

罗宁生物云-操作指南

分析项目：稀释曲线

版本号：V1.0

1. 分析项目简述：

通过计算样本的 Alpha-多样性指数可以反映微生物群落的丰度 (Richness)、均匀性 (Evenness) 及多样性等。为了探索样本物种丰富度随测序深度的变化趋势，进行稀释曲线 (Rarefaction Curve) 的制作。方法是从样本中随机抽取一定量的序列，并统计它们所代表的物种数目，以抽取的一系列序列数和相应的物种数来构建曲线。通过绘制稀释曲线可间接反映样本中物种的丰富程度。当曲线趋于平缓时可认为测序深度已经基本覆盖到样品中所有的物种，增大数据量只会产生少量的低丰度物种。

2. 工具用途及使用场景：

本工具用于绘制稀释曲线。

3. 您需要准备以下文件：

3.1 如您在罗宁生物进行测序分析，并获得分析结题报告数据，请在以下路径查找文件

必须文件：Rarefaction_curve_data_summary.csv

路径：\03.AlphaDiversity

可选文件：sample_info_v1.csv

路径：\02.OTUAnalysis\Core_files

如分析结题报告中无 Rarefaction_curve_data_summary.csv 文件，请准备以下文件：

必须文件：otu_table_v1.csv

路径：\02.OTUAnalysis\Core_files

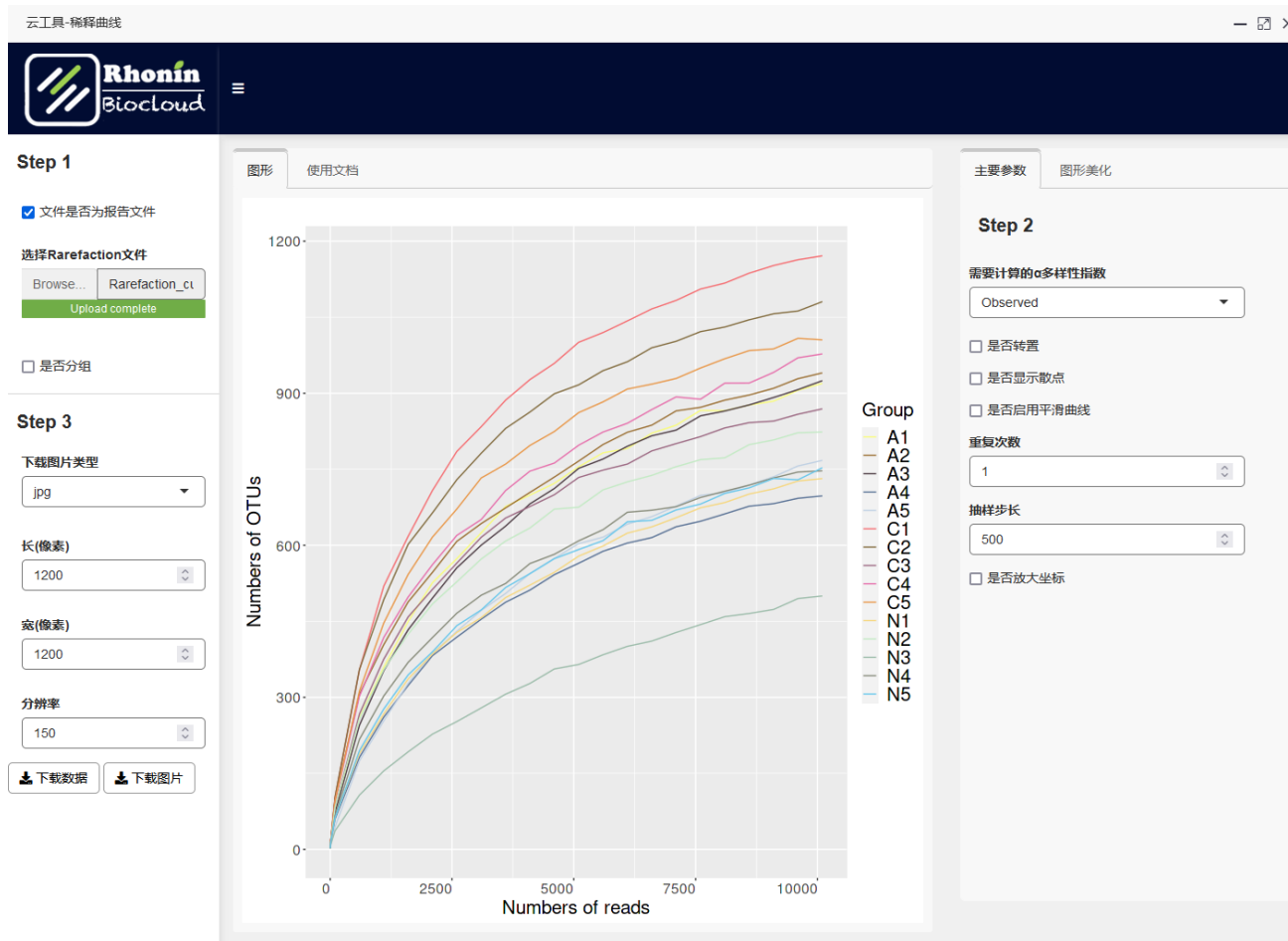
如需要分组请装备分组信息文件，请准备以下文件

必须文件：sample_info_v1.csv

路径：\02.OTUAnalysis\Core_files

3.2 如您未在罗宁生物获取数据分析报告，请准备 OTU 表（表格保存格式为*.csv），详细格式参见“使用文档”标签页中的说明。

4. 操作界面：



稀释曲线绘制工具界面如图所示分为左侧数据输入、输出图形文件参数设置区域，右侧图形类型参数设置、图例修改美化功能区和中间图形实时展示页面。

5. 操作步骤：

本项操作由数据上传、参数选择和图形输出共 3 部分组成。

第 1 步：数据上传

注意：本步骤所有输入文件格式必须为 CSV 格式文件（后缀*.csv，Comma delimited）。

a. 如您在罗宁生物测序并获取完整结题报告数据，请按以下操作进行数据导入：

请在 Step 1 区域勾选“文件是否为报告文件”，并在“选择 Rarefaction 文件”入口处点击浏览，选择 Rarefaction_curve_data_summary.csv 文件（路径：\03.AlphaDiversity），如下图所示：

Step 1

文件是否为报告文件

选择Rarefaction文件

Browse...	Rarefaction_ct
Upload complete	

是否分组

如果样本有分组信息，请勾选“是否分组”复选框，并在“选择分组信息文件”入口处点击浏览，选择 sample_info_v1.csv 文件（路径：\02.OTUAnalysis\Core_files），如下图所示：

Step 1

文件是否为报告文件

选择Rarefaction文件

Browse...	Rarefaction_ct
Upload complete	

是否分组

选择分组信息文件

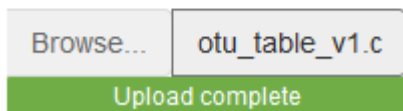
Browse...	sample_info_v
Upload complete	

- b. 如您已获取罗宁生物数据分析的部分分析结果（如仅有 OTU 表）或通过其他方式仅获得 OTU 表，请按以下方法进行数据导入：
- 首先，请在 Step 1 区域**去除勾选**“文件是否为报告文件”，并在“选择 otu 文件”入口处点击浏览，选择 otu_table_v1.csv 文件（路径：\02.OTUAnalysis\Core_files），如下图所示：

Step 1

文件是否为报告文件

选择otu文件

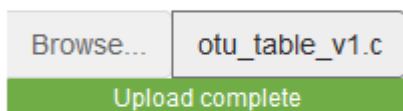


如果样本有分组信息，请勾选“是否分组”复选框，并在“选择分组信息文件”入口处点击浏览，选择 sample_info_v1.csv 文件（路径：\02.OTUAnalysis\Core_files），如下图所示：

Step 1

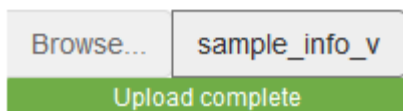
文件是否为报告文件

选择otu文件



是否分组

选择分组信息文件

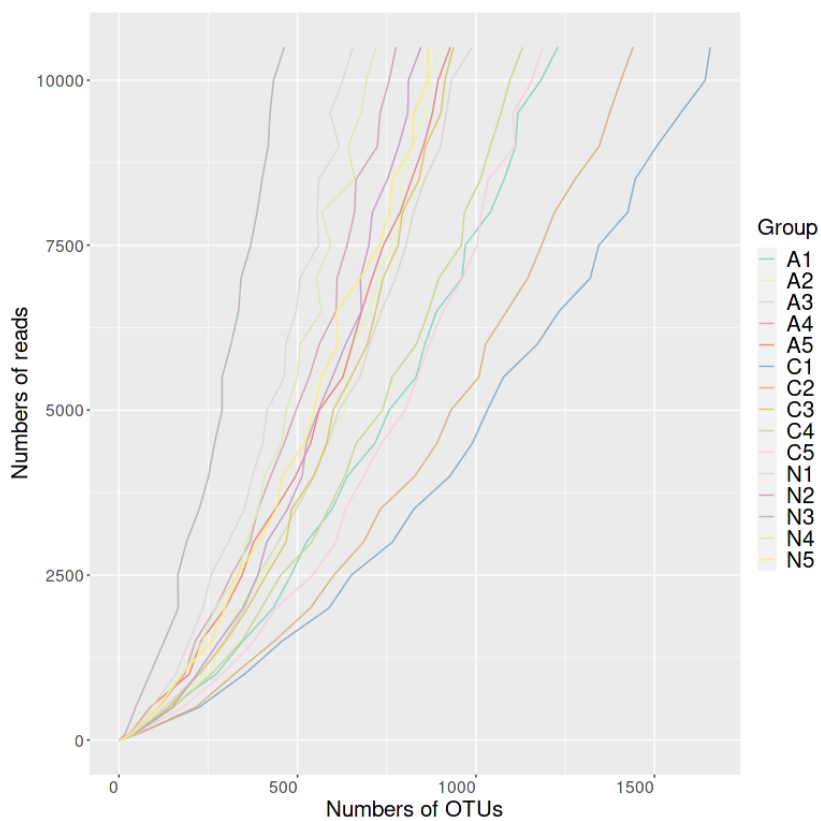


第 2 步：参数修改

完成第 1 步导入数据后，请在右侧图形参数设置、图例修改美化功能区根据作图需求选择或填写合适的参数进行图片修改，修改参数时中间区域会实时展示输出图片。

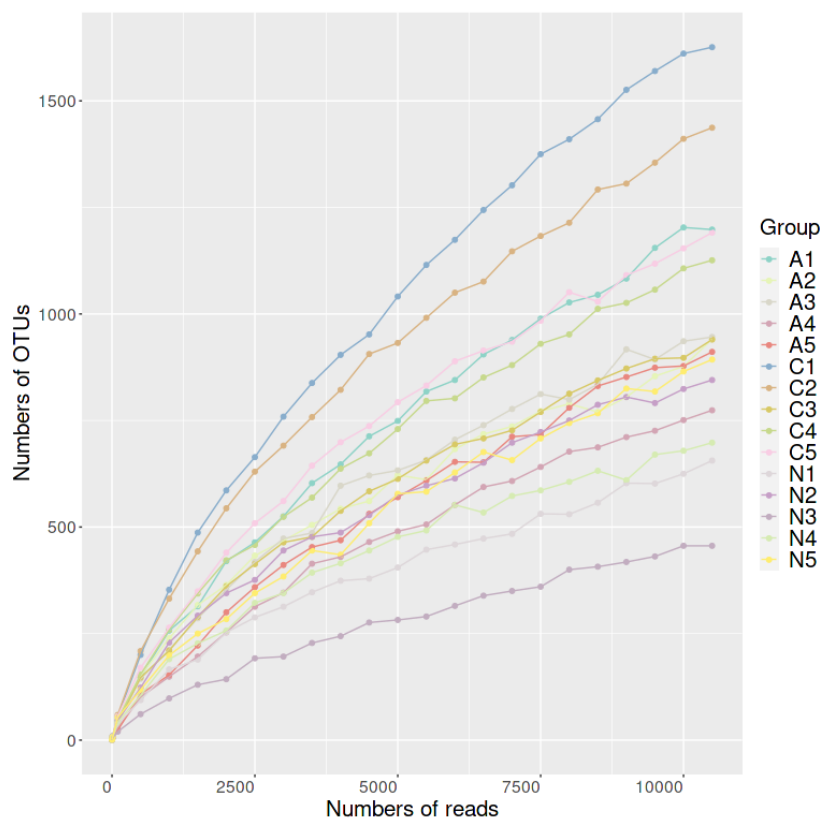
- a. 该步骤主要参数有：需要计算的 α 多样性指数、是否转置、是否加入散点、是否启用平滑曲线、重复次数、抽样步长、是否放大坐标。
- b. 需要计算的 α 多样性指数
该参数需要设置绘图所用多样性指数数据。如输入文件缺失所需 alpha 多样性指数信息，请前往云工具“alpha 多样性指数计算”重新计算结果并导入本工具输出图形。
- c. 是否转置
该参数的设置将转换图形 X 轴和 Y 轴的方向。

如勾选“是否转置”复选框，示例如下图：



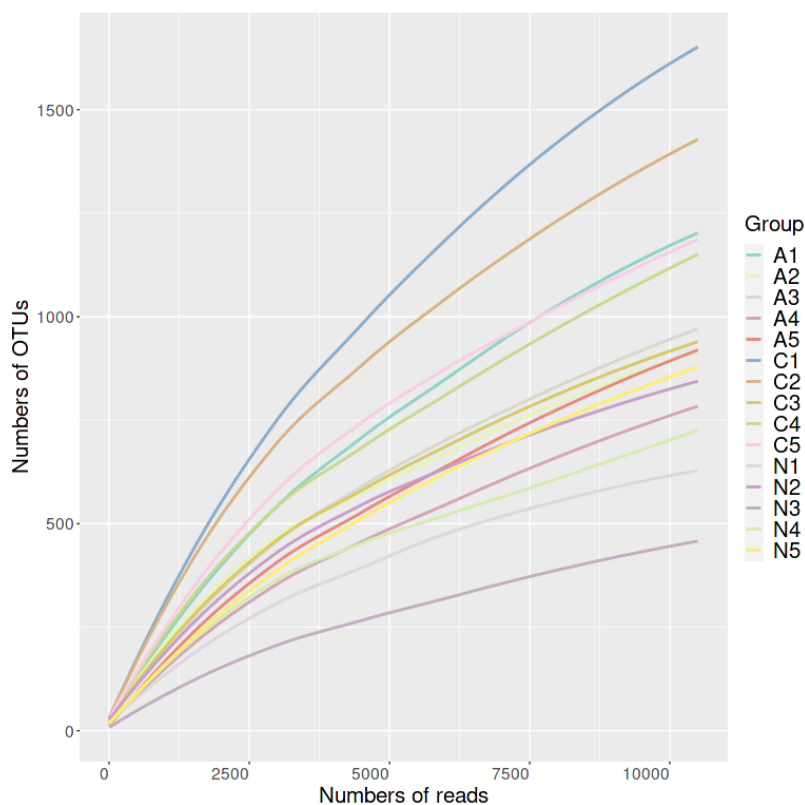
d. 是否加入散点

勾选“是否显示散点”复选框将显示散点图层，效果如下图所示：



e. 是否启用平滑曲线

勾选“是否启用平滑曲线”复选框将显示曲线使用 smooth 曲线进行拟合的结果，效果如下图所示：



f. 重复次数

该参数用于设定重复抽样次数 ($N \geq 1$)，结果将以重复抽样结果的均值展示。重复次数越多，计算时间越长。

g. 抽样步长

该参数用于设定抽样步长，抽样步长越小，计算时间越长。

h. 是否放大坐标

该参数用于设定图片横轴范围。勾选该复选框后需要在“设定横坐标最大值”中填写合适的数值，用以描述图中曲线到达平滑值。

第 3 步：图形输出

在完成数据上传、参数修改设置后，请在图片输出参数设置区域设置下载图片类型、图形分辨率、图片大小等设置，最终点击“下载图片”按钮，将图片文件保存至指定路径文件夹。

注意：本工具可输出位图和矢量图两种图片格式，其中位图提供 JPEG 或 PNG 格式文件，矢量图提供 PDF 或 SVG 格式文件。

第 4 步：数据输出

点击下载后会获得 csv 表格文件 (Rarefaction_curve_data_summary.csv)：

其中 Depth 列为抽样深度，Alpha_diversity_mean 为该深度的均值，Alpha_diversity_sd 为有分组的情况下的该深度标准差。

	A	B	C	D	E	F
1		Depth	SampleID	Alpha_dive	Alpha_dive	Measure
2	1	1	A1	1	NA	Observed
3	2	10	A1	7	NA	Observed
4	3	100	A1	44	NA	Observed
5	4	500	A1	139	NA	Observed
6	5	1000	A1	260	NA	Observed
7	6	1500	A1	321	NA	Observed
8	7	2000	A1	394	NA	Observed
9	8	2500	A1	466	NA	Observed
10	9	3000	A1	537	NA	Observed
11	10	3500	A1	611	NA	Observed
12	11	4000	A1	684	NA	Observed
13	12	4500	A1	678	NA	Observed
14	13	5000	A1	733	NA	Observed
15	14	5500	A1	792	NA	Observed
16	15	6000	A1	863	NA	Observed
17	16	6500	A1	891	NA	Observed
18	17	7000	A1	916	NA	Observed
19	18	7500	A1	995	NA	Observed
20	19	8000	A1	1026	NA	Observed
21	20	8500	A1	1071	NA	Observed
22	21	9000	A1	1112	NA	Observed

6. 术语简述：

7. 常见问题解答：

Q：稀释曲线是怎么计算的，横纵坐标有什么含义，如何解读稀释曲线？

A：稀释曲线随机抽取一定数量的序列，抽取的序列数目逐步增大，抽取的序列数目的集合即为 x 轴坐标；随着抽样量的增加，统计对应序列数目的 Alpha 多样性，即为 y 轴坐标。将 xy 轴形成的一系列值绘制成曲线，即为稀释曲线。稀释曲线可以用于反映 Alpha 多样性，并且帮助了解测序深度是否足够用于后续分析。

Q：为什么横坐标的最大值可能小于 OTU 表中的序列条数？

A：1、如果稀释曲线赢平缓了，继续抽样意义不大，所以可以终止抽样。最终抽样量将会少于序列条数；2、如果 OTU 表中的序列条数不完全一致，最大抽样量可以选择序列最少的样品的序列数，使得其他样品的抽样量小于其自身的总序列数。

Q：为什么有些点的多样性大于它的下一个点

A：因为抽样是随机进行的，如果抽样量之间差距不大，有可能随机抽取序列的多样性存在一些波动，使得抽样量大的点的多样性略微小于抽样量小的。

Q：我的稀释曲线没有完全平，数据能用么？

A：如果稀释曲线接近平缓，数据是可以用于后续分析的。因为低丰度物种可能导致稀释曲线需要大量的序列才能完全平缓，但低丰度物种对总体结果可能不存在较大的影响，所以适当舍弃一些低丰度物种所贡献

的多样性也是有必要的。