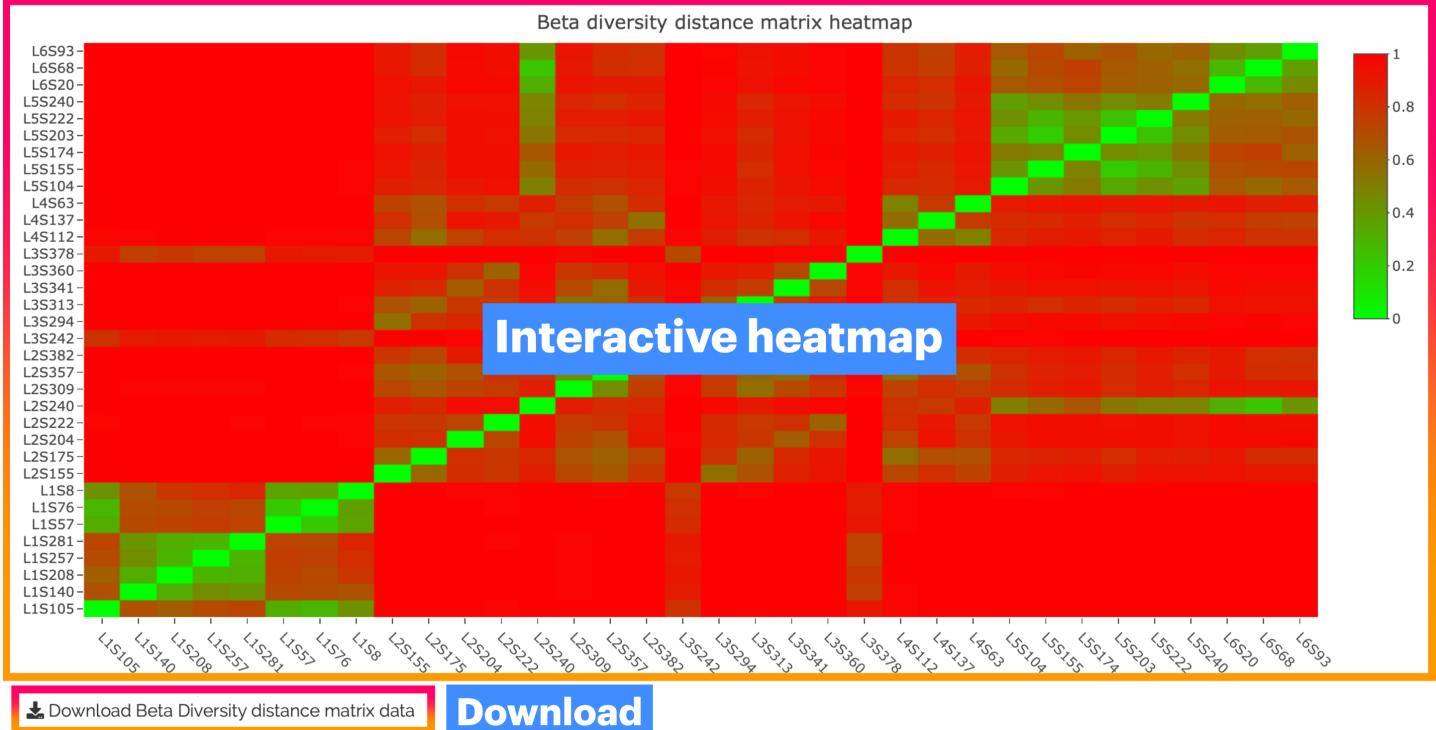
- \* If ANOVA is selected when creating the boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.
- \* Download Alpha diversity post-hoc test result:  
Click on the “Download Alpha Diversity statistical result” button to download the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



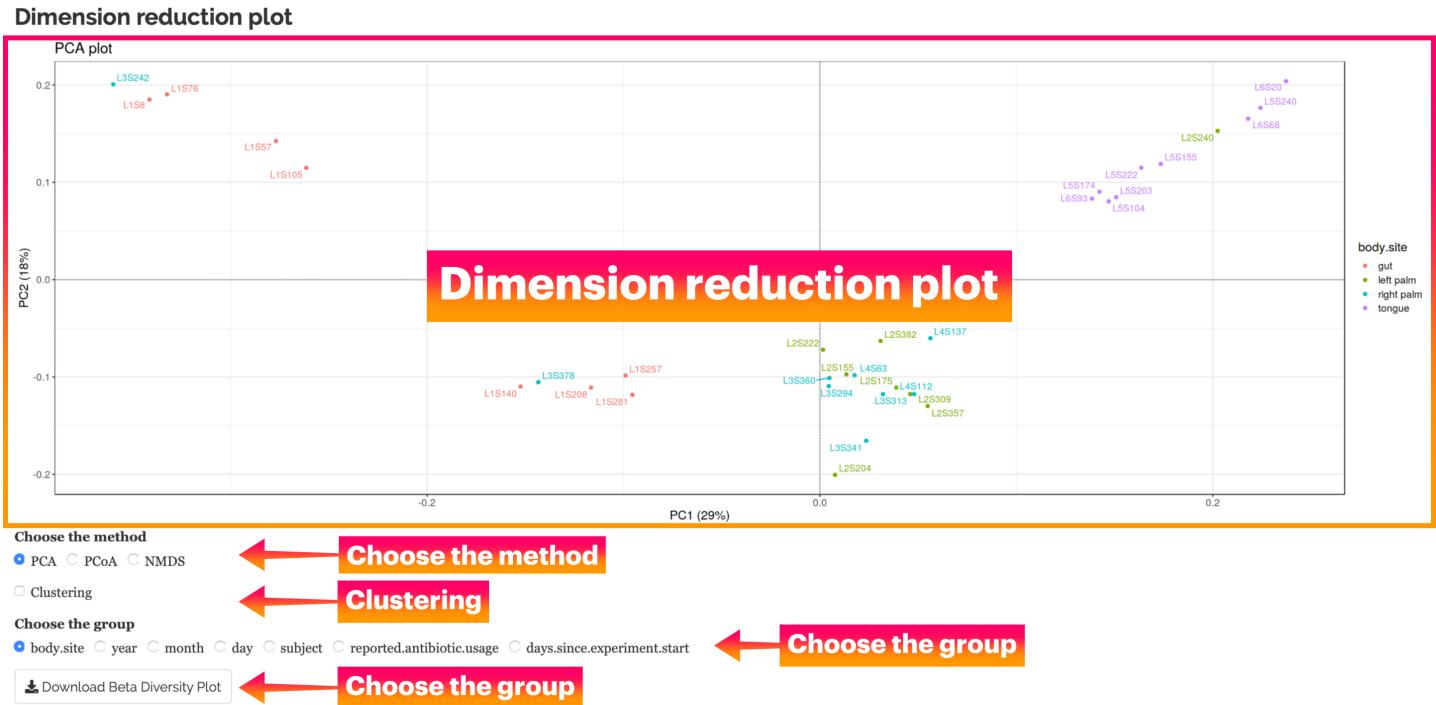
\* Interactive heatmap:

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 the table of statistical results:

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 i

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

Input the sampling depth

898  
learn more

Input the 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 the 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  2 3 4 Next

 Download Faith PD table

**Download**

\* Faith PD table:

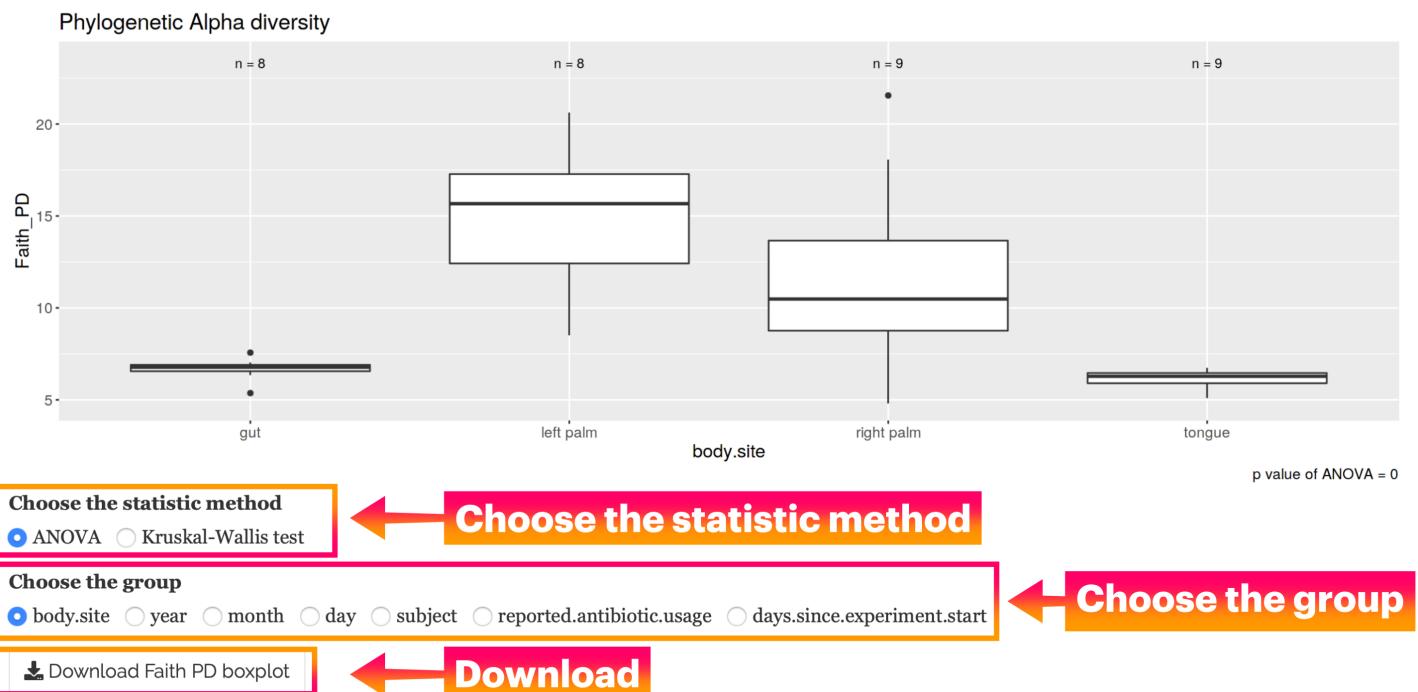
This table shows the Faith PD (phylogenetic diversity) of all samples.

\* 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

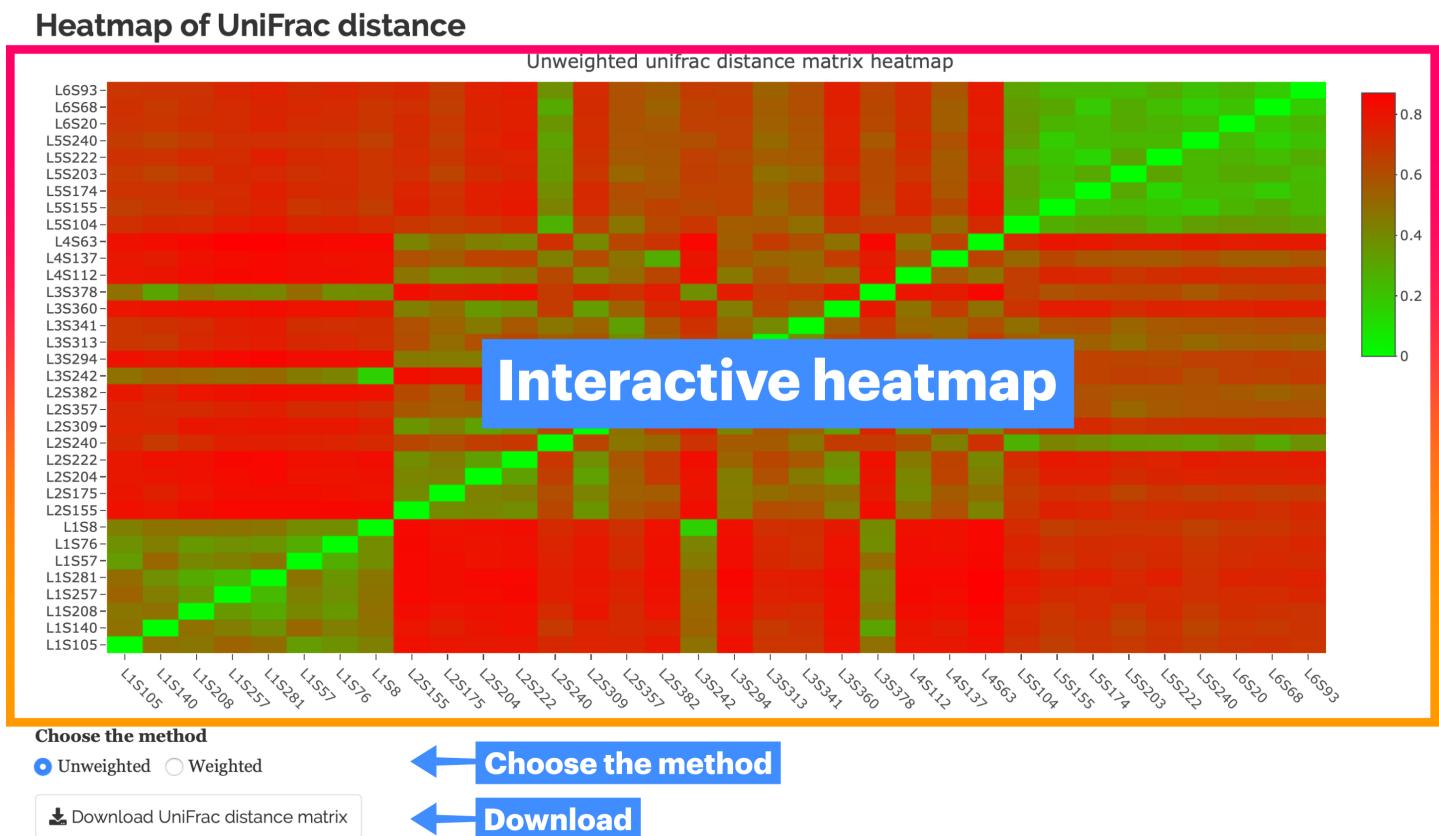
\* If ANOVA is selected when creating the Faith PD boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.

\* 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:**

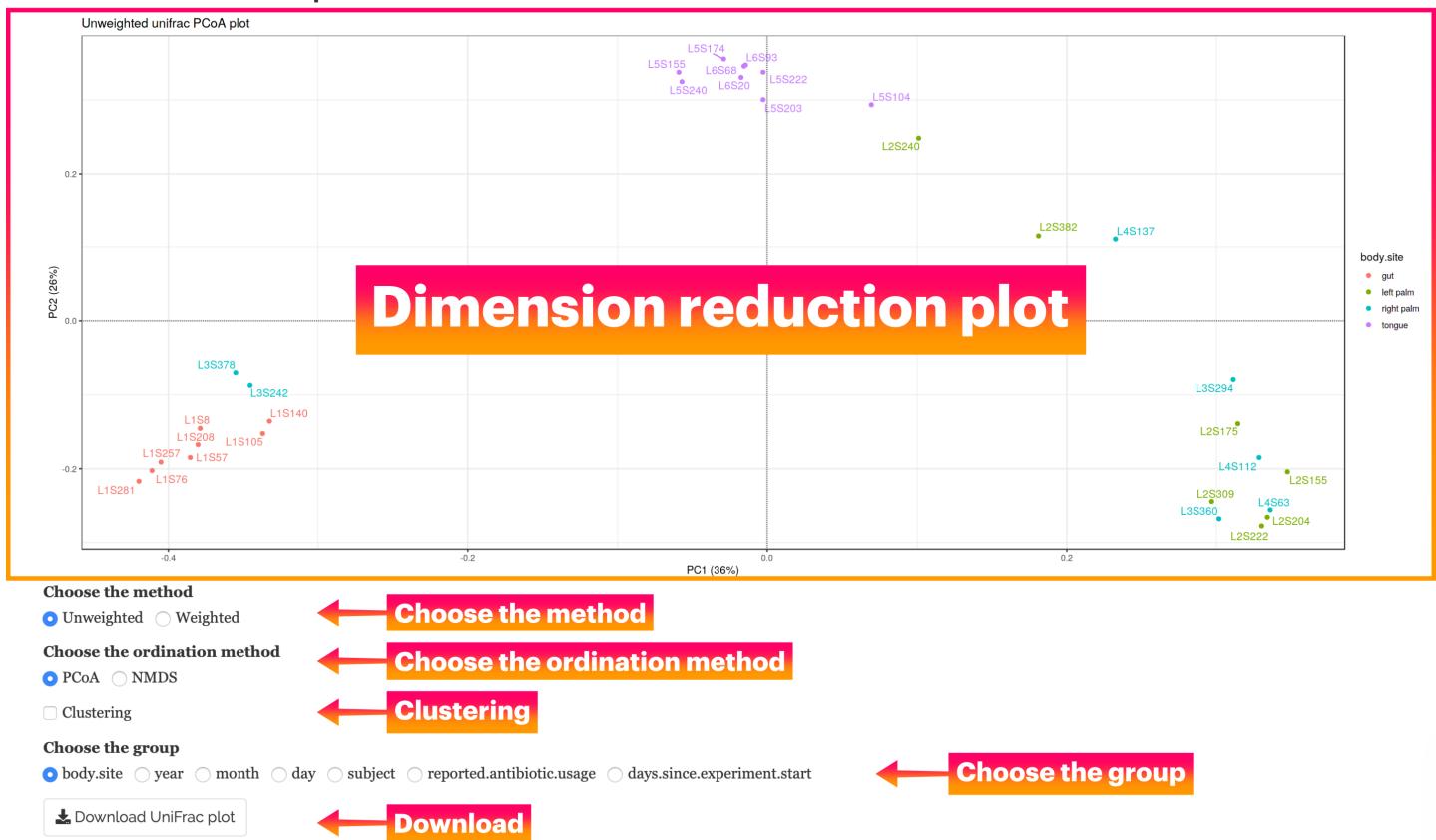
Select unweighted UniFrac (not consider the richness of taxa) or weighted UniFrac (consider the richness of taxa).

\* **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 QIIME2 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. A red box highlights the 'Choose the target comparison' section. Below it, a list of categorical data options is shown with 'body.site' selected. Two 'Start!' buttons are present: one in a grey box and one in a red box.

Choose the target comparison (Should be categorical data)

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

Start! Start

\* Choose the target comparison:

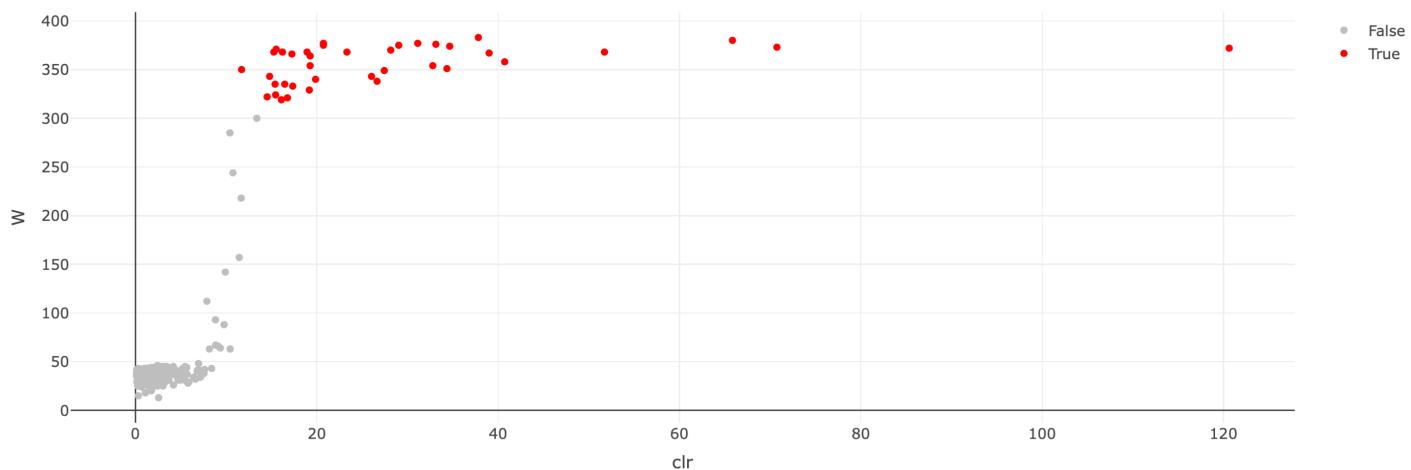
Select a group. ANCOM will then find significantly different abundant taxa among subgroups in that group.

\* 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 is a descriptive text box about FAPROTAX and two file upload sections: "Upload the metadata file" and "Upload the taxonomic table file". Each section has a "Browse..." button and a "No file selected" placeholder. To the right of these sections is a "Start!" button (highlighted with a red box and a red arrow pointing to it from below) and a "Demo" button. Above the "Start!" button is a large orange "Upload" button. To the right of the upload buttons are two preview options: "Function annotation table" and "Function barplot".

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

## (B) Inspect result

### 1. Function annotation table

Display reads of the function types in every sample.

Function annotation table  Function barplot

**Summary**  
0 out of 770 taxa (0 %) were assigned to at least one function type  
770 out of 770 taxa (100 %) could not be assigned to any function type  
0 function types were represented (i.e. associated with at least one taxa)

**Summary**

**Function annotation table**

Show 10  entries Search:

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	0
6 sulfur_respiration	0	0									0	0	0	0
7 thiosulfate_respiration	0	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	0
9 arsenate_detoxification	0	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	0

Showing 1 to 10 of 54 entries Previous  2 3 4 5 6 Next

\* Summary: Essential information regarding the function prediction.

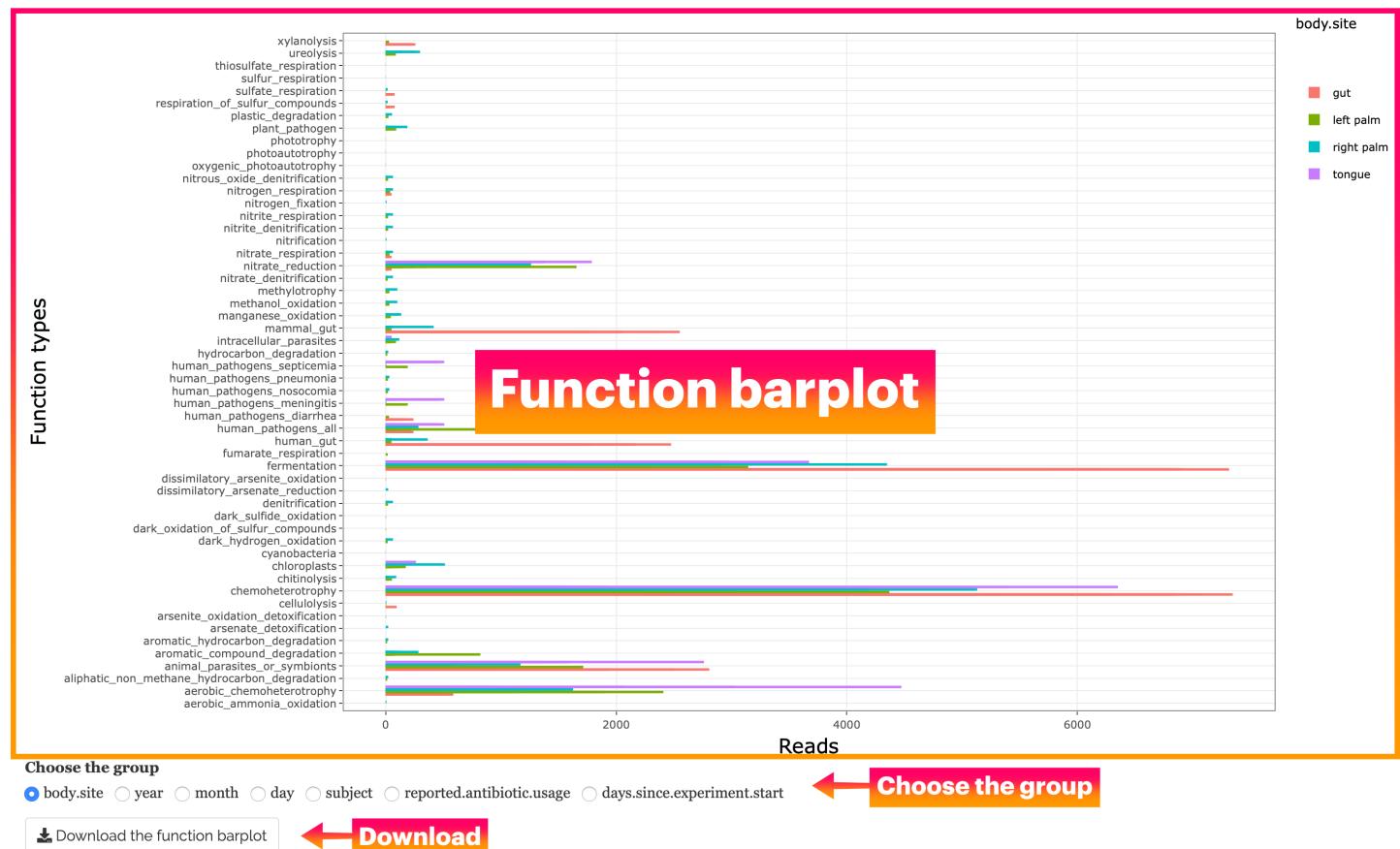
\* Download the function table:

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.