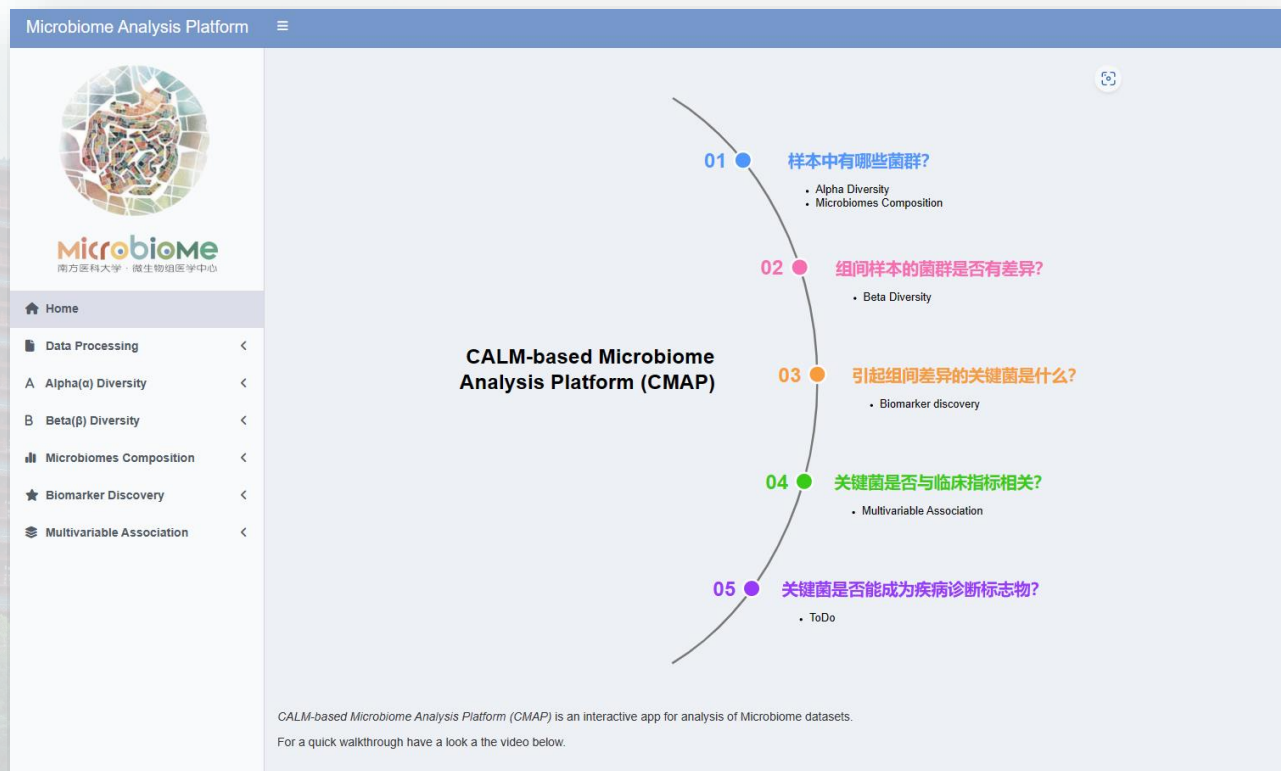


# CALM-based Microbiome Analysis Platform

## CMAP云平台使用说明

南方医科大学珠江医院检验医学部  
南方医科大学微生态医学中心

# CMAP主页



外网地址: <https://cmap.cpolar.top/>

南方医科大学校园网地址: 172.18.44.121

打开浏览器输入其中一个地址, 即可登录CMAP平台。

外网地址速度较慢, 校园网用户推荐在校园网范围内, 使用校园网地址登录。


# CMAP介绍



CMAP是本课题开发一个微生物组组学数据分析云平台，拥有功能强大且设计良好的交互式界面，综合了主流的微生物组学分析模块构建多样化的分析流程。

目的是为未经生物信息学系统培训的研究人员和临床医生能参考目前主流分析方法轻松挖掘微生物组学数据，包括数据预处理，统计分析和可视化。

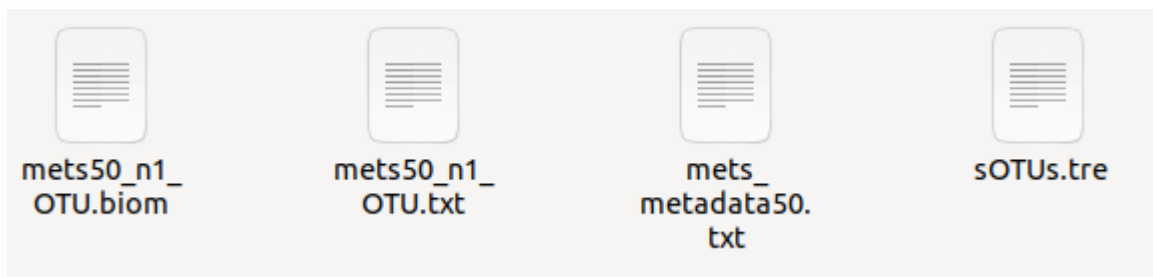
# 演示流程

- 
1. 下载和上传示例数据
  2. 预览和过滤数据
  3. 利用不同的分析模块挖掘微生物组学数据




## 示例数据

示例数据来源于GGMP（广东省肠道微生物组项目，Guangdong Gut Microbiome Project），从广东省随机抽样了14个地区，收集了7000例成年人的粪便，16S测序其V4区，并以问卷形式收集个人信息等数据。这里从中选择了25例代谢综合征患者，以及25例对照作为演示。



# 下载数据



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
  - Upload Data
  - Rarefaction
  - Meta Data Filter
- Alpha( $\alpha$ ) Diversity
- Beta( $\beta$ ) Diversity
- Microbiomes Composition
- Biomarker Discovery
- Multivariable Association

### Upload data

**Abundance profile (.txt or .biom):**



Browse... No file selected

**Metadata file (.txt):**

Browse... No file selected




**(Optional) Phylogenetic tree (.tre or .nwk):**

Browse... No file selected

Submit     Example Data     Tutorial

### Data preview

点击Example Data, 下载示例数据

-  demo\_metadata.txt
-  demo\_otu\_table.biom
-  demo\_phylogene.tre

# 上传数据



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
  - Upload Data
  - Rarefaction
  - Meta Data Filter
- Alpha( $\alpha$ ) Diversity
- Beta( $\beta$ ) Diversity
- Microbiomes Composition
- Biomarker Discovery
- Multivariable Association

### Upload data

**Abundance profile (.txt or .biom):**

Browse... demo\_otu\_table.biom

Upload complete

**Metadata file (.txt):**

Browse... demo\_metadata.txt

Upload complete

**(Optional) Phylogenetic tree (.tre or .nwk):**

Browse... demo\_phylogene.tre

Upload complete

Submit    Example Data    Tutorial

### Data preview

```
# A tbl_mpse (which can be converted to MPSE via as.MPSE) abstraction: 103,350 x 24
OTU      Sample Abund... RareA... distr... age gender occup... smoke... height weight BMI waist Brist... MetS Obesity Overw...
<chr>    <chr>    <dbl>    <dbl>    <chr>    <int> <chr>    <chr>    <chr>    <dbl>  <dbl>  <dbl> <dbl>    <int> <chr>    <chr>    <chr>
1 Seq4034 G4406... 0 0 G440606 65 f j n 147. 45.3 21.1 69.5 4 n n n
2 Seq4034 G4415... 0 0 G441502 53 f k n 146. 54.2 25.4 91.4 4 n n y
3 Seq4034 G4403... 0 0 G440305 63 f j n 148. 54.4 25 74.8 5 n n y
4 Seq4034 G4409... 0 0 G440981 52 f k n 152. 56 24.2 84.3 2 y n y
5 Seq4034 G4402... 0 0 G440205 70 m l n 160. 75.5 29.5 100. 4 y y y
6 Seq4034 G4414... 0 0 G441424 62 f j n 155. 56 23.2 80.5 4 n n n
7 Seq4034 G4453... 0 0 G445302 47 f k n 160. 72.2 28.3 85.2 4 y y y
8 Seq4034 G4401... 0 0 G440104 42 f c n 152. 54.6 23.6 77.7 5 n n n
9 Seq4034 G4415... 0 0 G441502 51 m a y 158. 57.5 23 77.6 5 n n n
10 Seq4034 G4408... 0 0 G440883 75 f k y 148. 50 22.7 84.6 3 y n n

# ... with 103,340 more rows, 7 more variables: Kingdom <chr>, Phylum <chr>, Class <chr>, Order <chr>, Family <chr>,
# Genus <chr>, Speies <chr>, and abbreviated variable names ^Abundance, ^RareAbundance, ^district, ^occupation, ^smoke_now,
# ^Bristol_stool_type, ^Overweight
# i Use `print(n = ...)` to see more rows, and `colnames()` to see all variable names
```

上传文件:

- 物种丰度表 (OUT/ASV表)
- 包含分组信息的meta表
- 进化树文件 (可选)

点击Submit后, Data preview出现  
数据预览即数据上传成功!

# 输入格式要求

1、**OTU/ASV表**为biom格式或txt格式的，OTU/ASV表的以样本名作为第一行(列名)，每一行为OTU/ASV的ID(行名)，最后一列的列名为taxonomy包含对应行OTU的注释信息，列与列之间必须以tab键分隔。

#OTU ID	G440606625	G441502228	G440305319	taxonomy
Seq4034	0	0	0	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__s__
Seq12622	214	14	2	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__Bacteroides; s__plebeius
Seq13264	122	273	255	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__Faecalibacterium; s__prausnitzii
Seq8703	1	9	0	k__Bacteria; p__Actinobacteria; c__Coriobacteriia; o__Coriobacteriales; f__Coriobacteriaceae; g__Collinsella; s__aerofaciens
Seq4307	110	13	26	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__Ruminococcus; s__bromii



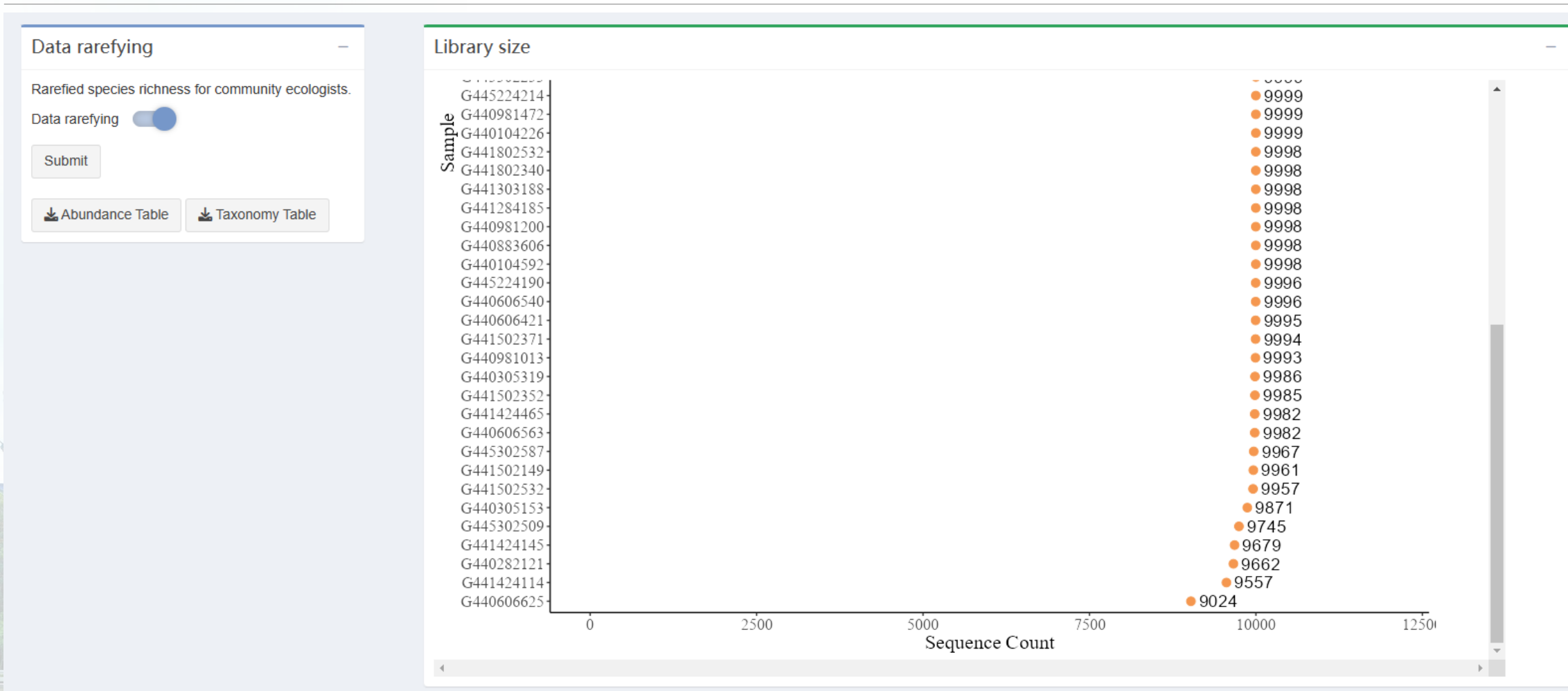
## 输入格式要求

2、含分组信息和/或环境因子信息的`metadata`，制表符分割的文本格式文件  
第一列列名必须为`#SampleID`，且样本名不能重复。

#SampleID	district	age	gender	MetS
G440104002	G440104	78	m	y
G440104226	G440104	42	f	n
G440104592	G440104	55	f	y
G440205073	G440205	47	f	n
G440205303	G440205	67	f	y


3、进化树文件（可选），该文件一般以`.tre`、`.nwk`为后缀名，包含了物种的进化关系。若缺少该文件将不能进行涉及物种进化关系的分析（如计算 $\alpha$ 多样性的PD指数、 $\beta$ 多样性的Unifrac距离等）

# 数据抽平



查看各样本的reads的总数，可以按最小样本进行数据抽平。

# 数据过滤



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
  - Upload Data
  - Rarefaction
  - Meta Data Filter**
- Alpha( $\alpha$ ) Diversity
- Beta( $\beta$ ) Diversity
- Microbiomes Composition
- Biomarker Discovery
- Multivariable Association

**Filter meta data**

Number of rows: 20 / 50

**Sample**  
50 items selected

**district**  
14 items selected

**age**  
21 52 81  
21 36 51 66 81

**gender**  
2 items selected

**occupation**  
8 items selected

**smoke\_now**  
2 items selected

**height**  
131 173  
131 142 152 163 173

**weight**  
35.5 78  
35.5 46.1 56.8 67.4 78

**BMI**  
17.8 32.8  
17.8 21.6 25.3 29 32.8

**waist**  
65 104.5

**Preview Filter:**

Sample	district	age	gender	occupation	smoke_now	height	weight	BMI	waist	Bristol_stool_type	MetS
G440981472	G440981	52	f	k	n	152	56	24.2	84.3	2	y
G445302253	G445302	47	f	k	n	159.6	72.2	28.3	85.15	4	y
G440104226	G440104	42	f	c	n	152	54.6	23.6	77.7	5	n
G441502371	G441502	51	m	a	y	158.2	57.5	23	77.6	5	n
G441502517	G441502	45	m	a	y	167	67	24	87	4	n
G441424145	G441424	33	f	f	n	151.5	57.5	25.1	75.5	2	n
G440282121	G440282	47	m	a	n	172	67.5	22.8	77.25	4	n
G440981013	G440981	50	m	a	y	173	74.2	24.8	103.6	4	y
G441284225	G441284	51	m	a	y	166	56.3	20.4	81	4	y
G440205358	G440205	44	f	f	n	152	56.4	24.4	78.05	4	y

1-10 of 20 rows

Previous 1 2 Next

[Save Filter Values](#)

**Filtered data:**

```

List of 14
$ Sample      : chr [1:50] "G440606625" "G441502228" "G440305319" "G440981472" ...
$ district    : chr [1:14] "G440606" "G441502" "G440305" "G440981" ...
$ age         : int [1:2] 21 52
$ gender      : chr [1:2] "f" "m"
$ occupation  : chr [1:8] "j" "k" "l" "c" ...
$ smoke_now   : chr [1:2] "n" "y"
$ height      : int [1:2] 131 173
$ weight      : num [1:2] 35.5 78
$ BMI         : num [1:2] 17.8 32.8
$ waist       : num [1:2] 65 104
$ Bristol_stool_type: int [1:2] 2 5
$ MetS        : chr [1:2] "n" "y"
$ Obesity     : chr [1:2] "n" "y"
$ Overweight  : chr [1:2] "n" "y"

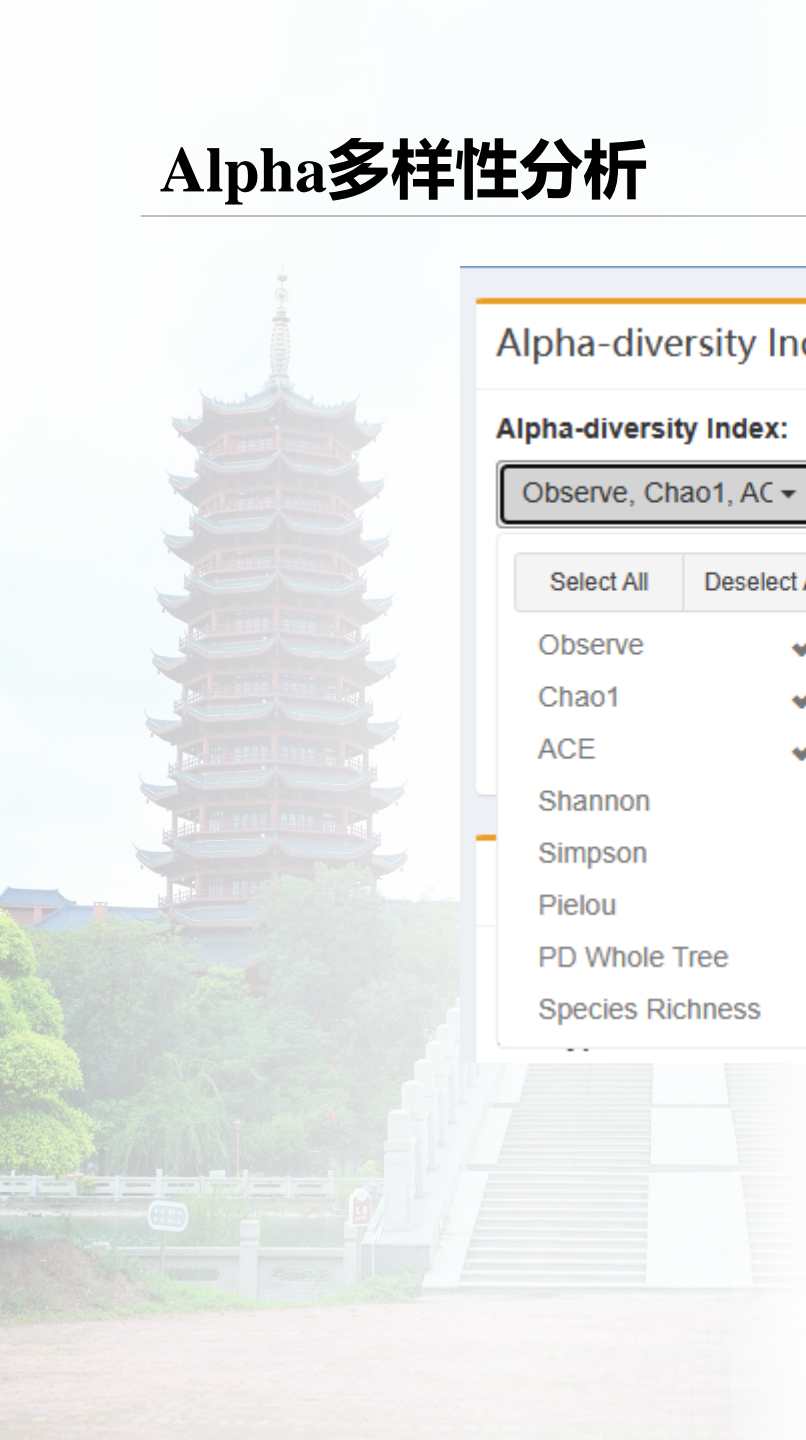
```

2. 筛选预览窗口，根据筛选器实时变动。

1. 数据筛选器

3. 筛选后数据，点击Save Filter Values, 保存筛选后的数据。

# Alpha多样性分析



Alpha-diversity Index

Alpha-diversity Index: **Observe, Chao1, AC**

Statistical method: **robust**

Type: **discrete**

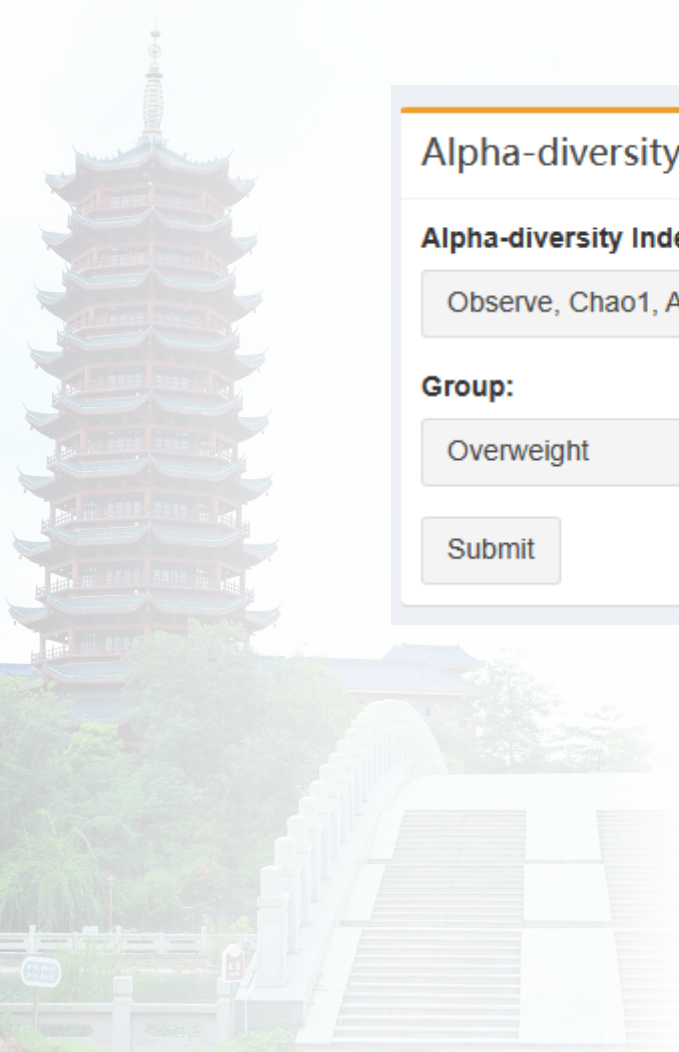
Select All Deselect All

- Observe ☒
- Chao1 ☒
- ACE ☒
- Shannon ☐
- Simpson ☐
- Pielou ☐
- PD Whole Tree ☐
- Species Richness ☐

1. 选择Alpha多样性指数
2. 组间差异检验方法
3. 选择分组
4. 点击Submit, 分析完成



# Alpha多样性分析



Alpha-diversity Index

Alpha-diversity Index: Observe, Chao1, AC ▼

Group: Overweight ▼

Submit

Statistical method:

robust ▼

parametric


nonparametric

robust

bayes

1. 选择Alpha多样性指数
2. 组间差异检验方法
3. 选择分组
4. 点击Submit，分析完成

# Alpha多样性分析



Alpha-diversity Index

Alpha-diversity Index: Observe, Chao1, AC

Statistical method: robust

Group: MetS

Type: discrete

district

age

gender

occupation

smoke\_now

height

weight

BMI

waist

Bristol\_stool\_type

MetS

Obesity

Overweight

boxviolin

Height: 8

DPI:

1. 选择Alpha多样性指数
2. 组间差异检验方法
3. 选择分组
4. 点击Submit, 分析完成

# Alpha多样性指数分析

Alpha-diversity Index

Alpha-diversity Index: Observe, Chao1, AC

Statistical method: robust

Group: Overweight

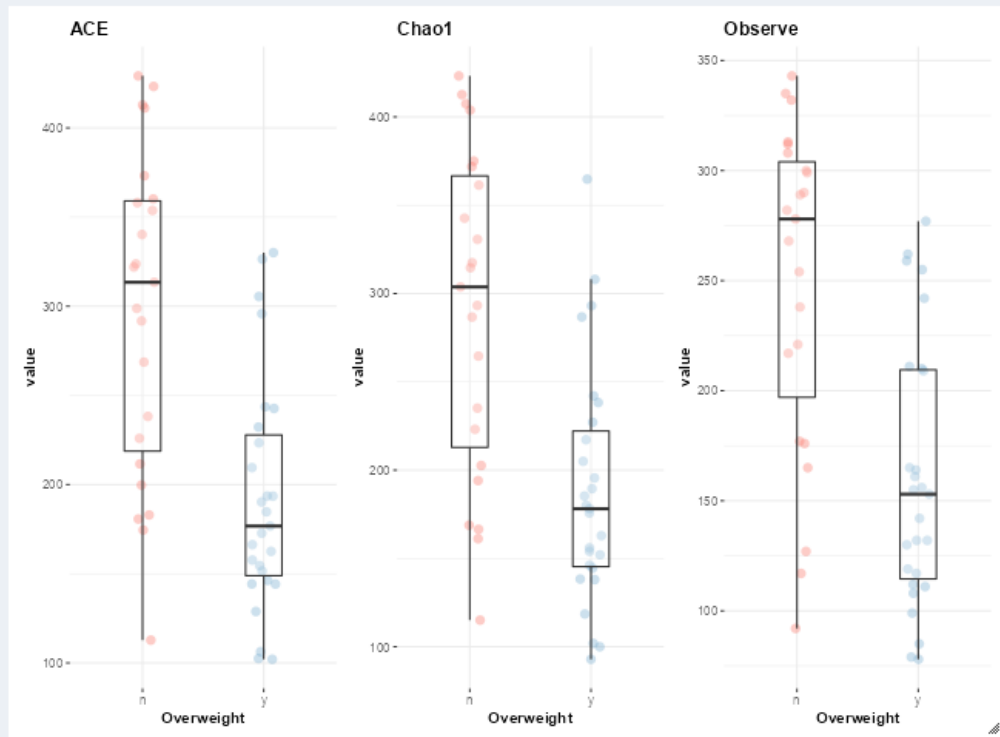
Type: discrete

Submit

Graphics Options Color Download

Whether to show:  
☐ statistical tests results

Plot type  
☒ box ☐ violin ☐ boxviolin



1. 选择Alpha多样性指数
2. 组间差异检验方法
3. 选择分组
4. 点击Submit, 分析完成

# 分析模块架构

1

Alpha-diversity Index

Alpha-diversity Index: Observe, Chao1, AC

Statistical method: robust

Group: Overweight

Type: discrete

Submit

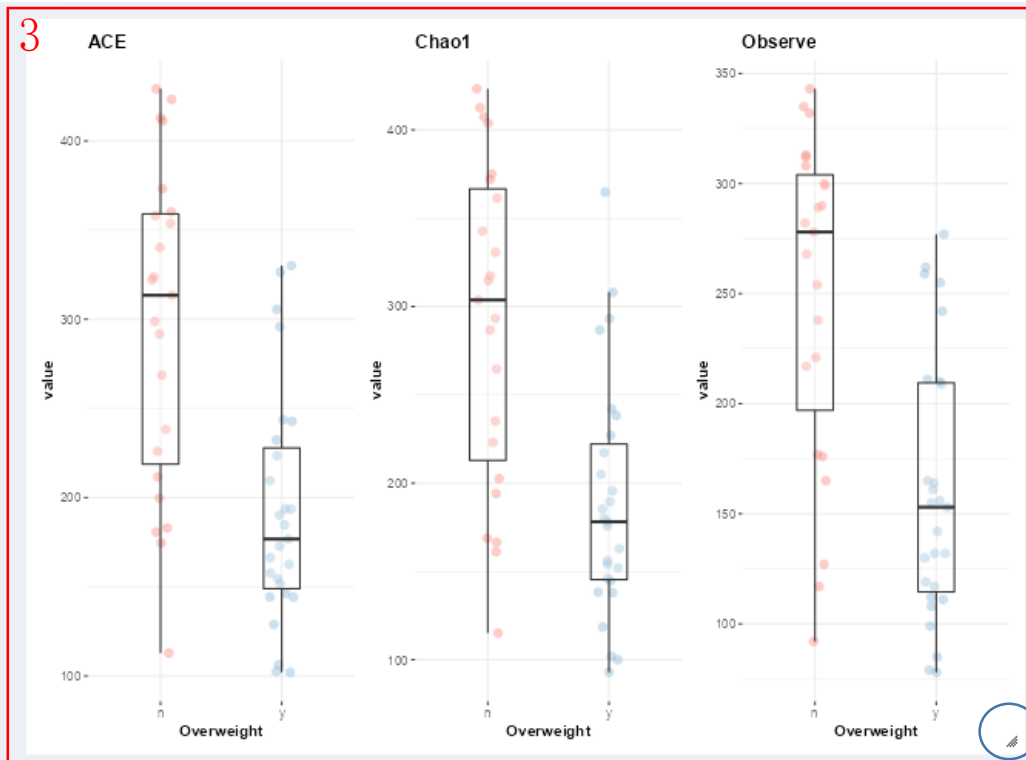
2

Graphics Options

Color Download

Whether to show:  
☐ statistical tests results

Plot type  
☒ box ☐ violin ☐ boxviolin



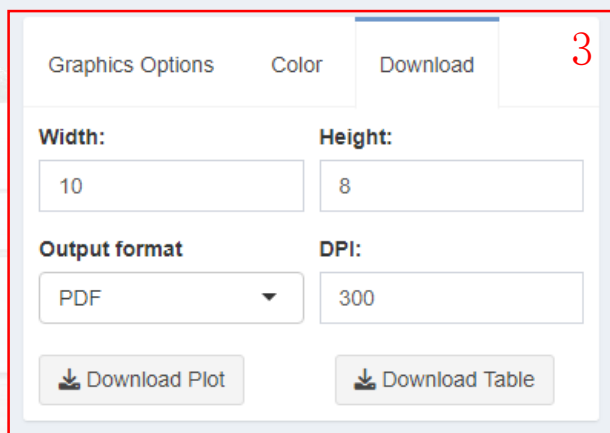
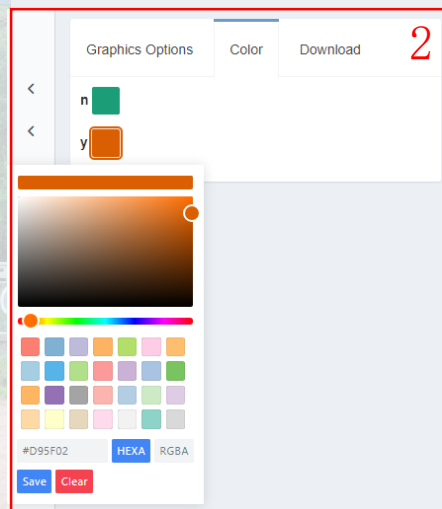
CMAF所有的分析方法都采用  
**模块化**设计:

1. 分析模块 -- 调整分析方法所用参数。
2. 图形设置 -- 调整输出结果的可视化效果。
3. 图形预览 -- 可视化输出的简单预览。

拖动右下角可  
缩放预览图形



# 分析模块架构



图形设置模块可以调整可视化输出结果：

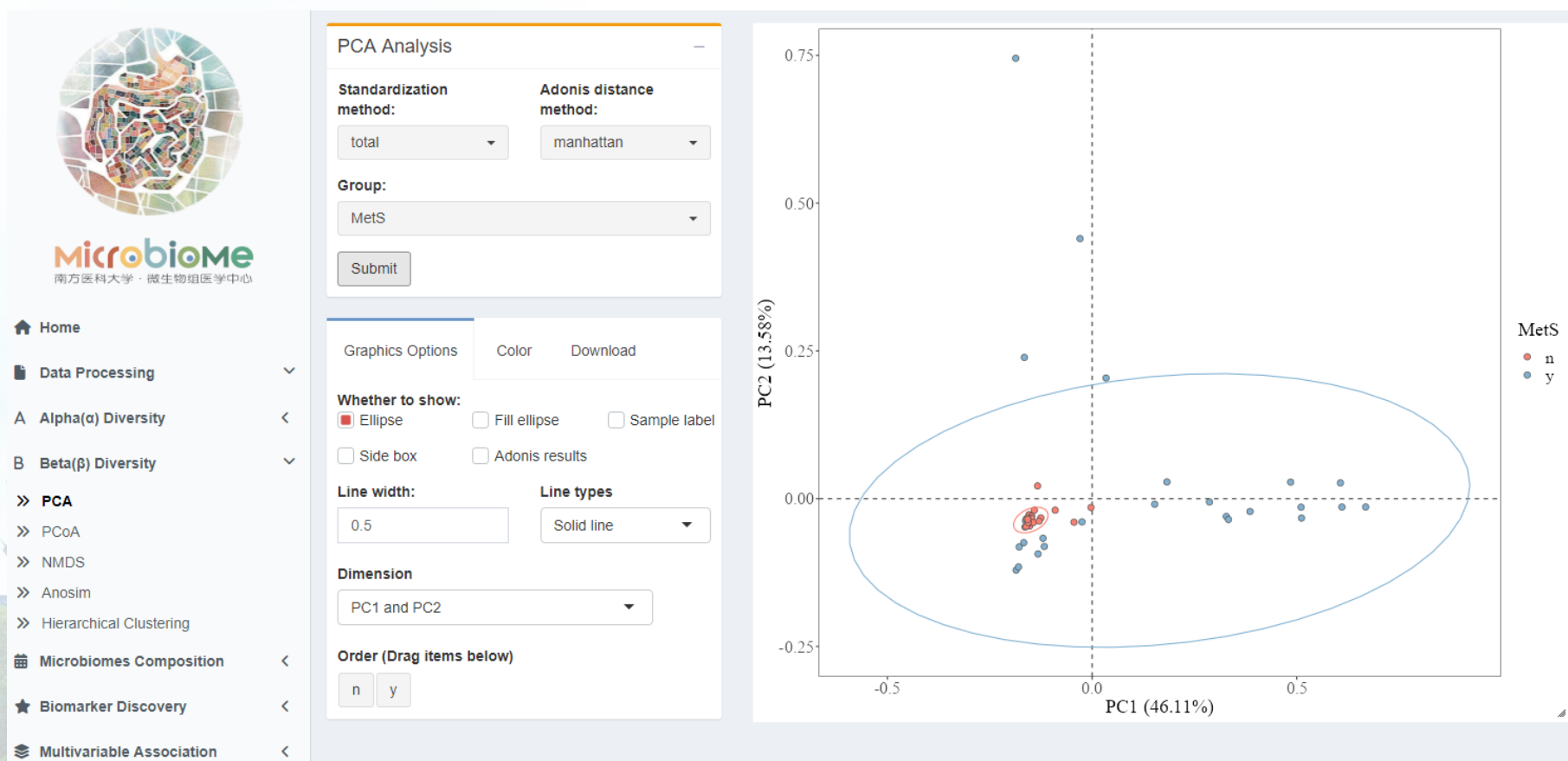
1. Graphics Options -- 微调可视化输出的类型。
2. Color -- 调整颜色。点击对应变量的颜色方块可以调出调色板。
3. Download -- 可以下载可视化结果的矢量图以及可视化结果的对应数据表格。可通过Width; Height等参数调整下载图形的宽高。

# 稀释曲线



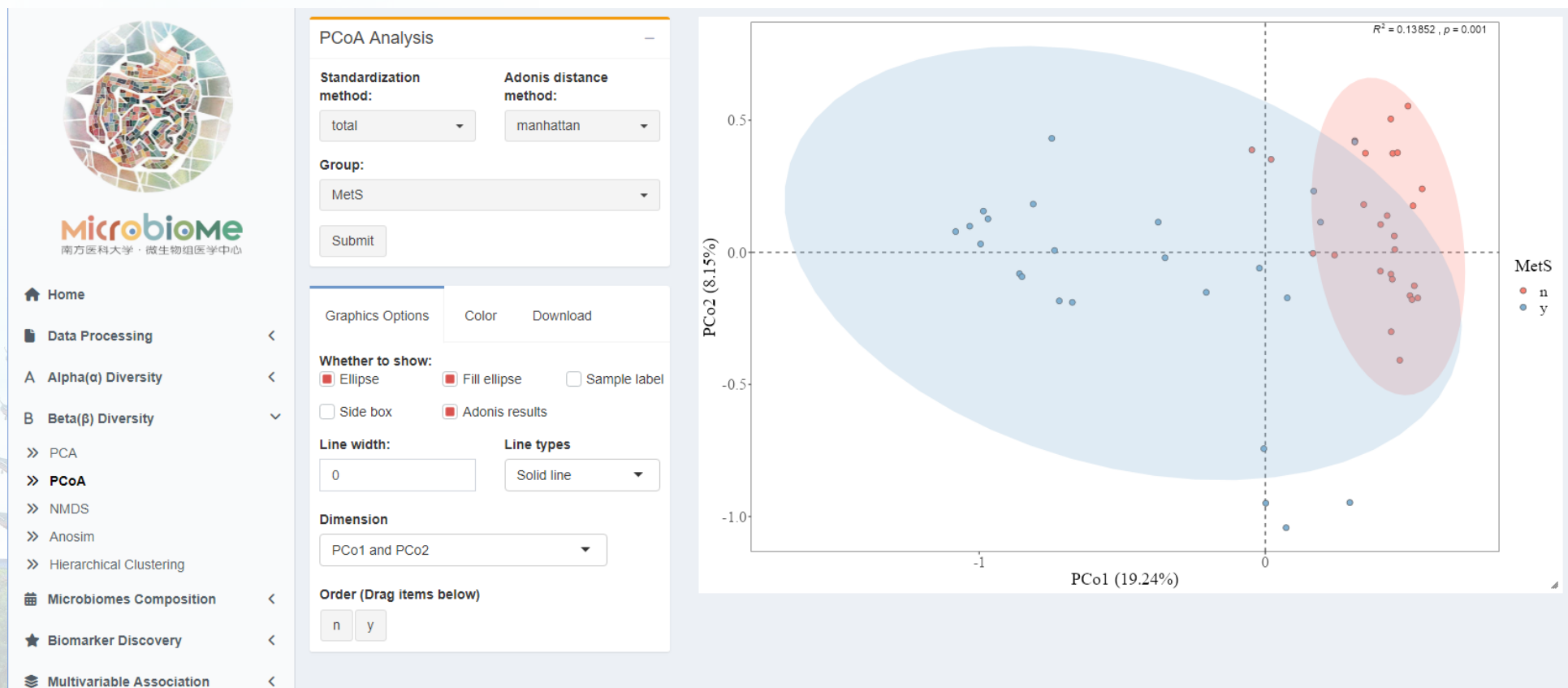
通过稀释曲线检查样本的测序深度是否足够（曲线是否趋于平缓）。

# PCA分析



PCA分析（Principal Component Analysis），是一种有效去除噪音和冗余，将原有的复杂数据降维分析的方法，揭示隐藏在复杂数据背后的简单结构。

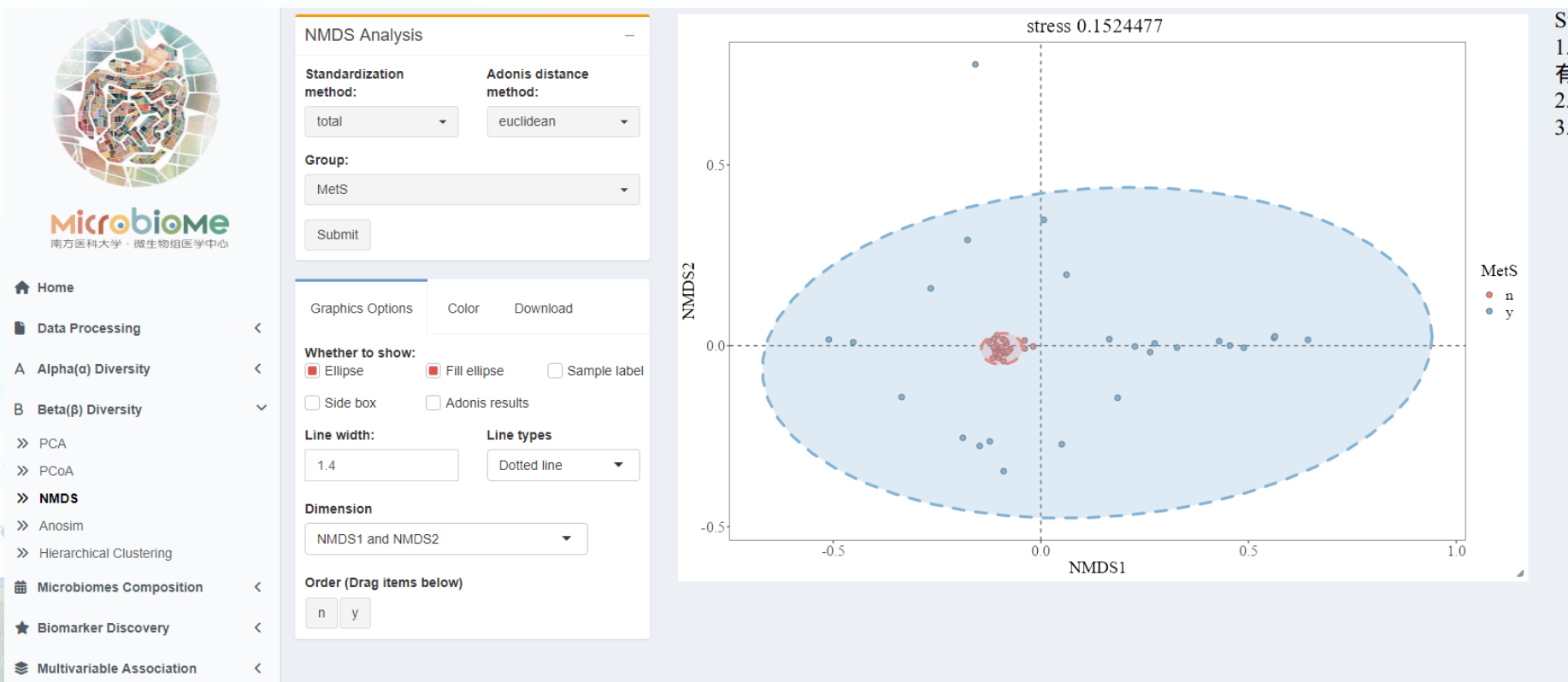
# PCoA分析



PCoA分析与PCA类似，二者都是通过降维找出影响样本群落组成差异的潜在主成分。主要区别在于，PCA利用物种丰度和基于欧氏距离直接作图，而PCoA是基于所选距离矩阵进行作图（请选择度量距离的方法）。



# NMDS分析




Stress值：检验NMDS分析结果的优劣。通常认为

1. 当stress < 0.2时可用NMDS的二维点图表示，其图形有一定的解释意义；
2. 当stress < 0.1时，可以认为是一个好的排序；
3. 当stress < 0.05时，则具有很好的代表性。

NMDS分析，即非度量多维尺度分析(non-metric multi-dimensional scaling)是一种将多维空间的研究对象（样本或变量）简化到低维空间进行定位、分析和归类，同时又保留对象间原始关系的数据分析方法。

# ANOSIM分析



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
  - Alpha( $\alpha$ ) Diversity
  - Beta( $\beta$ ) Diversity
    - PCA
    - PCoA
    - NMDS
    - Anosim**
    - Hierarchical Clustering
  - Microbiomes Composition
  - Biomarker Discovery
  - Multivariable Association

### Anosim Analysis

Standardization method: total

Distance method: manhattan

Group: Overweight

Permutations: 999

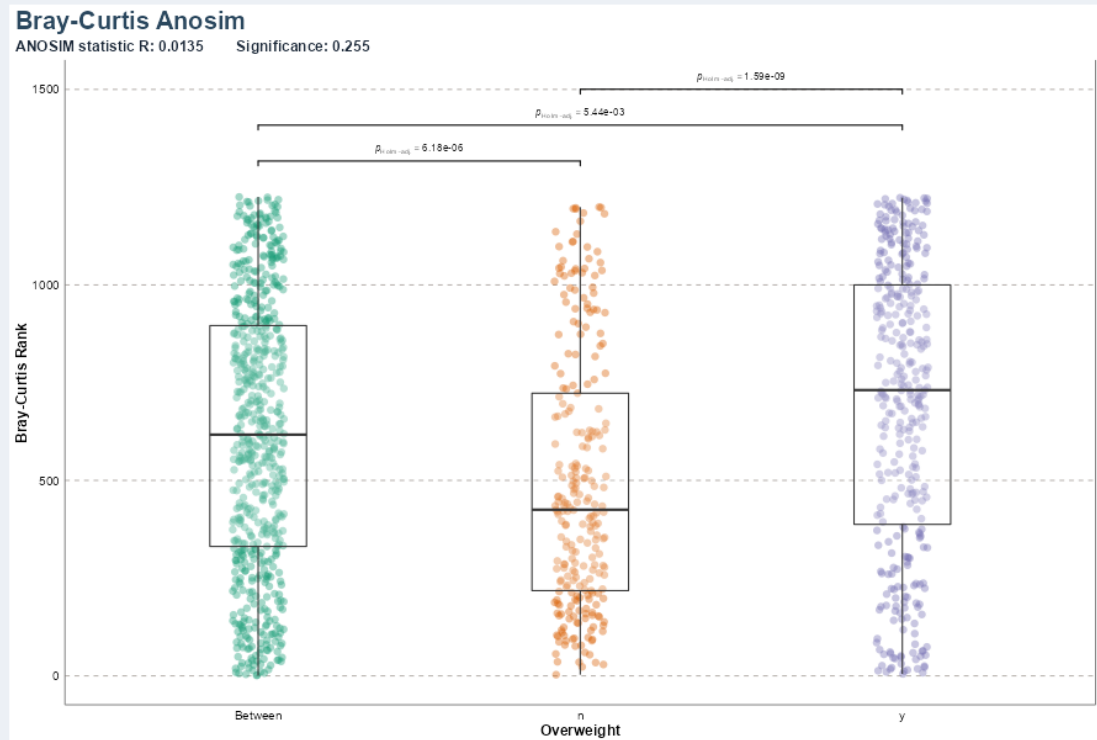
Submit

Color Download

Between

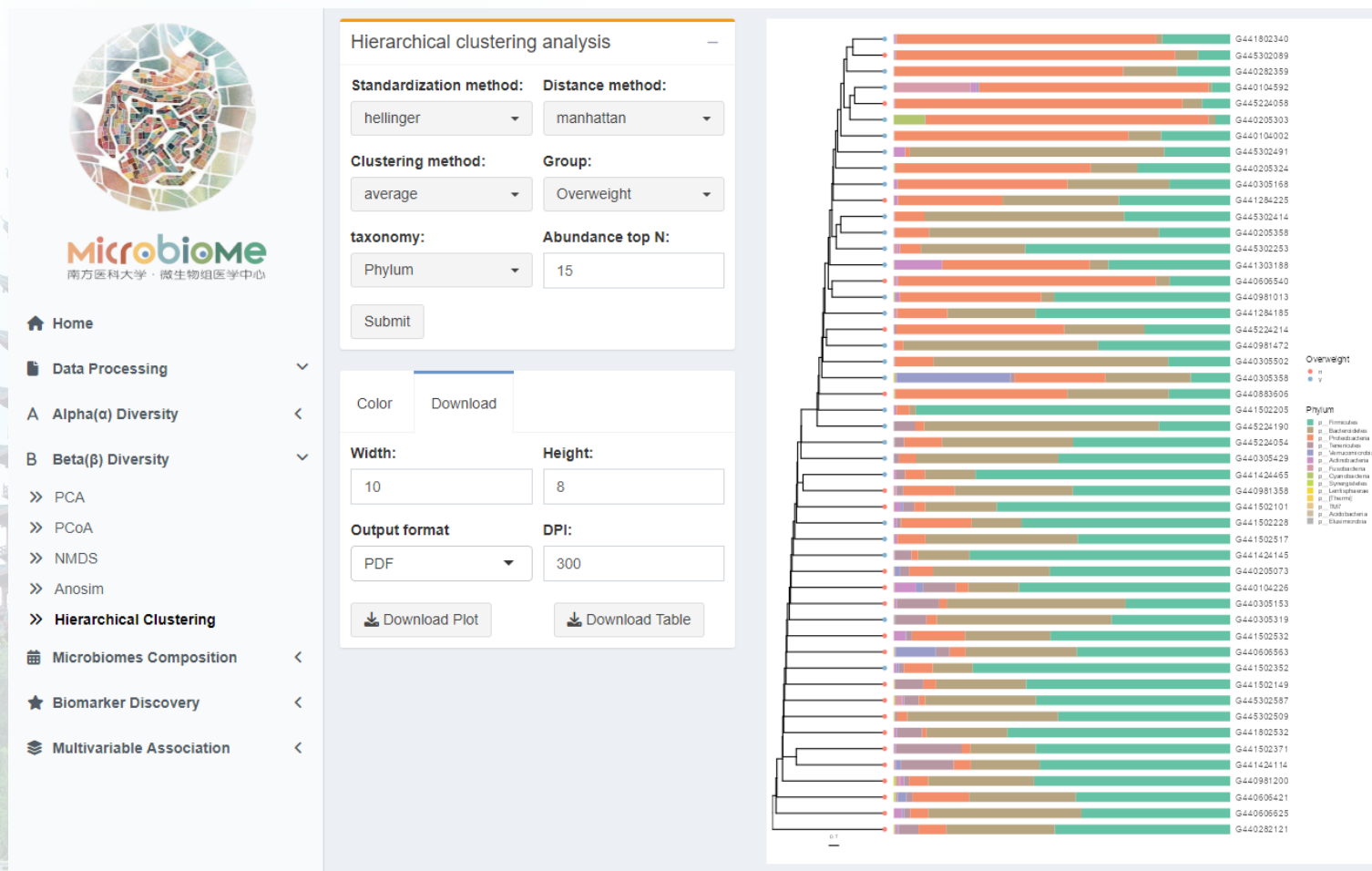
n

y



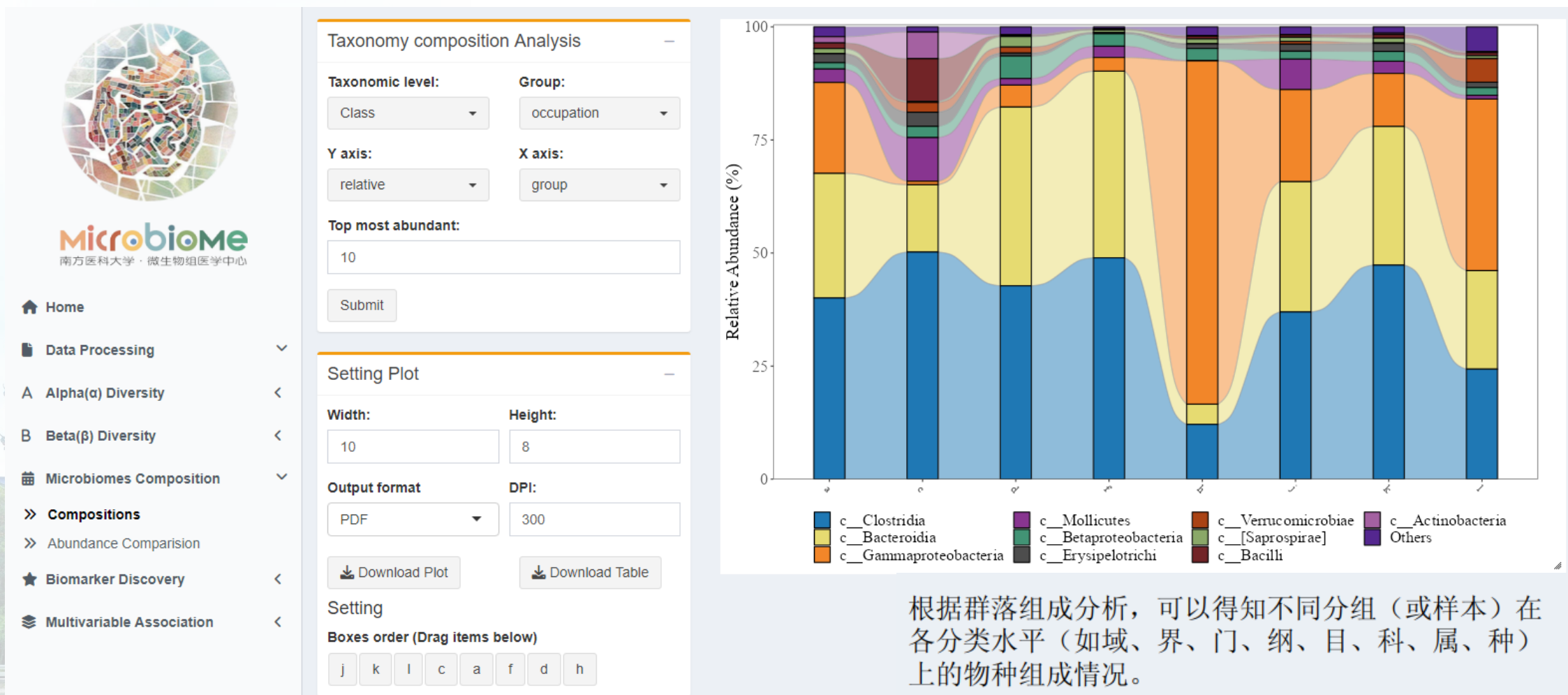
ANOSIM分析用来检验组间（两组或多组）的差异是否显著大于组内差异，从而判断分组是否有意义。

## 样本层次聚类分析



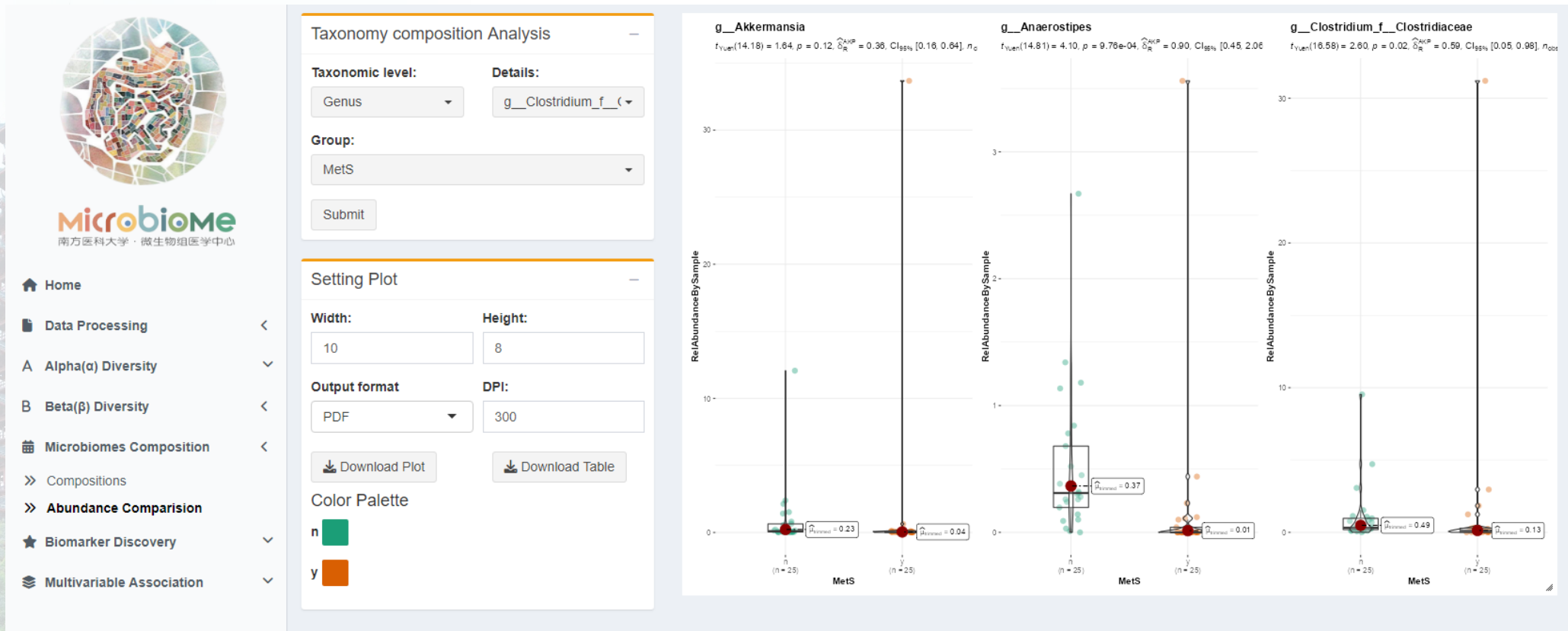
对距离矩阵进行层级聚类（Hierarchical clustering）可以清楚地看出样本分支的距离远近。

# 群落组成分析






# 群落组成分析



在Details中选择指定的菌种，可以比较不同分组在各分类水平上的特征菌群丰度。

# 探索生物标志物



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
- Alpha( $\alpha$ ) Diversity
- Beta( $\beta$ ) Diversity
- Microbiomes Composition
- Biomarker Discovery
  - » **LEfSe**
  - » Diff-Abundance plot
- Multivariable Association

### LEfSe Analysis

LDA score threshold:  P value cutoff:

P value adjust:  Taxonomy level:

Group:

### Setting Plot

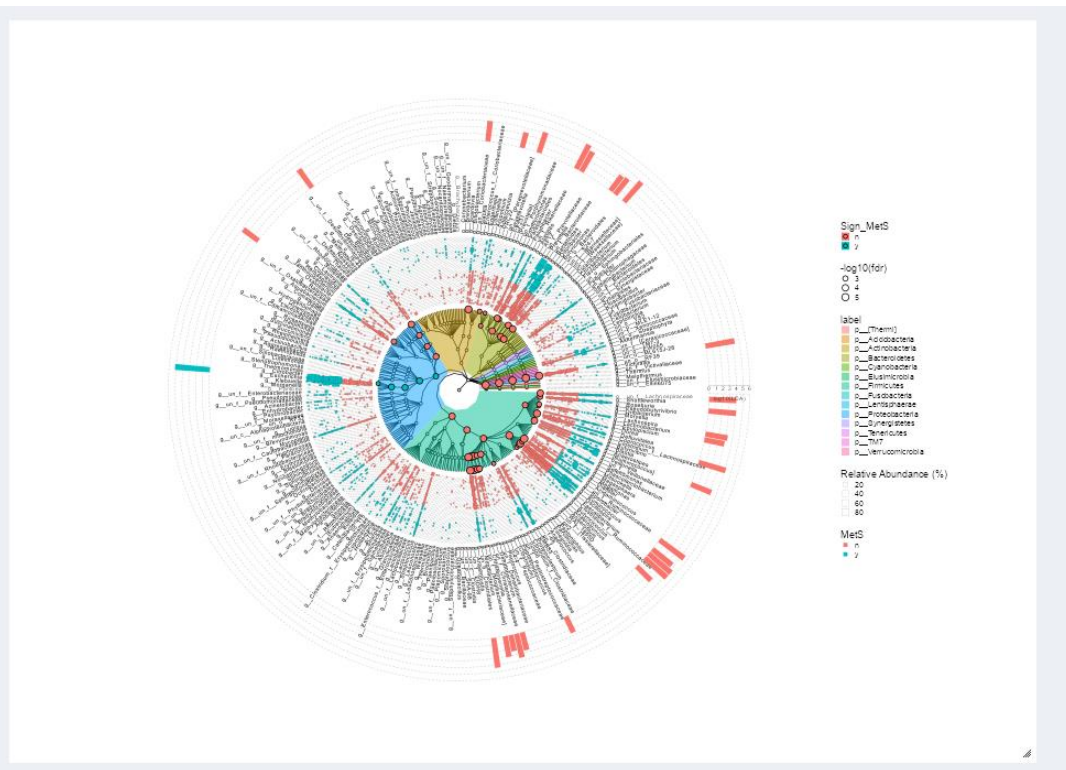
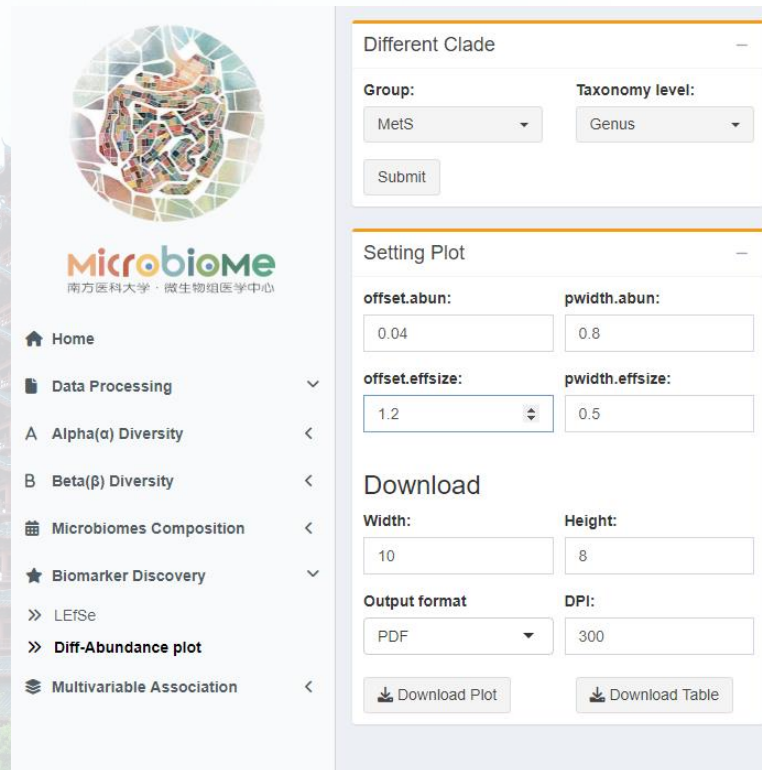
Width:  Height:

Output format:  DPI:



LEfSe分析首先使用非参数Kruskal-Wallis 秩和检测不同分组间丰度差异显著的物种，然后使用Wilcoxon秩和检验上一步的差异物种在不同组间子分组中的差异一致性，最后采用线性回归分析（LDA）来估算每个组分（物种）丰度对差异效果影响的大小。

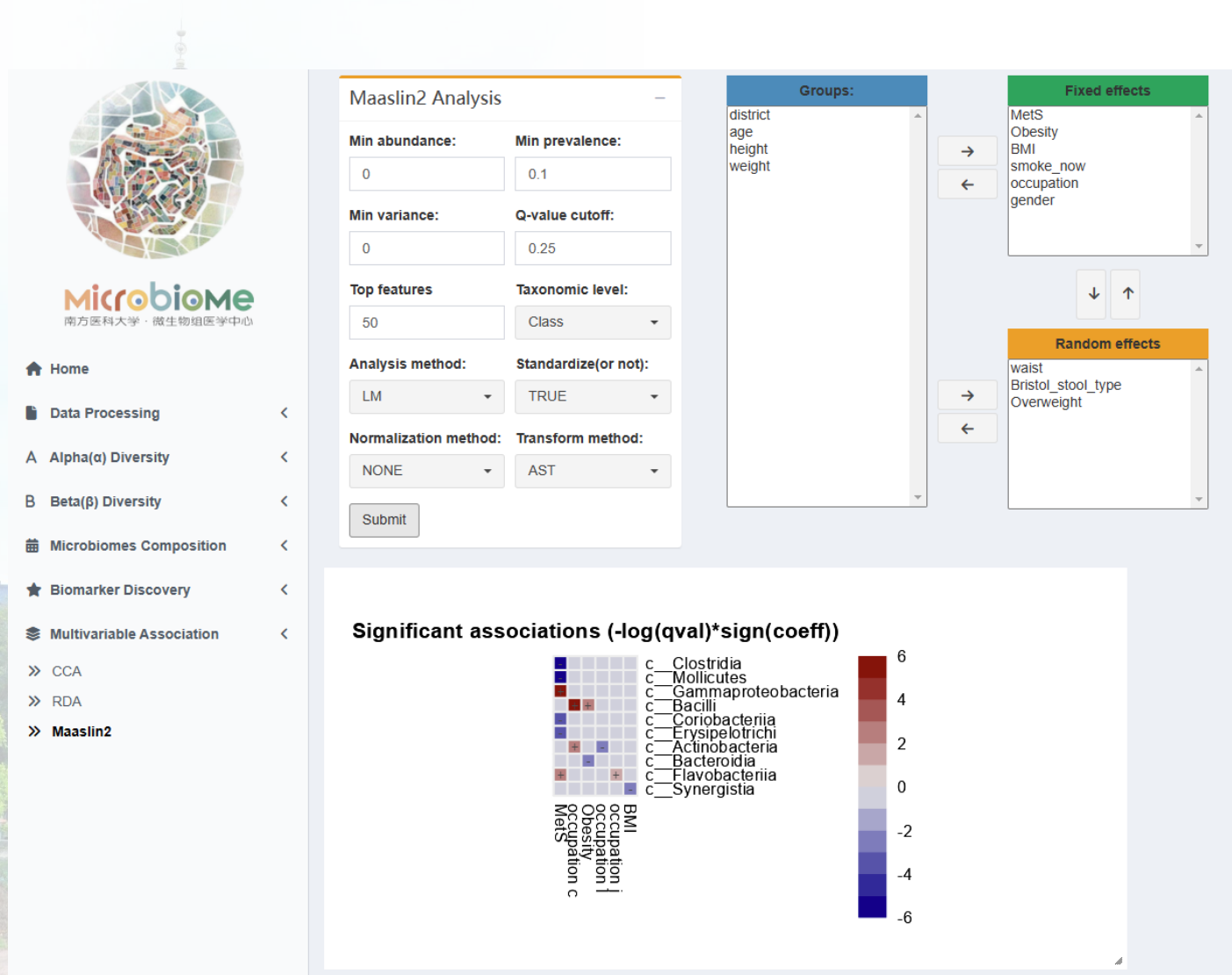
## 探索生物标志物



Different Clade的可视化方式是与Lefse稍微不同，内圈为差异物种树、中圈为各样本中菌群丰度热图、外圈为LDA值。



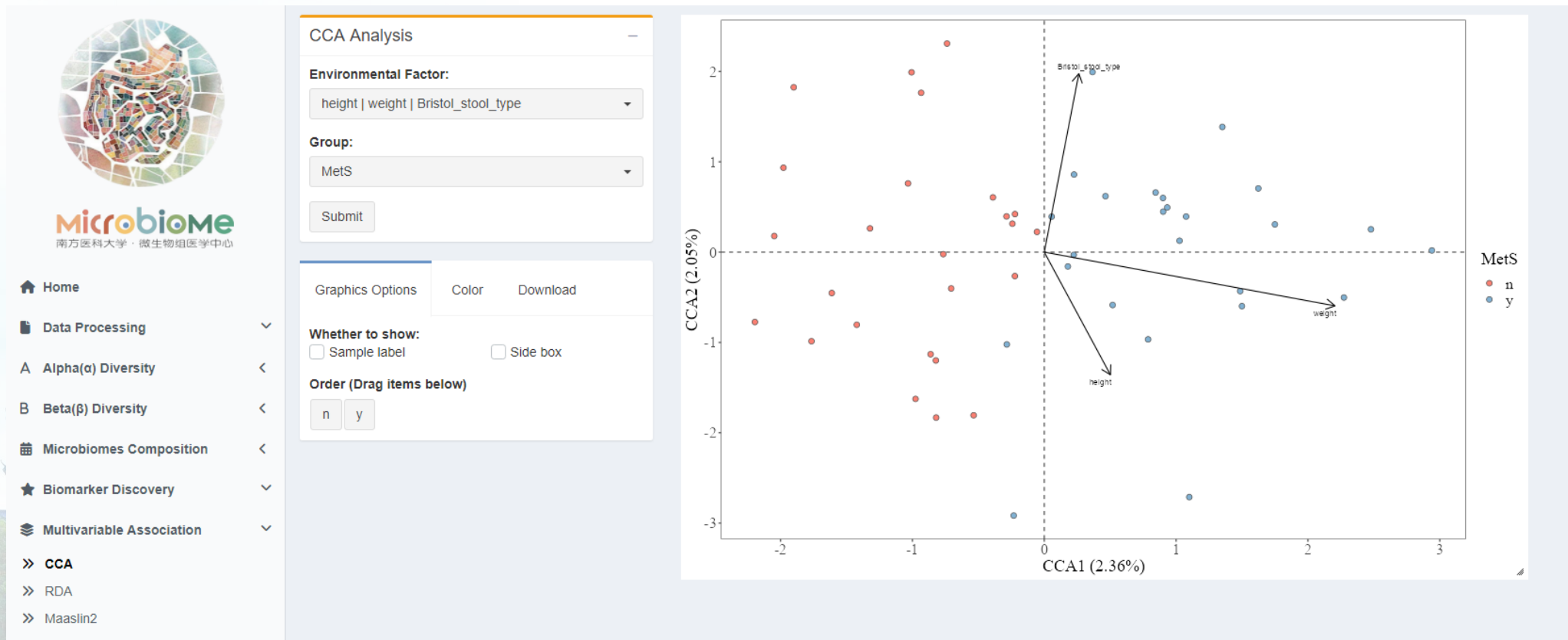
# 临床因子相关性分析



MaAslin (Multivariate Association with Linear Models) 分析是一种通过线性模型挖掘环境因子（如临床数据、理化指标等）与微生物群落物种或功能相对丰度数据间相关性的分析。




# 临床因子相关性分析



典范对应分析 (canonical correspondence analysis, CCA) 和冗余分析 (redundancy analysis, RDA) 是基于对应分析 (correspondence analysis, CA) 发展而来的一种排序方法, 将对应分析与多元回归分析相结合, 每一步计算均与环境因子进行回归, 又称多元直接梯度分析。此分析是主要用来反映菌群与环境因子之间关系。RDA是基于线性模型, CCA 是基于单峰模型。

# 阴道菌群CSTs分型



**Microbiome**  
南方医科大学·微生物组医学中心

- Home
- Data Processing
  - Upload Data
  - Rarefaction
  - Meta Data Filter
- A Alpha(α) Diversity
- B Beta(β) Diversity
- Microbiomes Composition
- Biomarker Discovery
- Multivariable Association
- Vaginal CSTs**

### Vaginal CSTs analysis

Based on a cross-sectional study of 394 healthy women at reproductive ages, Ravel et al classified the human vaginal microbial communities (HVMC) into five community state types (CSTs).

Powered by VALENCIA

Run CST architecture info

Download

Width: 10 Height: 8

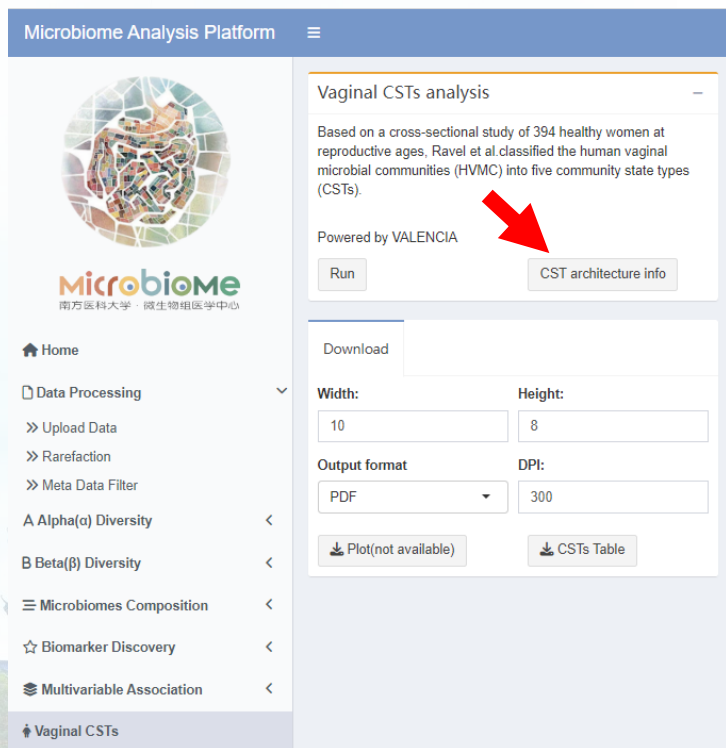
Output format: PDF DPI: 300

Plot(not available) CSTs Table

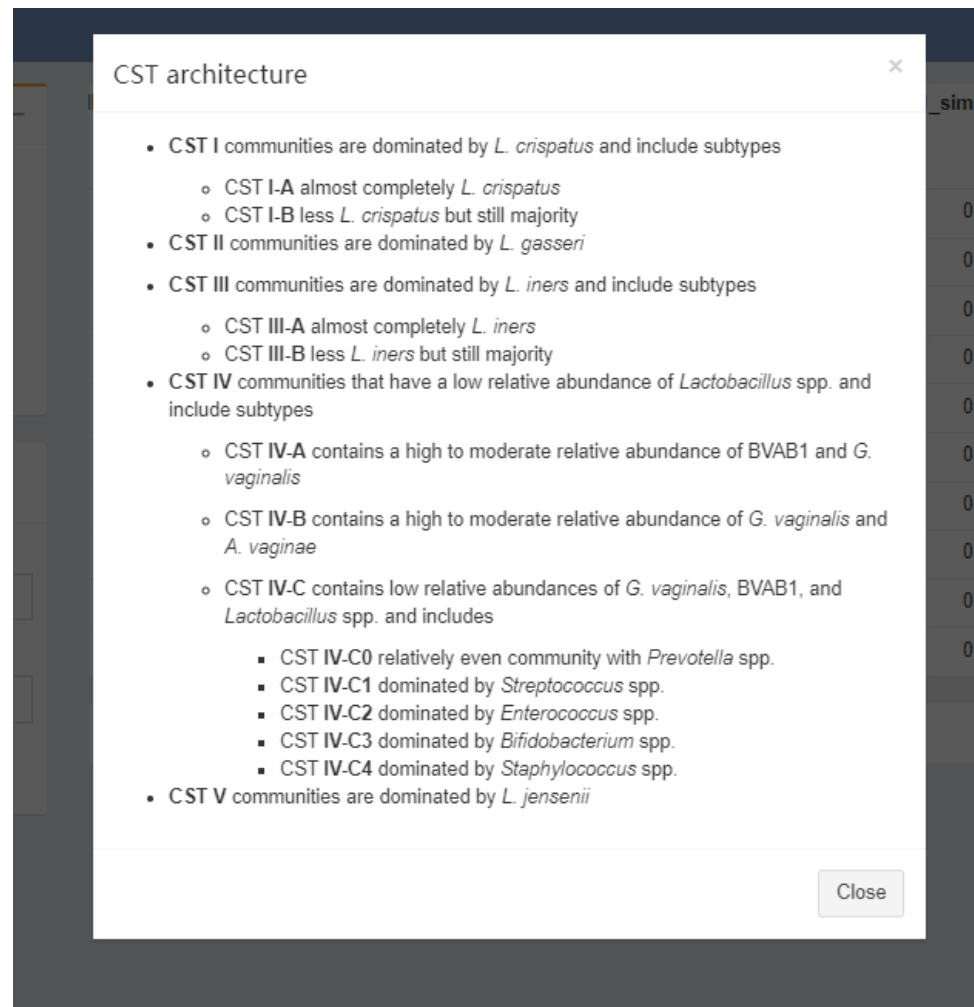
II_sim	III-A_sim	III-B_sim	IV-A_sim	IV-B_sim	IV-C0_sim	IV-C1_sim	IV-C2_sim	IV-C3_sim	IV-C4_sim	V_sim	subCST	score	CST
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I
0	0	0	0	0	0	0	0	0	0	0	I-A	0	I

针对阴道微生物生态数据，可以进行CSTs分型，以确定阴道菌群的主要类型。结果以可导出的表格展示。

# 阴道菌群CSTs分型



针对阴道微生物生态数据，可以进行CSTs分型，以确定阴道菌群的主要类型。详细CST类型解释可以点击CST architecture查看。





# Thank you

南方医科大学珠江医院检验医学部  
南方医科大学微生态医学中心

2023. 02. 02

珠江医院菌群分析平台



扫一扫二维码，加入群聊。

用户答疑QQ群：478765875