

Microbiota amplicOn
Characterization Implement

MOCHI

User Guide

TABLE OF CONTENTS

Chapter 1: Installation of local service

| | |
|-------------------------------|----|
| MacOS..... | 4 |
| (A) Install Docker | 4 |
| (B) Start MOCHI service..... | 6 |
| Windows 10..... | 8 |
| (A) Install Docker | 8 |
| (B) Start MOCHI service | 10 |
| Linux..... | 12 |
| (A) Install Docker..... | 12 |
| (B) Start MOCHI service | 12 |
| Additional Information | 13 |

Chapter 2: Analysis

| | |
|-----------------------------------|----|
| Sequence preprocessing..... | 15 |
| (A) Sequence summary | 15 |
| (B) Sequence denoising | 20 |
| (C) Taxonomy classification | 27 |
| Taxonomy analysis | 31 |
| (A) Upload files..... | 31 |
| (B) Inspect results | 33 |
| 1. Taxonomic table..... | 33 |
| 2. Taxonomic barplot | 34 |
| 3. Taxonomic heatmap | 35 |
| 4. Krona | 36 |
| 5. Alpha diversity | 37 |
| 6. Beta diversity | 40 |
| 7. Phylogenetic diversity | 43 |
| 8. ANCOM | 50 |
| Function analysis | 53 |
| (A) Upload files..... | 53 |
| (B) Inspect result | 54 |

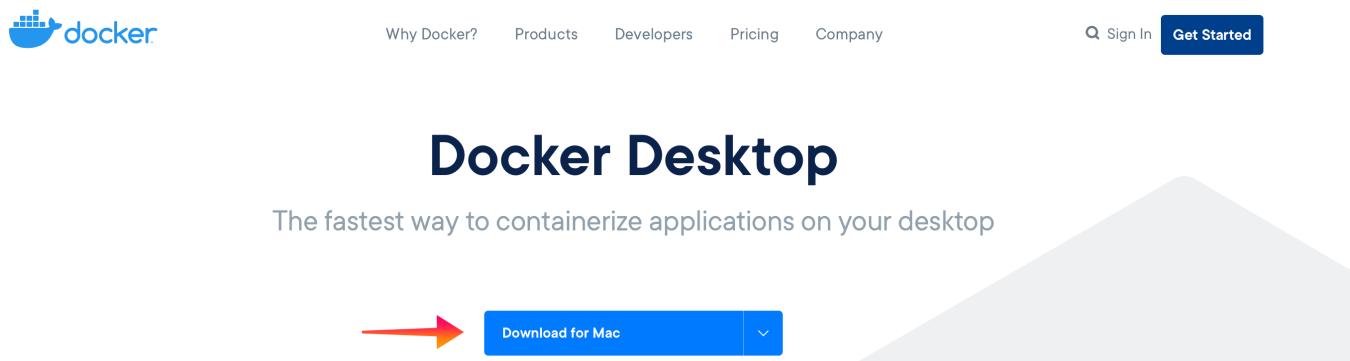
| | |
|------------------------------------|----|
| 1. Function annotation table | 54 |
| 2. Function barplot..... | 55 |

CHAPTER 1: INSTALLATION OF LOCAL SERVICE

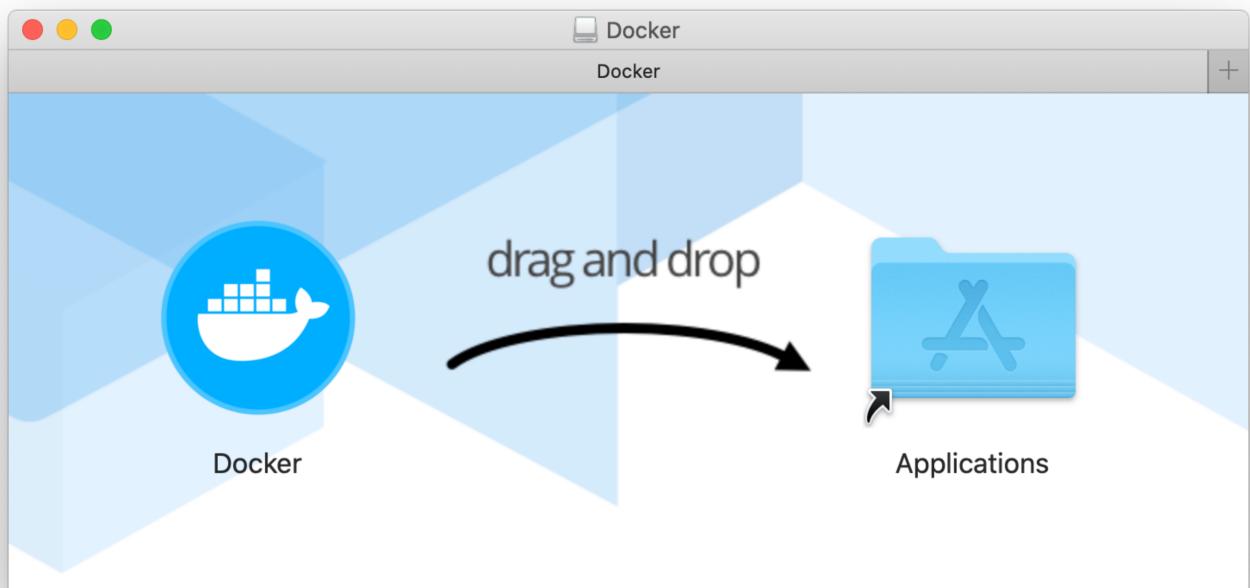
MacOS

(A) Install Docker

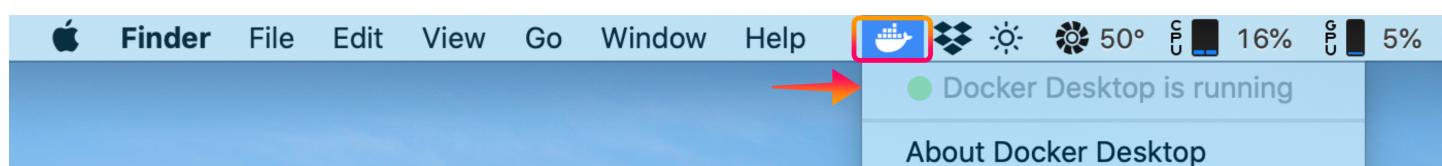
1. Download [Docker Desktop](#).



2. Open "**Docker.dmg**" file. Drag **Docker Desktop** app to the Applications folder.



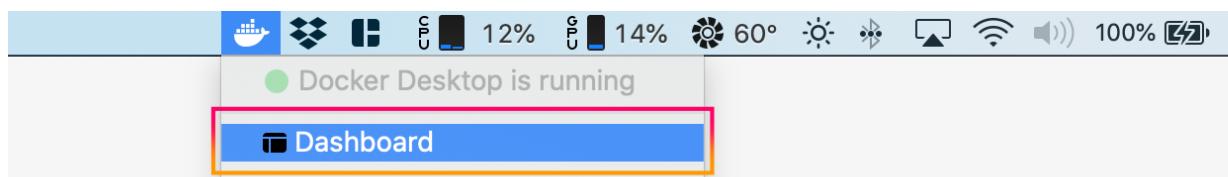
3. Start the Docker service by double clicking the Docker app. Wait for a few seconds to load, the docker icon should appear in the status bar.



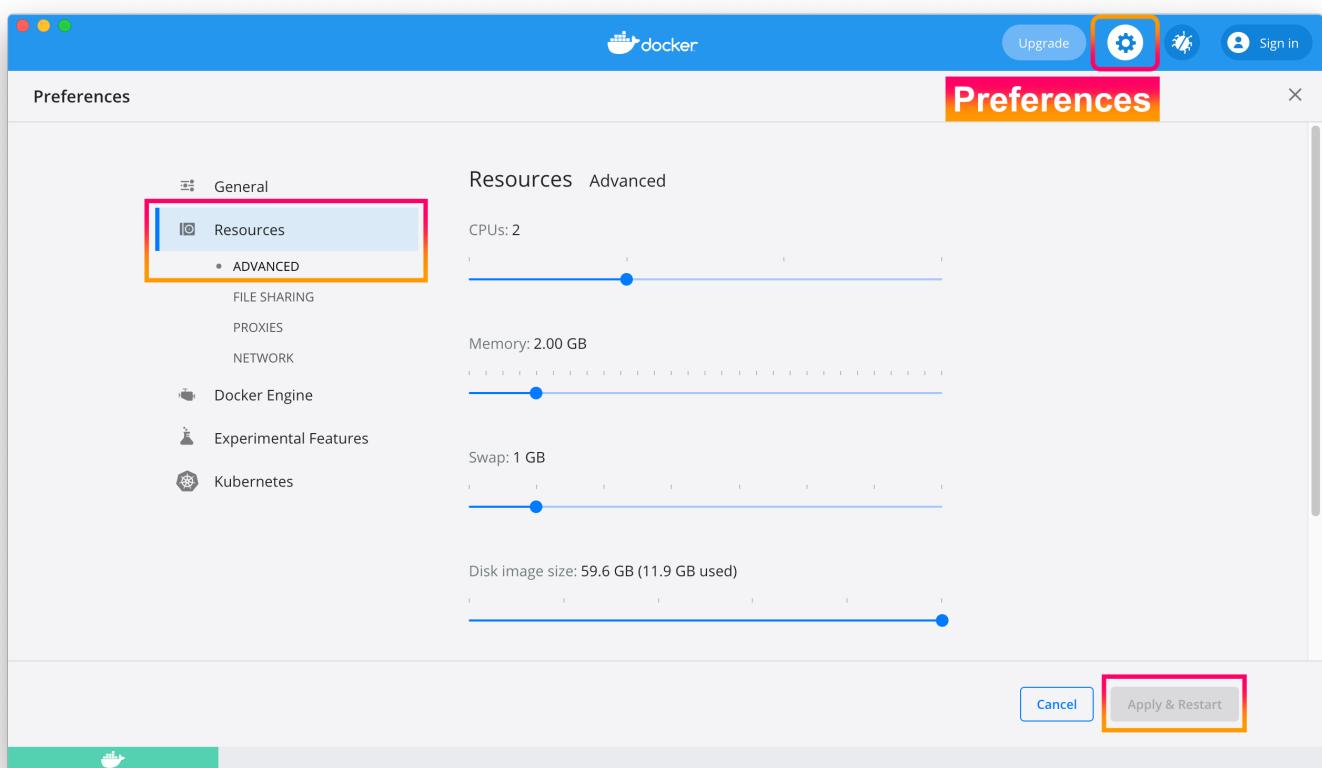
4. (Optional) In MacOS, the computational resources are preset in Docker app. To optimize the efficiency of analysis, the user can adjust the settings with the following instruction. We recommend settings **above 4 CPUs** and **8-16 GB memory** (by default, MOCHI only uses a maximum of 16 GB memory).

⚠ Assigning all of the resources to Docker may cause your system to delay or crash.

4.1. Open the Docker dashboard from the drop-down menu in the status bar.



4.2. Go to **Preference / Resources / Advanced**. Adjust the resources using the rolling bar. Press "Apply & Restart". Waiting for Docker to restart.

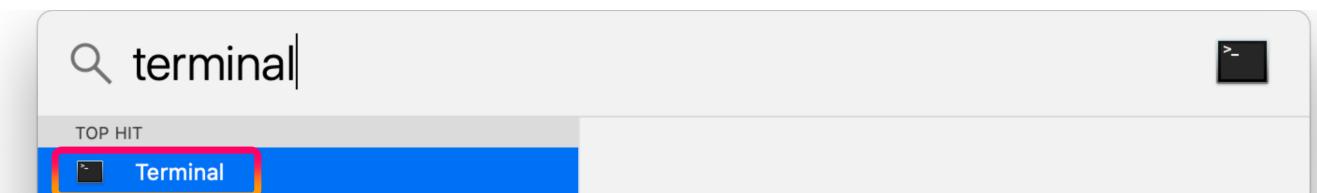


(B) Start MOCHI service

1. Download “**docker-compose.yml**” from [NCTU website](#), and place it under a folder named “**MOCHI**”.



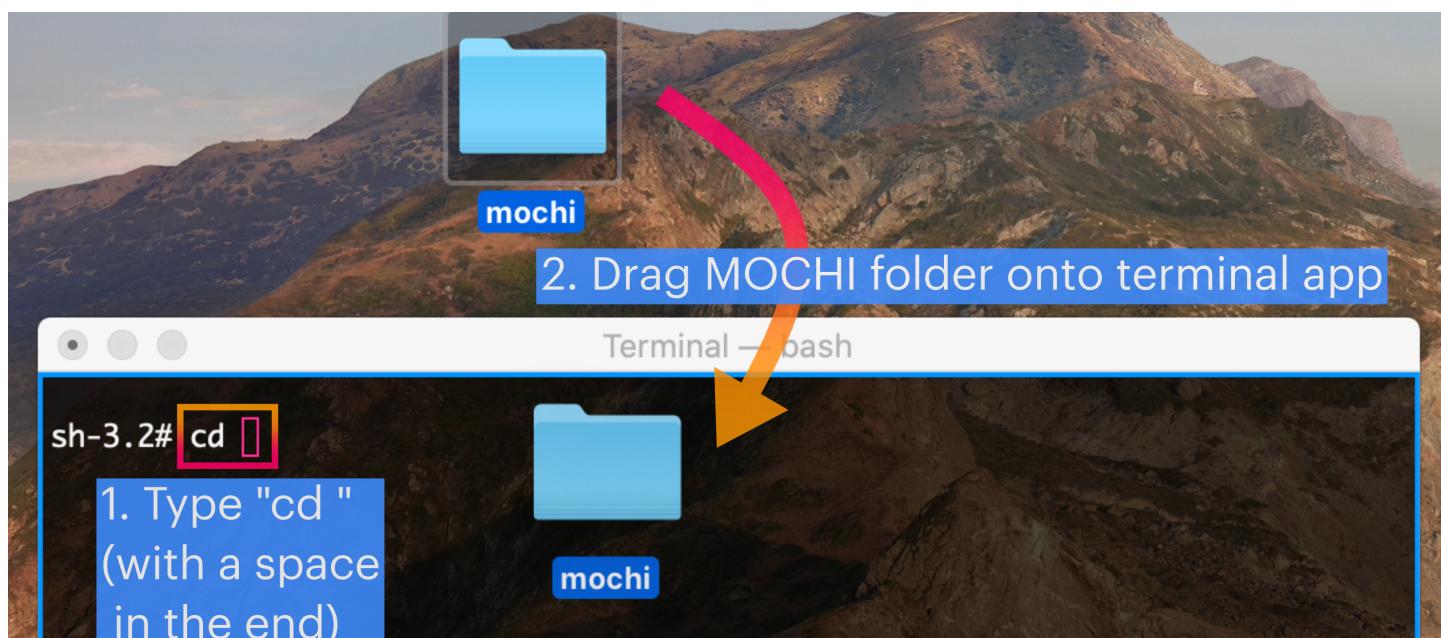
2. Open the **Terminal** app. (Press “Command + Space” to open Spotlight and type “terminal”.)



3. Run **cd /path/to/MOCHI** to navigate to the MOCHI directory.

```
sh-3.2# cd /Users/Mac/Desktop/mochi
```

Hint: If you do not know the folder path of MOCHI, type **cd** and then a space, and then drag the MOCHI folder into the terminal window. The folder path should appear in the terminal automatically.



4. Run `docker-compose up -d` to download and start the MOCHI image from Docker Hub. (The download process will only run during the first setup. The size of the MOCHI image is around 10 GB, and the running time depends on the download speed.)

```
[powang@MacBook-Pro--PoWang mochi % docker-compose up -d
WARNING: Some services (mochi_server) use the 'deploy' key, which will be ignored. Compose does not support 'deploy' configuration - use `docker stack deploy` to deploy to a swarm.
Creating network "mochi_default" with the default driver
Pulling mochi_server (dockerjjz/mochi_local:...
latest: Pulling from dockerjjz/mochi_local
f15005b0235f: Pull complete
1901fd813023: Pull complete
a92940affedf: Pull complete
dbebda29cb22: Pull complete
3c63b26b92fd: Pull complete
e0c15c0b4e0b: Pull complete
Digest: sha256:1501a145eb826f9f799239964eb064170fdc5be8abcd7b04fd9a61c888b9dee
Status: Downloaded newer image for dockerjjz/mochi_local:latest
Creating mochi_server ... done]
```

5. Open the browser, and type **127.0.0.1:3811** in the address bar. A MOCHI interactive webpage will appear and you will be able to begin your analysis.

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.

Overview of MOCHI

The MOCHI pipeline consists of three main steps:

- Sequence Preprocessing:** This step includes:
 - Step 1. Sequence summary:** Shows read counts and read quality.
 - Step 2. Sequence denoising (DADA2):** Shows sequence table and alpha rarefaction.
 - Step 3. Taxonomy classification:** Shows taxonomy prediction table.
- Taxonomy Analysis:** This step includes various plots and tables:
 - Krona chart.
 - Bar plot.
 - Heatmap.
 - Taxonomy table.
 - Alpha diversity plot.
 - Box plot.
 - PCoA plot.
 - Comparison between groups (ANCOM).
 - Post hoc analysis table.
- Function Analysis:** This step includes:
 - FAPROTAX table.
 - Functional annotation table.
 - Bar plot.

The advantages of MOCHI

- Friendly user interface: point-and-click and fill-in inputs, no programming required.
- Cross-platform: simple set-up with Docker containers on Linux, Windows, or Mac OS.
- Local computing resource: runs on your premise with privacy, not subject to network reliability and limitation.
- Compatible with other downstream analytical tools
- Publishable plots and charts generated with chosen parameters

Molecular Bioinformatics Lab, National Yang Ming Chiao Tung University, Taiwan 300, R.O.C. Last updated on 03/23/2021

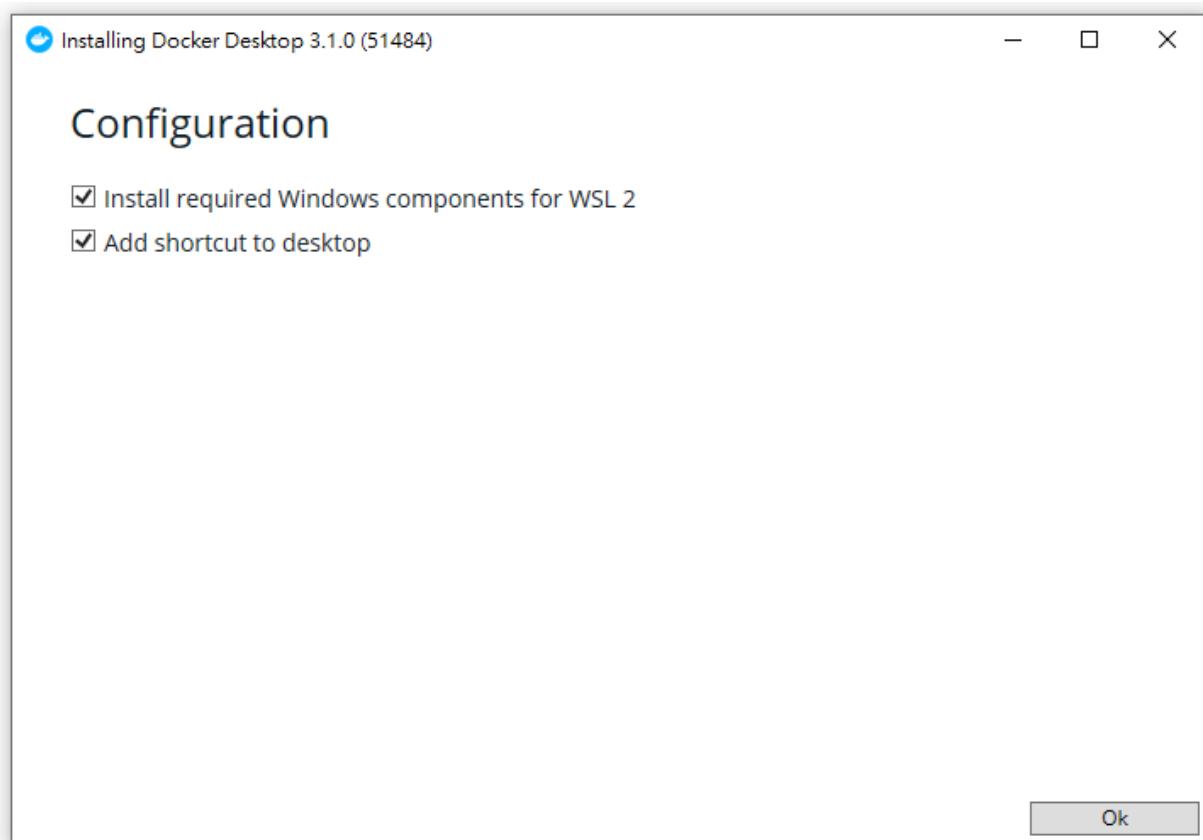
Windows 10

(A) Install Docker

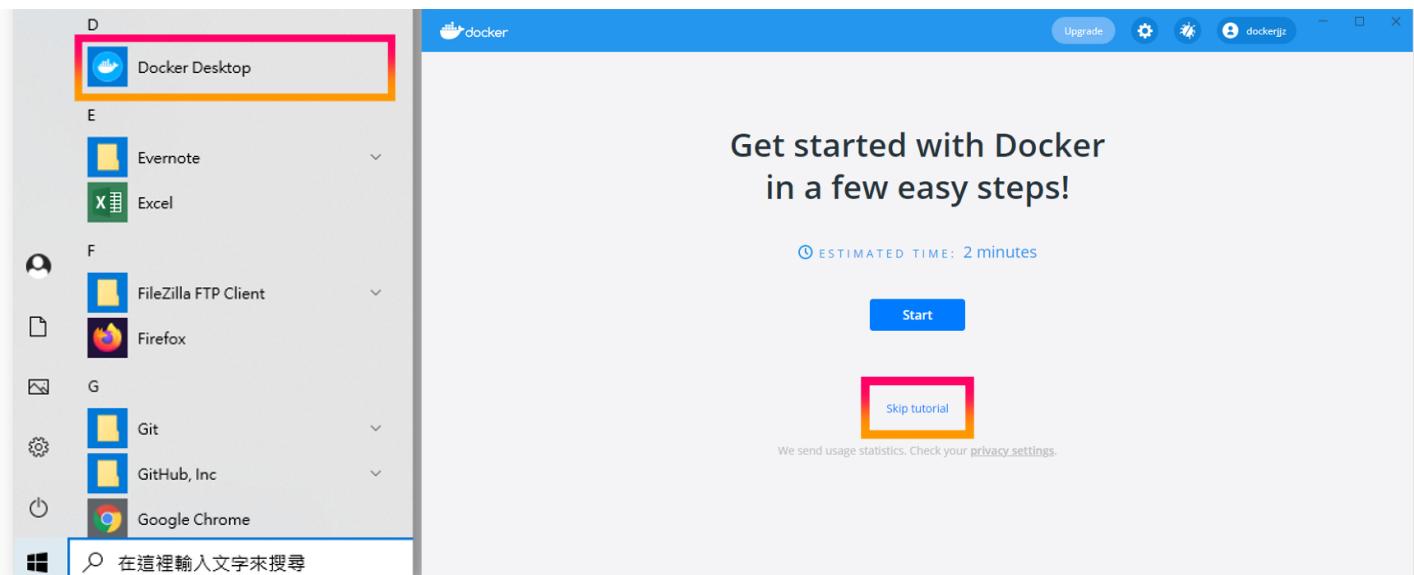
6. Download Docker Desktop.



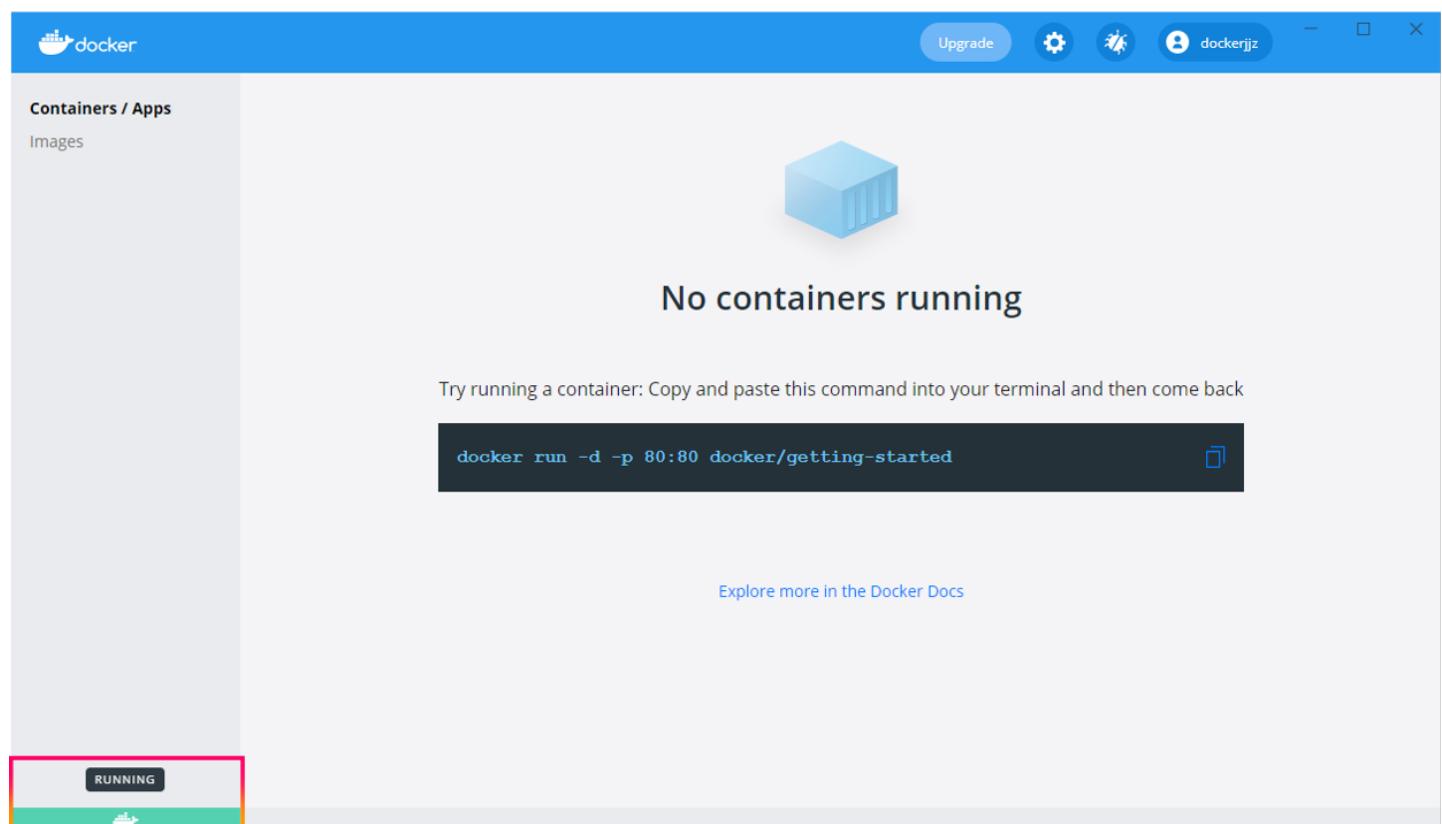
7. Run "**Docker Desktop Installer.exe**" and follow the instructions for setup. (You may be asked whether to install WSL2 Engine. By default, it is installed along with Docker to speed up the performance but is not required for running MOCHI. Use the default setting if you don't know what to choose.)



8. Start Docker by **clicking the Docker icon**. Press “Skip tutorial”.



9. If the Docker service has been successfully turned on, a green indicator will appear on the bottom-left in the Docker window.

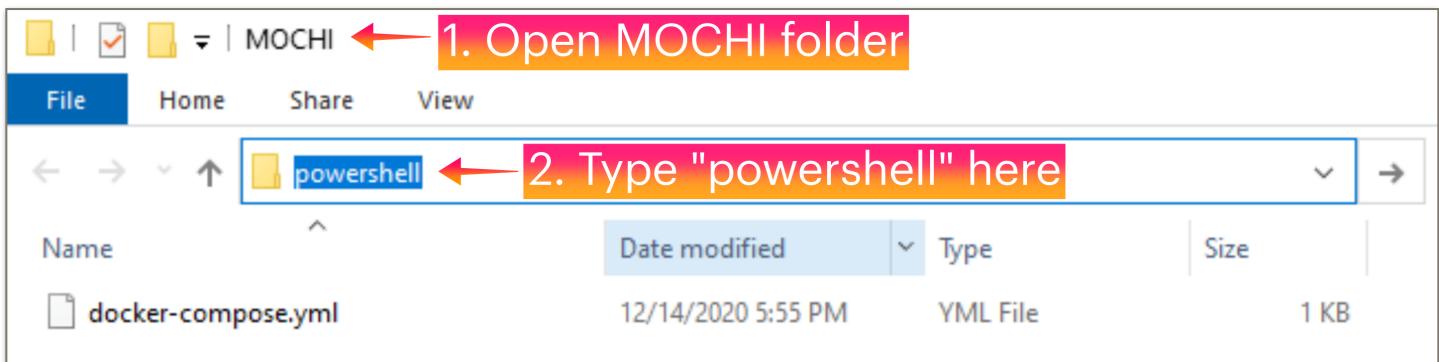


(B) Start MOCHI service

1. Download “**docker-compose.yml**” from the NCTU website, and place it in a folder named “**MOCHI**”.



2. Open **MOCHI** folder, type “**powershell**” in the address bar, and press enter. This will open a command-line shell under MOCHI directory.



3. In Powershell, run **docker-compose up -d** to download and start the MOCHI image from Docker Hub. (The download process will only run at the first setup. The size of the MOCHI image is about 10 GB, and the running time depends on the download speed.)

```
PS C:\Users\dodolab\Desktop\MOCHI> docker-compose up -d
Creating network "mochi_default" with the default driver
Pulling mochi_server (dockerjjz/mochi_local_version:)...latest: Pulling from dockerjjz/mochi_local_version
f15005b0235f: Pull complete
1901fd813023: Pull complete
a92940affedf: Pull complete
dbebda29cb22: Pull complete
3c63b26b92fd: Pull complete
e4191a297544: Pull complete
Digest: sha256:794909d921df9cc55edba44a3fe66701e43845c5c5578bf1e194064e071cbe
Status: Downloaded newer image for dockerjjz/mochi_local_version:latest
Creating mochi_server_version ... done
```

4. Open a browser, and type **127.0.0.1:3811** in the address bar. A MOCHI interactive webpage is now available to begin your analysis.

The screenshot shows the MOCHI web application interface. At the top, there's a navigation bar with links for Home, Sequence Preprocessing, Taxonomy Analysis, Function Analysis, and Tutorial. Below the navigation bar, a main title says "Welcome to MOCHI! (Microbiota amplicon CHaracterization Implement)". A brief description follows: "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 Shiny. MOCHI may also be downloaded and operated locally." The central part of the page is titled "Overview of MOCHI". It illustrates the three-step sequence preprocessing pipeline: Step 1 (Sequence summary) showing read counts and read quality; Step 2 (Sequence denoising (DADA2)) showing a sequence table and alpha rarefaction curves; and Step 3 (Taxonomy classification) showing a taxonomy prediction table. Arrows from these steps point to a large section on the right labeled "Taxonomy Analysis", which contains various plots: taxonomy table, bar plot, heatmap, PCA plot, box plot, Krone, (ANCOM), comparison between groups, post hoc analysis, and a Tukey test table. Finally, an arrow points from the taxonomy prediction table to a "Function Analysis" section, which includes a FAPROTAX icon, a functional annotation table, and a bar plot.

The advantages of MOCHI

- Friendly user interface: point-and-click and fill-in inputs, no programming required.
- Cross-platform: simple set-up with Docker containers on Linux, Windows, or MacOs.
- Local computing resource: runs on your premise with privacy, not subject to network reliability and limitation.
- Compatible with other downstream analytical tools
- Publishable plots and charts generated with chosen parameters

Molecular Bioinformatics Lab, National Yang Ming Chiao Tung University, Taiwan 300, R.O.C. Last updated on 03/23/2021

Linux

(A) Install Docker

- ▶ Please follow the [official guidance](#) for setup and start Docker Engine on Ubuntu.

(B) Start MOCHI service

1. Download “**docker-compose.yml**” from [NCTU website](#), and place it in a folder named “**MOCHI**”.



2. Navigate to the MOCHI directory.

```
$ cd /path/to/place/MOCHI
```

3. Download and start the MOCHI image from Docker Hub. (The download process will only run during the first setup. The size of the MOCHI image is about 10 GB, and the running time depends on the download speed.)

```
$ docker-compose up -d
```

4. Open a browser, and type **127.0.0.1:3811** in the address bar. A MOCHI interactive webpage is now available to begin your analysis.

Additional Information

* Stop, Restore or Remove MOCHI service

- ▶ To temporarily pause MOCHI service, please run `docker-compose stop`. This will create a paused MOCHI container, and can be restored with command `docker-compose start`.

NOTE: Please be aware that interrupting the MOCHI service when an analysis is running will not save the running process. The MOCHI container is automatically paused once you restart your computer or docker service.

- ▶ To permanently close MOCHI service, please run `docker-compose down`. This will remove all the data generated by MOCHI. Please save the results in advance. If you wish to open the service again, please run `docker-compose up -d`.
- ▶ To uninstall MOCHI image from your computer, please run `docker images`:

| REPOSITORY | TAG | IMAGE ID | CREATED | SIZE |
|-----------------------|--------|--------------|-------------|--------|
| dockerjjz/mochi_local | latest | 441fe987dce8 | 2 weeks ago | 9.05GB |

and look for “IMAGE ID” of the repository named “**dockerjjz/mochi_local**”. Then, run `docker rmi [IMAGE ID]` (e.g., `docker rmi 441`; type partial or complete image ID are both acceptable.)

* Remember to navigate under the MOCHI folder (where “`docker-compose.yml`” is located) before starting, stopping or restoring MOCHI service.

- ▶ For MacOS, Please see [MacOS / step B-3](#).
- ▶ For Windows 10, Please see [Windows 10 / step B-2](#).
- ▶ For Linux, Please see [Linux / step B-2](#).

* After installing MOCHI, two folders are created under the MOCHI folder:

1. **seqs_folder**

- ▶ This folder is used to store the user’s sequence data which will be loaded in “Sequence preprocessing - Sequence summary”.
- ▶ The file type of sequence data needs to be **fastq.gz** or **fq.gz**.
- ▶ The filename of sequence data needs to satisfy Casava 1.8 demultiplexed format or the following example format **[SampleID]_[direction of reads]**:
 - ✓ Forward read: LS105_R1 or LS105_1
 - ✓ Reverse read: LS105_R2 or LS105_2

2. **taxa_database**

- This folder is used to store the taxonomy database, such as Greengenes, Silva and PR2. Please see "Sequence preprocessing - Taxonomy classification".

- * If you wish to change the default path to the sequence data and the taxonomic database, open the "**docker-compose.yml**" file, replace the texts "**./seqs_folder**" and "**./taxa_database**" with new paths to the sequence data and the taxonomic database, respectively. Relative path is allowed.
- * The default maximum memory used by MOCHI is 16GB. To increase, please modify the resource limit in the "**docker-compose.yml**" file.

```
version: '3.7'

services:

  mochi_server:
    # build: .
    image: dockerjjz/mochi_local
    ports:
      - "3811:3838"
      - "8011:80"
    volumes:
      - ./seqs_folder:/home/imuser/raw_data/:rw
      - ./taxa_database/greengenes:/home/imuser/taxa_database/greengenes:rw
      - ./taxa_database/silva:/home/imuser/taxa_database/silva:rw
      - ./taxa_database/PR2/18S/seqs:/home/imuser/taxa_database/PR2/18S/seqs:rw
      - ./taxa_database/PR2/18S/taxonomy:/home/imuser/taxa_database/PR2/18S/taxonomy:rw
    container_name: mochi_server

    deploy:
      resources:
        limits:
          #cpus: '8'
          memory: 16G

      environment:
        - shiny_port=3811
        - nginx_port=8011
```

CHAPTER 2: ANALYSIS

Sequence preprocessing

(A) Sequence summary

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

The screenshot shows the MOCHI web interface. At the top, there is a blue header bar with several tabs: "Home", "Sequence Preprocessing" (which is currently selected and highlighted in white), "Taxonomy Analysis", "Function Analysis", and "Tutorial". Below the header, the main content area has a title "Welcome to MOCHI". On the left, there's a sidebar with three steps: "Step 1. Sequence summary", "Step 2. Sequence denoising", and "Step 3. Taxonomy classification". A red arrow points to the "Step 1. Sequence summary" link. To the right of the sidebar, there is some descriptive text about MOCHI and its underlying technology.

2. In the left panel, press the “Browse” button to upload sequence files in the “fastq.gz” or “fq.gz” format. The size limit is 20 MB per file.

Alternatively, you could press the “Example sequences” button to download the example files and upload. The parameters for example analysis are set once you pressed the “Example sequences” button.

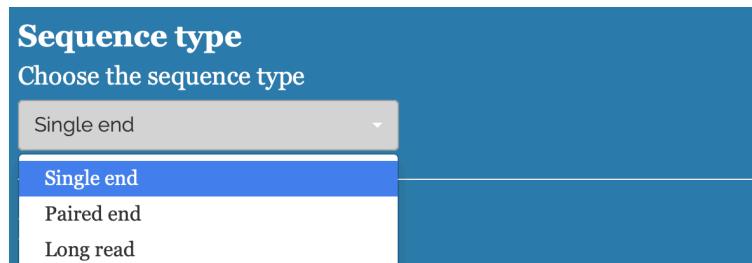
This screenshot shows the "Sequence files" section of the MOCHI interface. It includes a "Job ID: fc7c08f307" field with a "Not your job id ?" link. Below it is a "Sequence files" heading and a instruction to "Please select and upload the sequence files (*.fastq.gz or *.fq.gz)". There is a "Browse..." button and a "No file selected" message. At the bottom, there is an "Example sequences" button.

- * The Job ID is used to retrieve analyzed results, which will be stored on our server for two weeks. To recall the data, press “Not your job id ?” (above figure), and enter the Job ID.

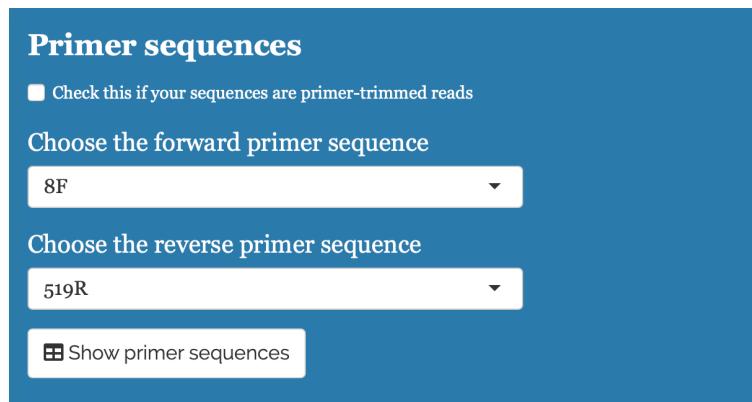
3. Sequence type: The settings for “**Step 1. Sequence summary**” and “**Step 2. Sequence denoising**” are different based on the sequence type chosen.

3.a. **Single-end or Paired-end**

- 3.a.1. The sequence type is automatically detected. If not correct, please choose manually.



- 3.a.2. Choose the **primers** or check the box when using primer-trimmed reads.



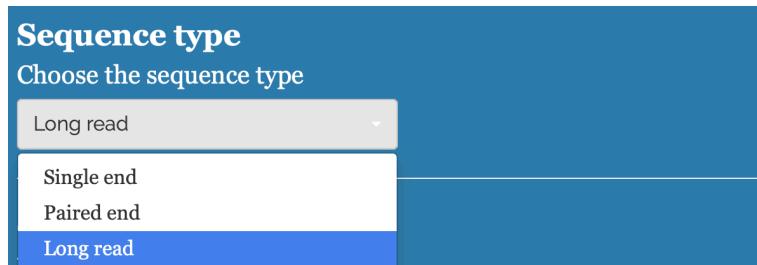
- 3.a.3. Set the number of threads for running the analysis. If zero, all available cores will be used. If you do not know the number to enter, leave it at the default number (all threads - 2).



- 3.a.4. Click on the “**Start!**” button.

3.b. Long-read

- 3.b.1. The sequence type will not automatically be detected. **Please select manually.**



- 3.b.2. Set the number of threads for running the analysis. If zero, all available cores will be used. If you do not know the number to enter, leave it at the default number (all threads - 2).



- 3.b.3. Click on the "**Start!**" button.

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

* running status

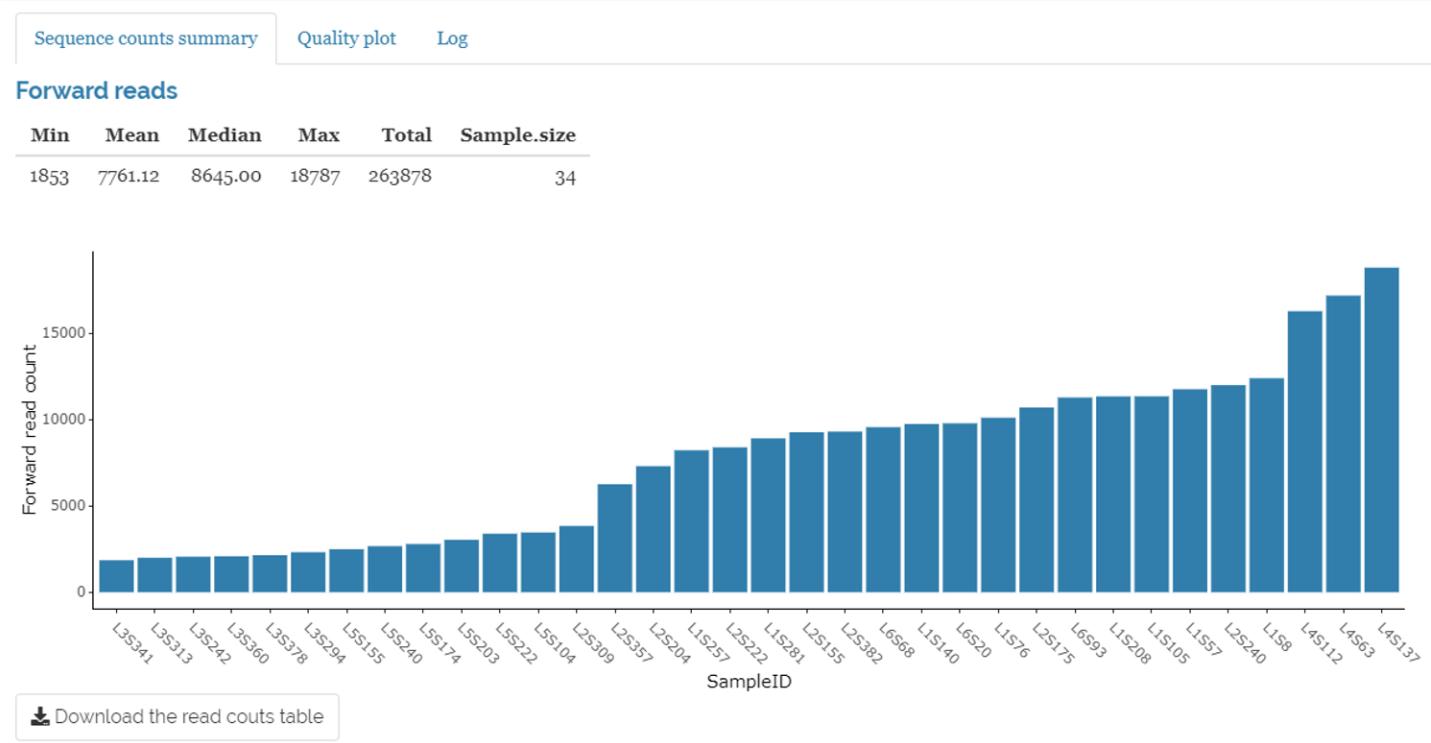


* complete status

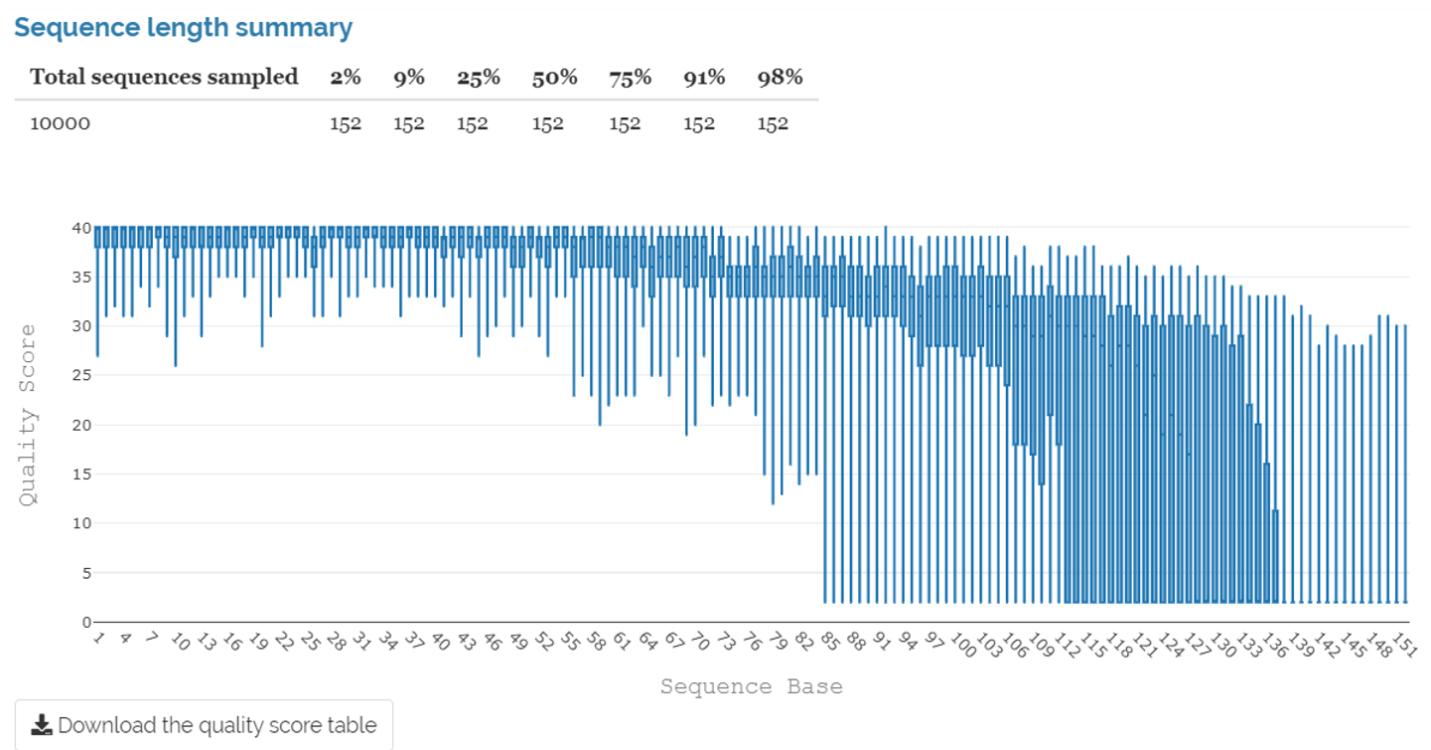
Successful!

This analysis took 11.5 secs. You can inspect the results now.

4.1. Demo results — “Sequence counts summary” summarizes the read counts of all samples.



4.2. Demo results — "Quality plot" summarizes the sequence length and show the distribution of quality score at each sequence base.



4.3. Demo results — "Log" records the used parameters and provides a button to download the table.

| Record | Value |
|-------------------|---------------------|
| time | 2021-05-31 01:44:22 |
| duration | 11.16 secs |
| sequence_type | Single end |
| sample_size | 34 |
| primer_trimmed | TRUE |
| forward_primer | 515F |
| reverse_primer | 806R |
| computing_setting | 6 |

 Download

(B) Sequence denoising

1. Select “**Sequence Preprocessing**” in 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 banner says "Welcome to MOCHI". Under the banner, there are three steps listed: Step 1. Sequence summary, Step 2. Sequence denoising, and Step 3. Taxonomy classification. A red arrow points to the "Sequence denoising" link. To the right of the steps, it says "built on QIIME2 with a friendly web interface powered by the R package of Shiny".

2. Depending on sequence type selected in “**Step 1. Sequence summary**”, the settings for denoising will be different.

2.a. Single-end or Paired-end

- 2.a.1. Set the position and the quality score for trimming the sequences.

The screenshot shows the "Sequence trimming" and "Quality score filtering" settings in the MOCHI interface. Under "Sequence trimming", there are fields for "The start position" (set to 0) and "The end position" (set to 0). There are also "learn more" and "Example" buttons. Under "Quality score filtering", there is a field for "Quality score threshold" (set to 0) and a "learn more" button.

* Starting and ending position:

Base pairs below the starting position and above the ending position will be trimmed off. 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 above setting, sequences less than 120 bp will be discarded. If the ending position is set to 0, no truncation or length trimming will be performed.

* Quality score:

Nucleotides with quality score less than or equal to specified value are trimmed off. The truncated reads shorter than the ending position are discarded.

- 2.a.2. Set the parameter of the **chimera**, **computing setting** and upload the **metadata**.

The screenshot shows two configuration sections within a blue-themed software interface:

- Chimeric reads filtering**: A section titled "The minimum fold-change value" with a text input field containing the value "1". Below the input is a "learn more" button.
- Computing setting**: A section titled "Number of threads MOCHI can use" with a text input field containing the value "2".

* **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.

* **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 specified, all available cores will be used. If you do not know the number to enter, leave it at the default number.

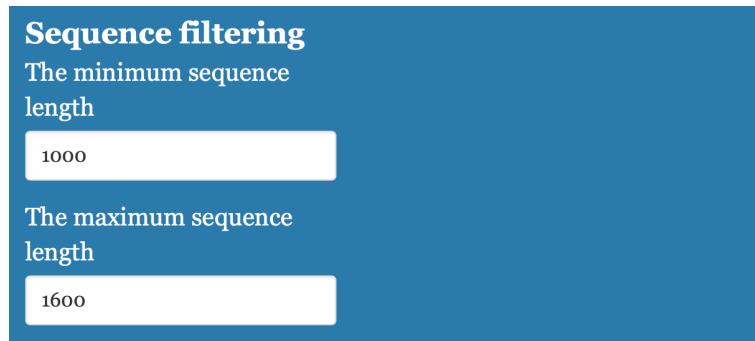
2.b. Long-read

- 2.b.1. Set the minimum and maximum sequence lengths allowed for analysis. Sequences below the minimum length and above the maximum length will be discarded.

Sequence filtering

The minimum sequence length
1000

The maximum sequence length
1600



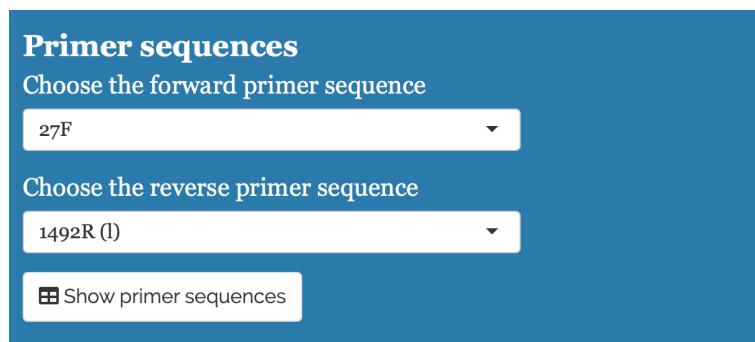
- 2.b.2. Choose the primers.

Primer sequences

Choose the forward primer sequence
27F

Choose the reverse primer sequence
1492R (l)

Show primer sequences



- 2.b.3. Assign the number of threads permitted for denoising.

Computing setting

Number of threads MOCHI can use
2



3. Click on the “**Start!**” button.

2. Sequence denoising (DADA2) for Single end

(1) Start to denoise.

 **Start!**

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

Denoising successfully!

This analysis took 1.25 mins. You can inspect the results!

- 4.1. Demo results — "Summary" summarizes the read counts of all samples.

| Summary | Filter info | Sequence info | Rarefaction plot | Table | Log | | | | | | | | | | | | |
|---|-------------|----------------------------|------------------|-----------|------------------------------|-----|------|--------|-----|-------|-------------|--------|---------|---------|---------|-----------|----|
| Sample read count summary | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Min</th><th>Mean</th><th>Median</th><th>Max</th><th>Total</th><th>Sample.size</th></tr> </thead> <tbody> <tr> <td>897.00</td><td>4523.03</td><td>4010.50</td><td>9820.00</td><td>153783.00</td><td>34</td></tr> </tbody> </table> | | | | | | Min | Mean | Median | Max | Total | Sample.size | 897.00 | 4523.03 | 4010.50 | 9820.00 | 153783.00 | 34 |
| Min | Mean | Median | Max | Total | Sample.size | | | | | | | | | | | | |
| 897.00 | 4523.03 | 4010.50 | 9820.00 | 153783.00 | 34 | | | | | | | | | | | | |
| Sample summary table | | | | | | | | | | | | | | | | | |
| Show <input type="text" value="10"/> entries | | | | | Search: <input type="text"/> | | | | | | | | | | | | |
| SampleID | Read count | Number of ASVs observed in | | | | | | | | | | | | | | | |
| 1 L1S105 | 9820 | 63 | | | | | | | | | | | | | | | |
| 2 L1S140 | 9743 | 65 | | | | | | | | | | | | | | | |
| 3 L1S208 | 8701 | 84 | | | | | | | | | | | | | | | |
| 4 L1S257 | 8339 | 81 | | | | | | | | | | | | | | | |
| 5 L1S281 | 8146 | 72 | | | | | | | | | | | | | | | |
| 6 L1S57 | 7866 | 70 | | | | | | | | | | | | | | | |
| 7 L1S76 | 7780 | 61 | | | | | | | | | | | | | | | |

4.2. Demo results — "Summary" summarizes the read counts of all ASVs.

ASV read count summary

| Min | Mean | Median | Max | Total | Number.of.ASVs |
|------|--------|--------|----------|-----------|----------------|
| 0.00 | 199.46 | 23.00 | 11371.00 | 153783.00 | 771 |

ASV summary table

| ASV | | Read count | Number of samples observed in |
|-----|----------------------------------|------------|-------------------------------|
| 1 | 4b5eeb300368260019c1fbc7a3e718fc | 11371 | 13 |
| 2 | fe3offof71a38a39cf1717ec2be3a2fc | 8929 | 16 |
| 3 | d29fe3c70564fc0f69f2c03e0d1e5561 | 8621 | 25 |
| 4 | 868528ca947bc57b69ffd83e6b73bae | 7660 | 10 |
| 5 | 154709e160e8cada6fb21115acc80f5 | 7410 | 13 |
| 6 | 1d2e5f3444ca750c85302ceee2473331 | 7185 | 23 |
| 7 | 0305a4993ecfd8ef4149fdfc7592603 | 5389 | 11 |
| 8 | cb2fe0146e2fbcb101050edb996aoee2 | 4645 | 15 |

4.3. Demo results — "Filter info" shows filtered read counts of each sample at every step of DADA2.

| Summary | Filter info | Sequence info | Rarefaction plot | Table | Log |
|-----------------|-------------|---------------|-------------------|---------------|-------------------|
| Show 10 entries | | | | Search: | |
| SampleID | Input read | Filtered read | Filtered read (%) | Denoised read | Denoised read (%) |
| | | | | | |
| 1 | L1S105 | 11340 | 8571 | 75.58 | 8499 |
| 2 | L1S140 | 9736 | 7676 | 78.84 | 7604 |
| 3 | L1S208 | 11335 | 9260 | 81.69 | 9146 |
| 4 | L1S257 | 8216 | 6705 | 81.61 | 6627 |
| 5 | L1S281 | 8904 | 7066 | 79.36 | 6975 |
| 6 | L1S57 | 11750 | 9298 | 79.13 | 9259 |
| 7 | L1S76 | 10100 | 8394 | 83.11 | 8336 |
| 8 | L1S8 | 12386 | 7662 | 61.86 | 7623 |
| 9 | L2S155 | 9261 | 4112 | 44.4 | 3932 |

4.4. Demo results — "Sequence info" summarizes the length and the bases of denoised sequences.

Summary Filter info Sequence info Rarefaction plot Table Log

Sequence Length Statistics

| Sequence.count | Min.length | Mean.length | Max.length | Range | Standard.deviation |
|----------------|------------|-------------|------------|-------|--------------------|
| 770 | 120 | 120 | 120 | 0 | 0 |

Seven-Number Summary of Sequence Lengths

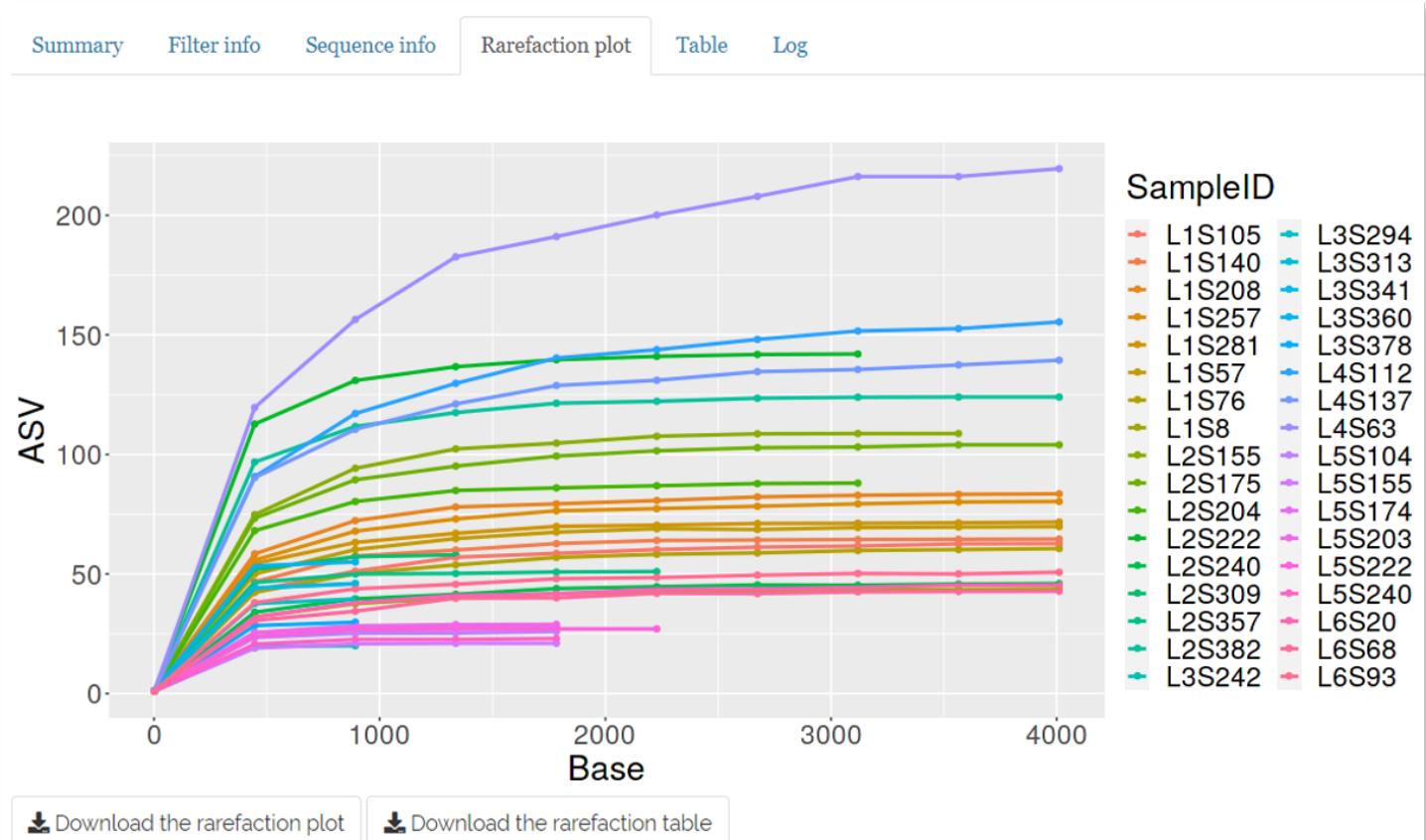
| Percentile: | 2% | 9% | 25% | 50% | 75% | 91% | 98% |
|---------------|-----|-----|-----|-----|-----|-----|-----|
| Length (nts): | 120 | 120 | 120 | 120 | 120 | 120 | 120 |

Sequence table

Show 10 entries Search:

| ASV | Sequence.length | Sequence |
|------------------------------------|-----------------|---|
| 1 4b5eeb300368260019c1fb7a3c718fc | 120 | TACGGAGGATCCGAGCGTTATCCGGATTATGGGTTAAAGGGAGCGTAGATGGATGTTA |
| 2 fe3offff71a38a39cf1717ec2be3a2fc | 120 | TACGTAGGGTGCAGCGTTAACCGAATTACTGGCGTAAAGCGAGCGCAGACGGTTACTTA |
| 3 d29fe3c70564fc0f69f2c03e0d1e5561 | 120 | TACGTAGGTCCCGAGCGTTGTCCGGATTATGGCGTAAAGCGAGCGCAGGCGGTTAGATA |
| 4 868528ca947bc57b69fdf83e6b73bae | 120 | TACGGAGGATCCGAGCGTTATCCGGATTATGGGTTAAAGGGAGCGTAGATGGATGTTA |
| 5 154709e160e8cada6fb2115acc80f5 | 120 | TACGGAGGATCCGAGCGTTATCCGGATTATGGGTTAAAGGGAGCGTAGGTGGATTGTTA |
| 6 1d2e5f3444ca750c85302ceee2473331 | 120 | TACGGAGGGTGCAGCGTTAACCGAATAACTGGCGTAAAGGGCACCGCAGGCGGTGACTT |
| 7 0305a4993ecf2d8ef4149fdfc7592603 | 120 | TACGGAGGATCCGAGCGTTATCCGGATTATGGGTTAAAGGGAGCGTAGGCGGACGCTTA |

4.5. Demo results — "Rarefaction plot" shows the ASV number of every sample at different sampling depth.



4.6. Demo results — "Table" shows the read counts of each ASV in every sample. =

| Summary | Filter info | Sequence info | Rarefaction plot | Table | Log | | | | |
|-------------------------------------|-------------|---------------|------------------|--------|------------------------------|-------|-------|------|--------|
| Show 10 entries | | | | | Search: <input type="text"/> | | | | |
| ASV | L1S105 | L1S140 | L1S208 | L1S257 | L1S281 | L1S57 | L1S76 | L1S8 | L2S155 |
| 1 4b5eeb300368260019c1fbc7a3c718fc | 2175 | 0 | 0 | 0 | 0 | 2806 | 3308 | 2594 | 10 |
| 2 fe3offof71a38a39cf1717ec2be3a2fc | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 d29fe3c70564fc0f69f2c03e0d1e5561 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 353 |
| 4 868528ca947bc57b69ffd83e6b73bae | 0 | 2248 | 2107 | 1177 | 1721 | 0 | 0 | 0 | 0 |
| 5 154709e160e8cada6fb21115acc80f5 | 802 | 1174 | 694 | 406 | 242 | 1081 | 930 | 1623 | 0 |
| 6 1d2e5f3449ca750c85302ceee2473331 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 5 | 27 |
| 7 0305a4993ecf2d8ef4149fdfc7592603 | 1197 | 647 | 964 | 909 | 531 | 414 | 334 | 243 | 0 |
| 8 cb2fe0146e2fbcb101050edb996a0ee2 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 82 |
| 9 997056ba80681bbbdd5d09aa591eadco | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |
| 10 3c9c437f27aca05f8db167cd08off1ec | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Showing 1 to 10 of 770 entries

Previous [1](#) [2](#) [3](#) [4](#) [5](#) ... [77](#) Next

4.7. Demo results — "Log" records the used parameters and provides a button to download the table.

| Summary | Filter info | Sequence info | Rarefaction plot | Table | Log |
|--|-------------|---------------|------------------|-------|---------------------|
| Record | | | | | Value |
| time | | | | | 2021-05-31 01:55:54 |
| duration | | | | | 1.87 mins |
| sequence_type | | | | | Single end |
| start_position_trim | | | | | 0 |
| end_position_trim | | | | | 120 |
| quality_score_truncate | | | | | 2 |
| chimeric_reads_min_fold_change | | | | | 1 |
| metadata_upload | | | | | TRUE |
| computing_setting | | | | | 6 |
|  Download | | | | | |

(C) Taxonomy classification

1. Select “**Sequence Preprocessing**” in the top bar, choose “**Step 3. Taxonomy 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 right side of the header, there is a dropdown menu with options: Step 1. Sequence summary, Step 2. Sequence denoising, Step 3. Taxonomy classification, and a link to QIIME2. A red arrow points to the "Step 3. Taxonomy classification" link. Below the header, there is some descriptive text about MOCHI being a 16S or 18S microbiota amplicon rRNA analysis tool.

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

The screenshot shows a "Database" selection page. It has a title "Database" and a subtitle "Select the reference database for taxonomy classification." Below this, there is a section titled "Choose the database" with a dropdown menu set to "Silva (Not detected)". There are two buttons below the dropdown: "Auto download database" and "Example".

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 the database:

- * Silva: Follow the [link](#). Choose a version to download. Decompress the downloaded file. Copy the two folders “**rep_set**” and “**taxonomy**” to the folder “**taxa_database/silva**”.
- * Greengene: Follow the [link](#). Choose a version and download the corresponding “**otus.tar.gz**”. Decompress the downloaded files. Copy the two folders “**rep_set**” and “**taxonomy**” into 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**”. Decompress the downloaded files. Copy the file “**pr2_version_X.XX.X_16S_mothur.fasta**” into the folder “**taxa_database/PR2/18S/seqs**” and the file “**pr2_version_X.XX.X_16S_mothur.tax**” into the folder “**taxa_database/PR2/18S/taxonomy**”.

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

Reference sequence filtering

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

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

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

4. Set the minimum and maximum lengths for filtering the reference sequence.

2. Filter the reference sequence by length

Minimum length

Maximum length

[learn more](#)

*** 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 lengths of the denoised sequences. To disable length filtering, set the values to zero.

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

Computing setting

Number of threads MOCHI can use

6. Click on the "**Start!**" button.

3. Taxonomy classification

(1) Classify taxonomy

 **Start!**

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

Taxonomy classification has been finished!

This analysis took 55.92 secs. You can inspect the results!

8. Demo results — “taxonomy results” shows the ASV and assigned taxonomy. The three buttons at the bottom are for downloading the files to conduct the subsequent analyses.

Inspect the taxonomy classification result.

Taxonomy result

[Log](#)

| ASV | Taxon | Confidence |
|--------------------------------------|--|-------------------|
| 1 4b5eeb300368260019cfbc7a3c718fc | k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__Bacteroides; s__ | 0.99340978592142 |
| 2 fe30ff0f71a38a39cf1717ec2be3a2fc | k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Neisseriales; f__Neisseriaceae; g__Neisseria; s__ | 0.89490583340447 |
| 3 868528ca947bc57b69ffdf83e6b73bae | k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__Bacteroides; s__ | 0.98368032171919 |
| 4 154709e160e8cada6fb21115acc80f5 | k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__Bacteroides; s__ | 0.980528382140841 |
| 5 d29fe3c70564fc0f69f2c03e0d1e5561 | k__Bacteria; p__Firmicutes; c__Bacilli; o__Lactobillales; f__Streptococcaceae; g__Streptococcus; s__ | 0.99999998690498 |
| 6 1d2e5f3444ca750c85302ceee2473331 | k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pasteurellales; f__Pasteurellaceae; g__Haemophilus; s__parainfluenzae | 0.96830683630633 |
| 7 0305a4993ecf2d8ef4149fdfc7592603 | k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__Bacteroides; s__uniformis | 0.996031457767898 |
| 8 997056ba80681bbbdd5d09aa591eadco | k__Bacteria; p__Fusobacteria; c__Fusobacteriia; o__Fusobacteriales; f__Fusobacteriaceae; g__Fusobacterium; s__ | 0.907315818016303 |
| 9 3e9e437f27aca05f8db167cd08off1ec | k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Prevotellaceae; g__Prevotella; s__melaninogenica | 0.999998818423878 |
| 10 bfbbed36e63b69fec4627424163d20118 | k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__Faecalibacterium; s__prausnitzii | 0.999996821028142 |

Showing 1 to 10 of 501 entries

Previous [1](#) [2](#) [3](#) [4](#) [5](#) ... [51](#) Next

[Download](#)

(3) Download the files for the next step.

[The taxonomic table](#)

[The ASV table](#)

[The seqs data](#)

9. Demo results — "Log" records the used parameters and provides a button to download the table.

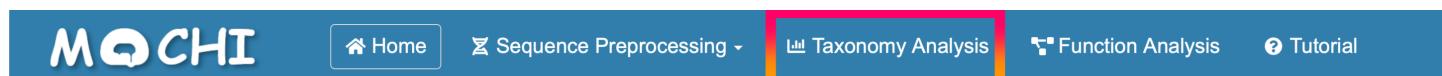
| Inspect the taxonomy classification result. | |
|---|---------------------|
| Taxonomy result | Log |
| Record | Value |
| time | 2021-03-18 05:53:21 |
| duration | 1.39 mins |
| sequence_type | Single end |
| database | Greengenes_16S_88 |
| forward_primer | 515F |
| reverse_primer | 806R |
| min_length | 100 |
| max_length | 400 |
| computing_setting | 6 |

 Download

Taxonomy analysis

(A) Upload files

1. Select “**Taxonomy Analysis**” in 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 the **metadata, taxonomic table** and **ASV table** files. Then, select the format of uploaded files, e.g., “.qza” or “.txt”.

These files can be downloaded from the “Sequence Preprocessing - Taxonomic classification” section. Please see [Sequence preprocessing / Taxonomic classification / step 8](#). Alternatively, you can press the “Demo” button to download the example files first and then upload the files. 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 can upload self-derived taxonomic table file (FeatureTable[Frequency], level: 7) from QIIME2.

* **ASV table (.qza):**

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

* **ASV table (.txt):**

The table should include the information of ASV in the first column and Taxon in the last column.

Upload the metadata file

Browse... No file selected

Choose file format

- MOCHI/QIIME2 output (.qza)
- Plain text table (.txt)

Upload the taxonomic table file

Browse... No file selected

Upload the ASV table file

Browse... No file selected

18S rRNA

Start! **Reset**

Example files

- Click on the “**Start!**” button to run the analysis. (Or, click on the “**Reset**” button to re-upload the files).
- The results will be displayed in the right panel after pressing the “**Start!**” button.

Upload the metadata file

Browse... Metadata_example.tsv
Upload complete

Upload the taxonomic table file

Browse... Taxonomic_table_example_from_MOCHI.qza
Upload complete

Upload the ASV table file

Browse... ASV_table_example_from_MOCHI.qza
Upload complete

18S rRNA

Start! **Reset**

Example files

| Kingdom (K=3) | Phylum (K=20) | Class (K=45) | Order (K=72) | Family (K=126) | Genus (K=190) | Species (K=198) | LiS8 | LiS57 | LiS76 | LiS105 | LiS155 | LiS175 | LiS204 |
|---------------|----------------|---------------------|-----------------|------------------|------------------|-----------------|------|-------|-------|--------|--------|--------|--------|
| Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Bacteroidaceae | Bacteroides | Unassigned | 4217 | 3887 | 4238 | 2977 | 10 | 10 | 3 |
| Bacteria | Proteobacteria | Betaproteobacteria | Neisseriales | Neisseriaceae | Unassigned | Unassigned | 0 | 0 | 0 | 5 | 51 | 109 | 43 |
| Bacteria | Firmicutes | Bacilli | Lactobacillales | Streptococcaceae | Streptococcus | Unassigned | 30 | 5 | 0 | 0 | 377 | 929 | 126 |
| Bacteria | Proteobacteria | Gammaproteobacteria | Pasteurellales | Pasteurellaceae | Gallibacterium | Unassigned | 5 | 0 | 0 | 0 | 27 | 51 | 0 |
| Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Bacteroidaceae | Bacteroides | uniformis | 260 | 553 | 530 | 1439 | 0 | 0 | 0 |
| Bacteria | Firmicutes | Bacilli | Bacillales | Bacillaceae | Unassigned | Unassigned | 0 | 0 | 0 | 0 | 82 | 227 | 96 |
| Bacteria | Fusobacteria | Fusobacteriia | Fusobacteriales | Fusobacteriaceae | Fusobacterium | Unassigned | 3 | 0 | 0 | 31 | 31 | 53 | 0 |
| Bacteria | Bacteroidetes | Bacteroidia | Bacteroidales | Prevotellaceae | Prevotella | Unassigned | 0 | 0 | 0 | 80 | 16 | 88 | 10 |
| Bacteria | Proteobacteria | Gammaproteobacteria | Unassigned | Unassigned | Unassigned | Unassigned | 0 | 0 | 0 | 0 | 97 | 41 | 1376 |
| Bacteria | Firmicutes | Clostridia | Clostridiales | Ruminococcaceae | Faecalibacterium | prausnitzii | 129 | 693 | 906 | 910 | 0 | 6 | 19 |

Showing 1 to 10 of 229 entries

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.

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:

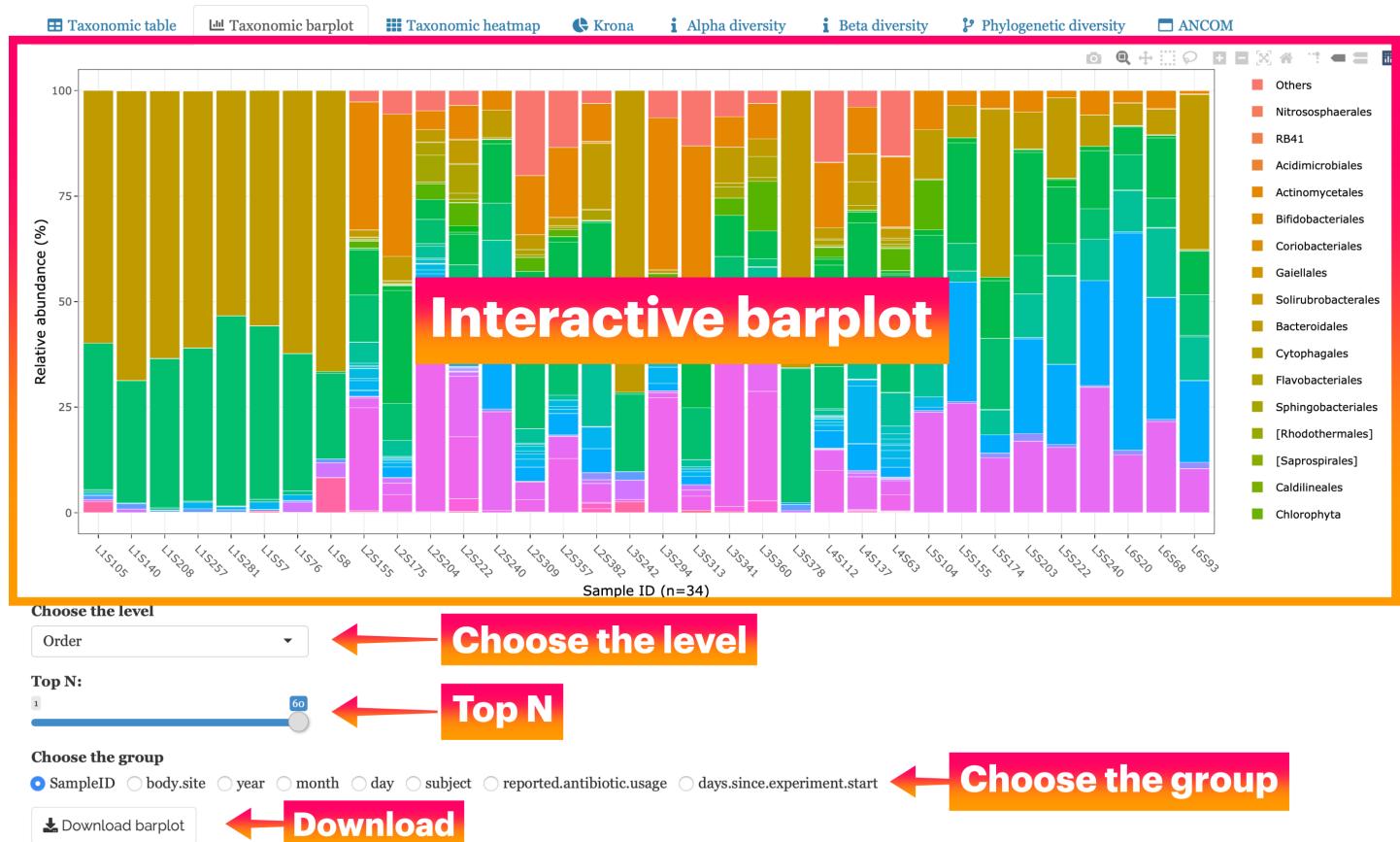
The right part of the table represents read count. The values of read count are categorized by the selected group. The column name indicates the variables of the selected group.

* Download the 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 the 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 in 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 the barplot:

Click on the “Download barplot” button to download the barplot. Alternatively, click on the camera icon in 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 logarithms 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 the 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 in 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:

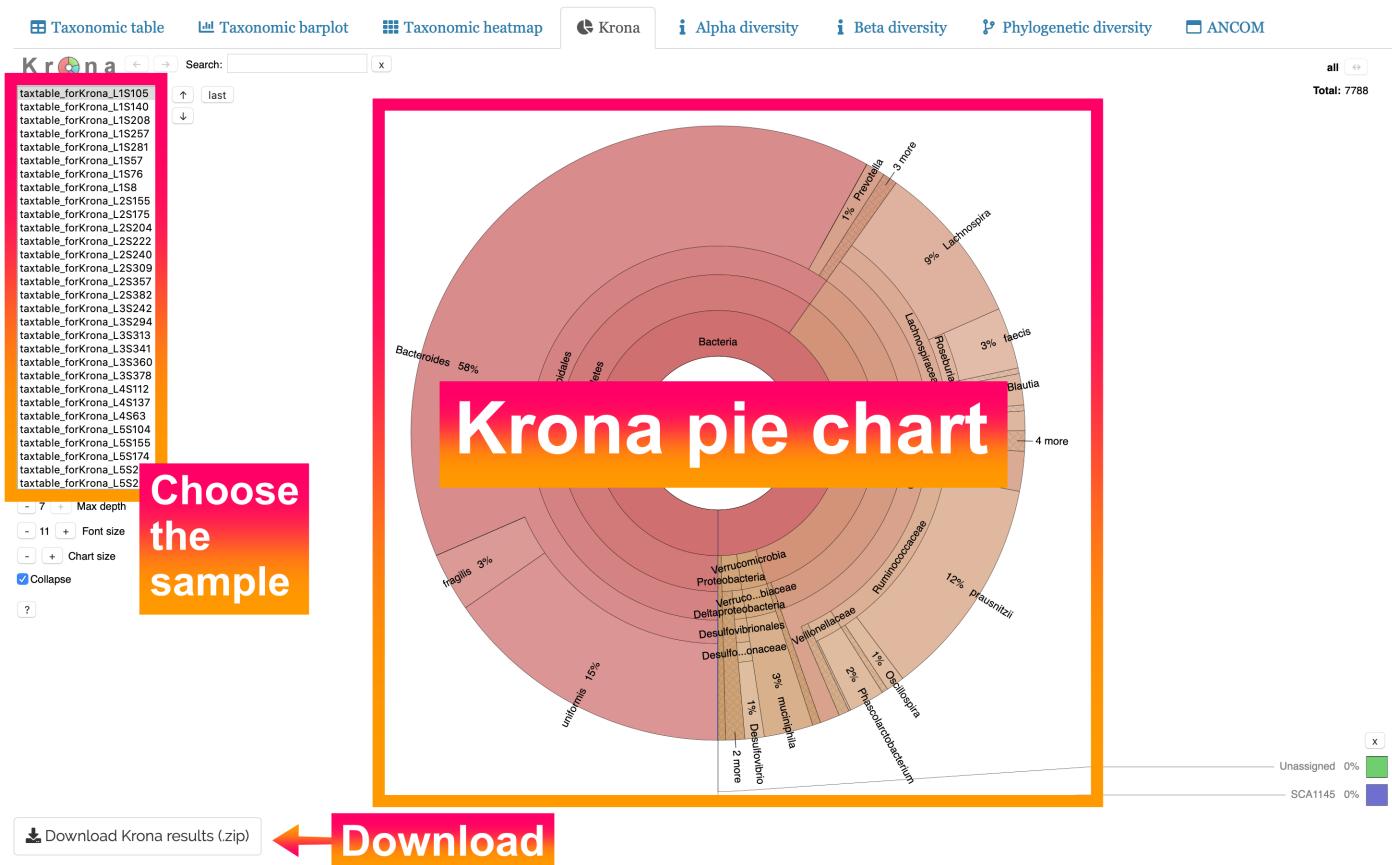
The heatmap will be grouped based on the selected metadata.

* Download the 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 to show the ratio of the selected taxon over different taxonomy level. Double click a taxon to zoom in the selected taxonomy level. To zoom back, click the backspace button in 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

[!\[\]\(842b9a660eb497650bab4e1dbd552bfc_img.jpg\) Download Alpha Diversity Table](#)

Download

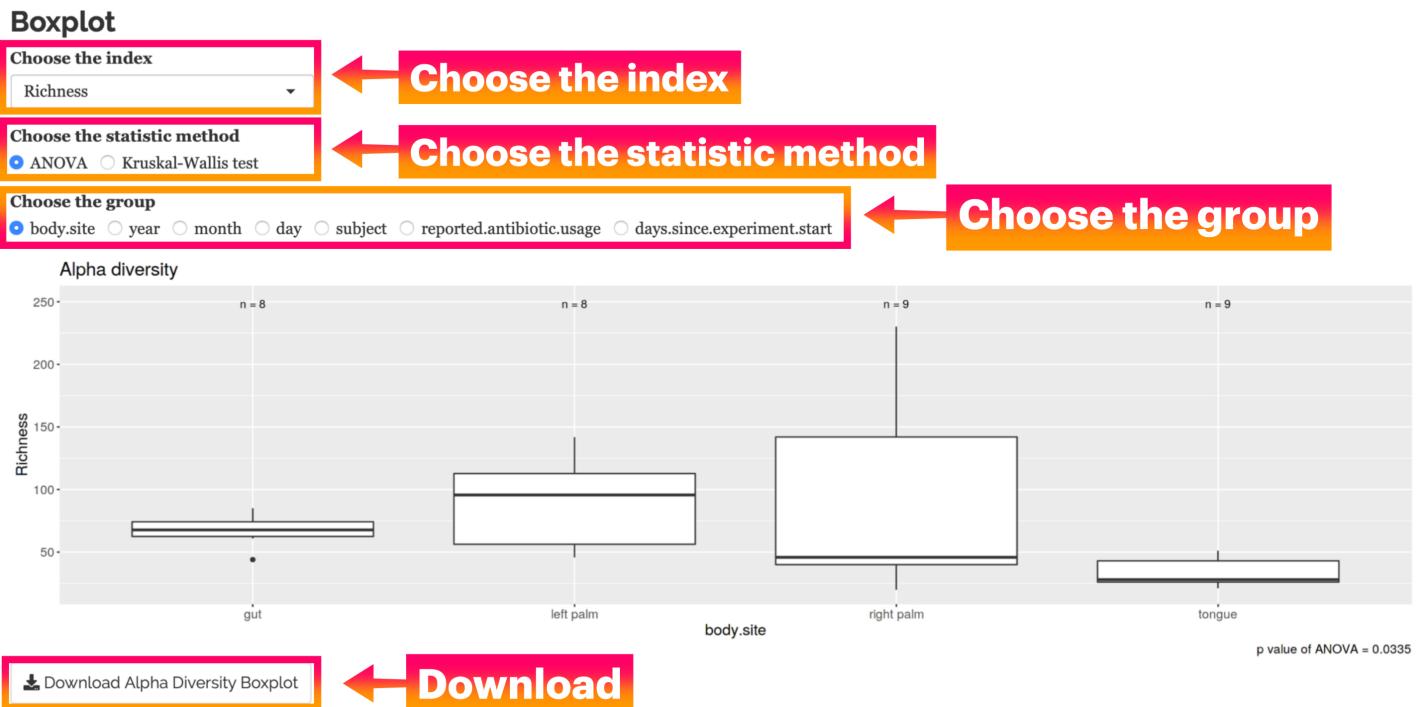
* Alpha diversity table:

This table shows the values of 8 alpha diversity indexes.

* Download the 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 the 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

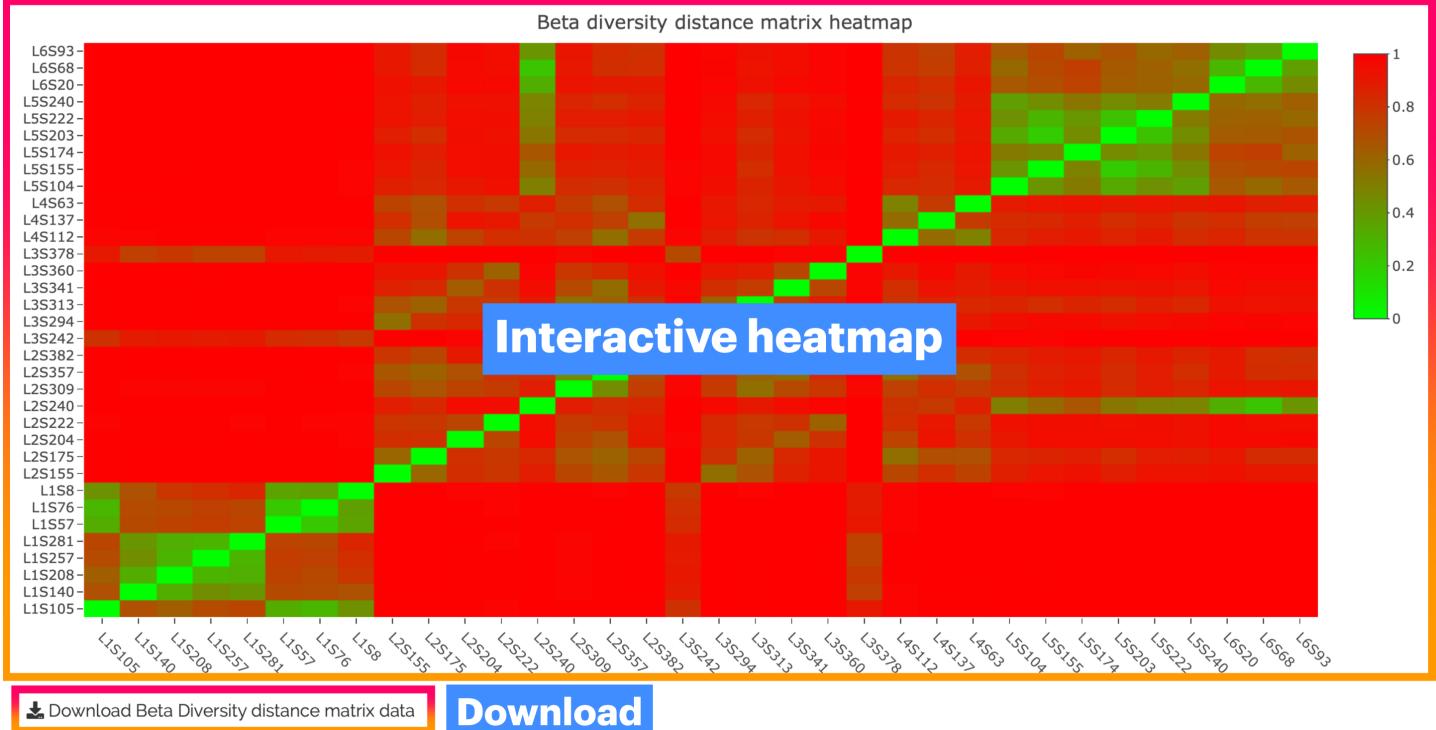
- * If ANOVA is selected when creating the boxplot, the Tukey test will be used for the post-hoc test. If Kruskal-Wallis is selected, then the Dunn test will be used.
- * Download the Alpha diversity post hoc test result:
Click on the “Download Alpha Diversity statistical result” button to download the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

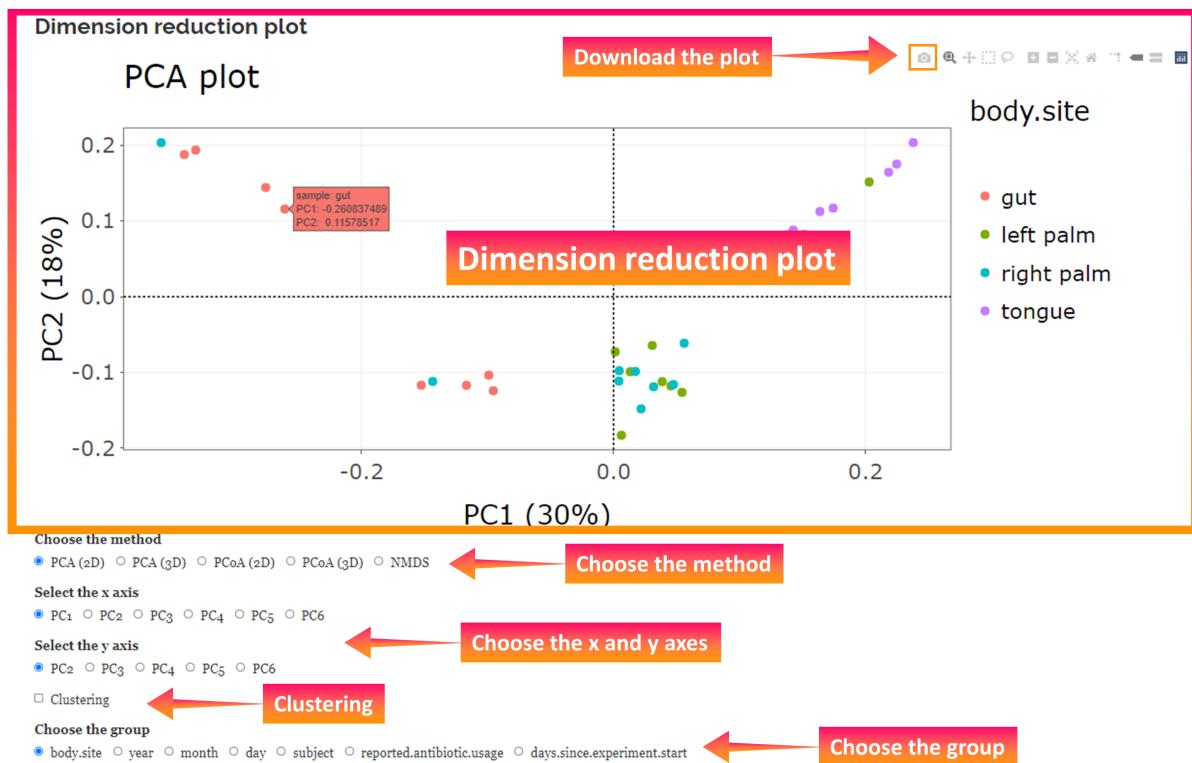
When a cursor hovers over the heatmap, the distance between species will be presented. Click and drag on the plot to zoom in and out. Double click on the plot to zoom back. Click on the camera icon in the top-right region of the heatmap to download the plot.

* 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



* Choose the method:

Select a dimension reduction method. We provide three dimension-reduction methods for visualization of beta diversity: PCA (Principal Component Analysis, 2D & 3D), PCoA (Principal Co-ordinates Analysis, 2D & 3D) and NMDS (Non-metric Multidimensional Scaling) for visualization of beta diversity.

* Choose the axes:

Individually select the PCs as x/y axes or x/y/z axes for 2D or 3D plots.

* 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 the Beta diversity plot:

Click on the camera icon to download the plot.

6.3. Statistical analysis

Statistical analysis

| PERMANOVA | | ANOSIM | | MRPP | | | | | | | |
|---|----------------|--|------------------|--|---------|---------|------------------|------------------------|---------|---------|------------------|
| R ² | P value | R | P value | A | P value | | | | | | |
| 0.3999 | 0.001 | 0.6855 | 0.001 | 0.2085 | 0.001 | | | | | | |
| Download PERMANOVA table | | Download ANOSIM table | | Download MRPP table | | | | | | | |
| Pairwise PERMANOVA | | Pairwise ANOSIM | | Pairwise MRPP | | | | | | | |
| Comparisons | R ² | P value | Adjusted P value | Comparisons | R | P value | Adjusted P value | Comparisons | A | P value | Adjusted P value |
| gut - left palm | 0.3983 | 0.001 | 0.0012 | gut - left palm | 1 | 0.001 | 0.0015 | gut - left palm | 0.2055 | 0.001 | 0.0012 |
| gut - right palm | 0.2834 | 0.001 | 0.0012 | gut - right palm | 0.6686 | 0.001 | 0.0015 | gut - right palm | 0.1456 | 0.001 | 0.0012 |
| gut - tongue | 0.5474 | 0.001 | 0.0012 | gut - tongue | 1 | 0.002 | 0.0024 | gut - tongue | 0.3046 | 0.001 | 0.0012 |
| left palm - right palm | 0.0585 | 0.544 | 0.544 | left palm - right palm | -0.0538 | 0.78 | 0.78 | left palm - right palm | -0.0018 | 0.543 | 0.543 |
| left palm - tongue | 0.2985 | 0.001 | 0.0012 | left palm - tongue | 0.6953 | 0.001 | 0.0015 | left palm - tongue | 0.1373 | 0.001 | 0.0012 |
| right palm - tongue | 0.276 | 0.001 | 0.0012 | right palm - tongue | 0.5343 | 0.001 | 0.0015 | right palm - tongue | 0.1412 | 0.001 | 0.0012 |
| Download pairwise PERMANOVA table | | Download pairwise ANOSIM table | | Download pairwise MRPP 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. The adjusted p-values are multiple testing corrected using the Benjamini-Hochberg method.
- * 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

897

learn more

Number of threads MOCHI can use

6

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

Start!

Start

example files

* 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 1 2 3 4 Next

 Download Faith PD table

Download

* Faith PD table:

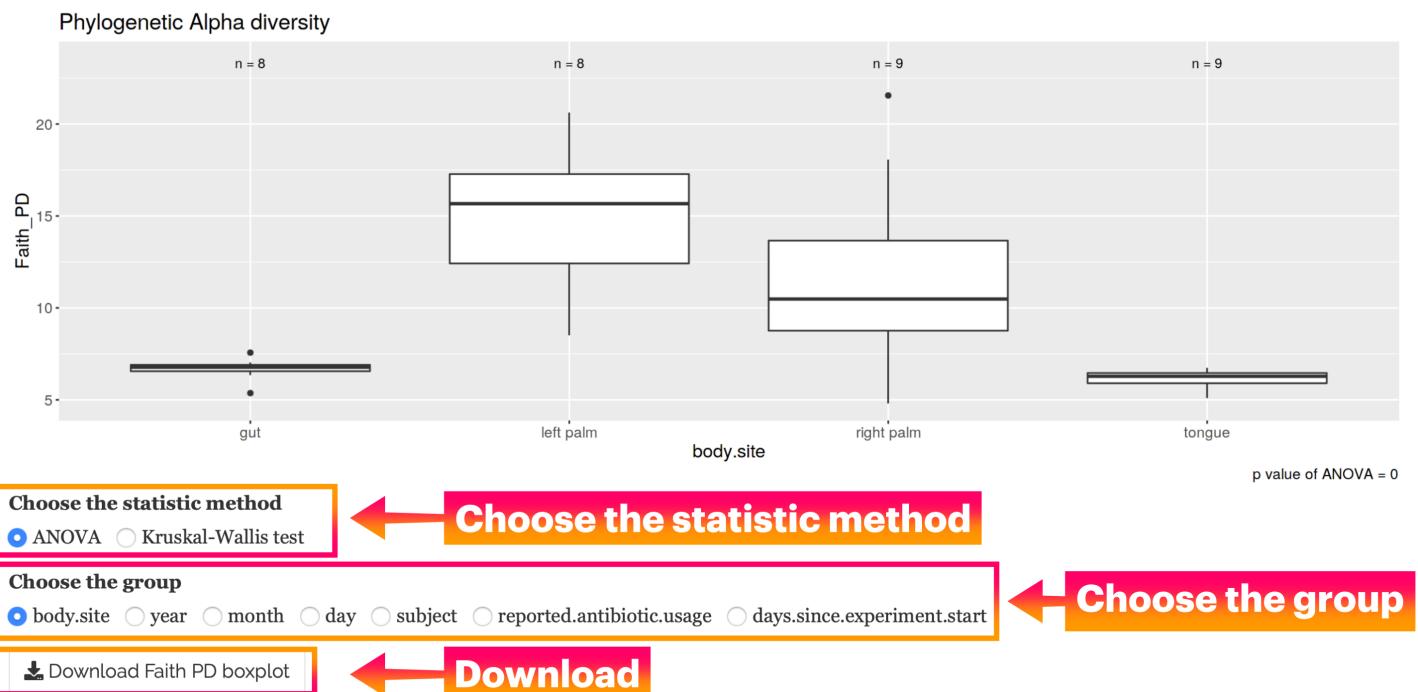
This table shows the Faith PD (phylogenetic diversity) of all samples.

* Download the 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 the Faith PD boxplot:

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

7.3. Post hoc analysis

Post hoc analysis

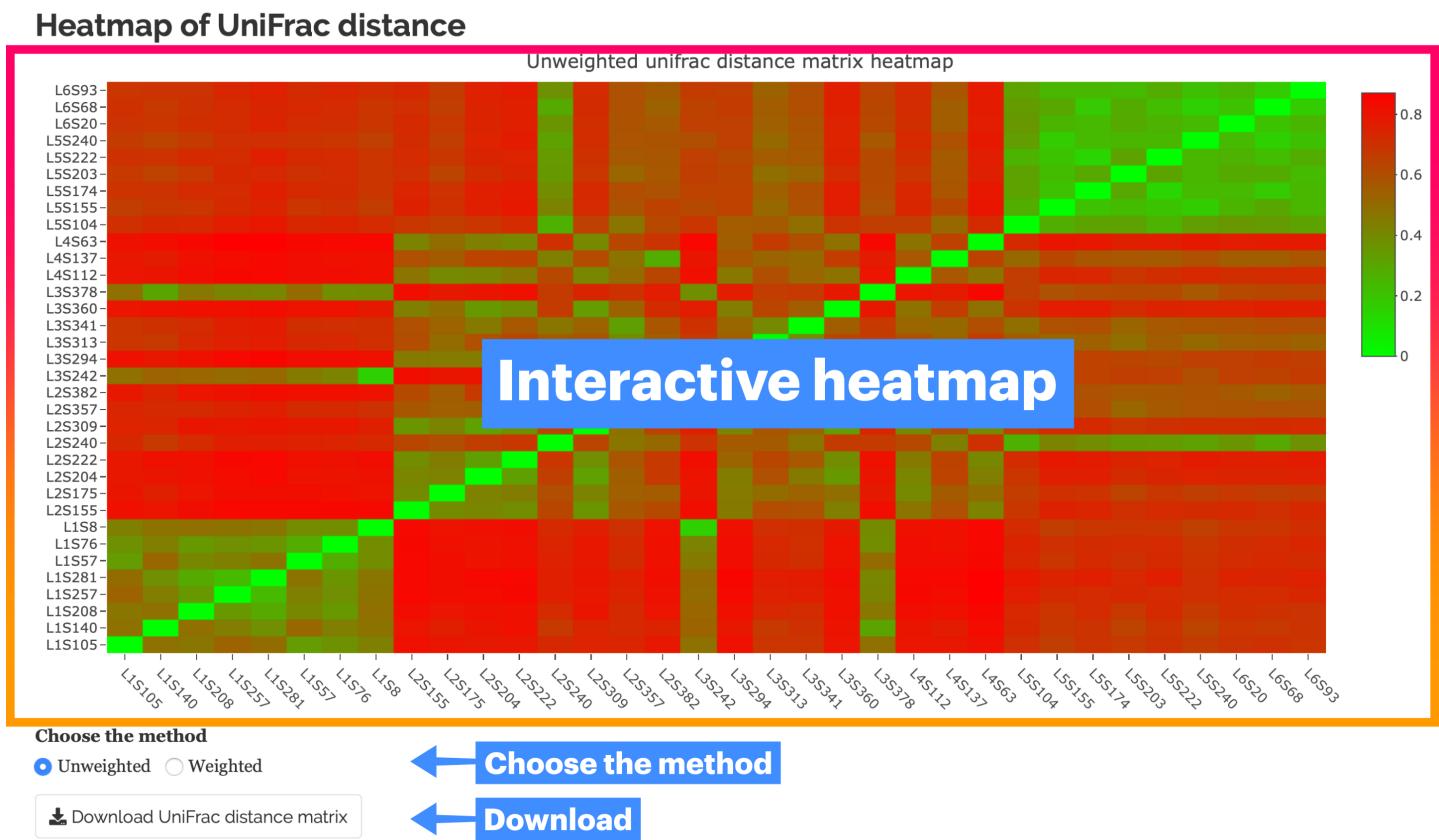
Tukey test

| 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 the 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 the 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 in the top-right region of the heatmap to download the plot.

* **Choose the method:**

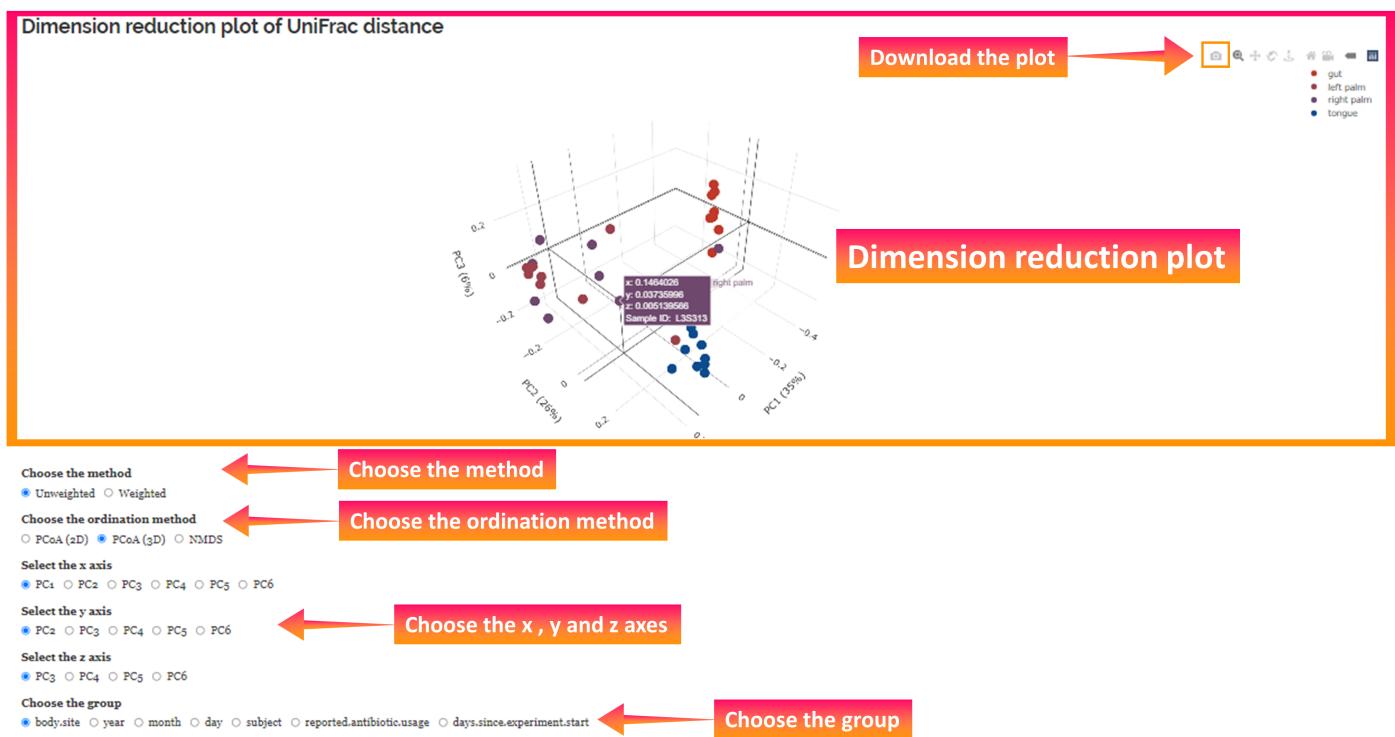
Select unweighted UniFrac (not consider the richness of taxa) or weighted UniFrac (consider the richness of taxa).

* **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



* 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. We provide two dimension-reduction methods for visualizing UniFrac distance: PCoA (Principal Co-ordinates Analysis, 2D & 3D) and NMDS (Non-metric Multidimensional Scaling).

* Choose the axes:

Individually select the PCs as x/y axes or x/y/z axes for 2D or 3D plots.

* 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 the UniFrac plot:

Click on the camera icon to download the plot.

7.6. Statistical analysis

Statistical analysis

PERMANOVA

| R^2 | P value |
|--------|---------|
| 0.1567 | 0.042 |

[Download PERMANOVA table](#)

ANOSIM

| R | P value |
|--------|---------|
| 0.1219 | 0.027 |

[Download ANOSIM table](#)

MRPP

| A | P value |
|--------|---------|
| 0.0488 | 0.02 |

[Download MRPP table](#)

Pairwise PERMANOVA

| Comparisons | R^2 | P value | Adjusted P value |
|------------------------|--------|---------|------------------|
| gut - left palm | 0.5049 | 0.001 | 0.0012 |
| gut - right palm | 0.3261 | 0.001 | 0.0012 |
| gut - tongue | 0.6587 | 0.001 | 0.0012 |
| left palm - right palm | 0.0649 | 0.394 | 0.394 |
| left palm - tongue | 0.4563 | 0.001 | 0.0012 |
| right palm - tongue | 0.3071 | 0.001 | 0.0012 |

[Download pairwise PERMANOVA table](#)

Pairwise ANOSIM

| Comparisons | R | P value | Adjusted P value |
|------------------------|---------|---------|------------------|
| gut - left palm | 0.9933 | 0.002 | 0.0024 |
| gut - right palm | 0.5742 | 0.001 | 0.0015 |
| gut - tongue | 1 | 0.001 | 0.0015 |
| left palm - right palm | -0.0191 | 0.504 | 0.504 |
| left palm - tongue | 0.7509 | 0.001 | 0.0015 |
| right palm - tongue | 0.4767 | 0.001 | 0.0015 |

[Download pairwise ANOSIM table](#)

Pairwise MRPP

| Comparisons | A | P value | Adjusted P value |
|------------------------|--------|---------|------------------|
| gut - left palm | 0.2643 | 0.001 | 0.0012 |
| gut - right palm | 0.1556 | 0.001 | 0.0012 |
| gut - tongue | 0.3811 | 0.001 | 0.0012 |
| left palm - right palm | 0.0008 | 0.45 | 0.45 |
| left palm - tongue | 0.2476 | 0.001 | 0.0012 |
| right palm - tongue | 0.1856 | 0.001 | 0.0012 |

[Download pairwise MRPP 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. The adjusted p-values are multiple testing corrected using the Benjamini-Hochberg method.

- * 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:

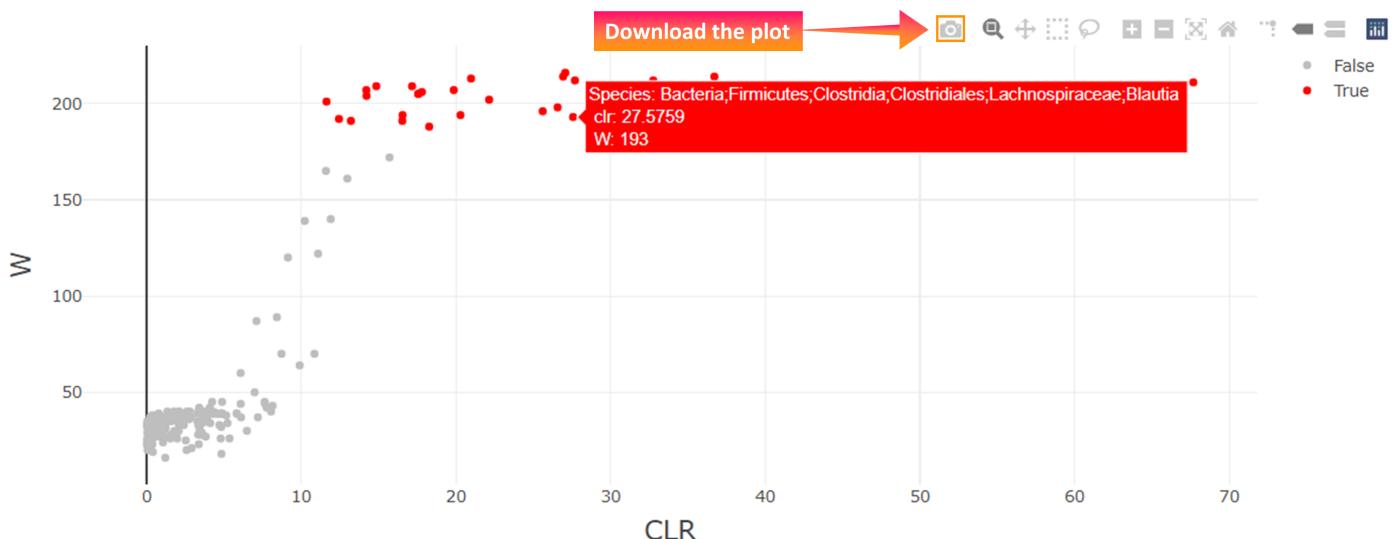
Select a taxonomic level for comparison.

* Start:

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

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

ANCOM Volcano Plot (Order)



The W value in the y axis is the number of sub-hypotheses that have rejected for a given taxon in ANCOM analysis.
The clr in the x axis 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**” in 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 file and taxonomic table. Alternatively, you can press the “Example files” button to download the example files first and then upload.

The image shows the MOCHI upload interface. A large red box highlights the left panel where users upload files. An orange arrow points from the "Upload" button in the right panel to the "Browse..." button in the left panel. The right panel contains links for "Function annotation table" and "Function barplot". The bottom left of the image has a yellow box containing the "Start!" button, which is also highlighted with a red box. Other buttons in the bottom left include "Reset", "learn more", and "Example files".

Upload the metadata file ⓘ

Browse... No file selected

Choose file format

- MOCHI/QIIME2 output (.qza)
- Plain text table (.txt)

Upload the taxonomic table file ⓘ

Browse... No file selected

← Upload

Start! Reset

learn more

Example files

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](#)

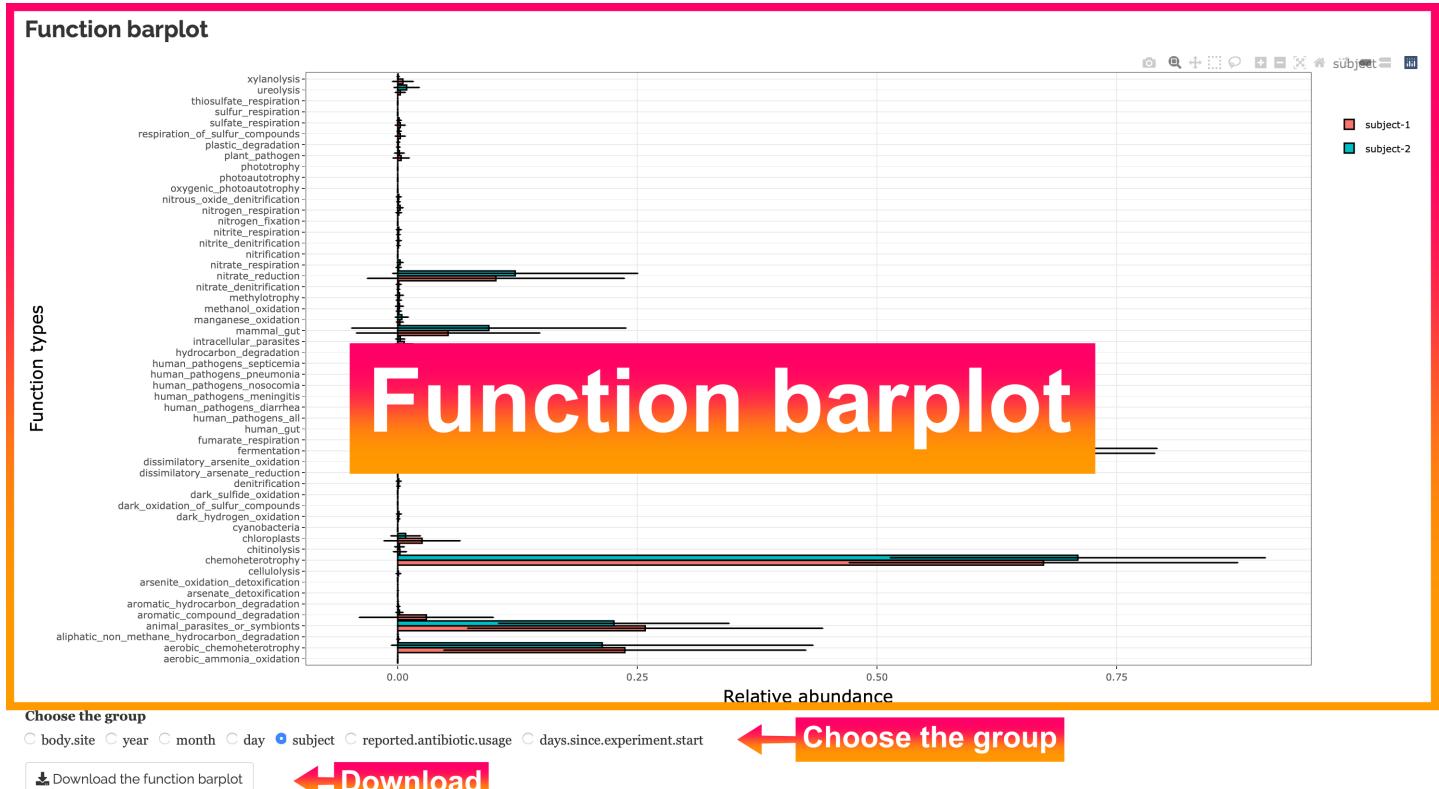
* Summary: Essential information regarding the function prediction.

* 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.