

Microbiota amplicOn
CHaracterization Implement

MOCHI

User Guide

TABLE OF CONTENTS

Chpater 2: 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	35
1. Function annotation table	35
2. Function barplot	36

CHAPTER 2: 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), 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 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 the selection window. Alternatively, you could press the “Demo” button to download the example files.

The screenshot shows the "Sequence files" section of the MOCHI interface. It asks to choose a directory containing sequence files (*.fastq.gz or *.fq.gz). There is a "Select the directory" button with a red box around it, and a "Demo" button below it. To the right, there is a section titled "1. Sequences summary (for single end)" with sub-sections "(1) Summarize the single-end sequences." and "(2) Inspect the result." with "Start!", "View!", "Download", and "log file" buttons.

! Important: The sequence files should be placed under “seqs_folder” where MOCHI was installed.

3. Select a directory containing sequence files.

The screenshot shows a "Please select a directory" dialog box. It has a search bar at the top with "raw_data" and buttons for "Create new folder", "Sort content", and a pencil icon. Below is a tree view of directories under "raw_data": "raw_data" and "single_seqs_demo". A red box highlights the "single_seqs_demo" folder, with a red arrow pointing to it labeled "Select folder". To the right is a "Content" table listing files with their sizes. At the bottom are "Confirm selection" and "Cancel" buttons, with a red box around the "Select" button.

	Content	
1	L1S105_9_L001_R1_001.fastq.gz	889.1 kB
2	L1S140_6_L001_R1_001.fastq.gz	675.3 kB
3	L1S208_10_L001_R1_001.fastq.gz	861.9 kB
4	L1S257_11_L001_R1_001.fastq.gz	633.7 kB
5	L1S281_5_L001_R1_001.fastq.gz	685.7 kB
6	L1S57_13_L001_R1_001.fastq.gz	896.4 kB
7	L1S76_12_L001_R1_001.fastq.gz	767.0 kB
8	L1S8_8_L001_R1_001.fastq.gz	887.4 kB
9	L2S155_25_L001_R1_001.fastq.gz	770.0 kB
10	L2S175_27_L001_R1_001.fastq.gz	873.6 kB
11	L2S204_1_L001_R1_001.fastq.gz	615.1 kB
12	L2S222_23_L001_R1_001.fastq.gz	736.1 kB
13	L2S240_7_L001_R1_001.fastq.gz	929.2 kB
14	L2S309_33_L001_R1_001.fastq.gz	314.7 kB
15	L2S357_15_L001_R1_001.fastq.gz	502.5 kB
16	L2S382_34_L001_R1_001.fastq.gz	792.2 kB
17	L3S242_19_L001_R1_001.fastq.gz	149.3 kB

4. **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

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

Choose the reverse primer sequence

6. 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

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)

1. Sequences summary (for single end)

(1) Summarize the single-end sequences.

8. 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 in a box, along with a note: "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 in a separate window.

9. 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. 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. A note below the buttons says: "On the View webpage, you can download the result by right click Save as ...".

(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, the main content area has a heading "Welcome to MOCHI". To the right of the heading, there is a vertical list of steps: Step 1. Sequence summary, Step 2. Sequence denoising, and Step 3. Taxonomy classification. A red arrow points from the text "Set the position and the quality score to trim the sequences." to the "Step 2. Sequence denoising" link.

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 the "Sequence trimming" section of the MOCHI interface. It includes fields for "The start position" (set to 0) and "The end position" (set to 0). There are also "learn more" and "Demo" buttons. Two orange arrows point to the "The start position" and "The end position" fields respectively.

2. Sequence denoising (DADA2) for Single end

(1) Start to denoise.

Start!

(2) Inspect the sequence denoising result.

Show summary table Show seqs info Show filter info Show alpha rarefaction log file

Example output for sequence denoising

Summary table Seqs info Filter info Alpha rarefaction

3. Set the parameter of **chimera**, **computing setting** 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.

* Metadata integrating (optional):

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

* Computing setting:

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.

The screenshot shows three main sections of the MOCHI interface:

- Chimeric reads filtering:** A text input field containing "1". A red arrow points to the left of this section.
- Metadata Integrating (optional):** A file upload input field with "No file selected". A red arrow points to the right of this section.
- Computing setting:** A text input field containing "6". A red arrow points to the left of this section.

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.)

This screenshot shows the Sequence trimming section of the MOCHI interface. It includes a text input field for "The start position" containing "5".

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.

This screenshot shows the Quality score filtering section of the MOCHI interface. It includes a text input field for "Quality score threshold" containing "0". To the right, a progress bar indicates "Please wait..." with a circular loading icon.

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

This screenshot shows the MOCHI interface after the analysis is completed. A success message "Denoising successfully!" is displayed in a white box. Below it, a message states "This analysis took 48.32 secs. You can inspect the results!". The navigation bar at the top includes Home, Sequence Preprocessing, Taxonomy Analysis, Function Analysis, and Tutorial.

(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 navigation bar down 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 "Choose the database" dropdown menu currently set to "Silva (Not detected)" and an "Auto download database" button below it. A red arrow points from the "Auto download database" button towards the "Database" selection page. At the bottom, there 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 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

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.

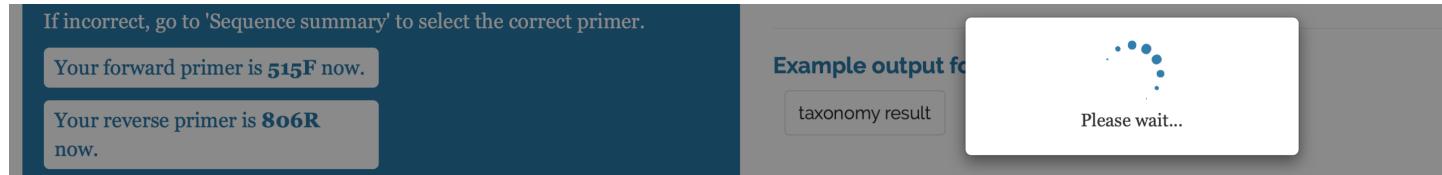
Choose the database

3. Taxonomy classification

(1) Classify taxonomy

Start!

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



MQCHI Home Sequence Preprocessing Analysis Taxonomy Classification Database

Taxonomy classification has been finished!
This analysis took 55.92 secs. You can inspect the results!

Database
Select the reference database for taxonomy classification.

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

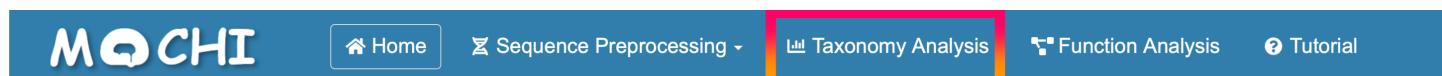
(3) Download the files for the next step.
The taxonomic table The ASVs table The seqs data

Example output for taxonomy classification
taxonomy result

Taxonomy analysis

(A) Upload files

9. 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.

10. 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.

The image shows the left panel of the MOCHI web interface, which contains three main sections for file uploads: "Upload the metadata file", "Upload the taxonomic table file", and "Upload the ASVs table file". Each section has a "Browse..." button and a "No file selected" placeholder. Below these sections is a checkbox for "18S rRNA" with an information icon. At the bottom are two buttons: "Start!" and "reset", and a "Demo" button.

- Click on the “**Start!**” button to run analysis. (Or, click on the “**reset**” button to re-upload the files).
- The results will be displayed on the right panel once the files are uploaded.

Upload the metadata file i

Browse... demo_metadata_single.tsv
Upload complete

Upload the taxonomic table file i

Browse... Taxonomic_table_example.qza
Upload complete

Upload the ASVs table file i

Browse... ASVs_table_example.qza
Upload complete

■ 18S rRNA i

● Start! ■ reset ■ Demo

■ Taxonomic table ■ Taxonomic barplot ■■■ Taxonomic heatmap ● Krona i Alpha diversity i Beta diversity p Phylogenetic diversity

■ ANCOM

Choose the group

SampleID 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
Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	0	0	0	0
Archaea	Crenarchaeota	Thaumarchaeota	Nitrosphaerales	Nitrosphaeraceae	Candidatus Nitrosphaera	Unassigned	0	0	0	0
Archaea	Crenarchaeota	Thaumarchaeota	Nitrosphaerales	Nitrosphaeraceae	Candidatus Nitrosphaera	SCA1145	0	0	0	0
Bacteria	Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	0	0	0	0
Bacteria	Acidobacteria	[Chloracidobacteria] RB41	Ellin6075	Unassigned	Unassigned	Unassigned	0	0	0	0
Bacteria	Actinobacteria	Acidimicrobia	Acidimicrobiales	Unassigned	Unassigned	Unassigned	0	0	0	0
Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Unassigned	Unassigned	Unassigned	0	0	0	0
Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	Unassigned	0	0	0	0
Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	Unassigned	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.

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

[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:

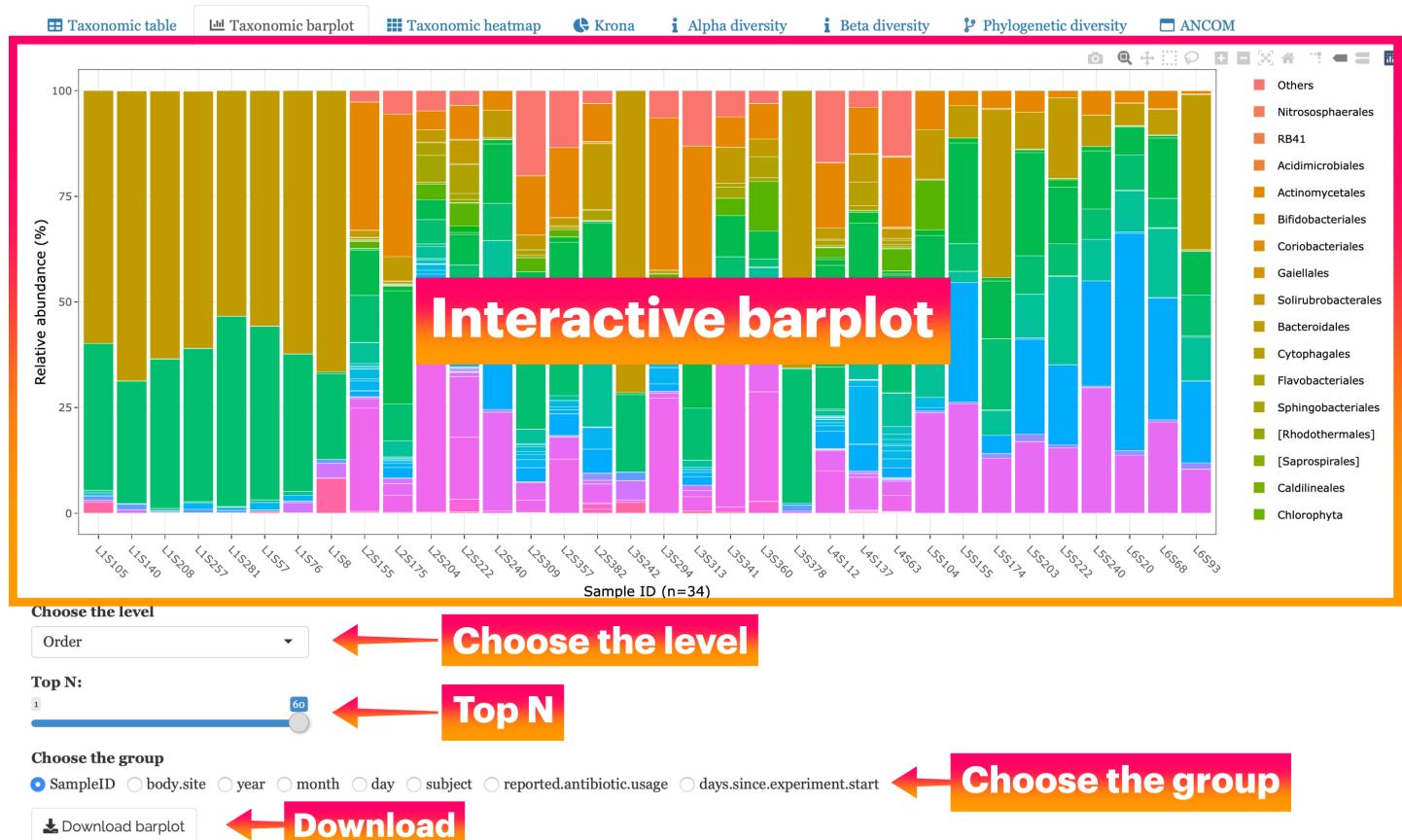
The right part of the table represents read count. The values of read count are categorized by the selected group. The column name indicates the variables of the selected group.

* 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 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:

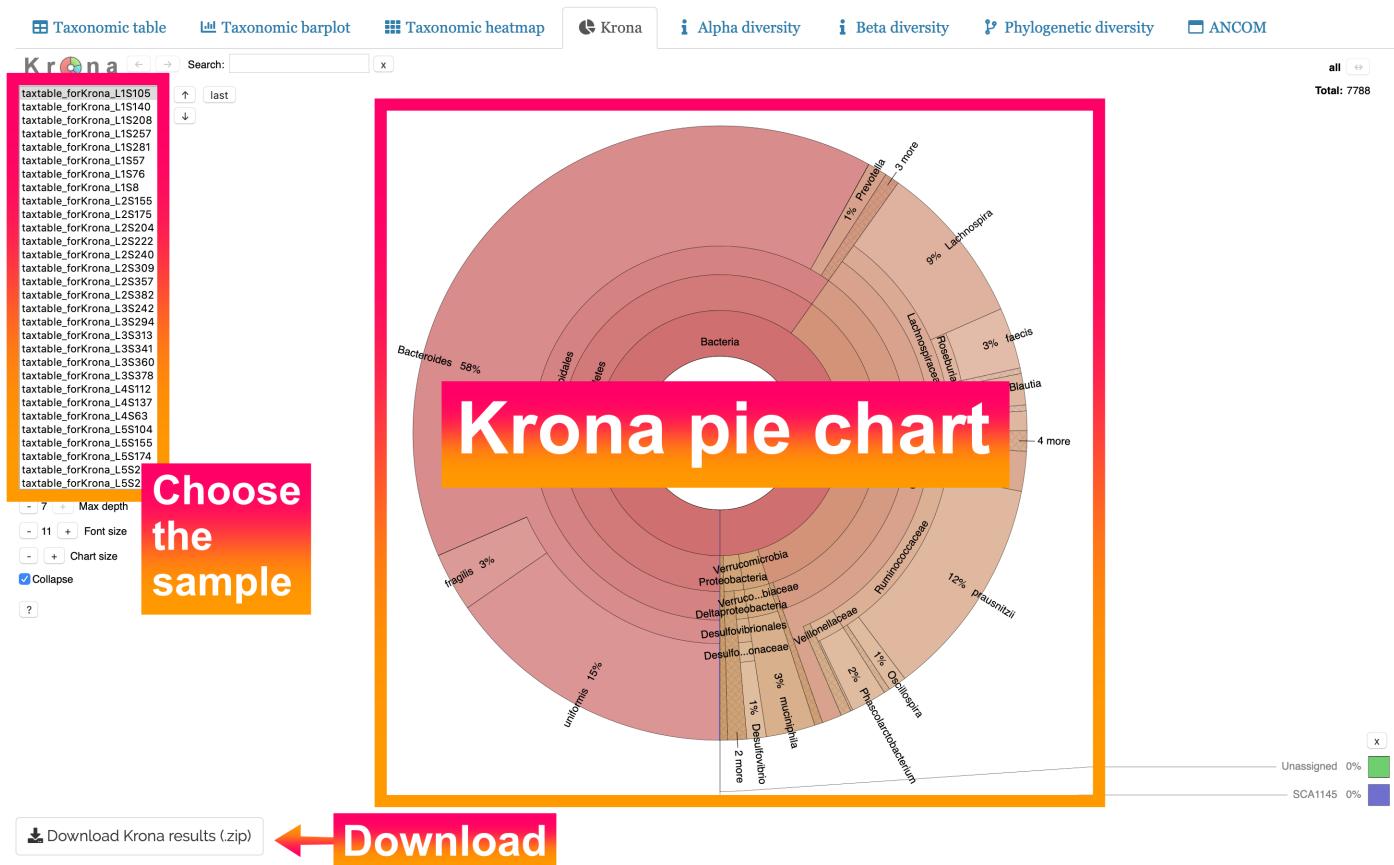
The heatmap will be grouped based on the selected metadata.

* 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 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](#)

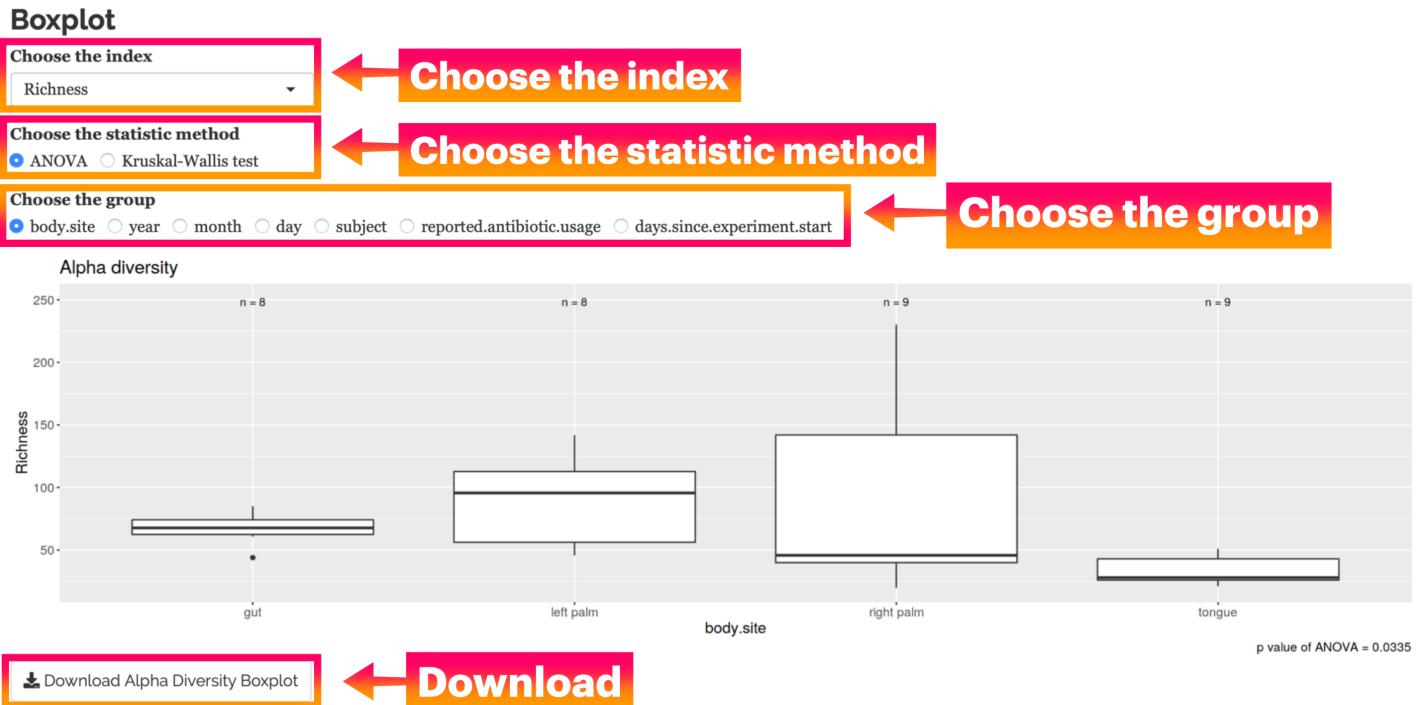
* Alpha diversity table:

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 (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

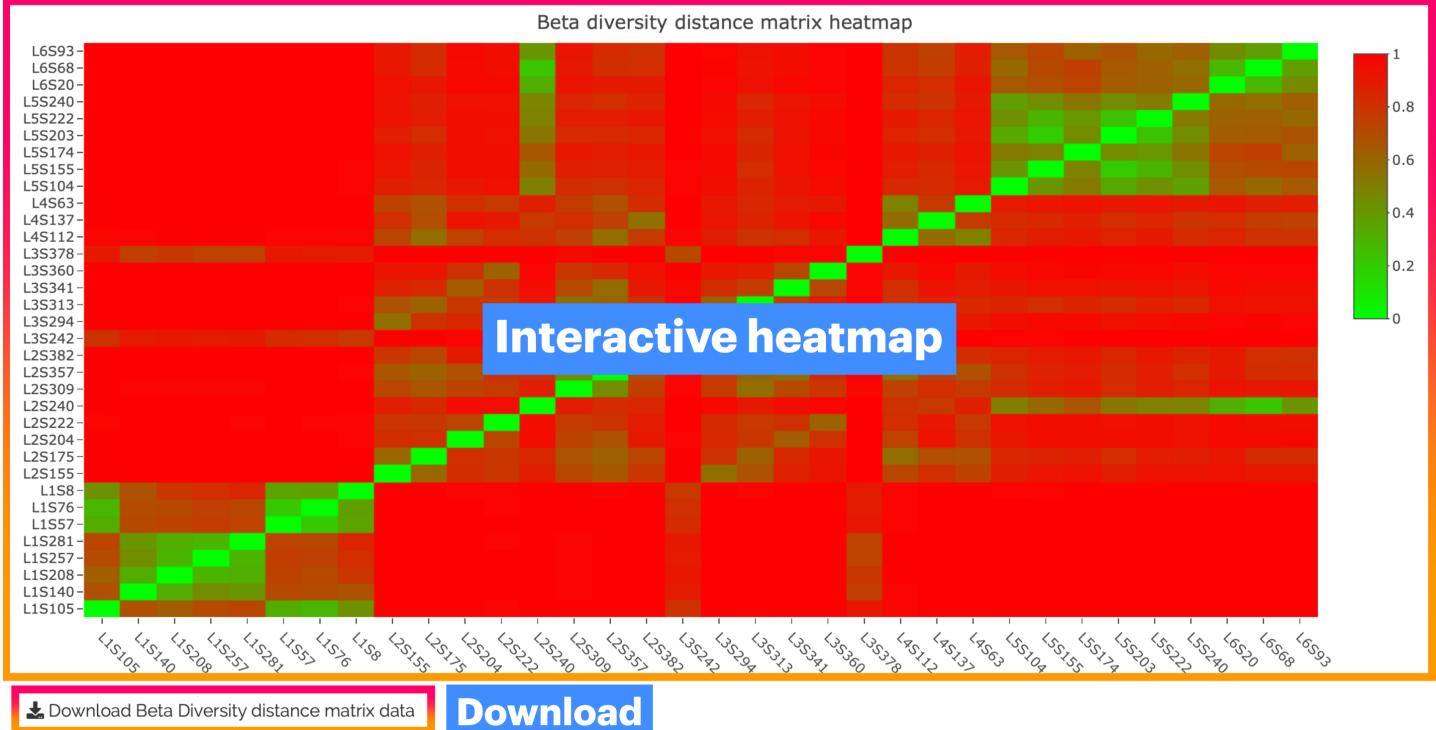
- * If ANOVA is selected when creating the boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.
- * Download Alpha diversity post-hoc test result:
Click on the “Download Alpha Diversity statistical result” button to download the post-hoc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

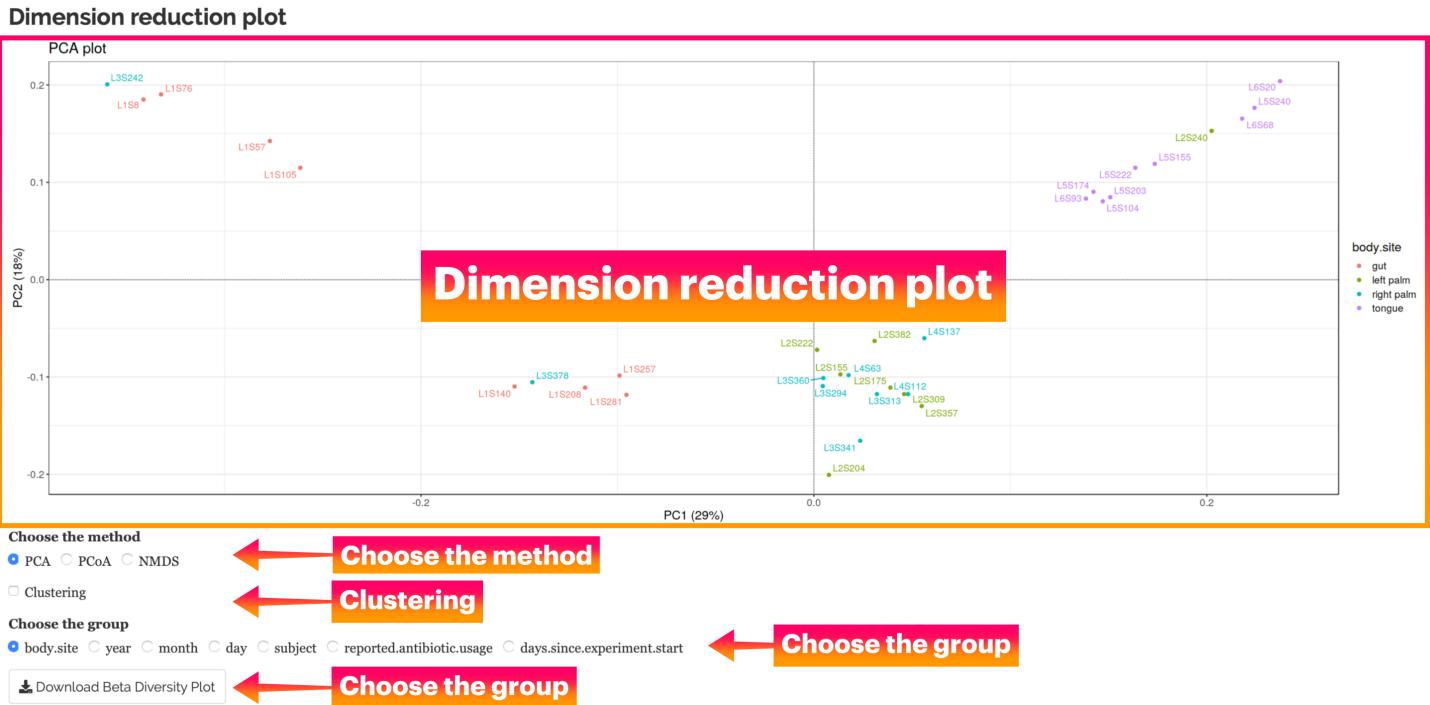
When a cursor hovers over the heatmap, the distance between species will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back. Click on the camera icon on the top-right region of the heatmap to download the plot.

* Download distance matrix:

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

* The values shown in the heatmap are the original values plus 0.01 and then natural log transformed.

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 ²	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 ²	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).

Phylogenetic diversity is a measure of diversity that take the genetic distance between species into consideration.

Upload the sequence file ?

Browse... No file selected

Sampling depth

898

learn more

Number of threads MOCHI can use

1

The default value is (number of threads on the system -2).

Start!

Demo

* Upload the sequence file:

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:

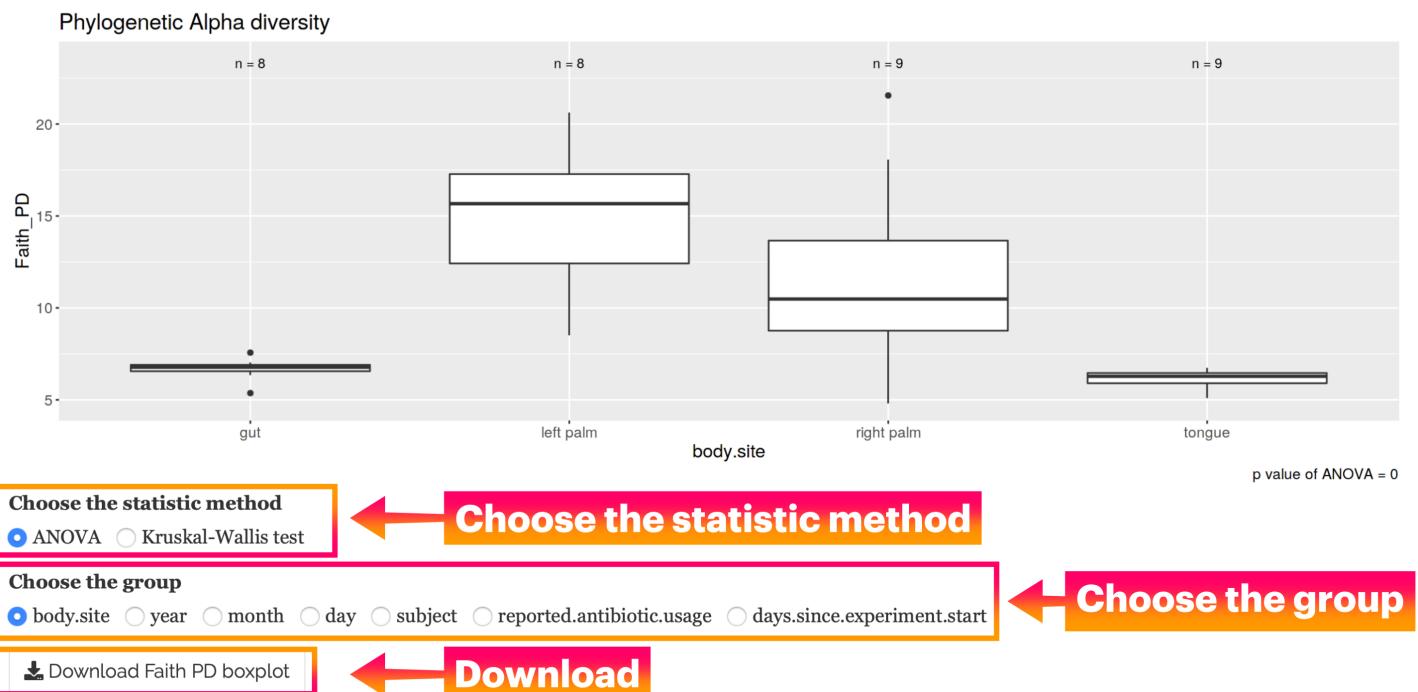
This table shows the Faith PD (phylogenetic diversity) of all samples.

* 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

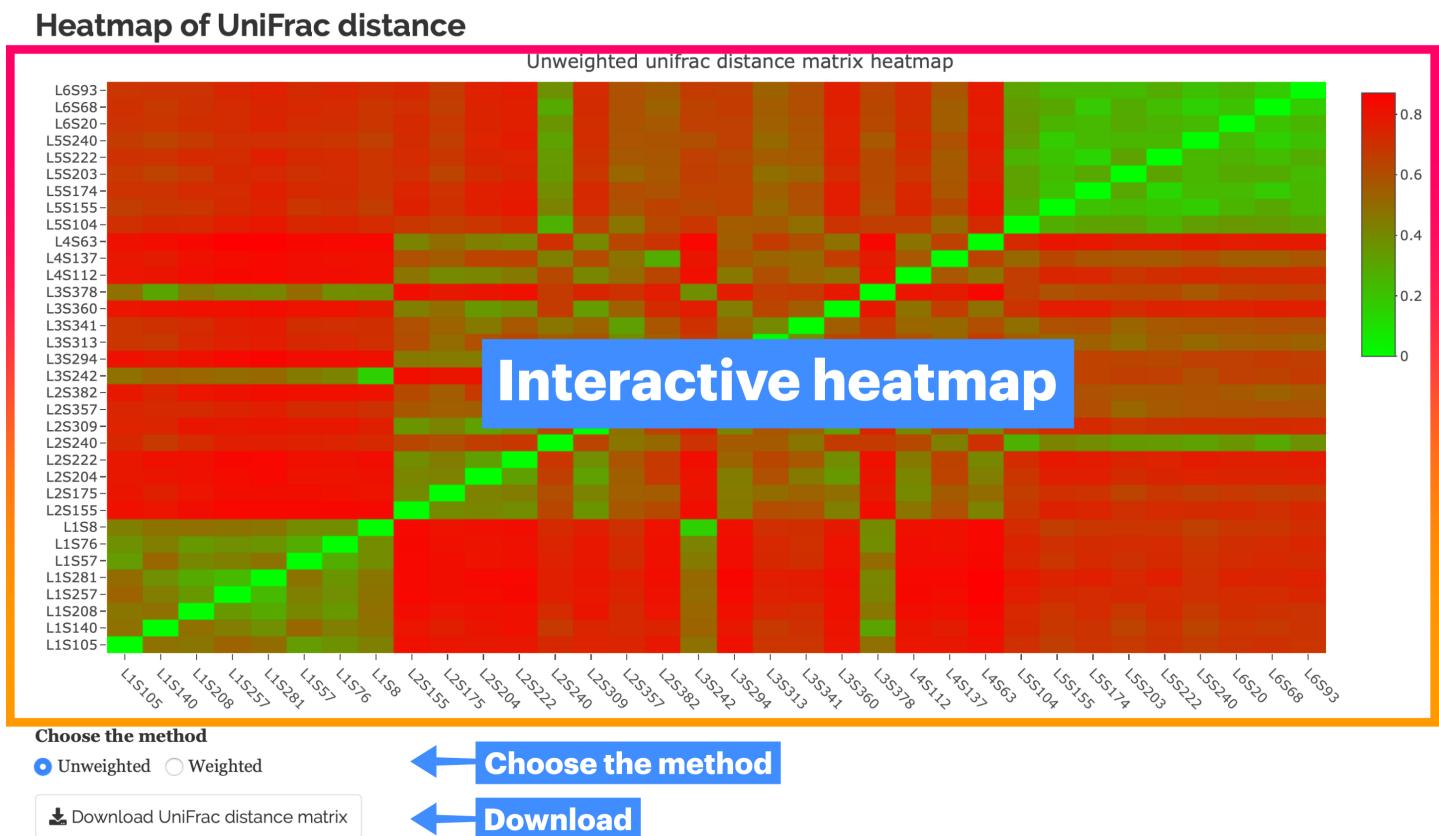
 Download Faith PD post hoc result

* 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).

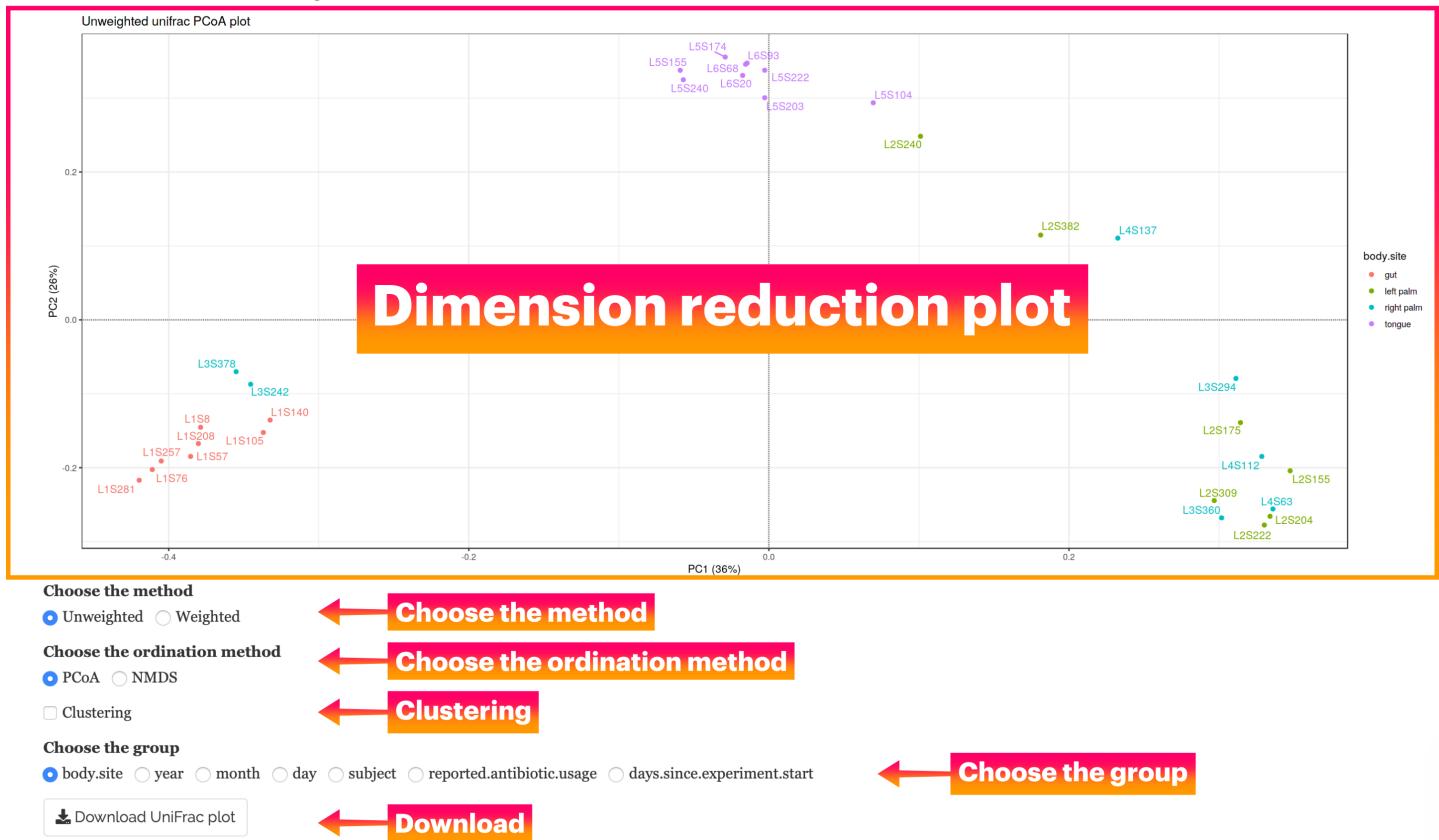
* Download heatmap matrix:

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

* The values shown in the heatmap are the original values plus 0.01 and then natural log transformed.

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 ²	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 ²	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](#).

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

ANCOM (Analysis of composition of microbiomes) is used for comparing the composition of microbiomes in two or more populations.

Select an attribute comparison

Select an attribute comparison.

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

Choose the level

Phylum

▼

Choose the level

 Start!

Start

* Select an attribute comparison:

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

* Choose the level:

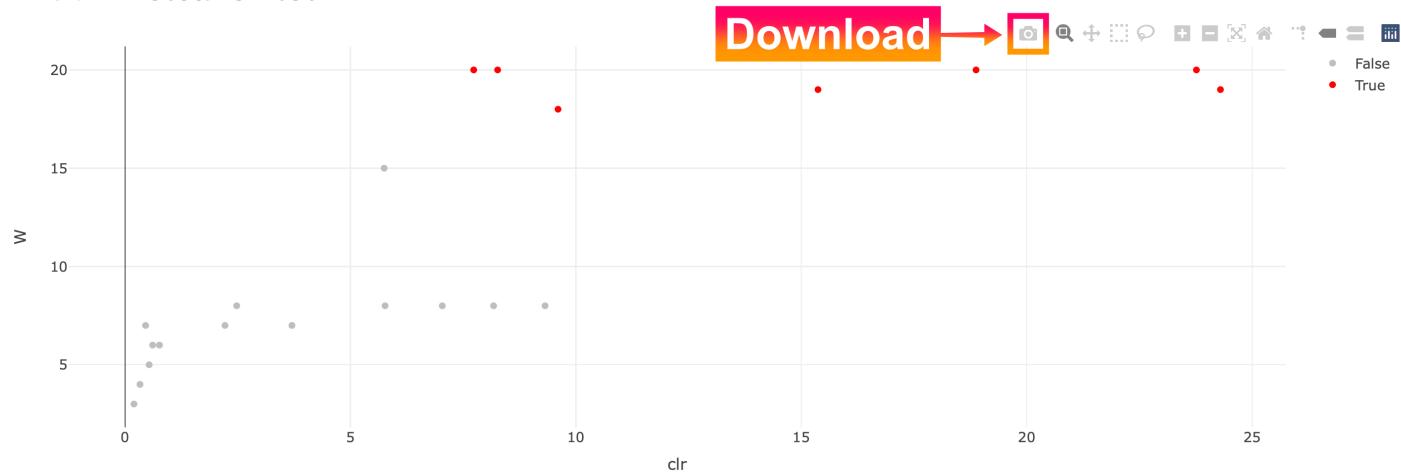
Select a taxonomic level for comparison.

* 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 the number of sub-hypotheses that have rejected for a given taxon in ANCOM analysis.
The clr represents log-fold change relative to the average microbe.

- * Download the ANCOM volcano plot:
Click on the camera icon to save the plot.
-

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

ANCOM results (Taxa with significant W value)

Show 10 entries		Search:						
	Kingdom	Phylum	Class	Order	Family	Genus	Species	W
1	Bacteria	Fusobacteria						20
2	Bacteria	Proteobacteria						20
3	Bacteria	Actinobacteria						20
4	Bacteria	Cyanobacteria						20
5	Bacteria	Firmicutes						19
6	Bacteria	Bacteroidetes						19
7	Bacteria	Verrucomicrobia						18

Showing 1 to 7 of 7 entries

Previous 1 Next

 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

2. 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.

3. 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" placeholder. On the right, there are two buttons: "Function annotation table" and "Function barplot". At the bottom, there is a large orange "Upload" button with a white arrow pointing left, and other buttons for "Start!", "reset", "learn more", and "Demo". A red box highlights the "Start!" button, and a yellow box highlights the "Upload" button.

4. Click on the "**Start!**" button to conduct the analysis. (Or, click on the "**reset**" button to re-upload the files).

(B) Inspect result

1. Function annotation table

Display reads of the function types in every sample.

Function annotation table

Show 10 entries Search:

Type	L1S105	L1S140	L1S208	L1S257	L1S281	L1S57	L1S76	L1S8	L2S155	L2S175	L2S204
1 methanol_oxidation	0	0	0	0	0	0	0	0	34	12	6
2 methylotrophy	0	0	0	0	0	0	0	0	34	12	6
3 aerobic_ammonia_oxidation	0	0	0	0	0	0	0	0	0	0	0
4 nitrification	0	0	0	0	0	0	0	0	0	0	0
5 sulfate_respiration	79								0	0	2
6 sulfur_respiration	0								0	0	0
7 thiosulfate_respiration	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
9 arsenate_detoxification	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

Showing 1 to 10 of 54 entries Previous [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) Next

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

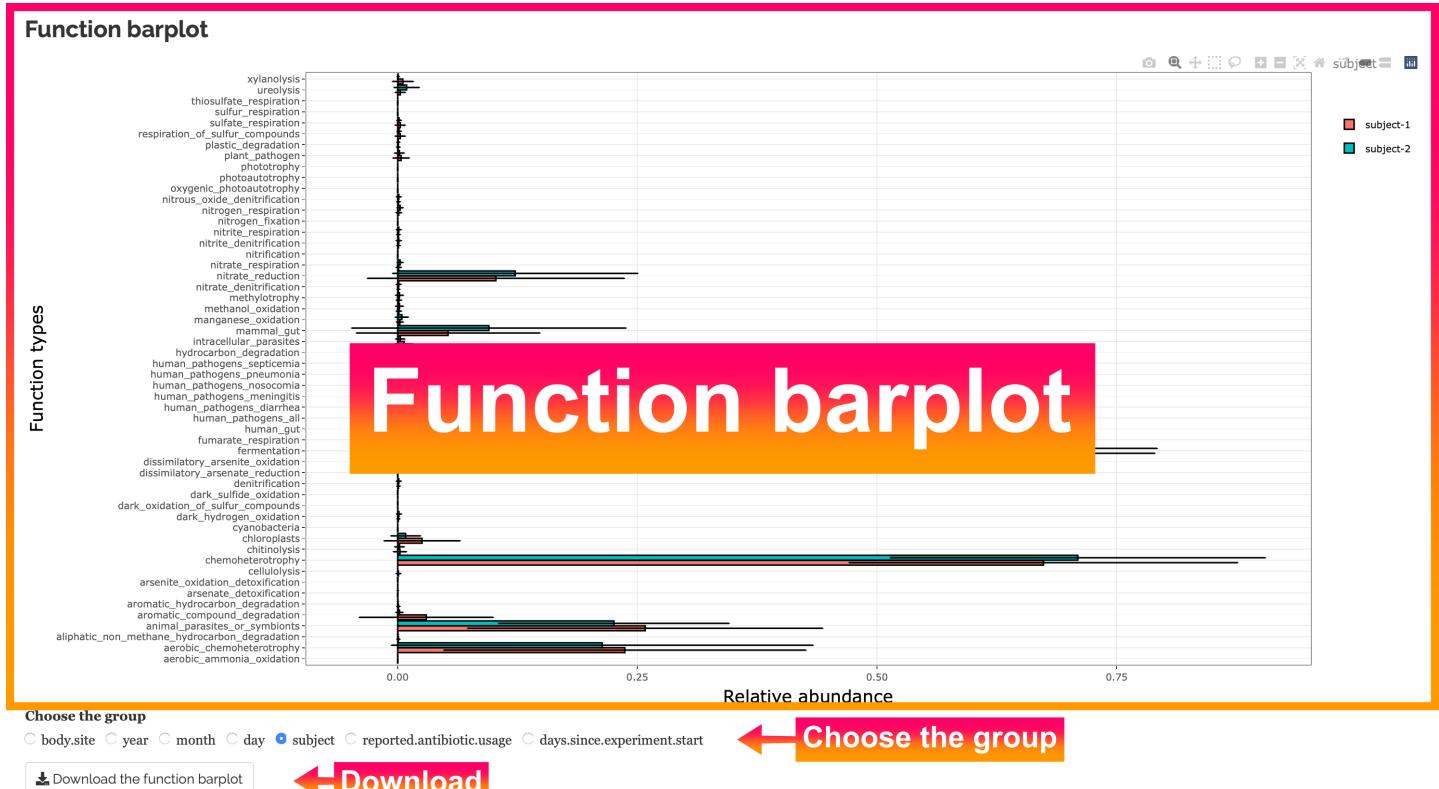
* Summary: Essential information regarding the function prediction.

* 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.



- * 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.