

扩增子结果文件补充说明

03.AlphaDiversity 【alpha 多样性分析结果】

1 Alpha_div 【组间 Alpha 多样性比较箱型图】

1.1 ACE 【ACE 指数】

1.1.1 ACE.{png,pdf} 【ACE 指数箱型图】

1.1.2 ACE_Tukey.txt 【ACE 指数多组间方差分析】

列数	列标题	含义
1	diff	组间差值
2	lwr	置信区间上限
3	upr	置信区间下限
4	p adj	P 值

1.1.3 ACE_wilcox.txt 【ACE 指数多组间非参数 wilcox 检验】

列数	列标题	含义
1	Difference	均值差
2	pvalue	P 值
3	sig.	是否显著，若 p 值<0.05,标*,p 值<0.01 标 **,p 值<0.001 标 ***
4	LCL	置信区间下限
5	UCL	置信区间上限

Tukey 和 wilcox 两种分析都是两两比较的方法，都是可以给出显著性检验，不同的是：

tukey 是参数检验，数据要求符合正态分布；wilcox 是非参数检验，不要求数据是否符合正

态分布，因此后者更为通用一些。

2 venn_Flower_figure 【基于 OTU 的韦恩花瓣图】

2.1 venn_flowerdata 【韦恩花瓣图数据】

2.1.1 1.*.uniq.otu.xls 【各样本独有 OTU 信息】

2.2 venn_flower_display.{png,pdf} 【韦恩花瓣图展示】

样本（或分组） ≤ 5 个，用 Venn 图展示；样本（或分组）5 个以上 15 个以下，用花瓣图展示。

3 alpha_diversity_index (group).txt 【样本（组间） α 多样性指数】

Alpha-diversity 主要关注局域均匀生境下的物种数目，因此也被称为生境内的多样性。在分析中，选取 Observed-species，Chao1，Shannon，Simpson，Good-coverage 几种不同的 Alpha 多样性指数，以表征样品中物种分布的多样性和均匀度，并直观展示测序深度和数据量情况。

Observed_species：直观观测到的物种数目（也即是 OTUs 数目）。Shannon：样品中的分类总数及其占比。群落多样性越高，物种分布越均匀，Shannon 指数越大。Simpson：表征群落内物种分布的多样性和均匀度。Chao1：估计群落样品中包含的物种总数。ACE：估计群落中 OTU 数目。Goods_coverage：测序深度指数。PD_whole_tree：群落内物种的亲缘关系。

simpson 有 3 种展示形式，即 Simpson's Index (D)，Simpson's Index of Diversity (1 - D) 和 Simpson's Reciprocal Index (1 / D)，它们对于反映群落多样性的效果相近但是计算的结果形式不同；我们用的是 Simpson's Index of Diversity (1 - D)。

4 (group) observed_species.{png,pdf} 【样本（组间）物种数目饱和度曲线】

稀释性曲线图中，当曲线趋向平坦时，说明测序数据量渐进合理，更多的数据量只会产生少量新的 OTUs。但是，随着测序数据量的不断增加，在 QC 条件相对宽松的情况下，曲线的增长趋势可能会比较大，这主要因为测序存在错误，测序产品的丰度信息也不断增加，导致曲线会保持一个增加的趋势。

5 (group)rank_abundance.{png,pdf} 【样本(组间)Rank Abundance 曲线】

Rank-abundance 曲线是分析多样性的一种方式。构建方法是：先统计每个样品中每个 OTU 的相对丰度（OTUs 包含的序列数除以总的序列数目），将 OTUs 按丰度由大到小进行排序，得到每个 OTU 的等级（从大到小排序的序号），再以 OTU 等级为横坐标，以 OTU 的相对丰度为纵坐标做图。

Rank-abundance 曲线可用来解释物种丰度和物种均匀度，物种的丰度由曲线的宽度来反映，物种的丰度越高，曲线在横轴上的范围越大；曲线的形状（平缓程度）反映了样品中物种的均度，曲线越平缓，物种分布越均匀，如果曲线迅速下降，表明少数优势物种，占据了样品中微生物数量的很大比例。

6 plot_observed_species.xls 【样本物种数目饱和度曲线数据文件】

列数	列标题	含义
1	series	抽取序列数
2	样本名称	各样本中各抽取序列数统计到的 OTUs 数目

稀释曲线是从样品中随机抽取一定测序量的数据，统计它们所代表物种数目（也即是 OTUs 数目），以数据量与物种数来构建的曲线。曲线趋向平坦时，说明测序数据量足够大，可以反映样品中绝大多数的微生物信息。

采用 Qiime 软件 (Version 1.9.1) 默认参数 , 设定最小抽取序列数为 10 , 最大抽取数为所有样品中最少的对应序列条数 , 步长为 (最大抽取数-最小抽取数) /10 , 每一步重复取样 10 次。