



微生物绘图入门

——Rstudio

Xinhu Guo

引言

- 针对大量且重复的微生物测序数据，使用R语言可快速绘制出图
- 初学者可以参考下文的示例，结合微信公众号和ChatGPT学习



数据获取



绘图



一、数据获取

1. 登陆美吉网站，查看报告

高级检索-中国知网

CNKI翻译助手

修改密码 - “党旗飘飘”...

Sci-Hub: 将知识带给...

Document search - ...

仪表盘 - 云深不知处

山东农业大学图书馆

国家税务总局全国增...

>> 其他书签

移动设备上的书签

美吉生物
Majorbio

首页云流程云工具云课堂云文库产品商城

开通VIP消息学习中心订单中心项目中心杨

项目中心

项目中心 > 概览

我的分析报告

任务信息	项目名称	操作
338F_806R-80samples 创建人: META 创建时间: 2021-03-02 15:50:(MJ20210114049-MJ-M-20210114118-杨海水-南.	查看报告
微微生物多样性-demo2017-01-21 创建人: 杨海水 创建时间: 2017-01-21 23:02:	微生物多样性示例数据-demo20170121	

我的工具结果

快捷入口

我的合同

我的送样订单

我的课程

待办事项

0 项 待处理事项

0 项 待接收项目

0 项 待领取课程

1 项 待领取福利课程

在线咨询



一、数据获取

2. 数据保存至美吉网盘（分类水平为一般为phylun，其它不用改变）

美吉生物 Majorbio

首页 云流程 云工具 云课堂 云文库 产品商城

开通VIP 消息 学习中心 订单中心 项目中心

分析记录 成功 185 失败 3 计算中

文献解析 视频教程

分析记录 工具分析记录

样本层级聚类分析 PCA分析 多维PCA分析 PCoA分析 **NMDS分析** NCM分析

NMDS分析, 即非度量多维尺度分析(non-metric multidimensional 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, 同时保持对象间原始关系的数据分析方法。其基本特征是将对象间的相似性或相异性数据看成点间距离的单调函数, 在保持原始数据次序关系的基础上, 用新的相同次序的数据列替换原始数据进行度量型数据降维分析。

OTU表: OTU_Taxon_Origin 分类水平: Phylum

分组方案: Wheat

RT_NS_NP RT_NS_P RT_S_NP RT_S_P PT_NS_NP PT_NS_P PT_S_NP PT_S_P NT_NS_NP NT_NS_P NT_S_NP NT_S_P Basic 注: 样本总数≥3

距离算法: bray_curtis

组间差异检验: ANOSIM 置换次数: 999

运行

NMDS

NMDS结果: NMDS_20230713_110815238

NMDS分析图

存报告 使用说明 下载

布图设置 编辑文字 编辑图形 文本框 标签 主题颜色 网格线 外边框 坐标轴 图例 更多参数 初始设置 保存调图 最新调图

图形全选: 未选 1:椭圆 2:散点 (开启编辑图形后使用)

新手指引

1. 记好文件名称是什么

2. 查看文件保存的途径

一、数据获取

2. 数据保存至美吉网盘(找到数据存放位置, 下载圈定的所有)

美吉网盘

杨海水

全部文件

原始数据

项目文件

传输文件

正在下载

传输完成 (2677)

下载失败

下载

刷新

MJ20210114049-MJ-M-20210114118-杨海水-南京农业大学-环境微生物多样性-80个样本-20210302 > 338F_806R-80samples > interaction_results > 4.Beta_diversity > NmDs > NMDS_20230713_110815238

<input type="checkbox"/>	文件名	推送时间	类型	大小	操作
<input type="checkbox"/>	anosim_results.txt	2023-07-13 11:11:37	txt	167 B	↓
<input type="checkbox"/>	bray_curtis_otu_file.xls	2023-07-13 11:11:37	xls	19.02 KB	↓
<input type="checkbox"/>	format_results.xls	2023-07-13 11:11:37	xls	99 B	↓
<input type="checkbox"/>	NMDS.pdf	2023-07-13 11:11:37	pdf	72.44 KB	↓
<input type="checkbox"/>	nmDs_sites.xls	2023-07-13 11:11:37	xls	1.59 KB	↓
<input type="checkbox"/>	nmDs_stress.xls	2023-07-13 11:11:37	xls	20 B	↓
<input type="checkbox"/>	运行参数.txt	2023-07-13 11:11:37	txt	522 B	↓

共 7 条 < 1 > 跳至 1 页

版本号: 1.0.5

1. 验证NMDS是否合理: P值

2. 所需数据

1. 验证NMDS是否合理: stress

Anosim R越大, $P < 0.05$, $Stress < 0.2$, 说明NMDS可以表征处理间的 β 多样性

Yang, LDD, 2020

二、绘图

1. 数据的准备（注意文件的后缀、英文的大小写、处理的标签）

1. 名称为：RICE-NMDS_sorted.xlsx，后缀很重要

sample	straw	tillage	root	NMDS1	NMDS2
RHF1a	S	PT	R	0.219	-0.1213
RHF1b	S	PT	NR	-0.1439	0.0117
RHF2a	S	PT	R	-0.0453	0.076
RHF2b	S	PT	NR	-0.1983	0.0381
RHF3a	S	PT	R	0.1417	-0.0487
RHF3b	S	PT	NR	-0.1879	-0.0258
RHM1a	S	NT	R	0.0214	0.0352
RHM1b	S	NT	NR	0.2288	-0.2016
RHM2a	S	NT	R	0.2367	-0.0941
RHM2b	S	NT	NR	-0.3227	0.0515
RHM3a	S	NT	R	0.0765	0.1073
RHM3b	S	NT	NR	-0.338	0.0463
RHX1a	S	RT	R	-0.0067	0.0886
RHX1b	S	RT	NR	0.2002	-0.0587
RHX2a	S	RT	R	0.0204	0.0538
RHX2b	S	RT	NR	0.2457	-0.0077
RHX3a	S	RT	R	-0.0811	0.1024
RHX3b	S	RT	NR	0.2518	-0.0654
RBF1a	NS	PT	R	-0.1669	-0.0992
RBF1b	NS	PT	NR	-0.018	-0.151
RBF2a	NS	PT	R	0.3422	0.12
RBF2b	NS	PT	NR	-0.0718	-0.1127
RBF3a	NS	PT	R	-0.0361	0.0746
RBF3b	NS	PT	NR	-0.0413	-0.132
RBM1a	NS	NT	R	0.1037	0.1261
RBM1b	NS	NT	NR	-0.2297	0.0049
RBM2a	NS	NT	R	-0.0215	0.1069
RBM2b	NS	NT	NR	-0.2378	-0.0019

2. 样本为：sample，大小写很重要，NMDS中的sample顺序不重要

3. 处理为：straw，tillage，root，可将数据按照处理进行分组

4. 样本sample与处理之间一定一一对应

二、绘图

2. R操作——工作路径设置

1. 数据需要与R文件放在同一个文件中，确保在一个工作路径

2. 否则会报错，如示例，提示找不到文件“could not find”

```
R 4.3.1 · C:/Users/XinhuGuo/Desktop/水稻季/水稻-NMDS/
> nmds_data <- read_excel("RICE-NMDS_sorted0.xlsx")
Error in read_excel("RICE-NMDS_sorted0.xlsx") :
  could not find function "read_excel"
> |
```

二、绘图

2. R操作——包的安装——install.packages(“xxx”)

The screenshot displays the RStudio interface with a script editor on the left and the Environment pane on the right. The script editor contains R code for installing and loading packages, and creating a scatter plot. Red annotations are overlaid on the code to explain the steps.

1. 初次使用，需要安装包，命令为install.packages(“xxx”), xxx为你需要的包，安装成功左下角会有提示

2. 加载包，为library(xxx)，成功后会有提示，蓝色字体即成功

```
1 install.packages("ggplot2")
2 library(ggplot2)
3 library(readxl)
4 library(scales)
5 #=====水稻秸秆=====
6 # 读取数据
7 nmds_data <- read_excel("RICE_NMDS.xlsx")
8
9 # 创建散点图
10 p1 <- ggplot(nmds_data, aes(x = NMDS1, y = NMDS2, color = straw, group = straw)) +
11   geom_point(shape = 16, size = 2.5, stroke = 1) +
12   stat_ellipse(aes(fill = straw), geom = 'polygon', level = 0.95, alpha = 0.1,
13   labs(x = "NMDS1", y = "NMDS2", color = "Rice") +
14   theme_minimal() +
15   scale_x_continuous(labels = number_format(accuracy = 0.1)) +
16   scale_y_continuous(labels = number_format(accuracy = 0.1)) +
17   scale_color_manual(values = c("blue", "red", "green")) +
18   theme(panel.border = element_rect(color = "black", fill = NA, size = 1),
19         axis.ticks.x = element_line(color = "black"),
20         axis.ticks.y = element_line(color = "black"),
21         axis.ticks.length = unit(-0.15, "cm"),
22         axis.text.x = element_text(face = "bold", color = "black", size = 16),
23         axis.text.y = element_text(face = "bold", color = "black", size = 16))
24
25 # 输出图形
26 p1
```

The Environment pane on the right shows the Global Environment with the message "Environment is empty". The Packages pane at the bottom shows the installed packages: ggplot2, readxl, and scales.

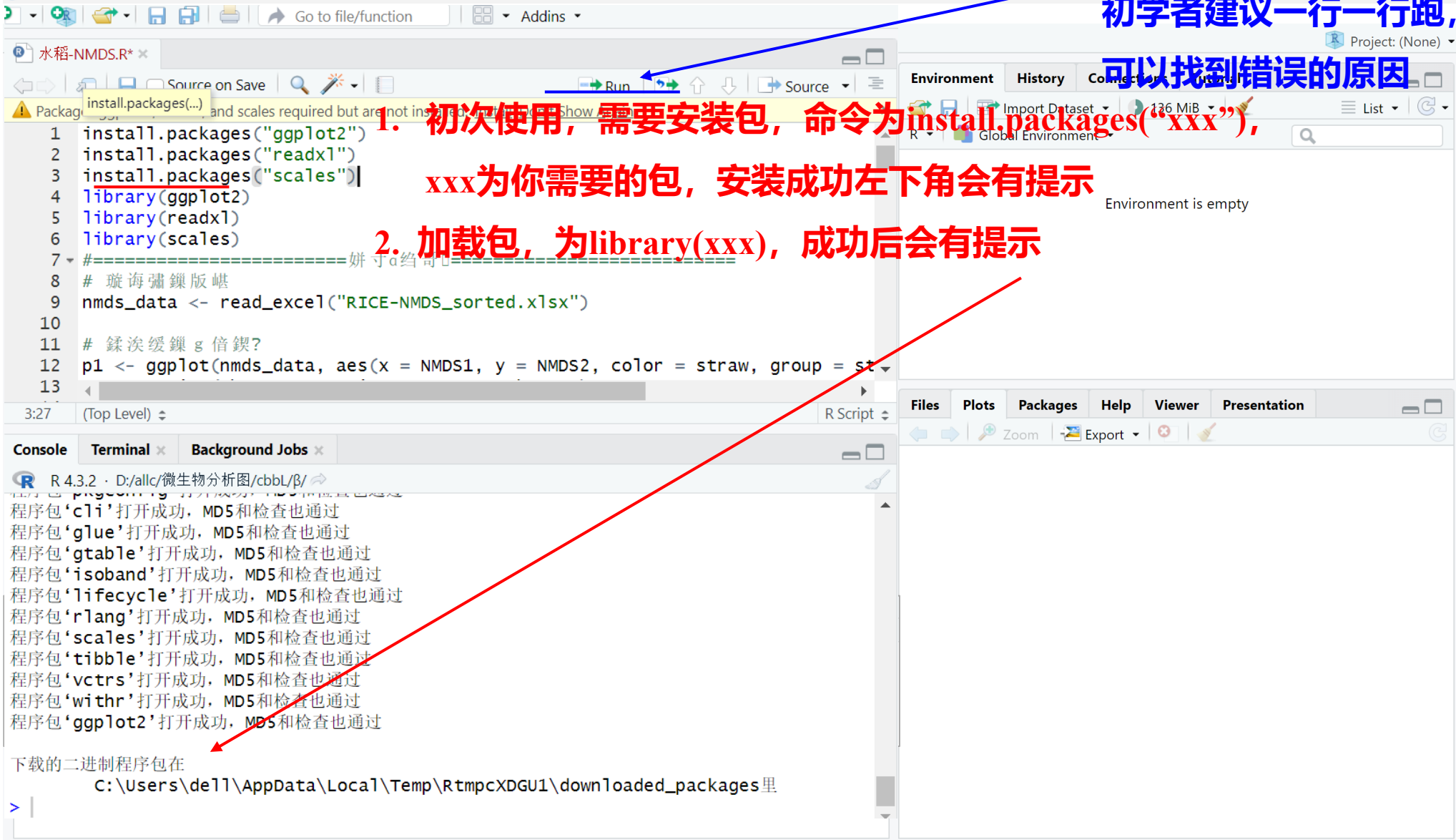
二、绘图

2. R操作——包的安装——install.packages(“xxx”)

开始：选定一行代码，点击run

初学者建议一行一行跑，

可以找到错误的原因



二、绘图

2. R操作——数据读取

1. 加载上面三个包后，输入自己的数据文件的名称，不能输错

2. Run，成功后右上会有提示，可点击查看

```
1 #install.packages("ggplot2")
2 library(ggplot2)
3 library(readxl)
4 library(scales)
5 #=====水稻秸秆=====
6 # 读取数据
7 nmds_data <- read_excel("RICE-NMDS_sorted.xlsx")
8
9 # 创建散点图
10 p1 <- ggplot(nmds_data, aes(x = NMDS1, y = NMDS2, color = straw, shape = straw)) +
11   geom_point(shape = 16, size = 2.5, stroke = 1) +
12   stat_ellipse(aes(fill = straw, geom = 'polygon', level = 0.95, alpha = 0.1, show.legend = FALSE)) +
13   labs(x = "NMDS1", y = "NMDS2", color = "Rice")
14   theme_minimal() +
15   scale_x_continuous(labels = number_format(accuracy = 0.1)) +
16   scale_y_continuous(labels = number_format(accuracy = 0.1)) +
17   scale_color_manual(values = c("blue", "red", "green")) +
18   theme(panel.border = element_rect(color = "black", fill = NA, size = 1),
19         axis.ticks.x = element_line(color = "black"),
20         axis.ticks.y = element_line(color = "black"),
21         axis.ticks.length = unit(-0.15, "cm"),
22         axis.text.x = element_text(face = "bold", color = "black", size = rel(1.2)),
23         axis.text.y = element_text(face = "bold", color = "black", size = rel(1.2)))
```

Environment: nmds_data 36 obs. of 6 v...

R 4.3.1 · C:/Users/XinhuGuo/Desktop/水稻季/水稻-NMDS/

```
> library(ggplot2)
> library(readxl)
> library(scales)
> nmds_data <- read_excel("RICE-NMDS_sorted.xlsx")
> |
```

二、绘图

2. R操作——数据读取

1. 这一行分别为 (数据名称: nmds_data; 坐标轴: aes, x, y; 颜色: color; 分组: group), 均可自己更改

2. color和group一般是同一个处理, 可更改为自己的处理

3. 颜色的填充, 可更改为自己的处理

4. 图像查看位置

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for installing ggplot2, reading data from 'RICE-NMDS_sorted.xlsx', and creating a scatter plot with ellipses. The code is as follows:

```
1 #install.packages("ggplot2")
2 library(ggplot2)
3 library(readxl)
4 library(scales)
5 #=====水稻秸秆=====
6 # 读取数据
7 nmds_data <- read_excel("RICE-NMDS_sorted.xlsx")
8
9 # 创建散点图
10 p1 <- ggplot(nmds_data, aes(x = NMDS1, y = NMDS2, color = straw, group = straw)) +
11   geom_point(shape = 16, size = 2.5, stroke = 1) +
12   stat_ellipse(aes(fill = straw), geom = 'polygon', level = 0.95, alpha = 0.1, show.legend = FALSE) +
13   labs(x = "NMDS1", y = "NMDS2", color = "Rice") +
14   theme_minimal() +
15   scale_x_continuous(labels = number_format(accuracy = 0.1)) +
16   scale_y_continuous(labels = number_format(accuracy = 0.1)) +
17   scale_color_manual(values = c("blue", "red", "green")) +
18   theme(panel.border = element_rect(color = "black", fill = NA, size = 1),
19         axis.ticks.x = element_line(color = "black"),
20         axis.ticks.y = element_line(color = "black"),
21         axis.ticks.length = unit(-0.15, "cm"),
22         axis.text.x = element_text(face = "bold", color = "black", size = rel(1.7))
23   )
```
- Environment:** Shows the loaded data frame 'nmds_data' with 36 observations.
- Plots:** The 'Plots' tab is active, showing a preview of the scatter plot.
- Console:** Shows the execution of the following commands:

```
> library(ggplot2)
> library(readxl)
> library(scales)
> nmds_data <- read_excel("RICE-NMDS_sorted.xlsx")
> view(nmds_data)
```

Red arrows indicate the following mappings:

- Arrow 1: Points from the instruction to the `nmds_data` variable in line 7.
- Arrow 2: Points from the instruction to the `aes(x = NMDS1, y = NMDS2, color = straw, group = straw)` in line 10.
- Arrow 3: Points from the instruction to the `scale_color_manual` function in line 17.
- Arrow 4: Points from the instruction to the 'Plots' tab in the right-hand pane.

三、代码获取



xy要加油鸭 ★

...

初来乍到，请多指教。因个人水平有限，文中内容请各位大佬酌情阅读，各取所需。特别声明!!! 推广非原创内容...



不懂绘图

学术推荐及绘图



清同趣科研 ★

...

清同趣科研是一伙在读博士创建的自由免费的生信交流平台! 我们专注于生信分析和科研趣事分享。生信分析聚焦...
展开



生态R学社 ★

...

我有文献、方法与君分享，愿君在科研路上一帆风顺、文章多多! R语言、生态为主



科研后花园 ★

专注于R语言绘图及数据分析!



R语言分析作图

...

励志让所有R语言零基础的同学，都可以轻松使用R语言做图以助力科研。



小田的生信随笔 ★

...

生信小农人; 分享一些数据分析类的文章等内容, 包括R语言、Python等。也会分享一些生活感悟, 散文随笔, 为大...



语言科学漫谈

...

传播语言研究方法及经验



小白鱼的生统笔记 ★

...

看心情写



土壤微生物组 ★

...

宏基因组、宏转录组和宏病毒组学是未来农业, 环境, 生态和海洋等领域挖掘和利用微生物组资源的重要方法, 土壤...