



# BMK微生物优势产品

—— 全长MCD+绝对定量

郑莉

2024年7月02日

# C 目录

CONTENTS

1

三代全长MCD

2

绝对定量

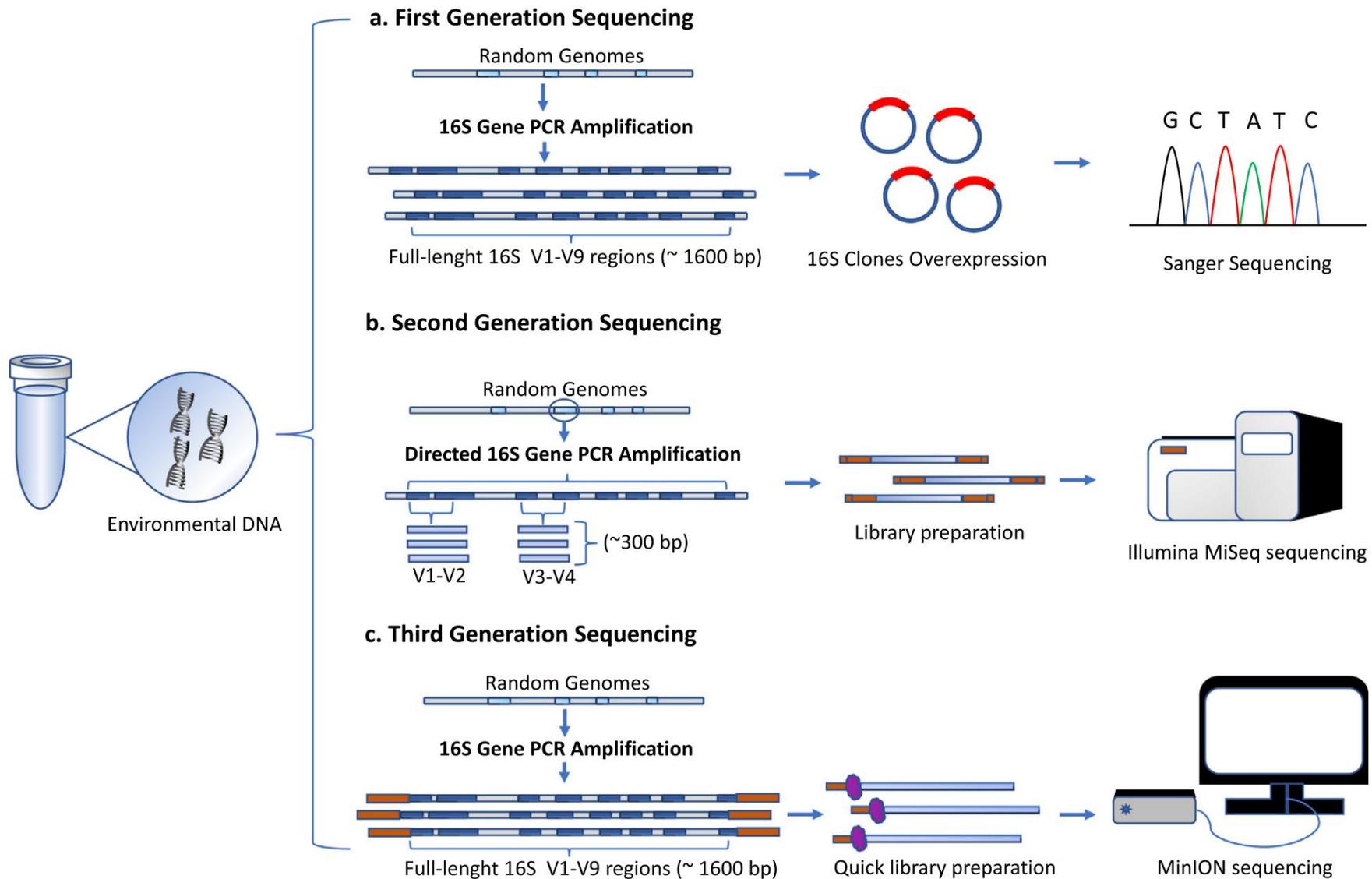
3

推广策略

1

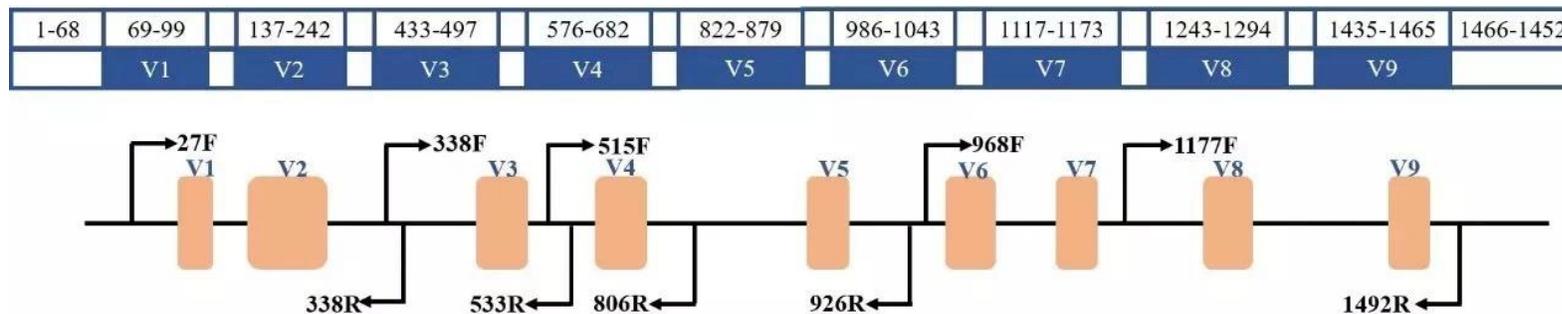
三代全长MCD

# 环境微生物群落研究方法



Comput Struct Biotechnol J. 2020;18:296-305

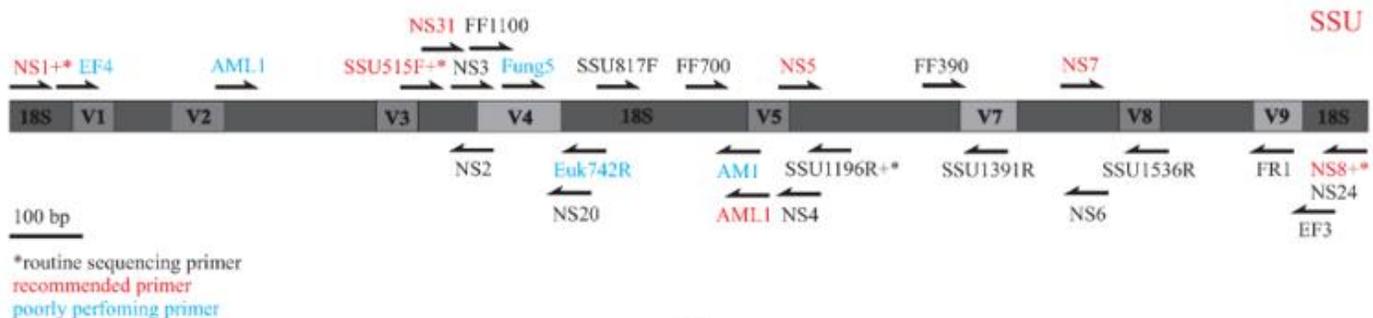
## 细菌多样性---16S



## 真菌多样性---ITS



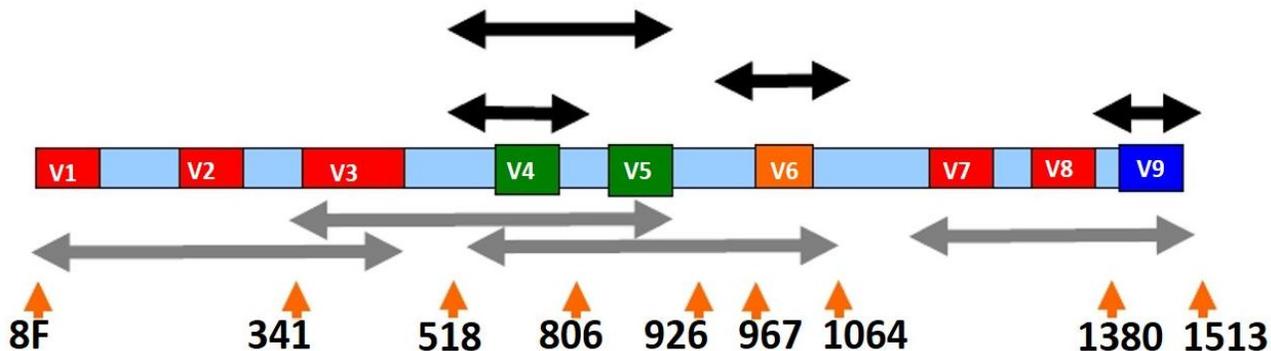
## 真核微生物多样性---18S



## 功能基因多样性

# 16s rRNA基因是微生物多样性研究最为重要的基因

## 可变扩增区域



V3	(341F--518R)	177 bp
V4	(518F--806R)	288 bp
V3V4	(341F--806R)	465 bp

- ① 16S rRNA 基因是编码原核生物核糖体小亚基的基因，长度约为1542bp，其分子大小适中，突变率小，是细菌系统分类学研究中最常用和最有用的标志。
- ② 16S rRNA基因序列包括9个可变区和10个保守区，保守区序列反映了物种间的亲缘关系，而可变区序列则能体现物种间的差异。
- ③ 16S rRNA基因测序以细菌16S rRNA基因测序为主，核心是研究样品中的物种分类、物种丰度以及系统进化。

# 16s rRNA基因9个高变区变异程度不一致，短读长无法覆盖

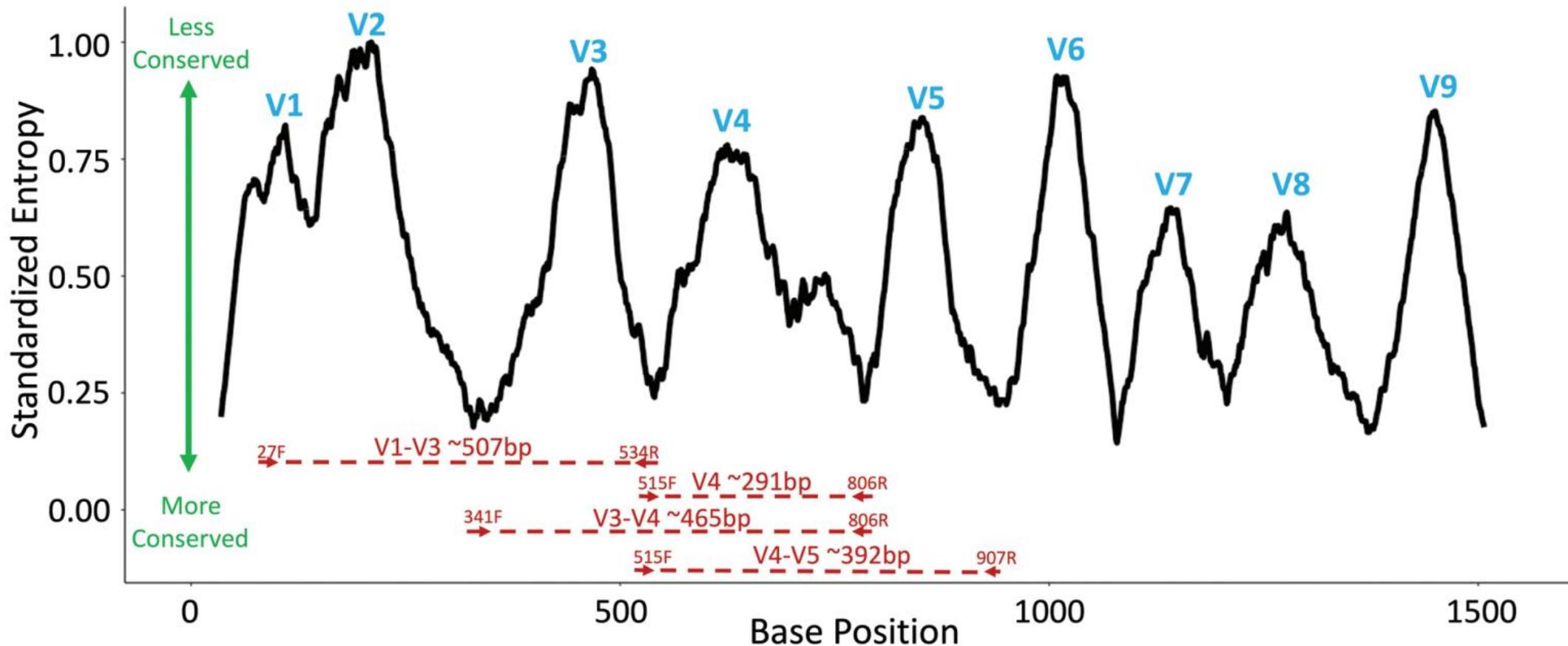
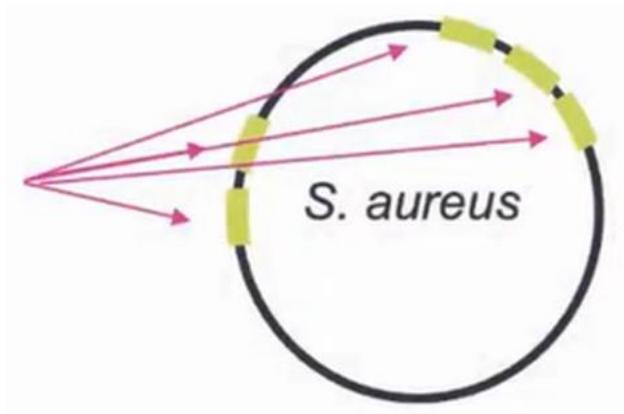


Illustration of conserved and variable regions of the 16S rRNA gene

Journal of Animal Science, 2022, 100, 1–18

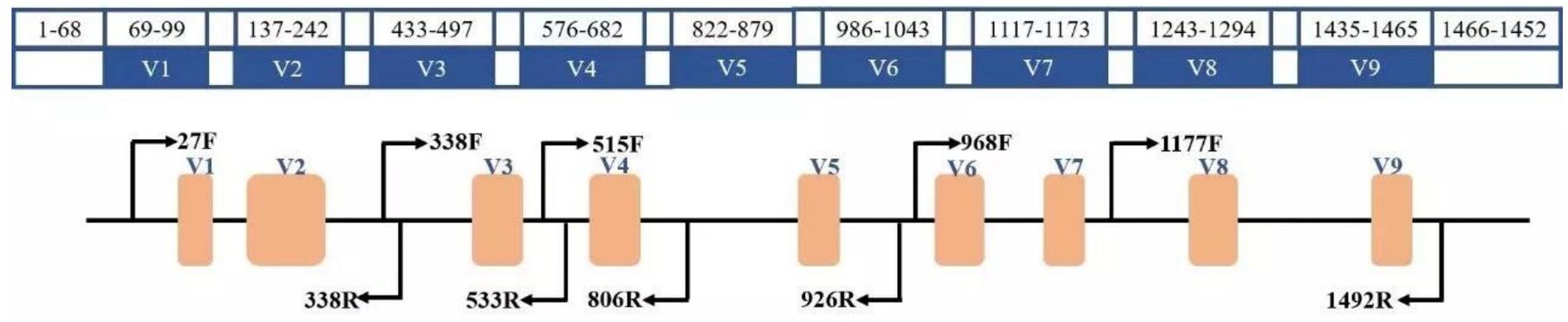
# 16s rRNA基因是微生物多样性研究最为重要的基因



- 16S rRNA 是原核核糖体小亚基的组成部分
- 所有细菌都有16S基因，大小适中，全长为1540bp
- 存在保守区和可变区，保守区序列反映了物种间的亲缘关系，而可变区序列则能体现物种间的差异。
- 这种保守基因的变异使进化枝鉴定成为可能



- 16S基因的靶向测序是评估宏基因组群落组成的一种经济的方法



- 通过**全长**、高度准确的扩增子测序，**降低序列相似性并检测更多核苷酸差异**，从而获得更好的分类分辨率从与 NGS方案相当的属水平的噪声中区分关键物种和菌株

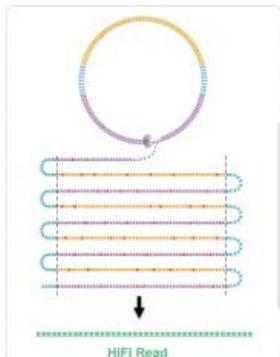
# 三代全长微生物多样性产品优势

## “种”水平精准注释全面突破助力科研新高度

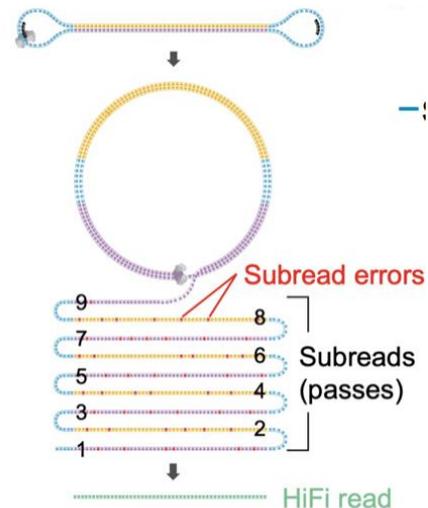
- \* “种”水平精准注释, 平均注释率 60%+
- \* PacBio 的 CCS 模式  $\text{minPasses} \geq 5$ , 自我矫正, 超高准确性
- \* 微生物多样性云平台全面开放, 主流程个性化均免费分析



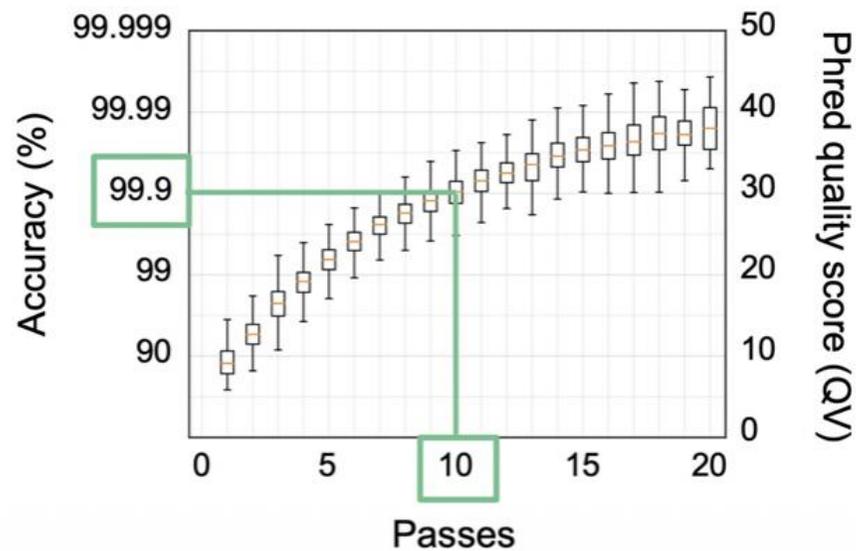
Pacbio 全长扩增子测序技术结合了 Sanger 测序的读长优势和 Illumina 测序的高通量优势, 在细菌、真菌群落结构分析中具有读长更长、通量更高、物种注释率更精准的优势。三代全长扩增子测序技术将开启微生物多样性种水平鉴定的新纪元。



Pacbio CCS 文库, 即采用如 8Kb, 10Kb 等小片段文库, 进行单一片段多轮测序的方式来提升准确性。由于 Pacbio 的原始错误为随机错误, 可通过 CCS 模式进行自身纠正, 来提升数据的准确性。据官方数据, 同一片段测序 4 次后, 单一 read 的准确性可达 99%。

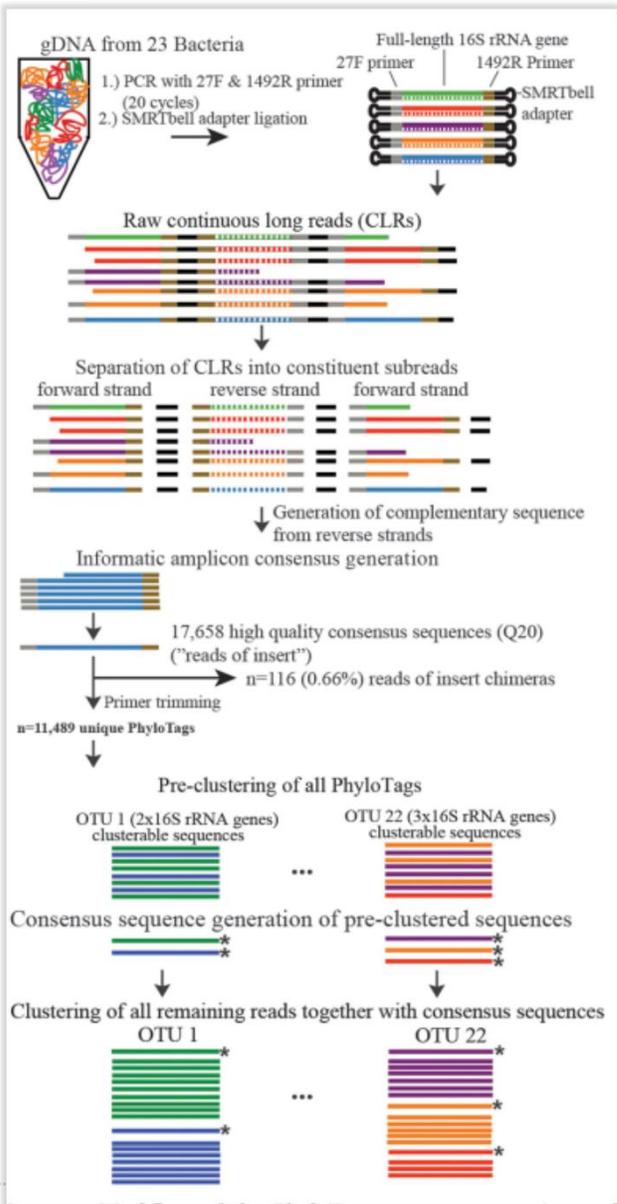


### HiFi Read Accuracy Improves with More Passes



# 技术对比文献解析：全长扩增子注释率高尤其是在Species Level

OPEN



The ISME Journal (2016) 10, 2020–2032  
© 2016 International Society for Microbial Ecology All rights reserved 1751-7362/16  
www.nature.com/ismej

## ORIGINAL ARTICLE

# High-resolution phylogenetic microbial community profiling

bioRxiv

High-resolution phylogenetic microbial community profiling

<i>Taxonomic level</i>	<i>% FL sequences classified</i>	<i>% V4 sequences classified</i>
Phylum	94.6	82.9
Class	92.9	80.7
Family	88.8	71.5
Genus	85.8	62.2
Species	74.5	49.4

Abbreviation: FL, full length.

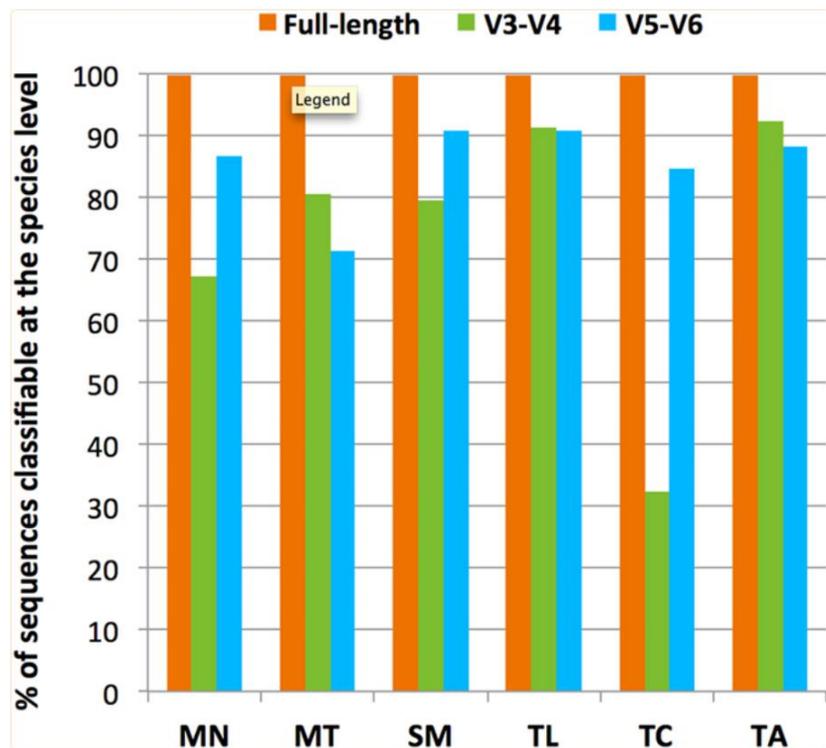
结论1:

全长扩增子注释率高  
尤其是在Species Level

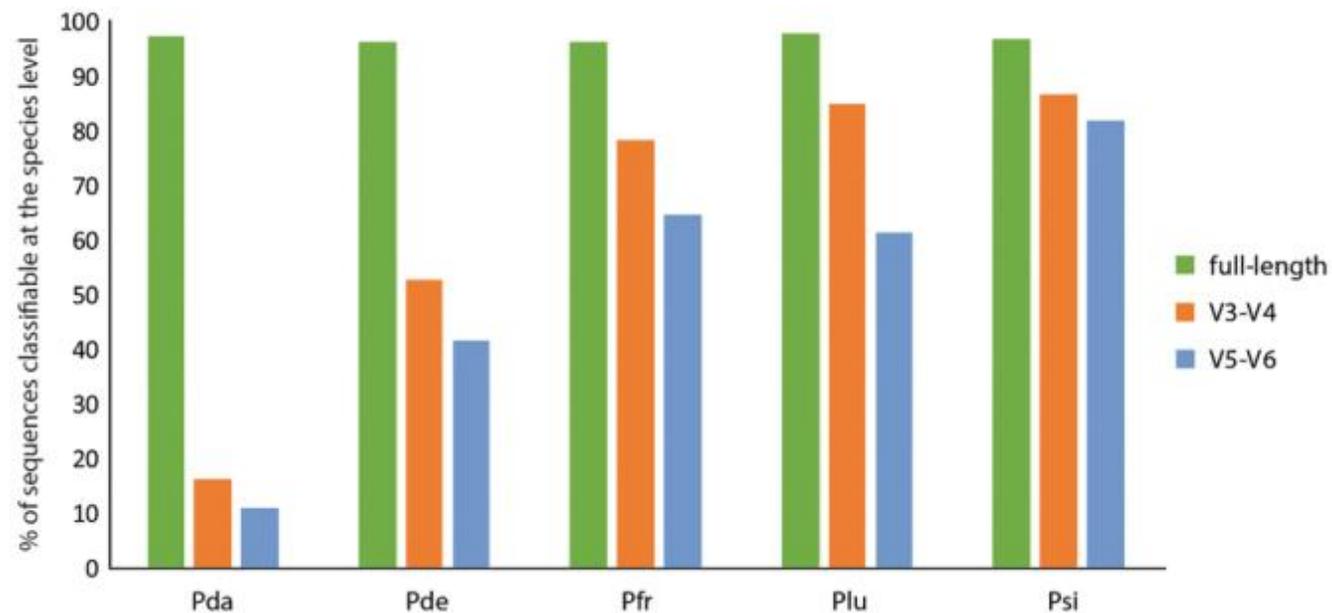


# 技术对比文献解析：全长扩增子注释率高尤其是在Species Level

文献表明，与短读长测序相比，全长16s测序注释到种水平的序列比例最高



Sci Rep. 2017;7(1):2774.

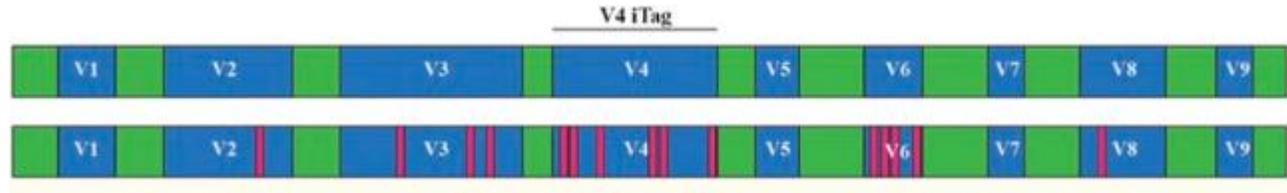


<https://doi.org/10.1016/j.ygeno.2021.06.001>

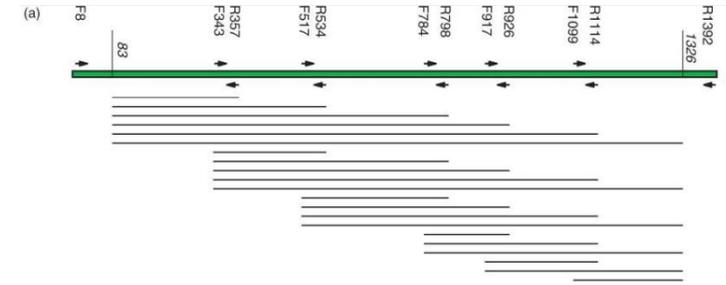
横坐标代表不同样本，纵坐标表示注释到注释到种水平的序列占总序列的百分比

# 技术对比文献解析：全长扩增子真实还原菌群结构

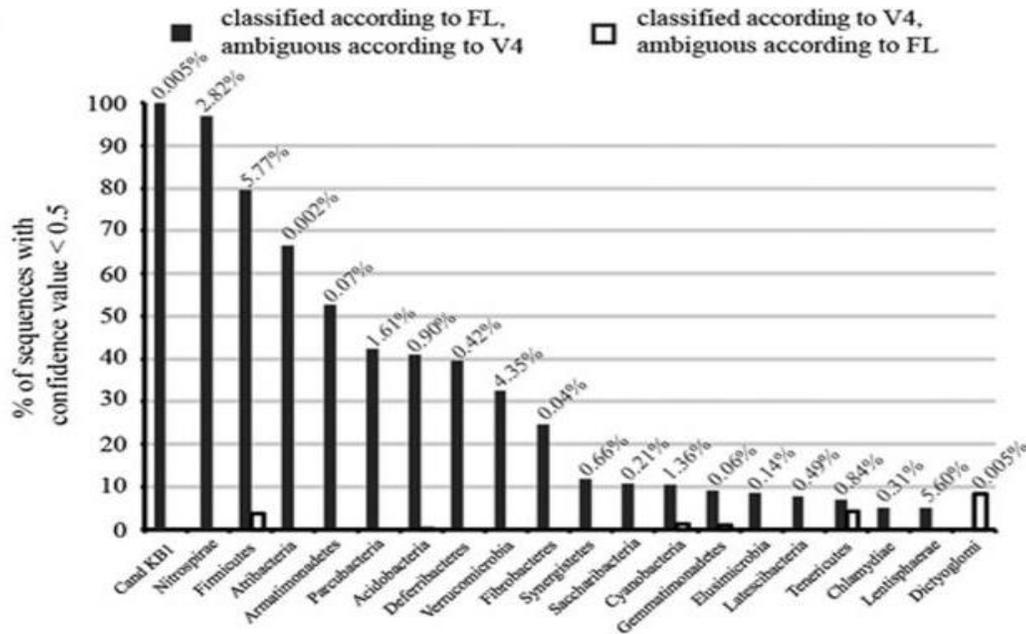
1. 不同物种不同高变区的变异程度不同，部分高变区研究可能严重**低估或者高估群落中微生物的多样性**；



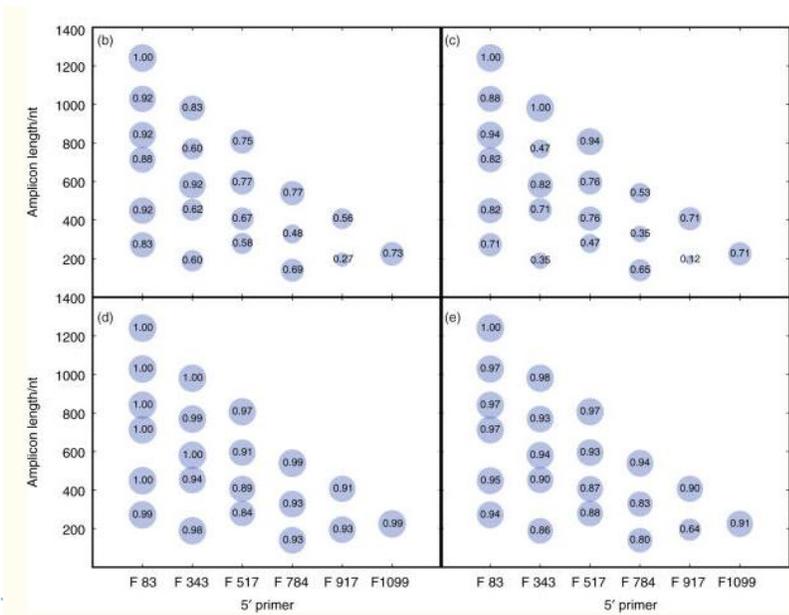
2. 不同高变区短读长扩增子的PCR引物选择会影响推断的**群落准确性**和某些细菌类群的敏感性，使得微生物组研究很难在全局水平上进行比较；



横轴表示物种，纵轴表示丰度，黑色柱子表示全长16S鉴定准确而V4区测序失真；白色柱子表示V4区鉴定准确而全长16S结果失真，对比可以看出，全长扩增子的覆盖度和准确度更高。



E Singer et al.  
ISME J

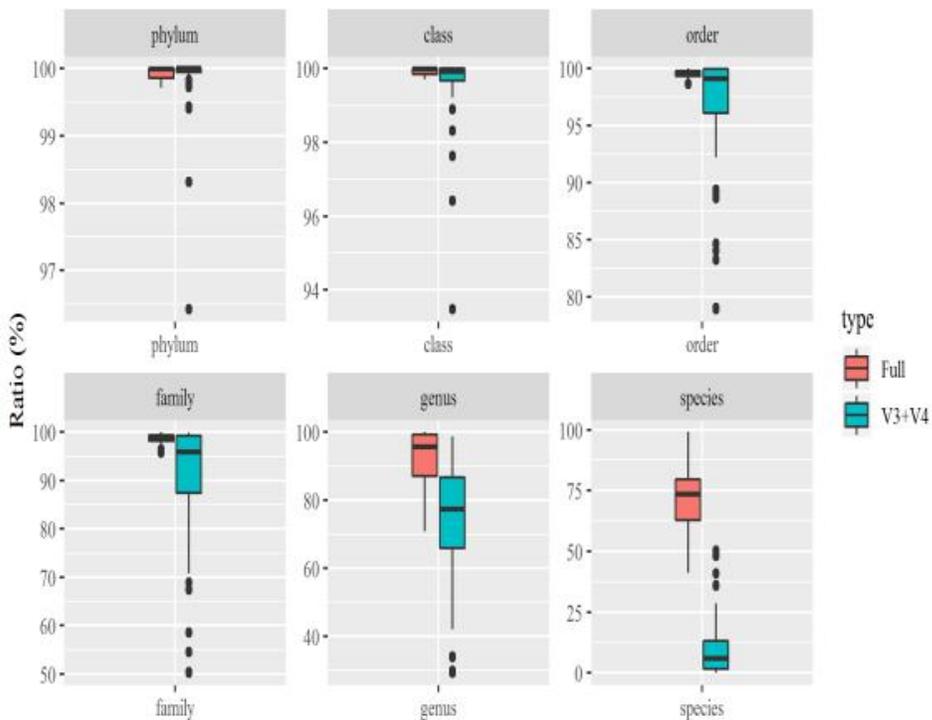


不同引物的扩增子数据对真实群落的恢复率及和相关性，**全长片段能更好的恢复群落结构**。

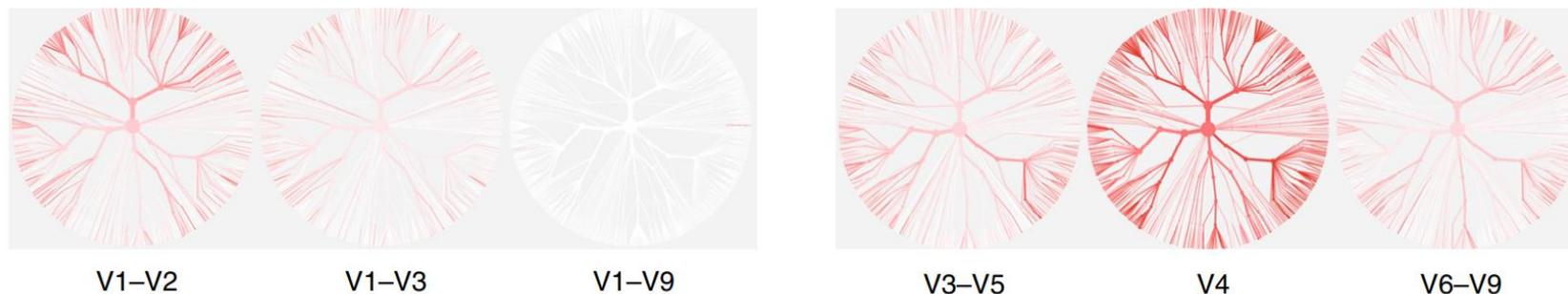
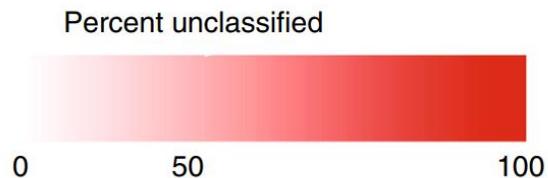
Nucleic Acids Res.

# 实测展示：优势1-显著提升种水平注释最高可提升至85%

三代全长扩增子显著提升了序列在种属水平注释率

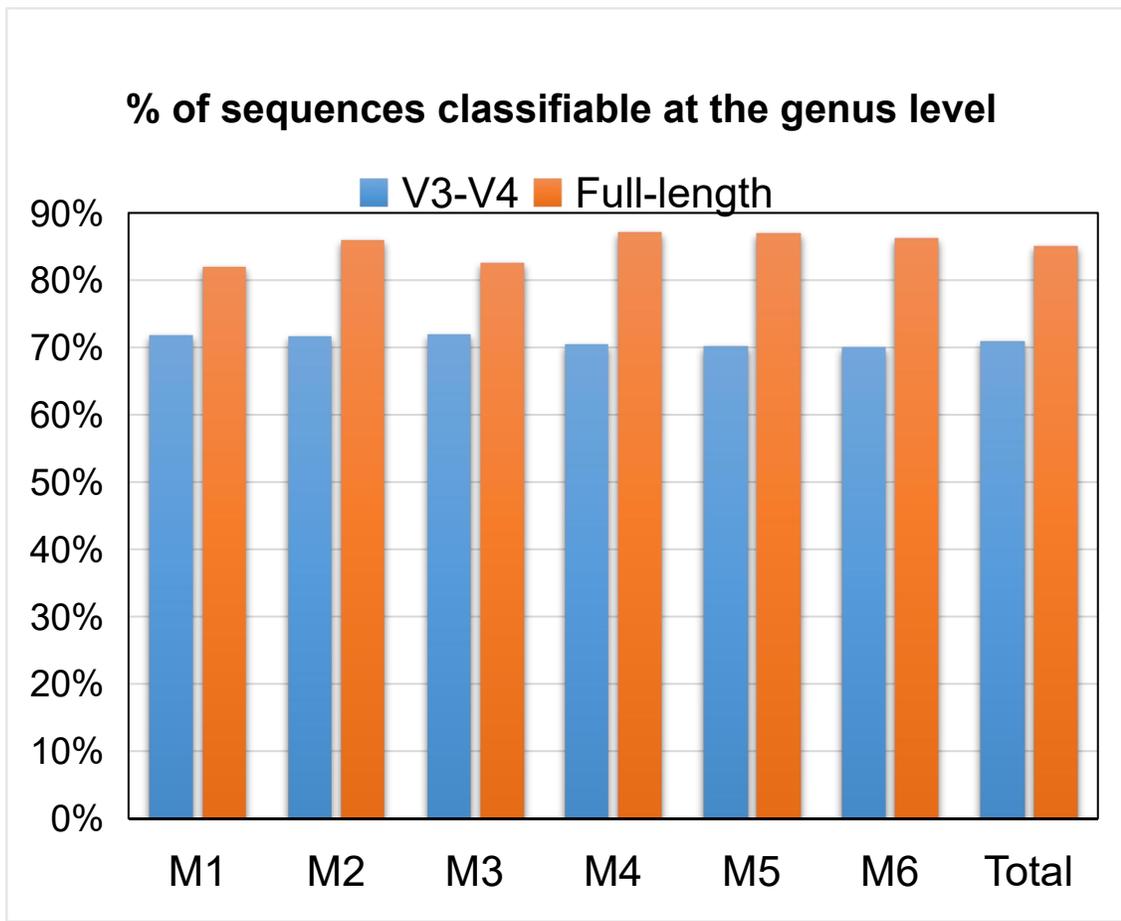


样品类型	二代属水平平均注释率	二代种水平平均注释率	三代属水平平均注释率	三代种水平平均注释率	种水平提升倍数
粪便	75%	6.4%	88%	72%	12
发酵	78%	20%	97%	85%	4
肠道	67%	9.4%	82%	62%	6
瘤胃	40%	5.2%	71%	11%	2
淤泥	43%	5.5%	63%	28%	5
土壤	48%	3.8%	68%	36%	11
水体	64%	7%	77%	56%	8

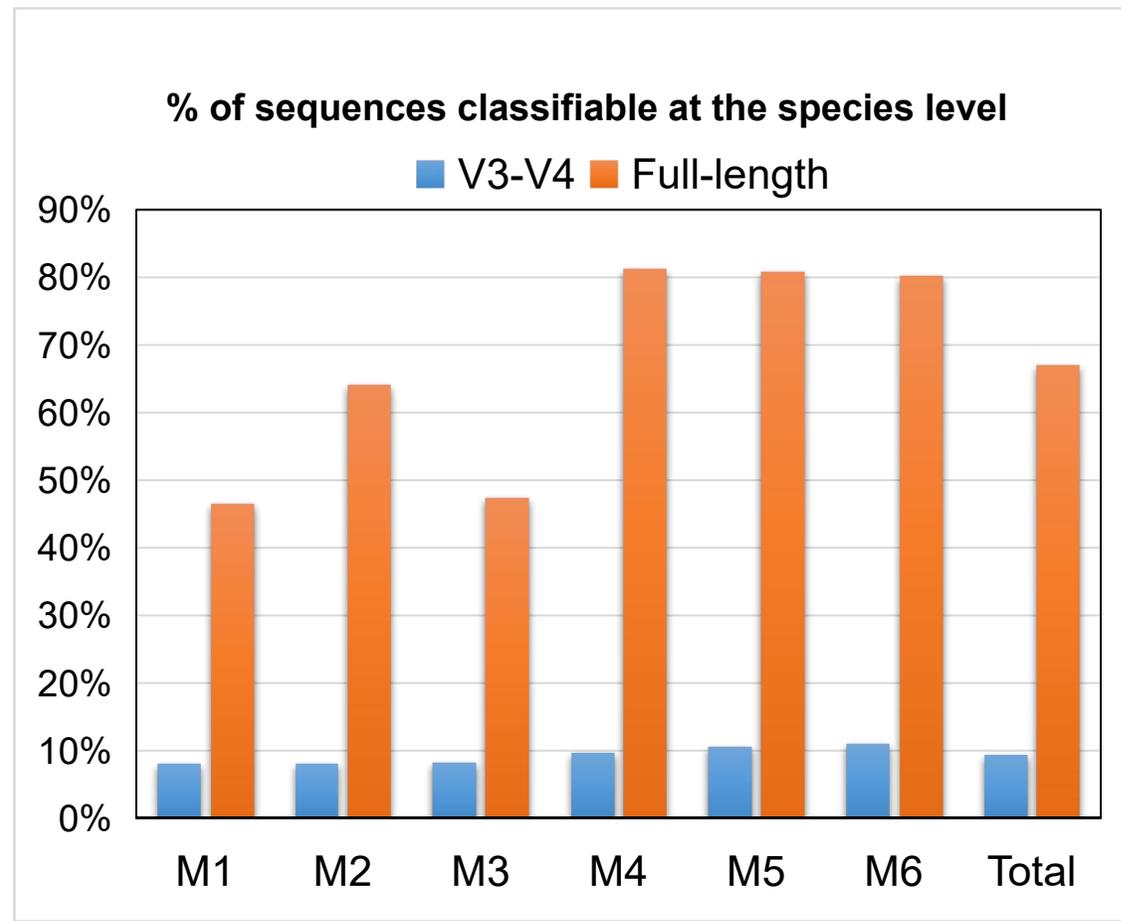


只有全长序列 (V1-V9) 才能提供可能存在的所有物种的无偏分辨率。

# 实测展示：优势1-显著提升种水平注释率

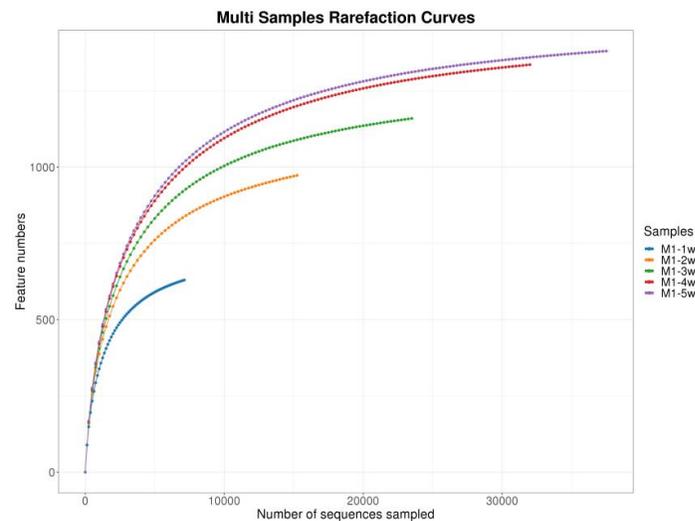
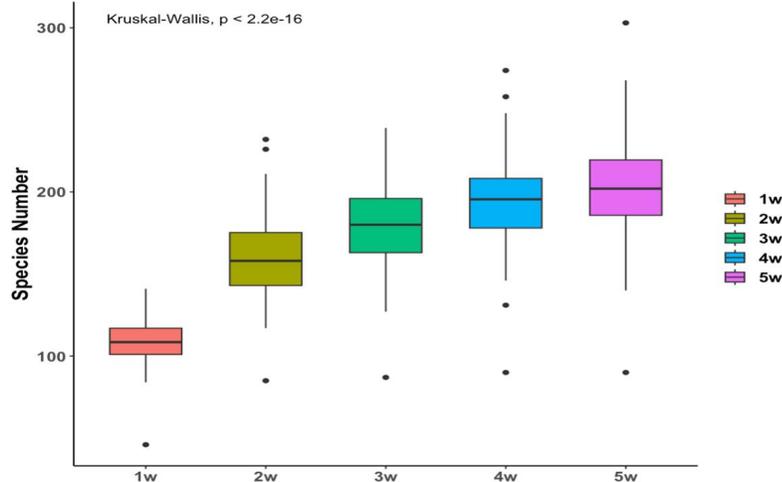
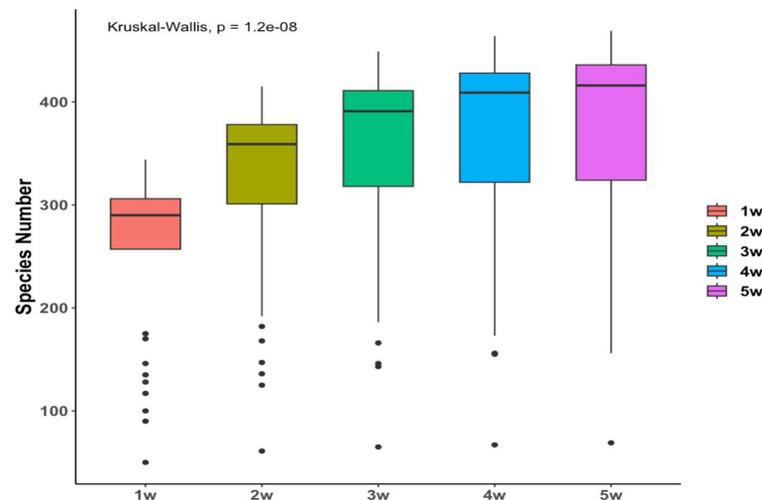
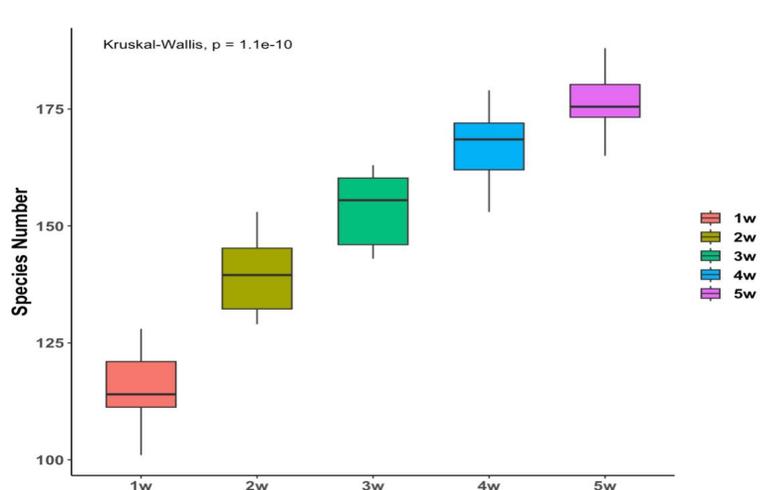


百迈客实测项目1



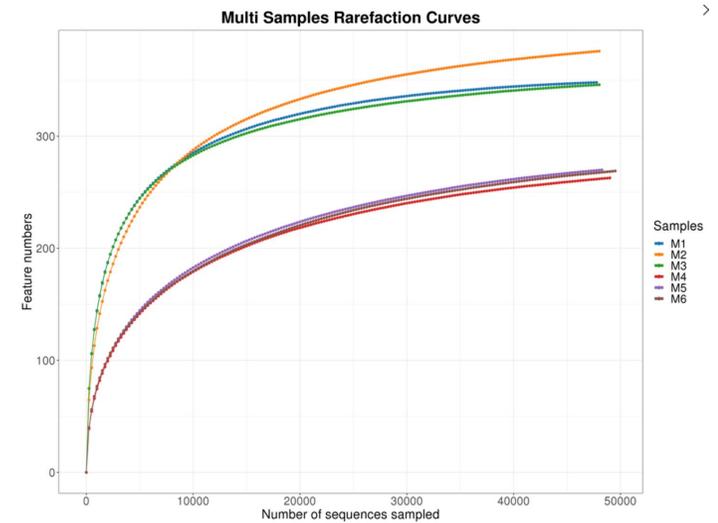
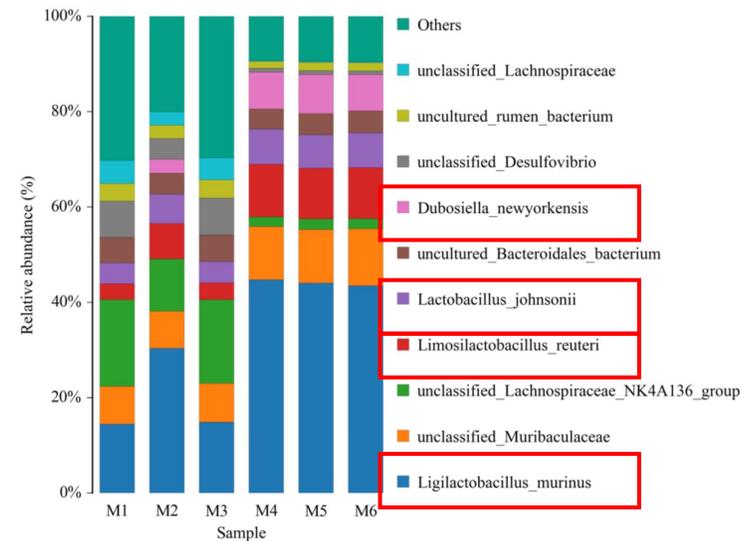
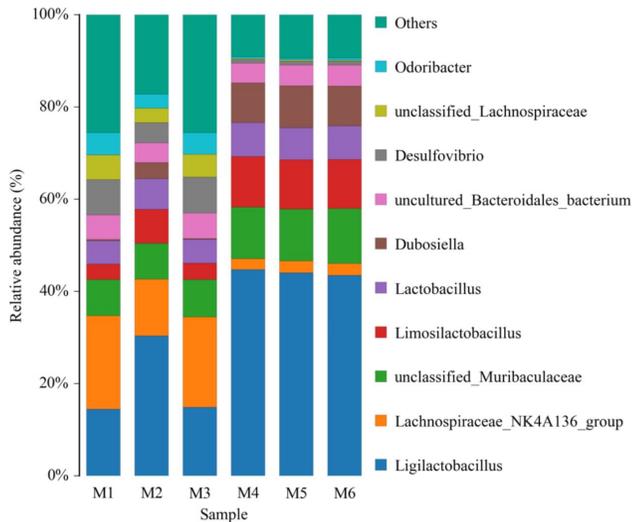
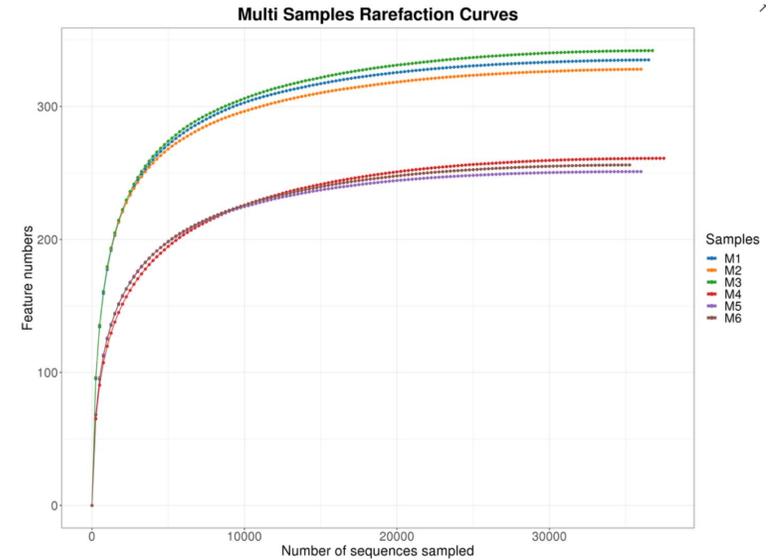
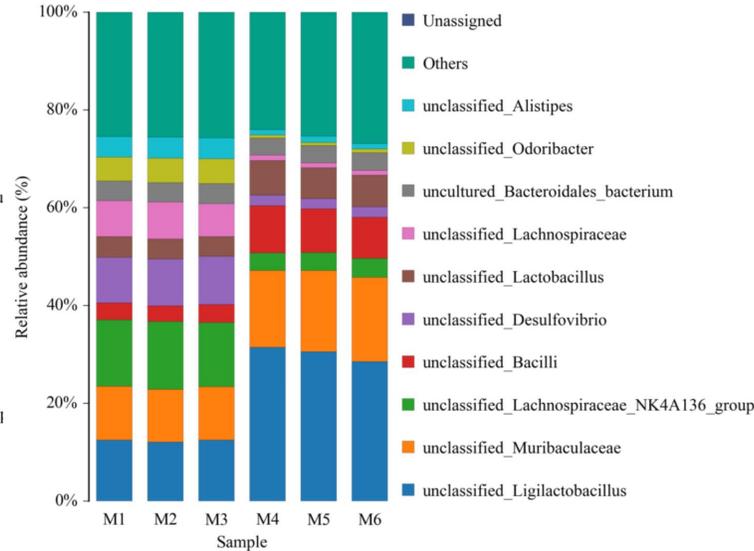
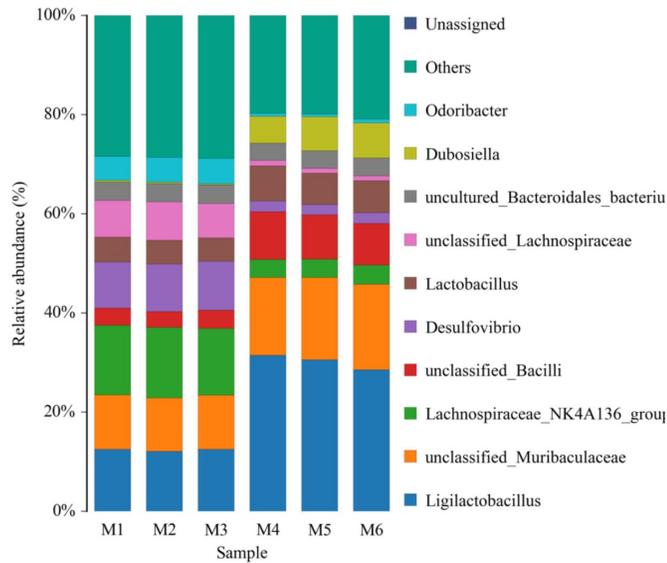
百迈客实测项目2

# 实测展示：优势2-低丰度物种检出，物种检出提升40%



图注：分别使用1w、2w、3w、4w、5w条CCS数据量进行种水平注释，5w注释到物种最多。

# 实测展示：优势3-种水平注释助力下游实验验证，提升文章水平



二代测序种水平注释都为未分类物种unclassified，通过三代测序注释到了4个明确的菌种，可用于下游分析

# 优势4-BMKCloud云平台兼容二三代，可选ASV分析

<https://international.biocloud.net/>

项目名称

[操作说明](#)

[视频教程](#)

## 第①步：综合选项

该平台目前支持二代和三代的测序数据，文件格式为fastq格式，双端测序数据文件名通常为XXX\_1.fq、XXX\_2.fq。

示例数据：[微生物多样性极速体验项目](#)  ▲示例数据不支持修改参数

流程版本：

数据类型： 单端  双端

报告名称 ：

## 第③步：分析参数选择

数据库选择：

分类方法选择：

聚类方法选择： 相似性聚类  序列相似性必须大于等于0.8，小于等于1.0

去噪(dada2)

过滤参数选择： 保留序列数总和占总序列数的比例大于等于十万分之五的特征。

保留至少在  个样品中的序列数都大于等于  的特征。

保留序列数总和大于等于  的特征。

## 第④步：扩增子类型

分子标记类型-----请选择和测序数据相匹配的分子标记类型，否则会导致分析错误。

16s  ITS  18s  功能基因

可变区-----请选择和测序数据相匹配的可变区或功能基因，否则会导致分析错误。

V3+V4(细菌)  V4(细菌)  V4+V5(细菌)  V3+V4(古菌)  V3+V4(内生菌)  全长



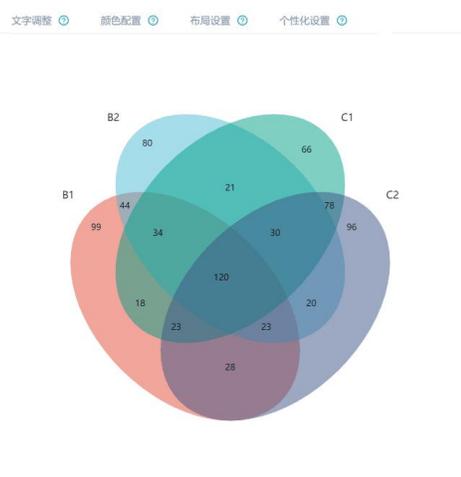
微生物多样性APP(二三代全兼容)

Silva138/UNITE/NCBI/FUNGENE

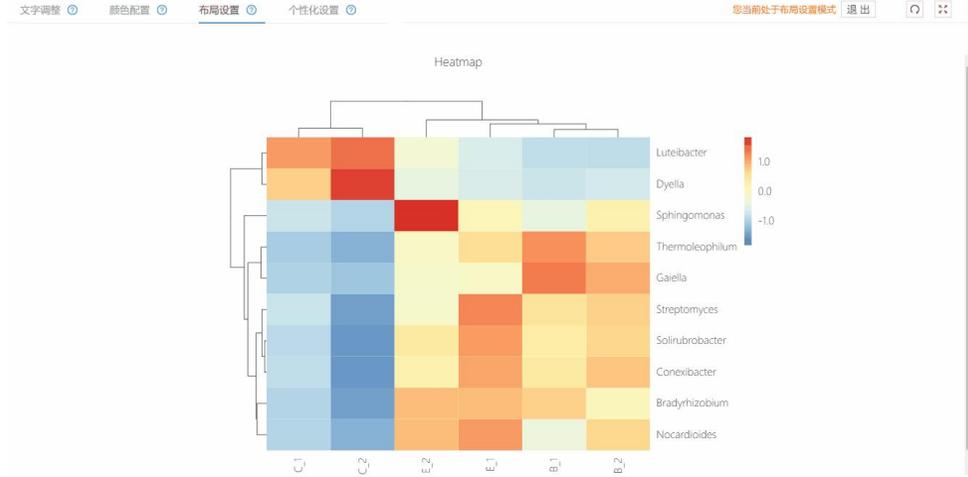
**32项** 标准分析 **42项** 个性化分析

# 优势5-百迈客云交互式让分析更便捷

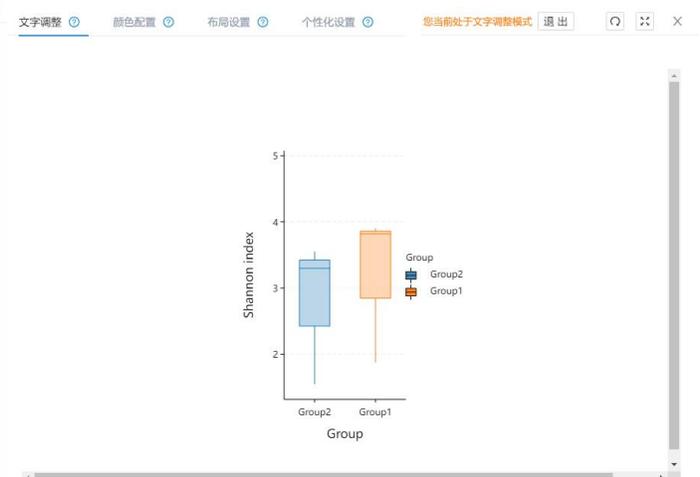
可动态调整的，**免费不限次数**修改图片配色，字体，**个性化**选项的功能



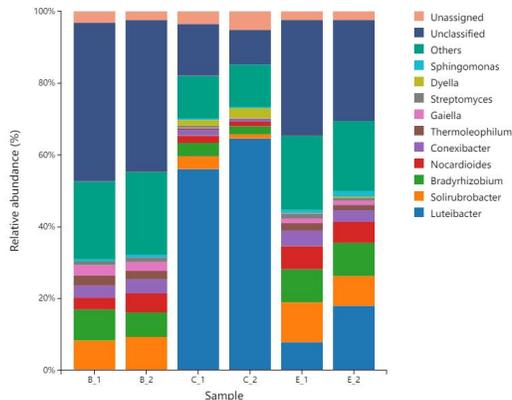
venn图



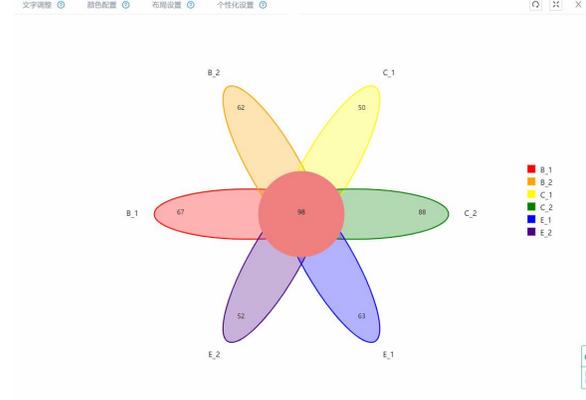
热图



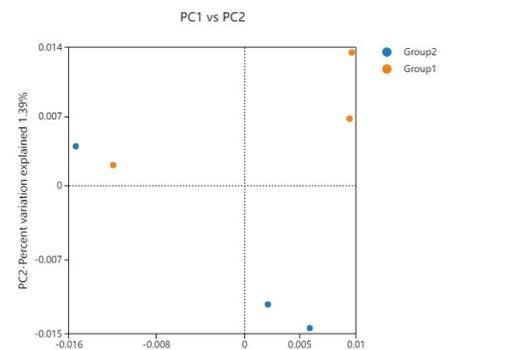
箱线图



堆叠柱状图



花瓣图

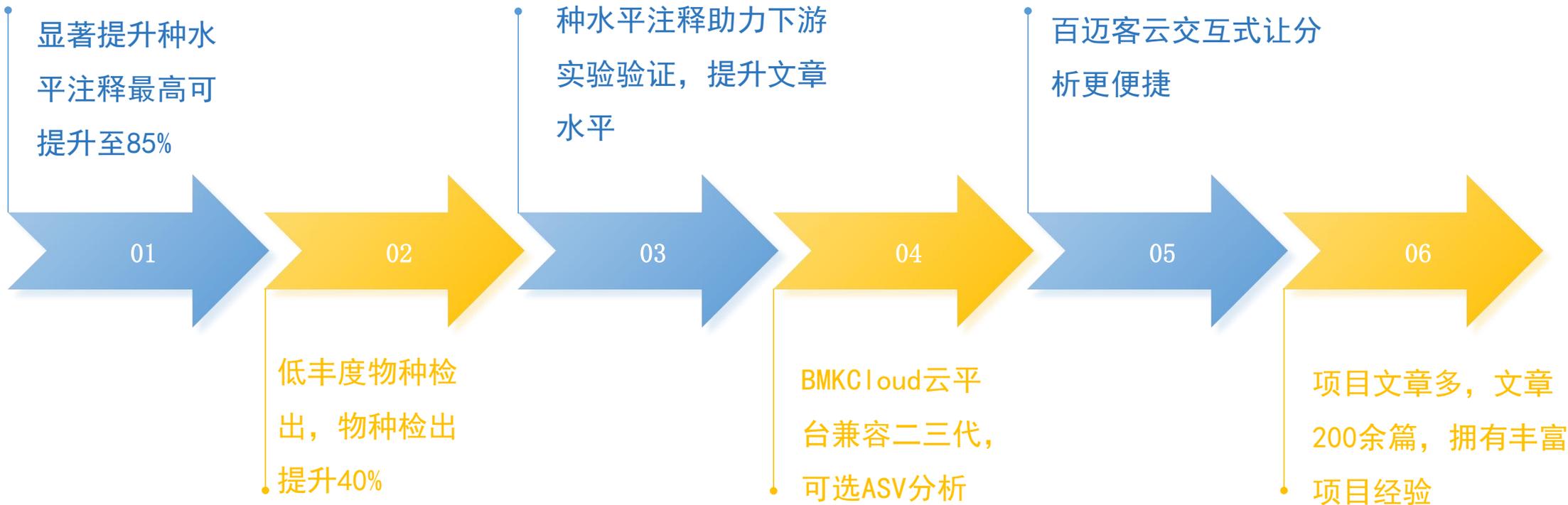


散点图

# 优势6-项目文章多，文章200余篇，拥有丰富项目经验

年份	期刊	影响因子	合作单位	题目
2023	Microbiome	16.837	中国农业大学	Gut microbiota-derived metabolites mediate the neuroprotective effect of melatonin in cognitive impairment induced by sleep deprivation
2023	Journal of Hazardous Materials	13.6	首都医科大学	The gut-brain axis involved in polystyrene nanoplastics-induced neurotoxicity via reprogramming the circadian rhythm-related pathways
2022	Environment International	13.353	首都医科大学	Polystyrene micro-/nanoplastics induced hematopoietic damages via the crosstalk of gut microbiota, metabolites, and cytokines
2022	Science of The Total Environment	10.753	南方医学院	Chronic exposure to polyvinyl chloride microplastics induces liver injury and gut microbiota dysbiosis based on the integration of liver transcriptome profiles and full-length 16S rRNA sequencing data
2022	Nano Research	10.269	首都医科大学	Melatonin and probiotics ameliorate nanoplastics-induced hematopoietic injury by modulating the gut microbiota-metabolism
2023	Ultrasonics Sonochemistry	9.336	五邑大学生物科技与大健康学院	Ultrasonic assisted extraction, characterization and gut microbiota-dependent anti-obesity effect of polysaccharide from Pericarpium Citri Reticulatae 'Chachiensis'
2023	Microbiology Spectrum	9.043	中国农业大学	Disruption of the Intestinal Mucosal Barrier Induced by High Fructose and Restraint Stress Is Regulated by the Intestinal Microbiota and Microbiota Metabolites
2022	Chemosphere	8.943	首都医科大学	Early hematopoietic injury triggered by benzene characterized with inhibition of erythrocyte differentiation involving the mollicutes_RF39-derived citrulline
2023	Oxidative Medicine and Cellular Longevity	7.3104	长春中医药大学	Rice Water-Fried Atractylodis Rhizoma Relieves Spleen Deficiency Diarrhea by Regulating the Intestinal Microbiome
2022	Ecotoxicology and Environmental Safety	7.129	南方医科大学	Polyvinyl chloride microplastics induced gut barrier dysfunction, microbiota dysbiosis and metabolism disorder in adult mice
2022	Life Sciences	6.78	成都中医药大学	Electroacupuncture promoted intestinal defensins and rescued the dysbioticcecal microbiota of high-fat diet-induced obese mice
2022	Nutrients	6.706	中国农业科学院	2'-Fucosyllactose Remits Colitis-Induced Liver Oxygen Stress through the Gut-Liver-Metabolites Axis

# 百迈客三代MCD优势总结

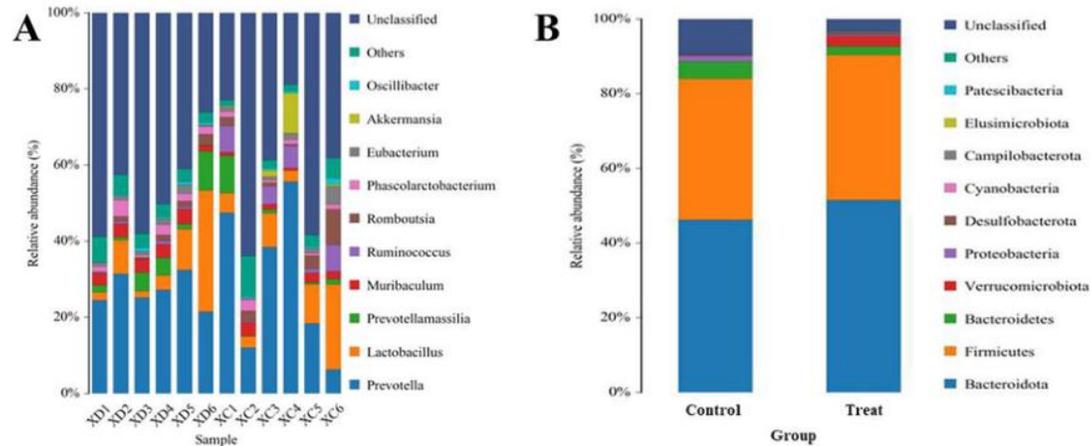


2

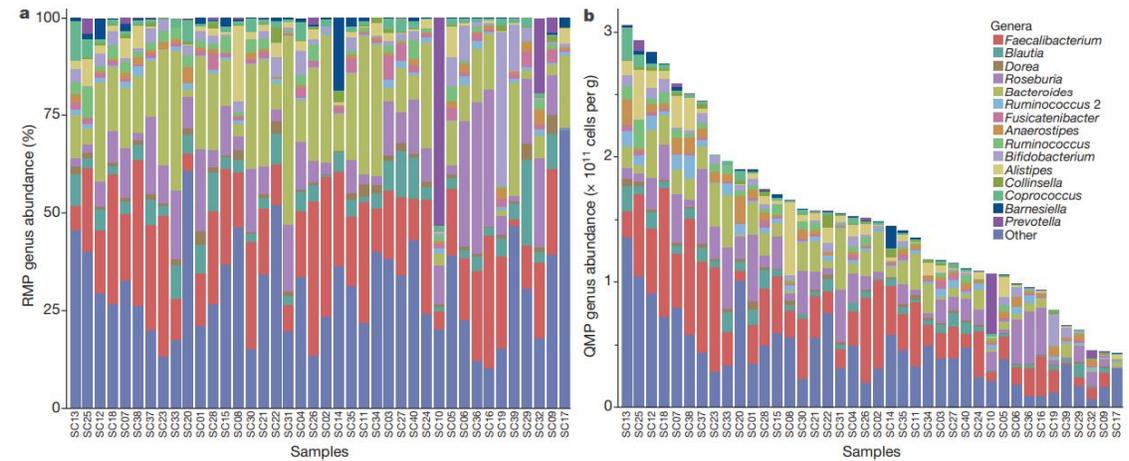
绝对定量

# 常规微生物多样性测序局限性

- 获取样品中物种的相对丰度，从而进行样品间、组间进行差异及相关性分析等，不能反映样品中物种真实的绝对丰度和真实细菌载量情况。
- 传统QPCR微生物定量通过标准品构建标准曲线，样本需单独扩增，扩增效率存在差异从而造成定量的不准确性。



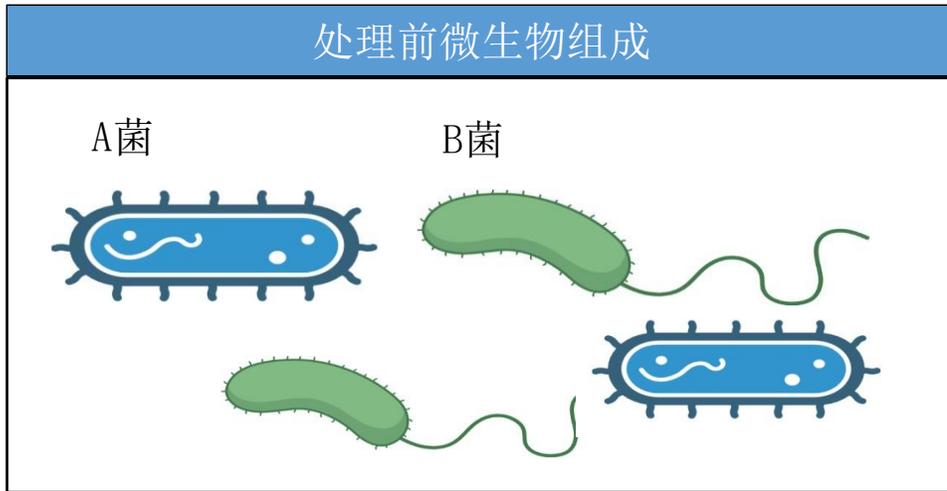
X. Dong et al. *Science of the Total Environment*, 2022



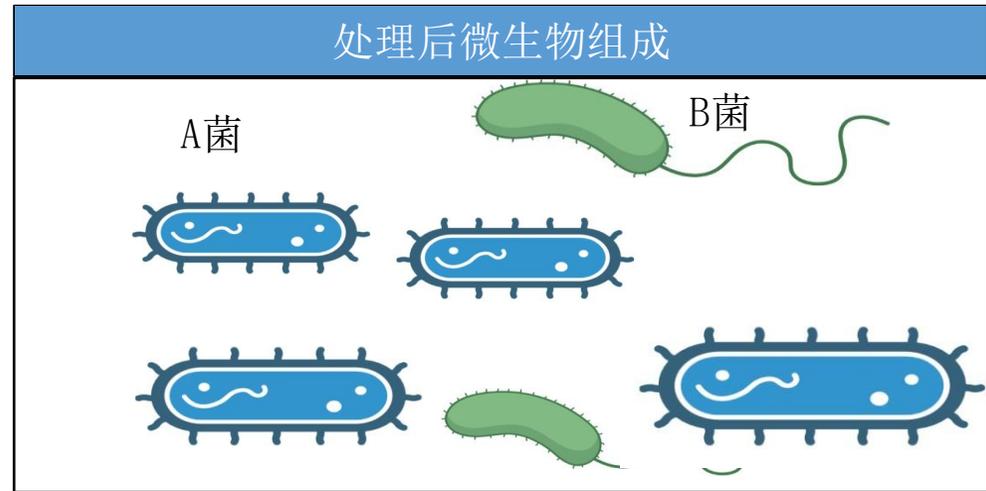
Vandeputte D et al. *Nature*, 2017

# 常规微生物多样性相对丰度局限性

- 相对丰度分析忽略了**不同样本之间总体微生物量存在的现实差异**，造成了分析结果的偏差。
- 绝对定量分析则是计量样本每种微生物的16S基因拷贝数或者细胞数目，从而实现绝对定量，因此相对于常规16S相对定量扩增子测序，**绝对定量分析能反映样本每种微生物的真实数量和组间样本的真实差异**。



处理前相对定量A:50% B:50%  
绝对定量: cell数: A:2 B:2



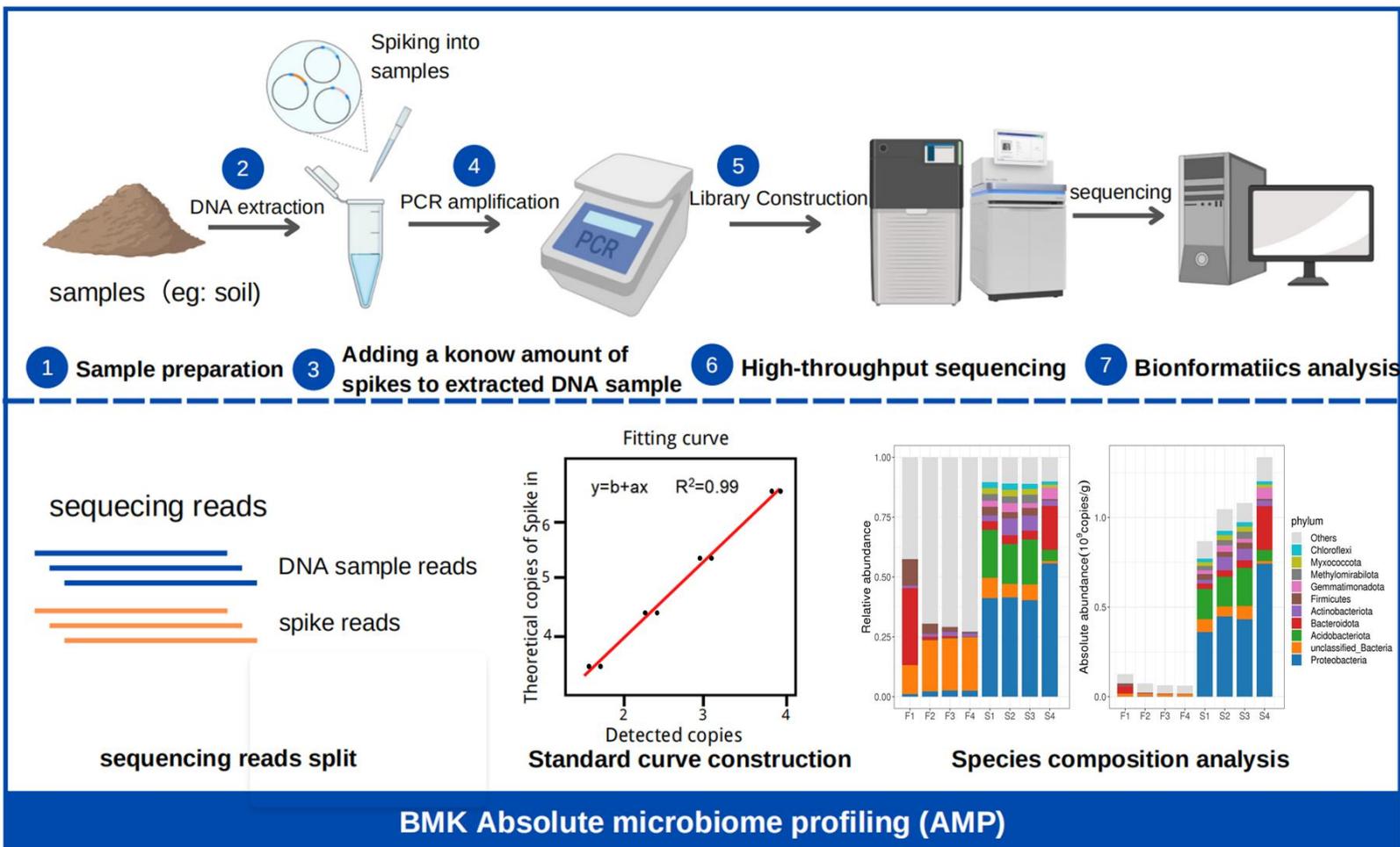
相对定量结果: A:66.7% B:33.3%  
绝对定量: cell数: A:4 B:2

相对定量结果为: A菌丰度增加, B菌丰度降低, 而**真实情况B菌并未减少**, A菌增加

绝对定量方法	主要应用(已发表)	优势	限制
荧光光谱法	水、土壤、食品和饮料以及空气	高亲和力；多种染料选择以区分活细胞和死亡细胞	无法对 DNA 完全降解的死细胞进行染色；一些染料同时结合DNA和RNA
流式细胞术	粪便、水和土壤	快速；单细胞计数；基于生理特征的灵活参数；能区分活细胞和死亡细胞	可能需要排除背景噪声；可能需要控制策略；可能需要稀释；对于复杂系统/不同种类的样品不理想
Spike-in内参序列法	土壤、淤泥和粪便	快速；易于整合到高通量测序中；高灵敏度；易于处理	内参、加入量和时间点对准确度有很大影响，可能需要16S rRNA拷贝数校正。
16S qPCR	粪便、临床(肺)、土壤、植物、空气和水	直接量化特定的分类群；经济有效且易于处理；高灵敏度；与低生物量样本兼容	可能需要16S rRNA拷贝数校正；存在与PCR相关的偏差；需要标准曲线

- 注：当不需要进行活细胞检测，无靶向特定分类群，无需检测形态学和低生物量样本时，建议选择使用 spike-in 进行绝对定量，且 spike-in 方法适合大样本量测序研究。
- 传统QPRC技术通过标准品构建标准曲线，样本需单独扩增，扩增效率存在差异而造成定量的不准确性。

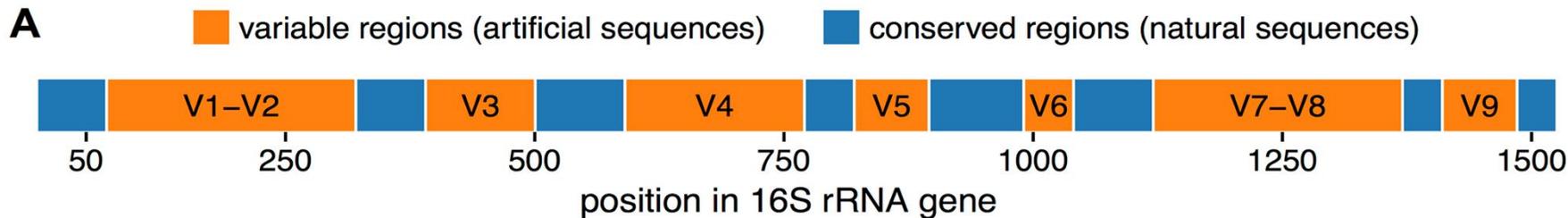
# 百迈客微生物绝对定量技术原理



- 运用高通量测序技术和spike-in内标法实现对细菌的绝对定量。
- 通过在样品中加入已知拷贝数的spike-in内参序列，与样品同时进行扩增、建库测序，然后通过内参的理论拷贝数与测序tags数，绘制标准曲线，结合标准曲线及样品的组织、核酸、扩增用量等计算得到OTUs/ASVs的绝对拷贝数表。
- 能够一次获得所有细菌的绝对定量结果，包括总丰度和每种菌的丰度。

# Spike-in 内参序列设计原则

什么是 Spike-in ?



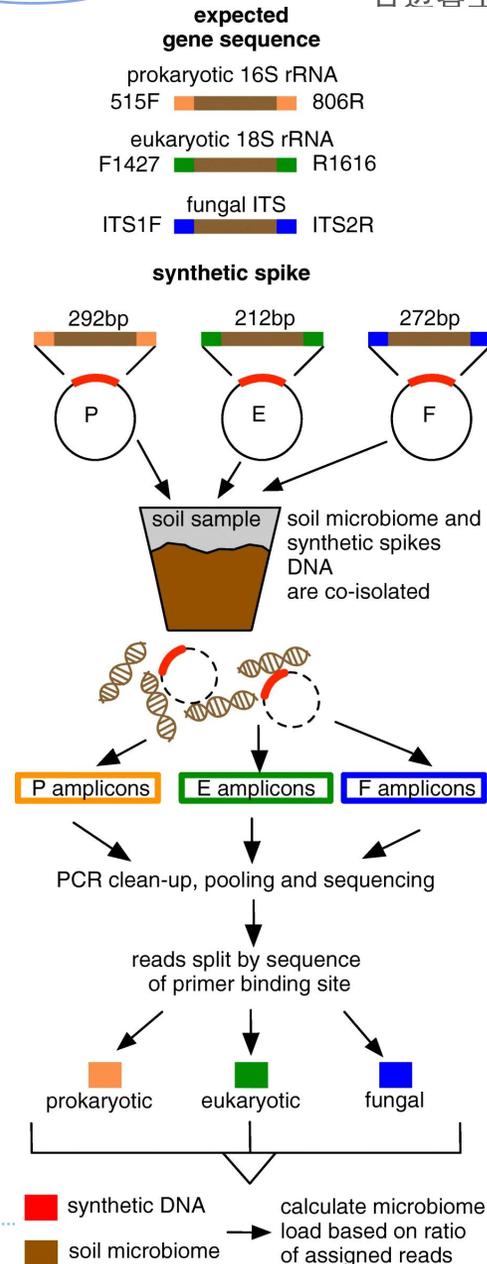
- spike-in是人工合成的一段核酸片段(Synthetic spike-in), 由与选定的天然 16S rRNA 基因相同的保守区和人工可变区组成。
- Spike in 序列与样本中的微生物基因组序列缺乏一致性。

简单理解就是它在样本中是独一无二的，不会影响到其他序列，易于区分，后续分析可以根据已知序列信息进行提取序列。目前我们是以80%以下同源性作为设计原则。

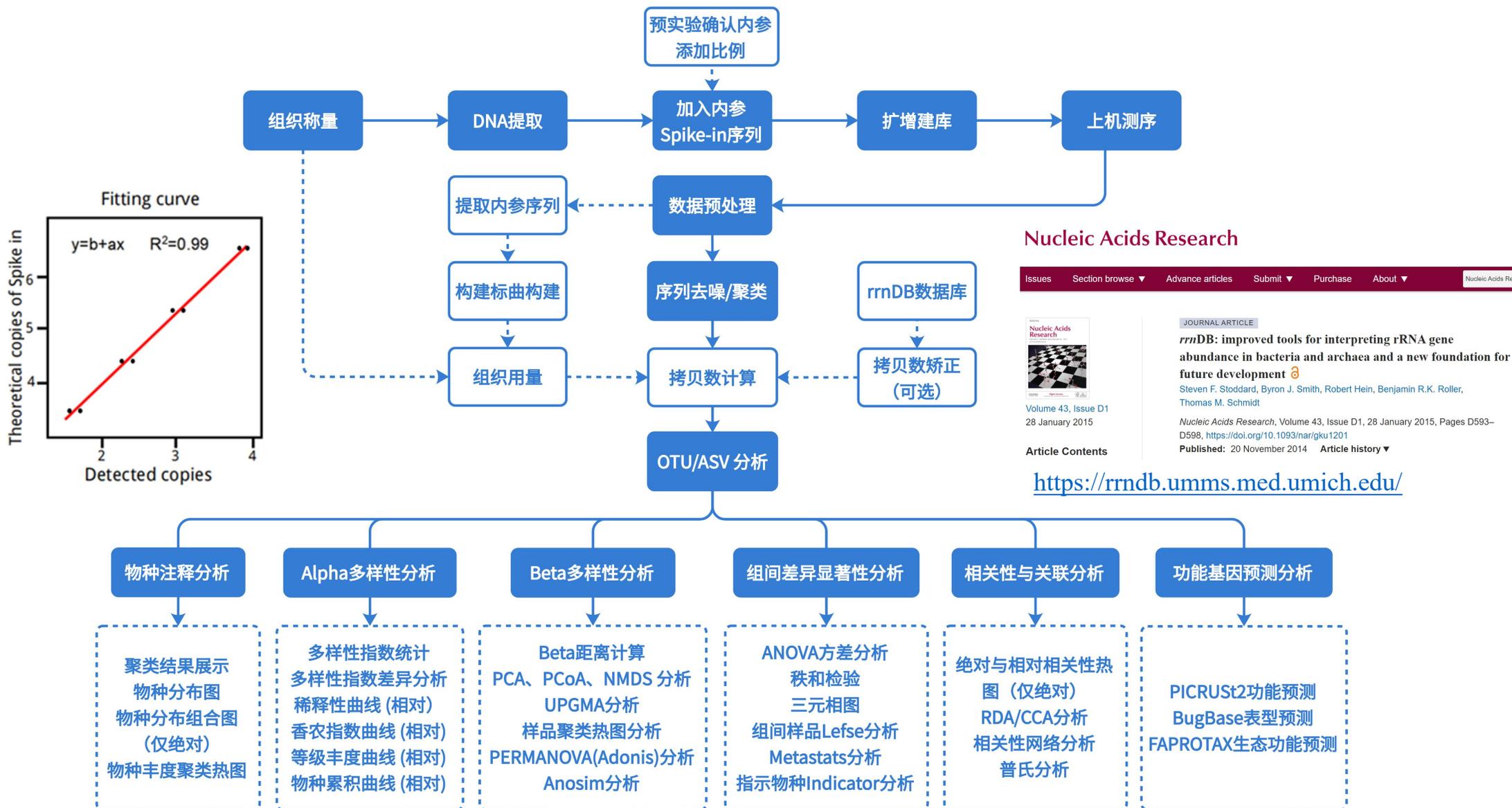
spike-in的设计有三个关键因素：

- ① 具有共同的基因引物结合位点(primer binding sites, PBSs);
- ② 与待测靶标序列相似的长度和GC含量;
- ③ 容易获得、易于处理;

将合成的序列克隆到质粒载体中，将其转化到大肠杆菌中进行扩繁得到大量的 spike-in 序列。

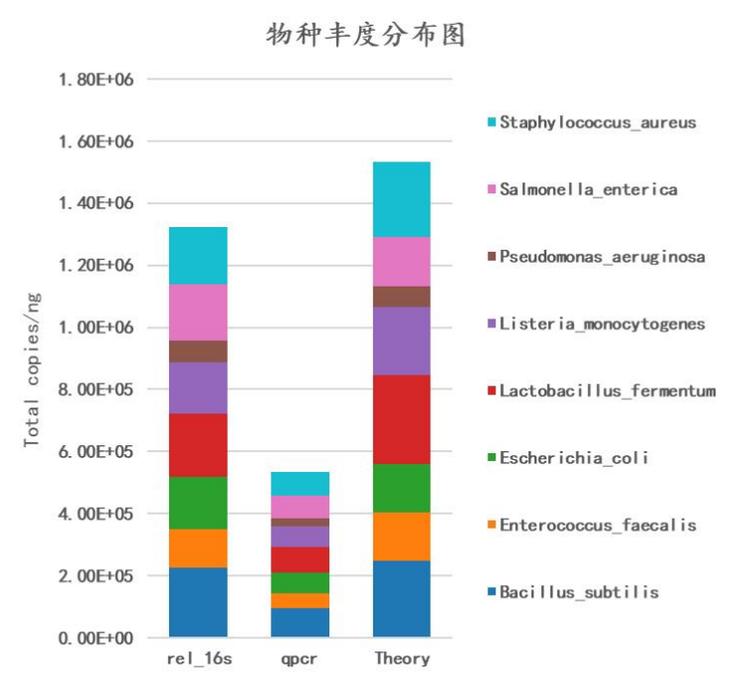
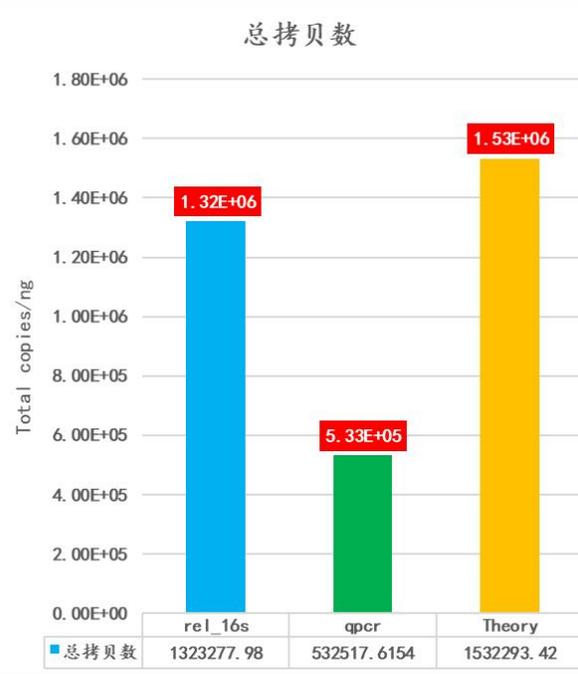


# 百迈客微生物绝对定量技术路线



# 绝对定量 vs QPCR

菌种拉丁文	rel_16s	qpcr	Theory
Bacillus_subtilis	225073.925	93590.51	248477.9692
Enterococcus_faecalis	125307.474	48442.03	153890.3978
Escherichia_coli	169347.0917	67968.11	157165.5012
Lactobacillus_fermentum	201124.3085	82468.68	287282.2715
Listeria_monocytogenes	165684.2939	66317.28	219494.4093
Pseudomonas_aeruginosa	71960.83478	25963.11	64460.86305
Salmonella_enterica	179028.4863	72352.76	160962.5668
Staphylococcus_aureus	185751.5657	75415.12	240559.4406
	rel_16s	qpcr	Theory
总拷贝数	1323277.98	532517.6	1532293.42



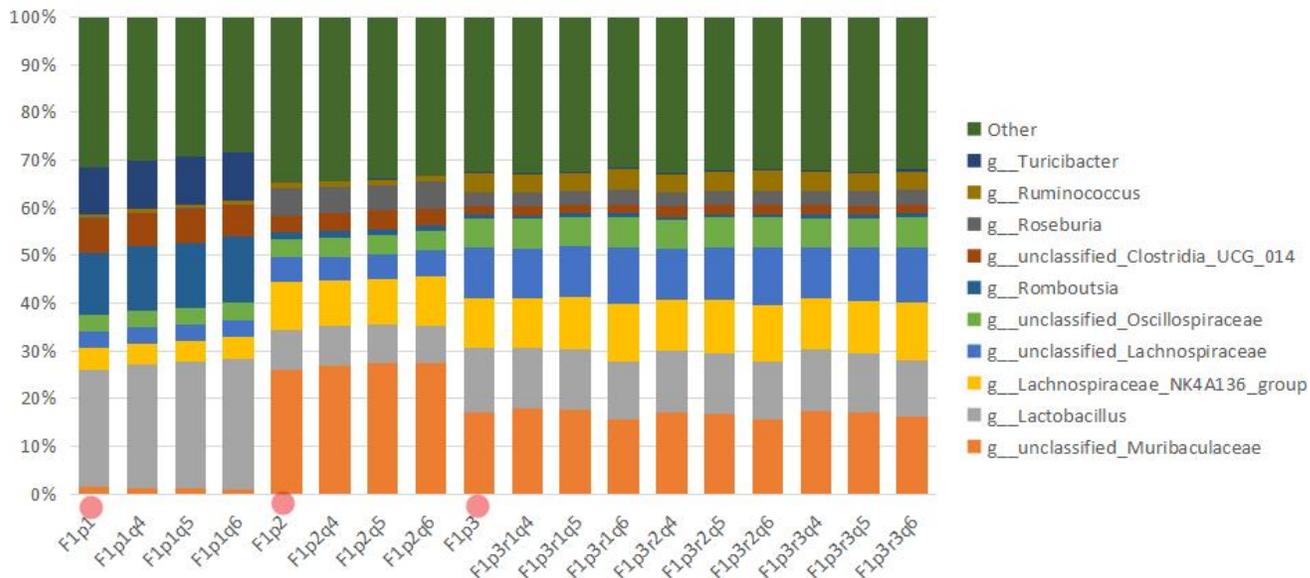
微生物绝对定量与qpcr定量结果对比图表

rel\_16s: 微生物多样性绝对定量结果; qpcr: qpcr经过相对丰度换算结果; Theory: 理论拷贝数

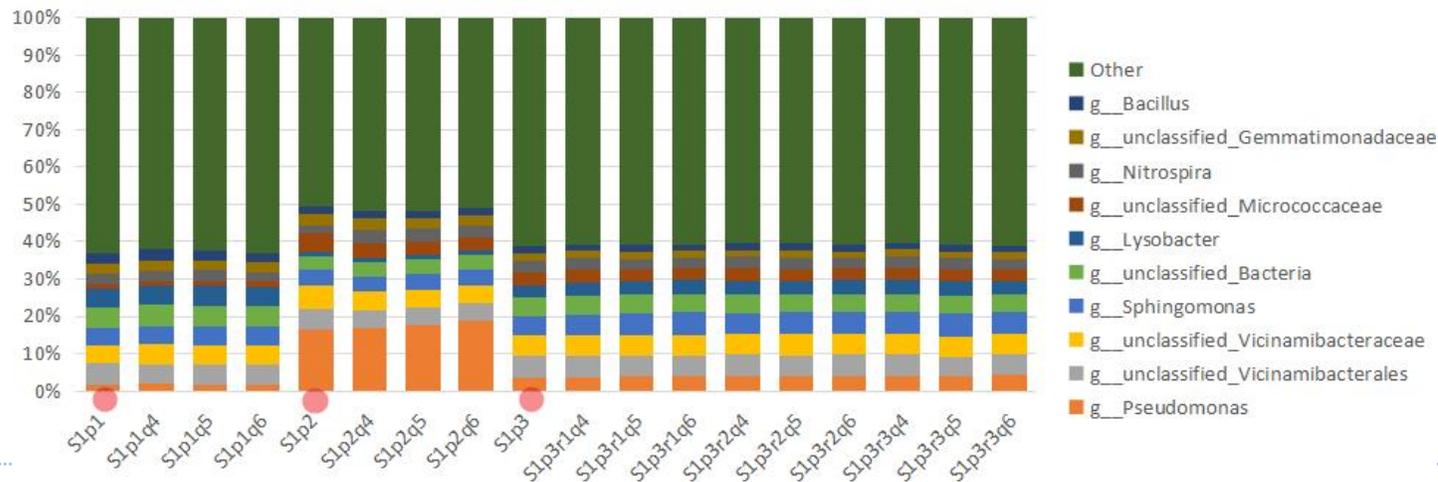
从测试结果整体看，微生物绝对定量结果更接近理论值，而QPCR的结果则存在一定的偏差

# 实测数据展示

F1 Genus(FAU575) 粪便



S1 Genus(SAU215) 土壤



内参序列设计规则:

不影响实际样品物种, 序列之间相似度低于80%。

内参涵盖7种不同浓度的内参序列, 以便监控到样品中高中低物种的情况。

3个样品(p1/p2/p3), 其中p3采用**技术重复**, 判断重复性情况。

均采用**3种**浓度的插入序列(q4 q5 q6)

控制掺入内参序列比例, 不会影响真实样品

# 实测结果展示-物种组成分析

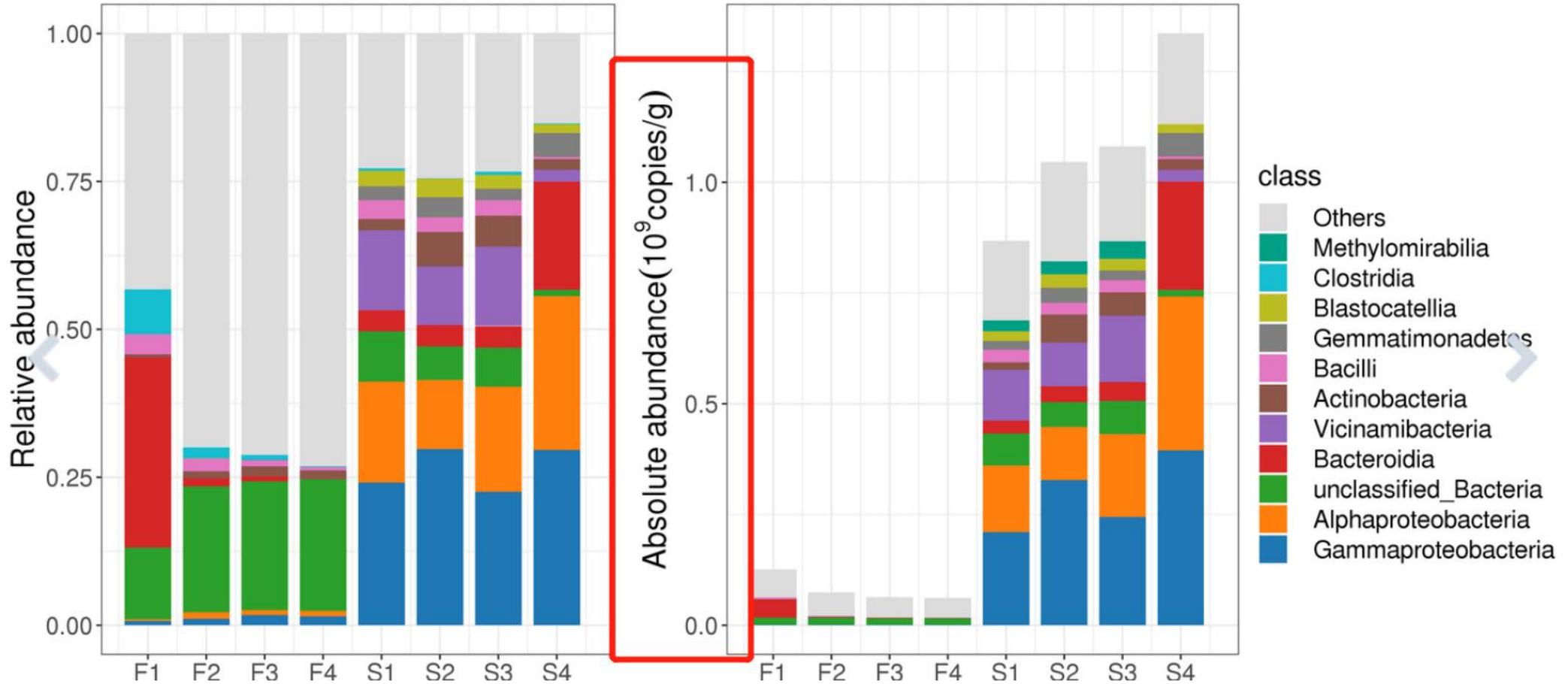


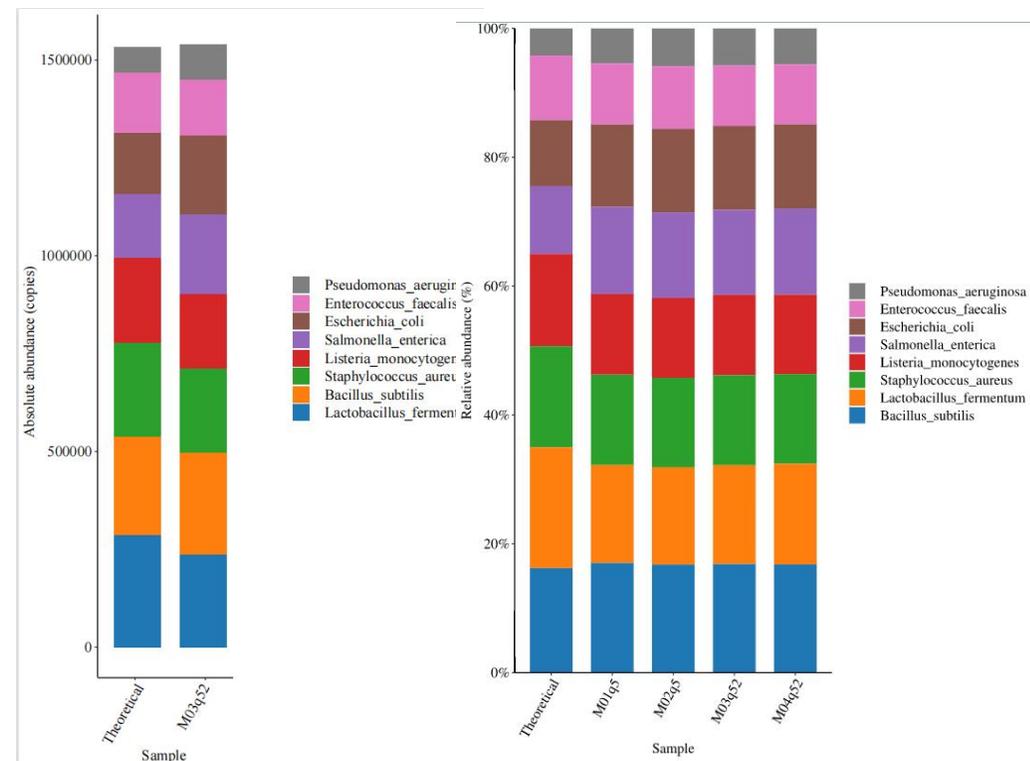
图5 genus水平物种分布图

左：横坐标为样品名称；纵坐标为相对丰度图

右：横坐标为样品名称；纵坐标为物种拷贝数(copies/g)。

# 百迈客优势一：技术成熟，数据稳定可靠

Mock菌种	真实比例(%)	未添加内标 M01(%)	添加内标1 M02(%)	添加内标2 M03(%)	添加内标3 M04(%)
<i>Bacillus subtilis</i>	17.4	17.58	17.66	17.51	17.29
<i>Enterococcus faecalis</i>	9.9	9.10	9.14	8.89	8.91
<i>Escherichia coli</i>	10.1	12.76	12.97	13.05	13.08
<i>Lactobacillus fermentum</i>	18.4	15.49	15.59	15.74	15.98
<i>Listeria monocytogenes</i>	14.1	12.45	12.28	12.37	12.24
<i>Pseudomonas aeruginosa</i>	4.2	4.88	4.98	5.10	5.05
<i>Salmonella enterica</i>	10.4	13.59	13.32	13.22	13.41
<i>Staphylococcus aureus</i>	15.5	14.16	14.07	14.12	14.04
与M01一致性	/	/	0.999	0.998	0.997
与理论比例一致性	/	0.896	0.897	0.9	0.897

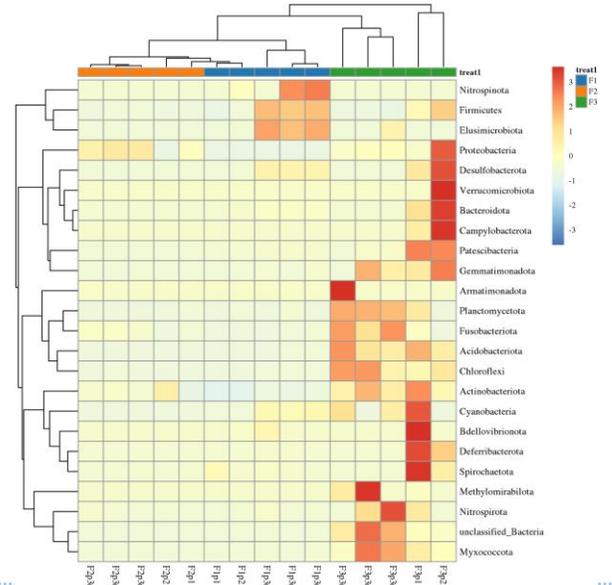
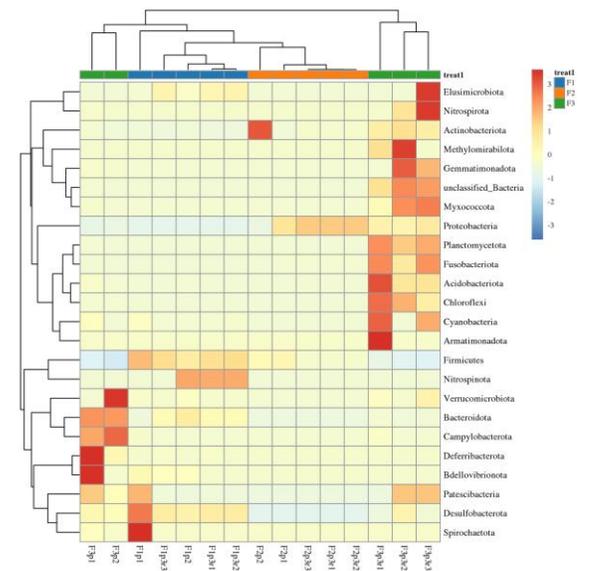
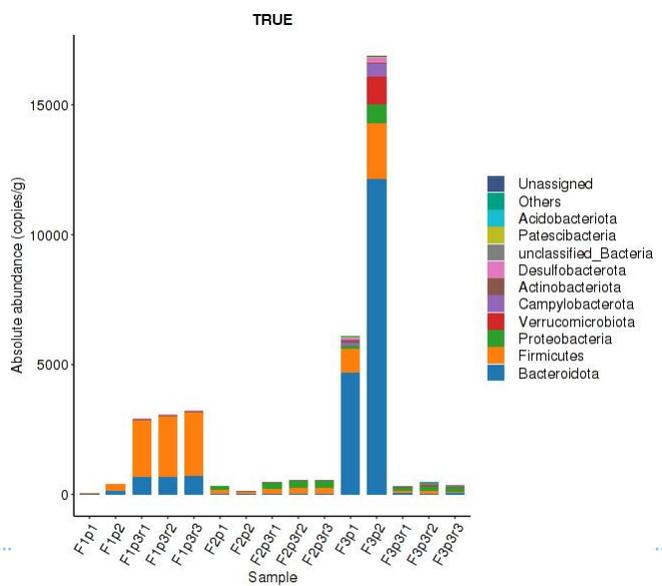
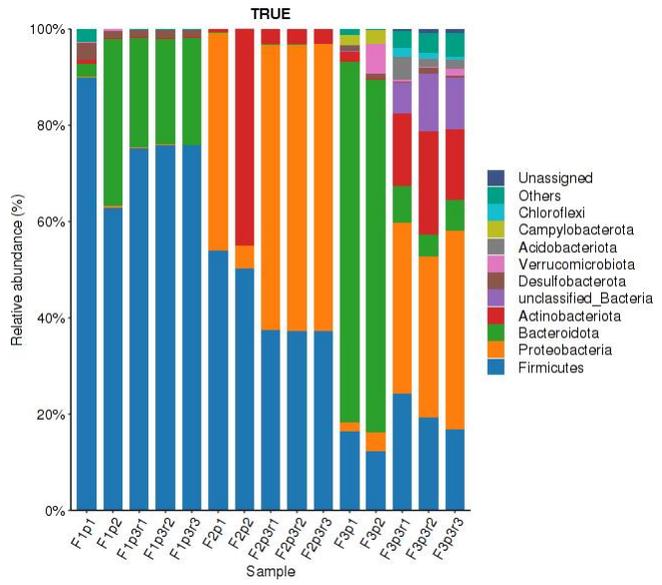
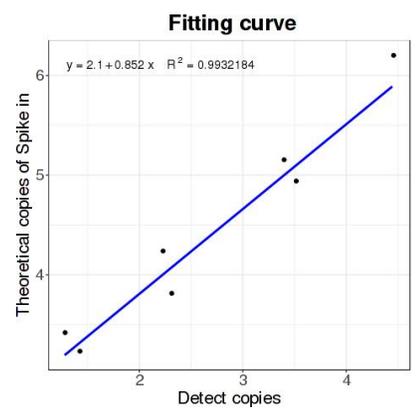
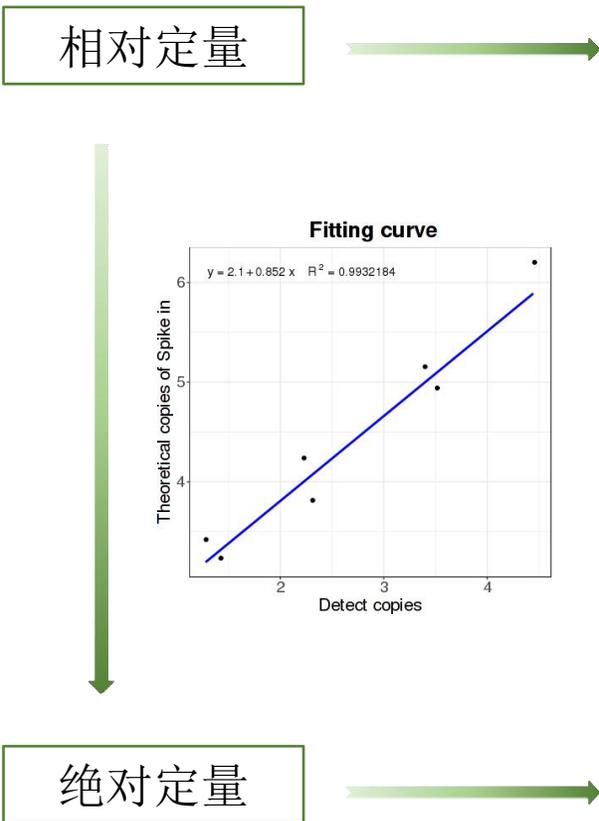


- 1、添加内参序列不会对样品内物种比例产生影响
- 2、标准品无论是否掺入内参序列与理论值相关性一致

注：M01为标准品样本，M02/M03/M04为标准品样品添加不同浓度的内参序列测试；标准品样品为 ZymoBIOMICS™ Microbial Community DNA Standard, Catalog Nos. D6305

# 百迈客优势二：同时实现相对定量和绝对定量

一次测序  
两份结果



# 百迈客优势三：rrnDB数据库拷贝数矫正，拷贝数更接近细胞数

## Nucleic Acids Research



Volume 43, Issue D1  
28 January 2015

*rrnDB*: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development 

Steven F. Stoddard, Byron J. Smith, Robert Hein, Benjamin R.K. Roller, Thomas M. Schmidt

*Nucleic Acids Research*, Volume 43, Issue D1, 28 January 2015, Pages D593–D598, <https://doi.org/10.1093/nar/gku1201>

Published: 20 November 2014 [Article history](#)

<https://rrnodb.umms.med.umich.edu/>

Last update:rrnDB, version 5.8,  
June 23, 2022

- 16S rRNA基因在一个物种中会有多拷贝，在PCR后会有多个扩增产物，比如物种A只有1个16S基因，物种B有2个16S基因，在原始群落中，二者丰度相同，在经过相同循环次数的PCR，理论上最终测序得到的reads中，物种B的reads数会是物种A的2倍；
- 而我们通常使用 reads 数表征某个OTU的丰度，尽管在群落中物种A和B相同，但是由于拷贝数的差异，所以定量的结果，不能正确的反映在群落中二者的丰度比例；丰度定量不完全准确，对于后续的 alpha 和 beta 多样性的分析都会有影响；
- rrnDB数据库中收录了16S基因为多拷贝的物种；经过矫正后的数据更接近微生物的细胞数，生信分析结果更**真实可靠**。

利用rrnDB数据库进行拷贝矫正，使拷贝数更接近细胞数

Stoddard S.F, Smith B.J., Hein R., Roller B.R.K. and Schmidt T.M. (2015) *rrnDB*: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research* 2014



# 百迈客微生物绝对定量测序优势总结

## 超高性价比

高

- 同时获得相对定量和绝对定量结果
- OTU聚类和ASV去噪可选
- 独有特色分析

## 高准确度

准

- ◆ rrnDB数据库进行拷贝数校正
- ◆ 16s拷贝数更接近真实细菌细胞数

## 种水平定量

种

- ◆ 结合三代测序实现物种水平定量
- ◆ Silva138等多种数据库任选

## 新发现

新

- 技术成熟，目前已有多篇文章发表
- 从绝对定量角度阐述问题，新发现

3

## 推广话术总结

产品	适用样本/研究方向	学校或学院类型	话术
三代全长	动物粪便 > 肠道 > 瘤胃、口腔等	医院、公卫学院、检验科、微生物研究说 (高频: 人、鼠、等 模式生物: 斑马鱼、蜜蜂等 保护动物: 熊猫等)	①三代我们现在引入了最新仪器revio, 自主研发了相关专利, 提高了测序有效数据量, 降低了成本, <b>5W条全长序列的价格才两百出头, 非常接近二代, 所以现在做三代性价比很高</b>
	土壤样	资环、环境、农学院、地学院、植物类	②三代的测序结果相比二代, 不仅在种和属水平上的注释准确率更高, 而且注释出来的真实物种更多, 后续的功能预测也更准, 特别适合您研究的XXX样本
	发酵类 (食品、酒类)	乳制品、食品院、葡萄酒学院、饲料加工等	③我们有 <b>简洁易操作的云平台</b> 可以使用, 兼容二代和三代, 还配套了专有的全长16S数据库(如: 人肠道、小鼠肠道等) 进一步提高注释率, 既有一键化式的结题报告可查看下载, 老师还可以在关注的模块下选择参数进行重新分析, 生成对应的结果
			④针对粪便样本, 更适合客户做下游菌种验证, 分离培养/商业菌株补充实验处理等, <b>提升文章水平</b>

## 小结

- **知道大概的原理** (新仪器revio是PB平台的最新仪器, 具有通量高的特点。是对单个DNA 分子进行测序, 边合成边测序, 并以SMRT芯片为测序载体, 捕捉光信号转换成碱基信息。)
- **知道三代和二代比优势在哪里** (从数据实测角度、文献角度、性价比角度说给老师听)
- **知道适用的样本和场景**
- **积极到针对性的学院、课题组, 疯狂安利推广, 去抢在别的公司做二代、做全长的客户**

## 客户群：

1. 对新技术感兴趣；
2. 想看到真实的菌群载量，从而挖掘不同层面的规律，常见于医学组织、植物组织、发酵、食品之类的，对实际菌群含量比较在意的研究需求

## 推广话术：

老师好，绝对定量是我们公司刚研发出来的新产品，通过添加spike-in内参DNA序列进行标定，获得目标样本菌群的绝对丰度信息。

仅仅根据16S rRNA测序产生的相对丰度进行分析可能会导致对微生物群动力学的非真实解释，有很多CNS期刊都用绝对定量来揭示真实的菌群结构和载量，从而挖掘更真实的规律。

绝对定量也只是比传统的扩增子贵了几十块钱而已，一次测序既能获得相对丰度，又能获得绝对丰度，性价比超级高。我们的云平台也支持分析这个产品的个性化。老师可以考虑尝试下，对文章可能是个锦上添花或者能发现不一样的规律。



# Thank you for your attention!

