

# 微生物组分析方法、应用 与科学传播

CGM第99期在线沙龙

刘永鑫  
中国科学院遗传与发育生物学研究所 工程师  
宏基因组公众号 创始人兼主编  
2020年6月17日



# 个人简介

2004-2008 东北农业大学  
2008-2011 东北农业大学  
2011-2014 中科院遗传发育所  
2014-2016 中科院遗传发育所  
2016-今 中科院遗传发育所

学士 微生物学  
硕士 作物遗传育种  
博士 生物信息学  
博士后 遗传学  
工程师 微生物组分析





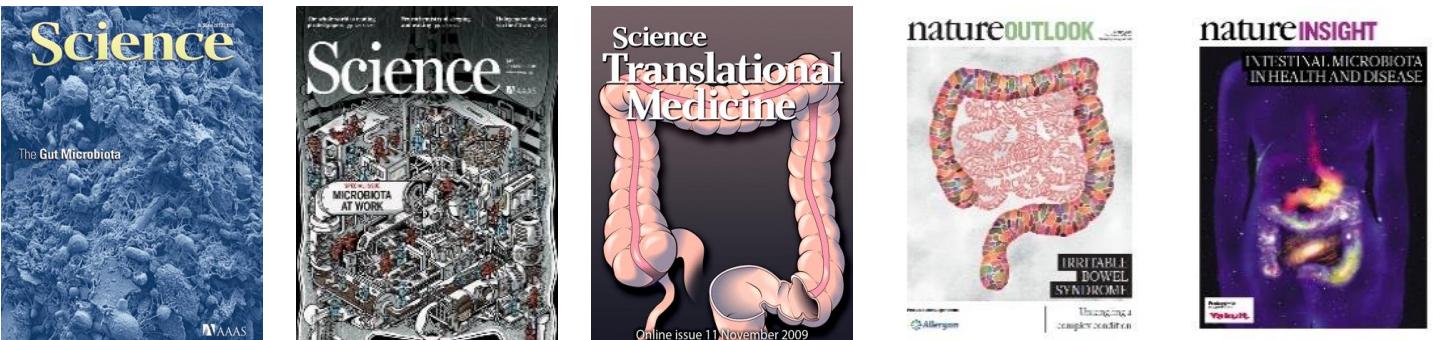
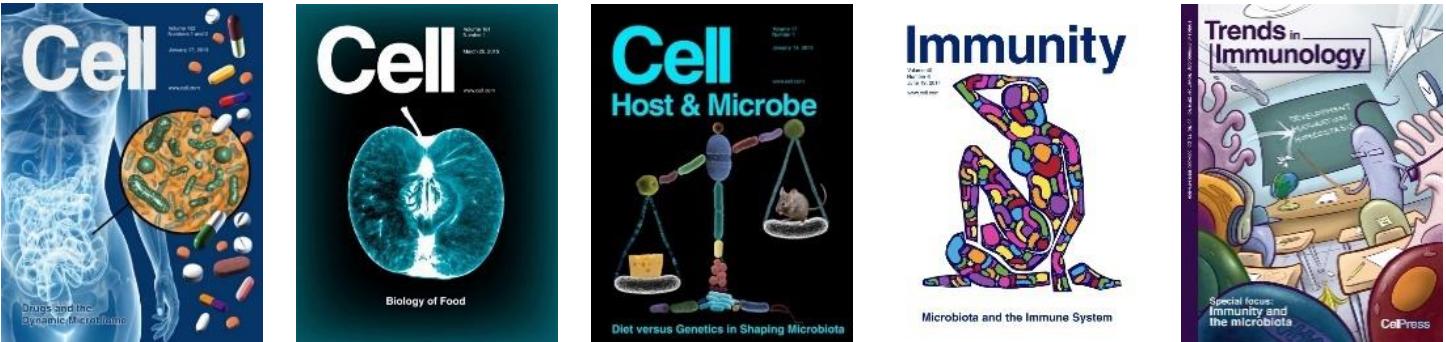
# 发表文章情况

- 本人合作在 **Science, Nature Biotechnology** 等杂志发表论文25篇，其中第一作者(含共同)论文12篇。应邀在 **Protein Cell, Current Opinion in Microbiology**、**遗传** 等杂志发表方法学综述文章。

杂志	发表时间	主要贡献	影响因子	引用	作者
<b>Nature Biotechnology</b>	2019.6 封面文章	发现微生物组在籼粳稻间差异和与氮利用效率的关系	<b>35.7</b>	49	共同一作(2)
<b>Science China Life Sciences</b>	2018.6 封面文章	挖掘水稻根系微生物组全生育期动态变化规律	<b>3.5</b>	31	共同一作(3)
<b>Protein Cell</b>	2020.5	微生物组分析方法	<b>7.6</b>	0	第一兼通讯
<b>Current Opinion in Microbiology</b>	2019.11	植物微生物组研究方法与应用	<b>6.9</b>	5	第一作者
<b>Science</b>	2019.5	参与解析三萜调控微生物组的规律	<b>41.0</b>	53	第三作者
<b>Nature Biotechnology</b>	2019.8	参与微生物组分析软件QIIME 2的开发与测试、并撰写十万字中文教程	<b>31.8</b>	447	参与作者

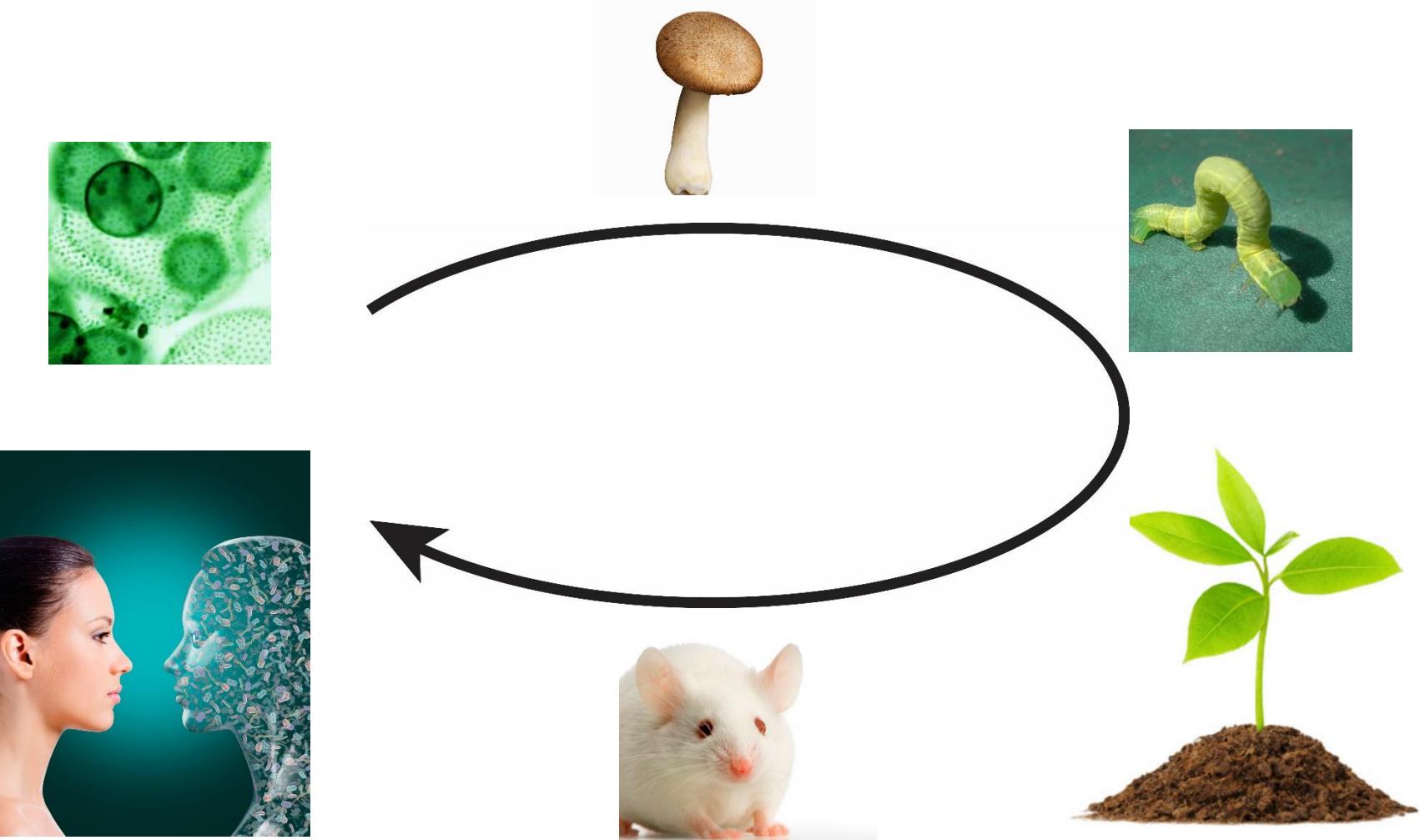


# 为什么微生物组这么热?





# 动植物生活环境充满微生物



脱离了微生物的生物学研究是不完整的

# 报告提纲

- 方法优化和开发
- 合作分析
- 科学传播



# 报告提纲

- 方法优化和开发
- 合作分析
- 科学传播





# 方法学的进展和存在问题

- 目前微生物组分析领域的处于高速发展阶段，方法和软件百花齐放、百家争鸣，很多方向软件形成三足鼎立，并没有出现一家独大；
- 对于绝大多数科研人员，开展分析中的主要难题不是没有软件可有，而是软件太多，不知如何选择，对软件方法的优缺点理解有限，结果准确性无法把握。

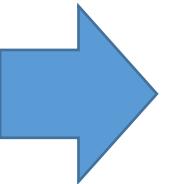
# 生物信息分析平台的搭建

## 分析代码

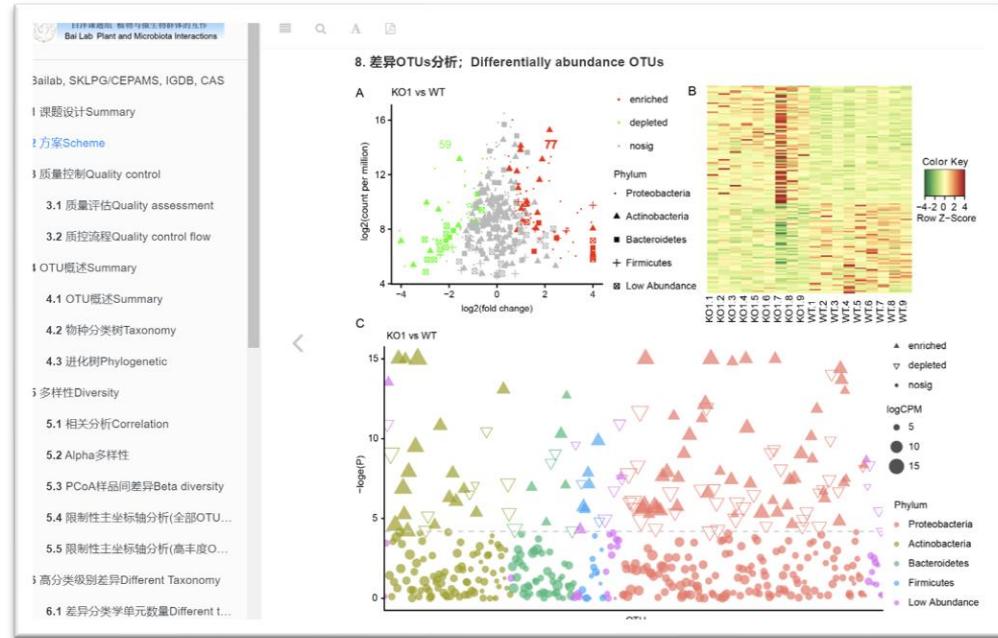
```
# Default parameter
xha=alpha.txt
a=beta
spare_group=compare_group.txt
sign=design.txt
execute='TRUE'
# otu=otu_table.css.txt
# FALSE # install package, default FALSE
ge_group='FALSE' # combine group1,group2 as new group name, not allow TRUE
output='result' # default work directory, remove low abundance < .1% and p_C
r_group='TRUE' # when pair_group have value turn TRUE
genotype='
list="WT", "DM1", "DM2", "D01", "D02"
batch=
list="1"
list="1"
up_order='FALSE' # order group by input list, not allow TRUE with merge_group
ect1='FALSE' # filter first group1 info
ect2='FALSE' # filter first group2 info
lth=5
lgt=3
lt_size=75
lblue=0.05
lch='FALSE' # Batch in shape
mutations=10000
alleles=9
安某一列批量CPCoA, 初始为一组, 不画
list="1"
"1"
function for script description and usage
ige()
ige()

ename: .. diversity.sh
<<EOF >&2>
ige:

```



## 结果报告



测试300+软件  
书写10000+行代码

30+种统计分析方法,  
标准分析周期仅需3天



# 参与主流微生物组分析平台开发



**QIIME** allows analysis of high-throughput community sequencing data

[JG Caporaso, J Kuczynski, J Stombaugh, K Bittinger... - Nature ..., 2010 - nature.com](#)

To the Editor: High-throughput sequencing is revolutionizing microbial ecology studies. Efforts like the Human Microbiome Projects 1 and the US National Ecological Observatory Network 2 are helping us to understand the role of microbial diversity in habitats within our ..

★ 99 [Cited by 20224](#) Related articles All 16 versions



Reproducible, interactive, scalable and extensible microbiome data science using  
**QIIME 2**

[E Bolyen, JR Rideout, MR Dillon, NA Bokulich... - Nature ..., 2019 - nature.com](#)

To the Editor—Rapid advances in DNA-sequencing and bioinformatics technologies in the past two decades have substantially improved understanding of the microbial world. This growing understanding relates to the vast diversity of microorganisms; how microbiota and ...

★ 99 [Cited by 443](#) Related articles All 15 versions

实现了可重复、可交互和可扩展的微生物组分析平台  
今年8月发表于**Nature Biotechnology** 杂志



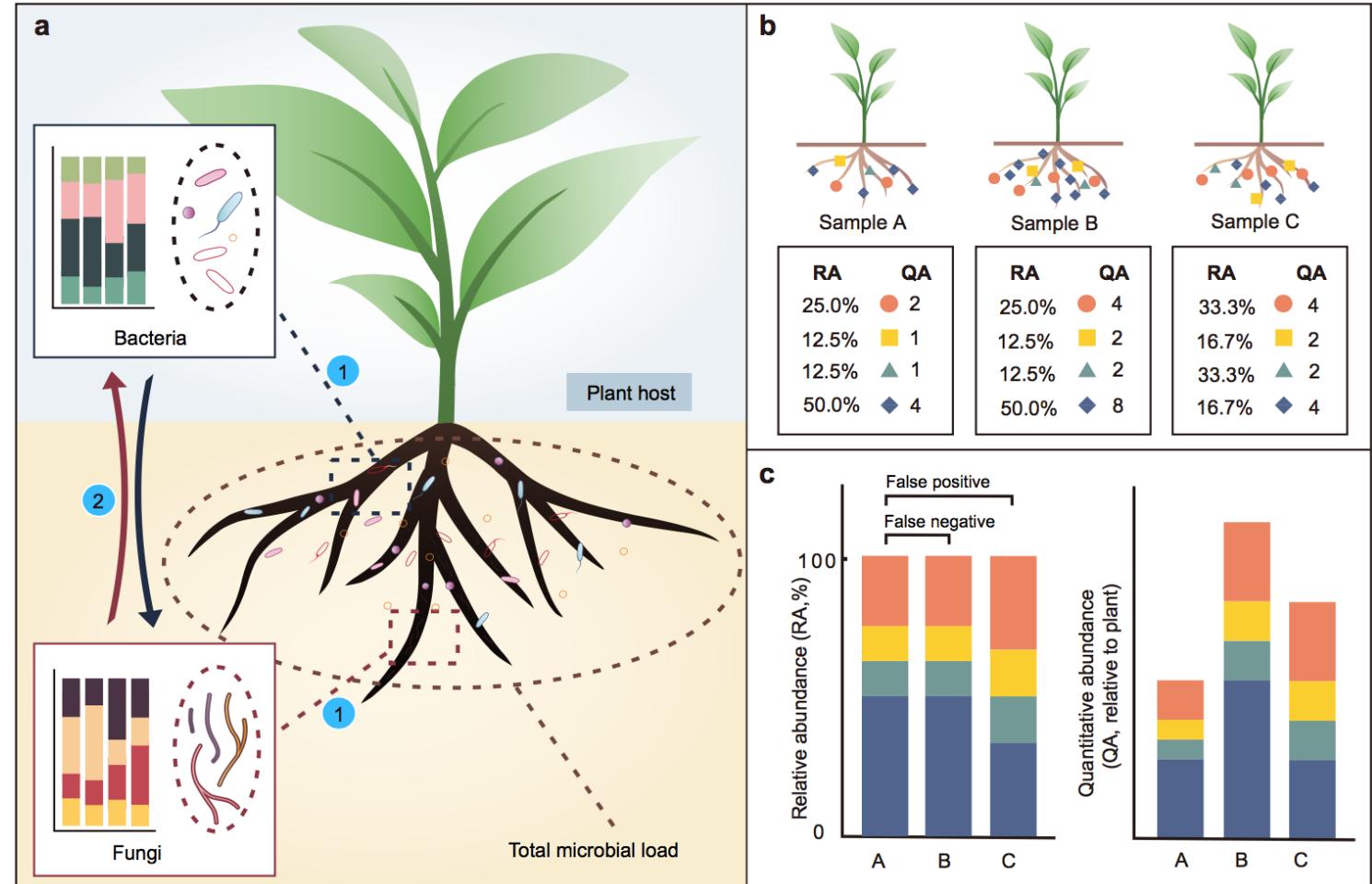
CEPAMS



# 植物扩增子定量方法开法

解决植物扩增子分析中  
只有相对丰度的问题

实现宿主植物、细菌、  
真菌三者间的定量，提  
供了绝对丰度的角度对  
微生物进行观察和比较





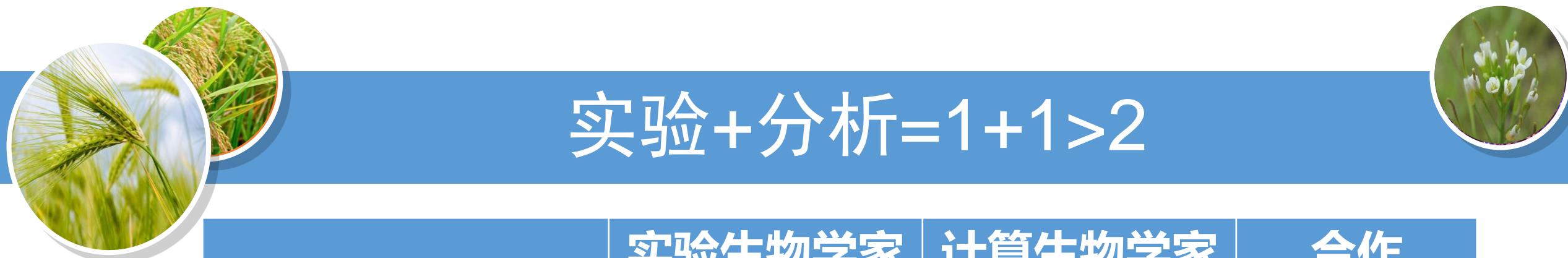
# 下一步工作重点

- 进一步优化和提高方法和软件的易用性
- 建立我国主导的、有影响力的分析流程
- 建立微生物组数据库的国内备份、更新站点
- 原创性方法的开发

# 报告提纲

- 方法优化和开发
- 合作分析
- 科学传播





# 实验+分析=1+1>2

	实验生物学家	计算生物学家	合作
丰富的实验材料	✓		✓
精湛的实验操作	✓		✓
强大的背景知识	✓		✓
数据分析的思路		✓	✓
方法的测试与比较		✓	✓
分析结果描述		✓	✓
生物学意义解读	✓	✓	✓ + ✓

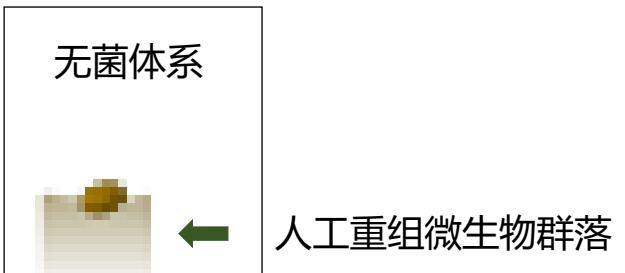
# 研究手段



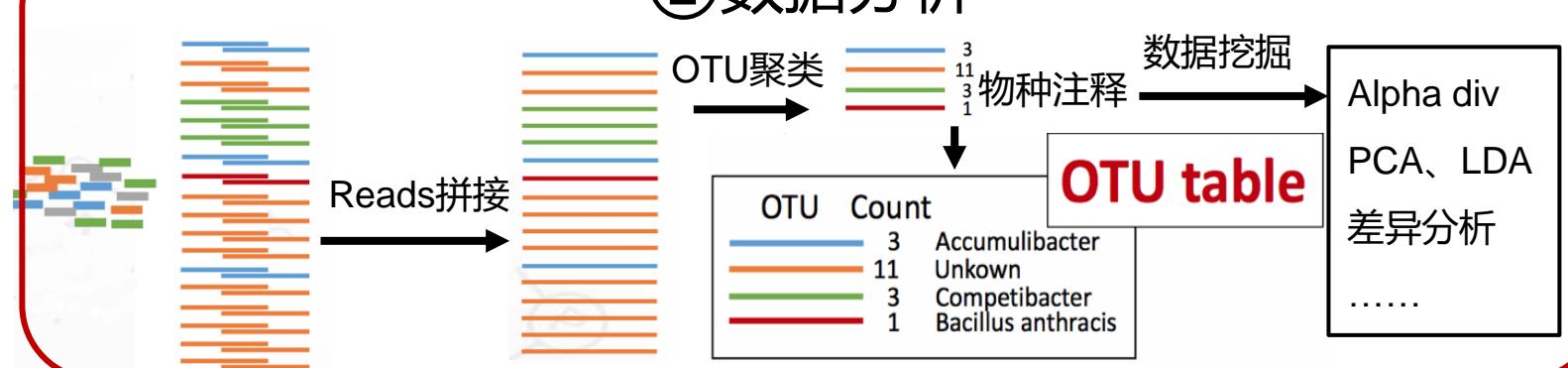
## ① 实验数据



## ③ 功能验证



## ② 数据分析



实验数据

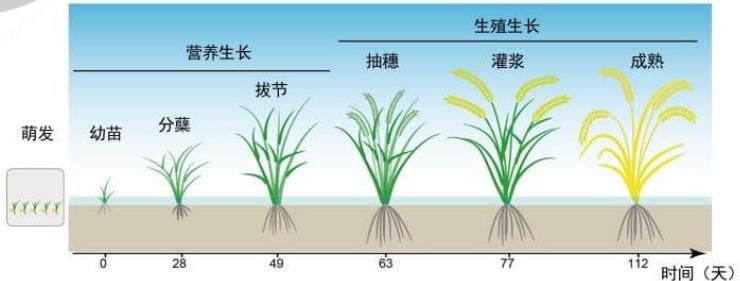
数据分析

功能验证

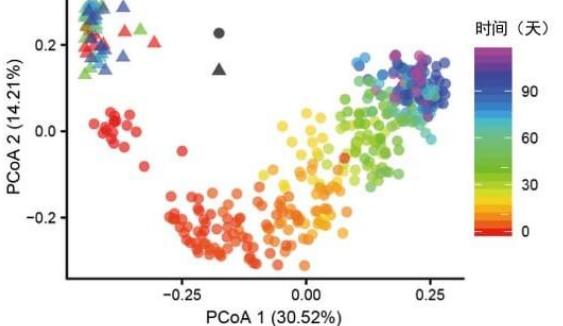
# 探索水稻全生育期微生物组



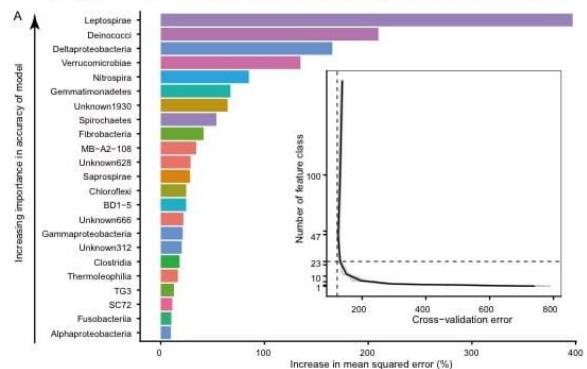
水稻田间全生育期示意图



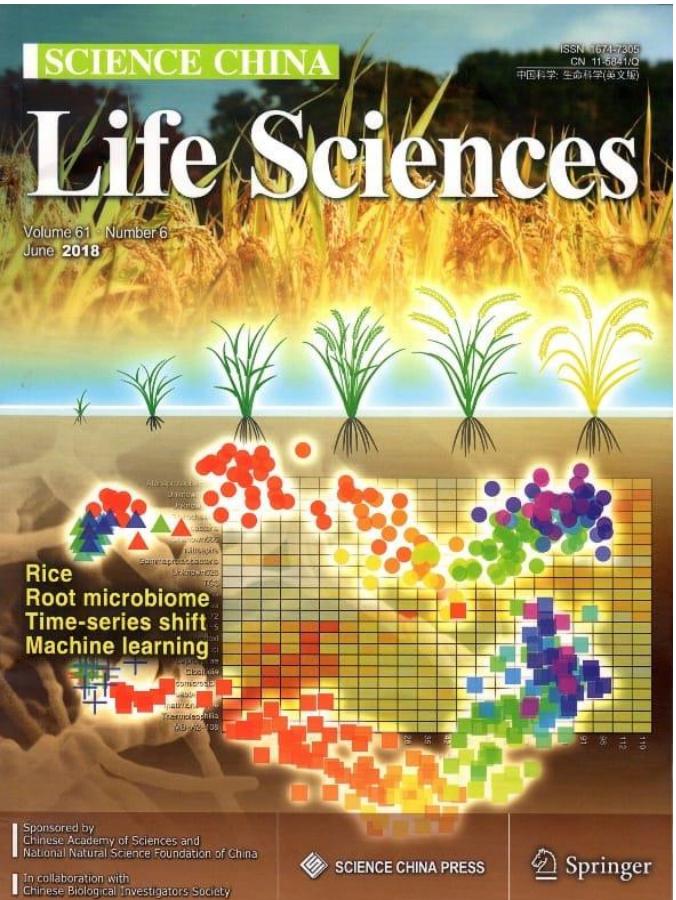
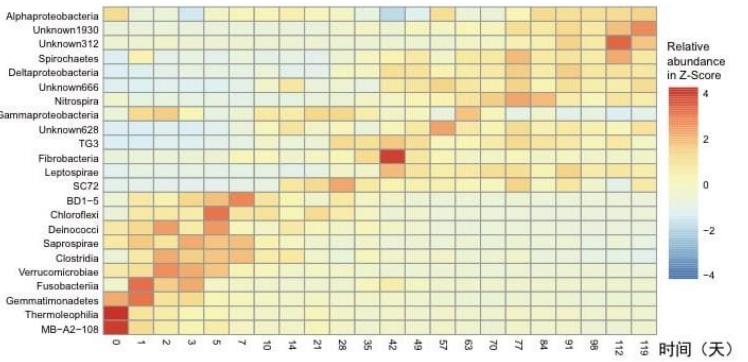
水稻进入生殖生长期后  
根系微生物组开始稳定



通过机器学习的方法鉴定  
水稻生长发育相关的根系代表性细菌



根系代表性细菌随水稻生长的变化规律



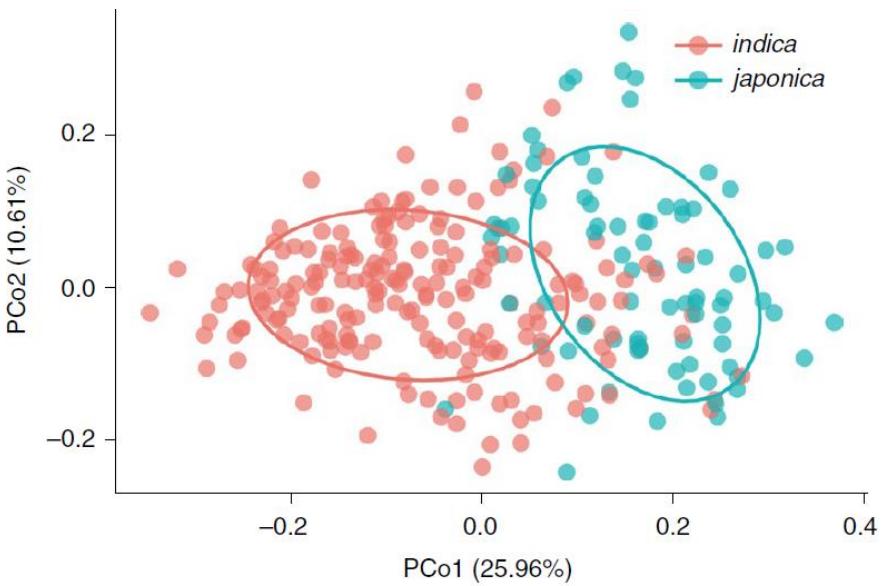
为农业益生菌的施用时间提供指导和理论基础

Jingying Zhang#, Na Zhang#, Yong-Xin Liu#, Xiaoning Zhang, Bin Hu, Yuan Qin, Haoran Xu, Hui Wang, Xiaoxuan Guo, Jingmei Qian, Wei Wang, Pengfan Zhang, Tao Jin, Chengcai Chu\* & Yang Bai\*. Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Science China Life Sciences*. 2018, 61: 613-621. doi:10.1007/s11427-018-9284-4

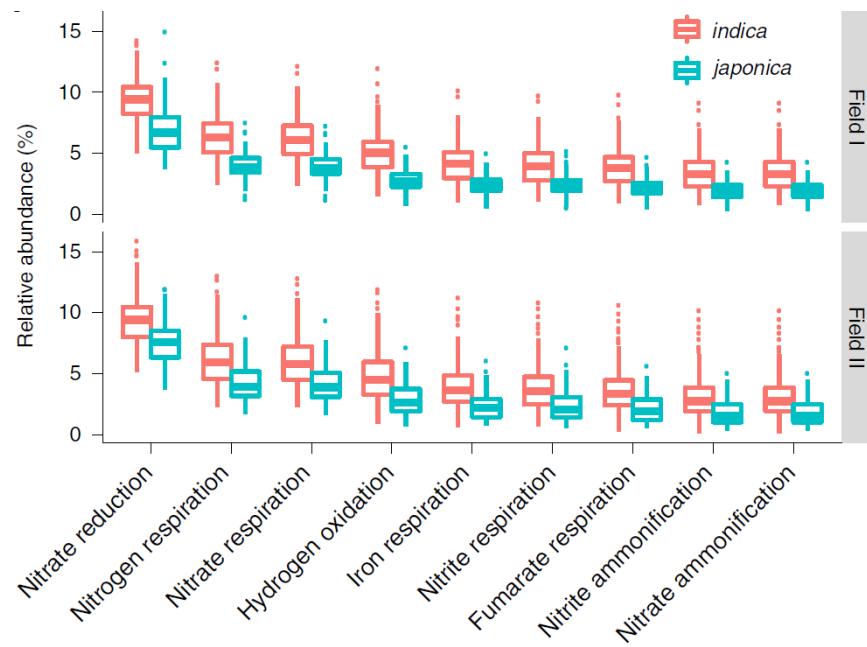


# 解析根系微生物组在氮利用中的作用

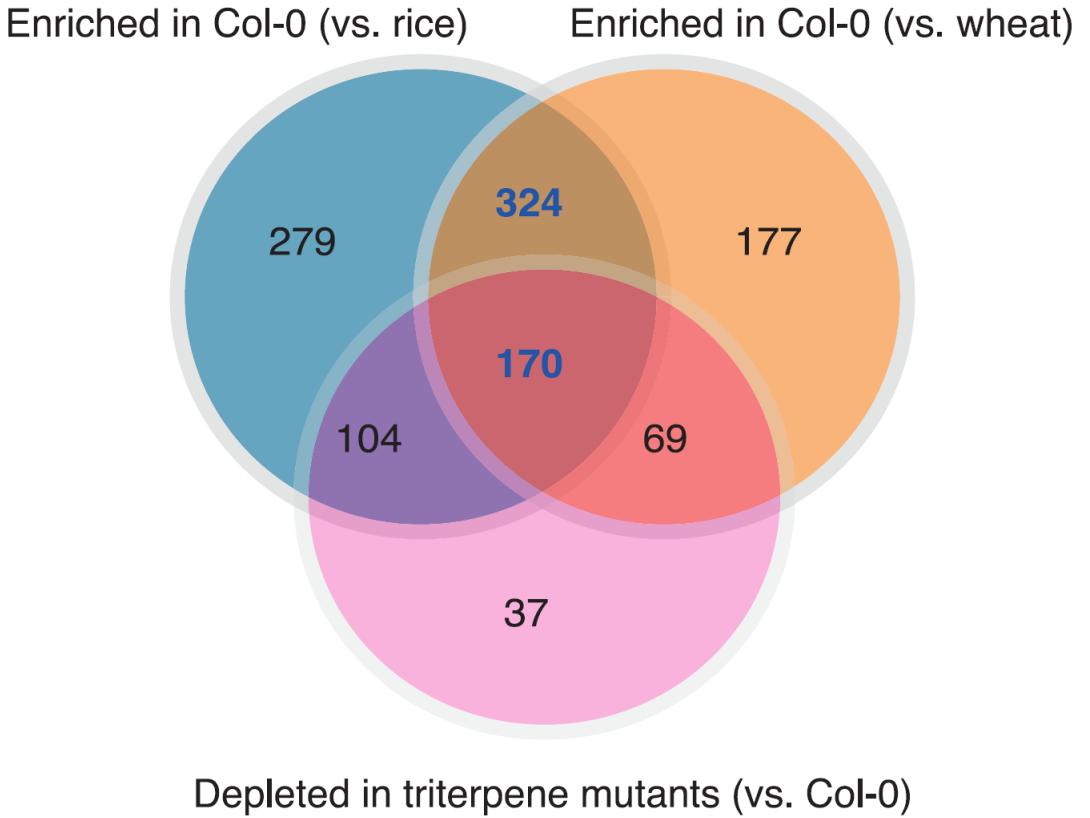
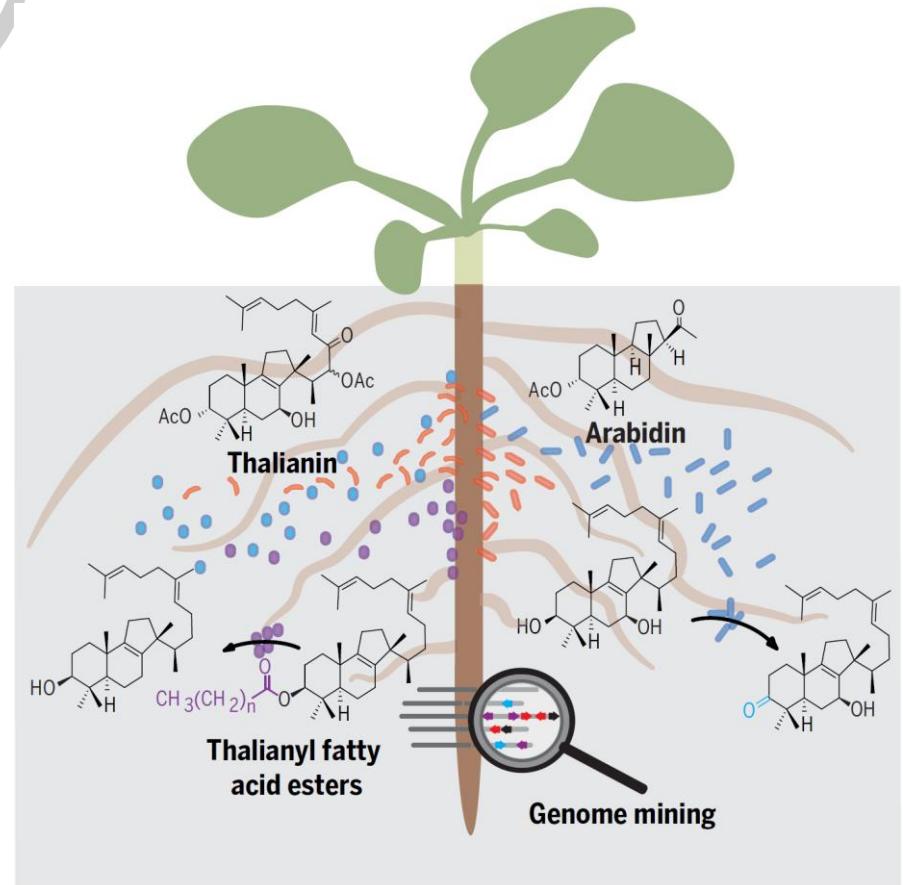
籼粳稻根系形成不同的微生物组



籼稻比粳稻富集更多  
氮循环相关细菌



为农业益生菌在氮肥高效利用中的应用提供了理论和材料基础



二半萜和三萜类化合物塑造拟南芥根系特异的微生物组结构



# 开展合作的主要问题和对策

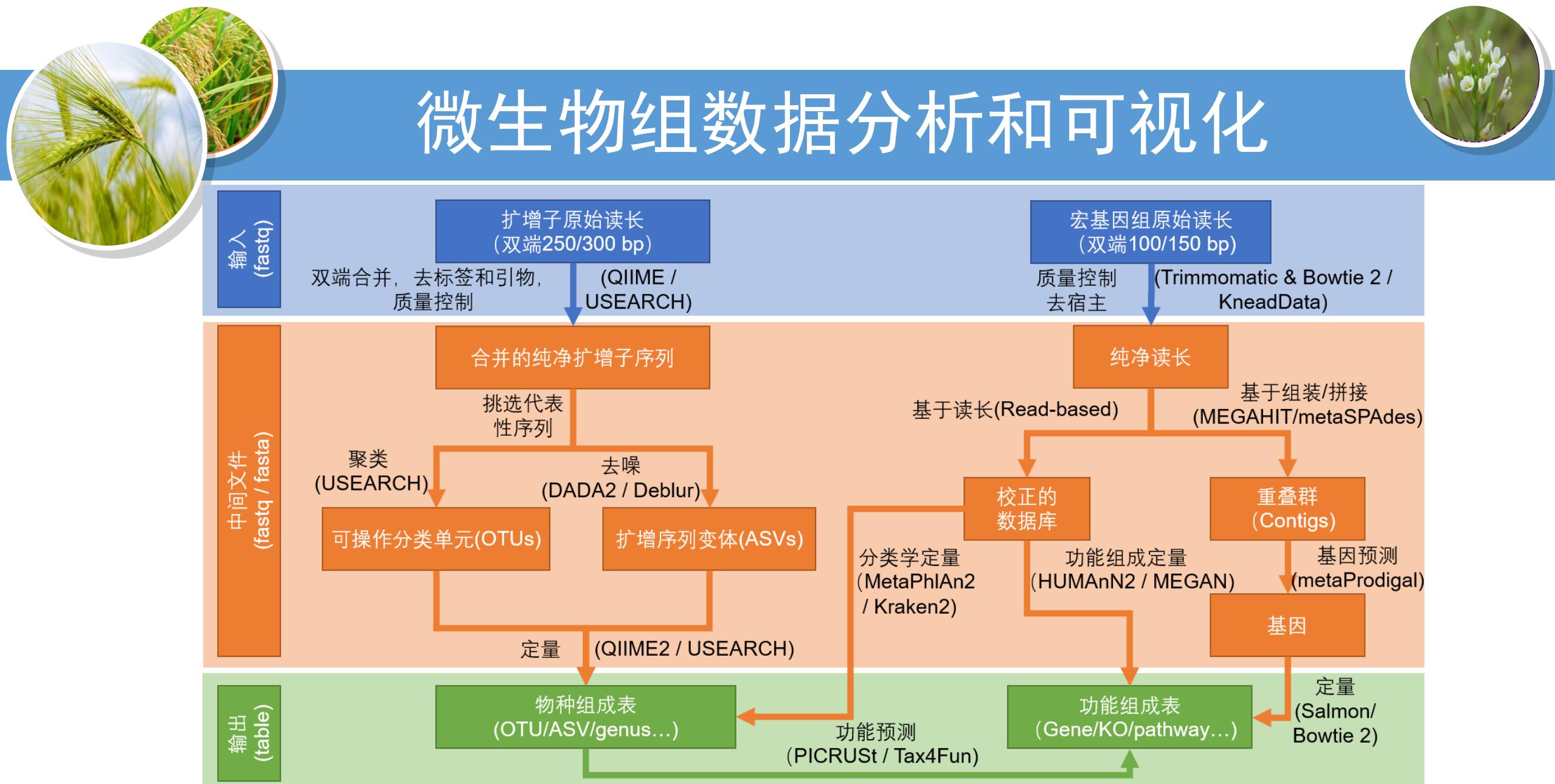
- 沟通：通常实验和分析人员间缺少了解和沟通，需要课题背景和需求清晰，执行前要多沟通
- 耐心：真正的合作是几个月到几年的长期磨合，切忌急功近利、无背景基础、公司式标准化分析离生物学规律相差甚远
- 利益：不同评价体系对共同作者认可程度不同，根据规则设定目标并根据最终贡献谈好利益分配问题
- 舍得：有舍才有得，否则很可能开展合作实现优势互补，达到突破自己的天花板，冲击国际领先或开创交叉新领域的目标

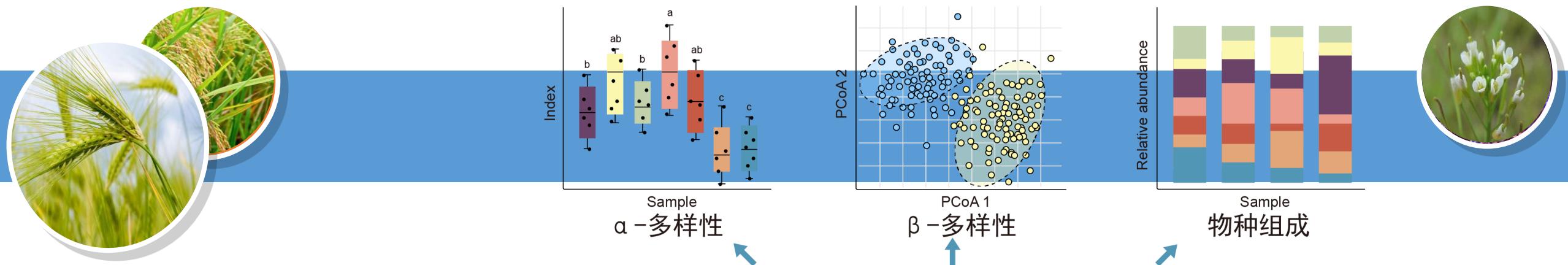
# 报告提纲

- 方法优化和开发
- 合作分析
- 科学传播

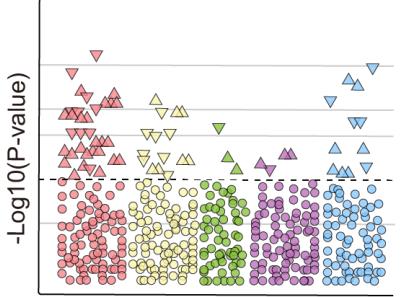


# 微生物组数据分析和可视化





# 基于特征表的常见统计分析方法 可视化样式

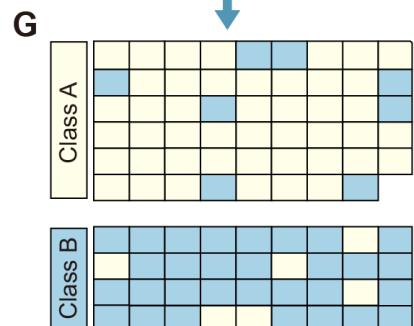
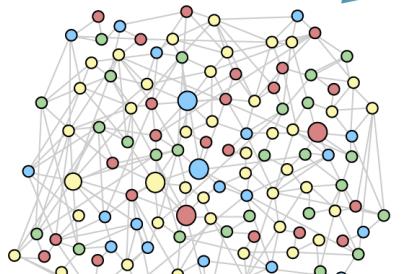
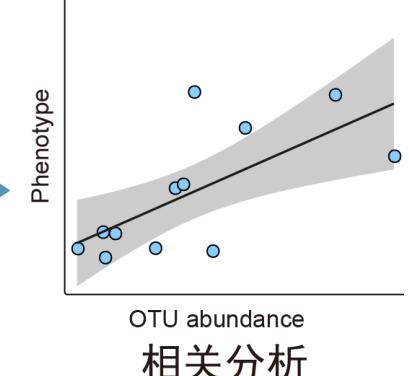


物种组成表

	Sample ID			
	S1	S2	S3	S4
OTU_1				
OTU_2				
OTU_3				
OTU_4				
OTU_5				
...				
OTU_n				

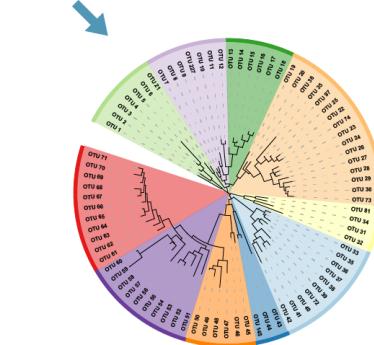
功能组成表

	Sample ID			
	S1	S2	S3	S4
KO_01				
KO_02				
KO_03				
KO_04				
KO_05				
...				
KO_n				



网络分析

机器学习-分类



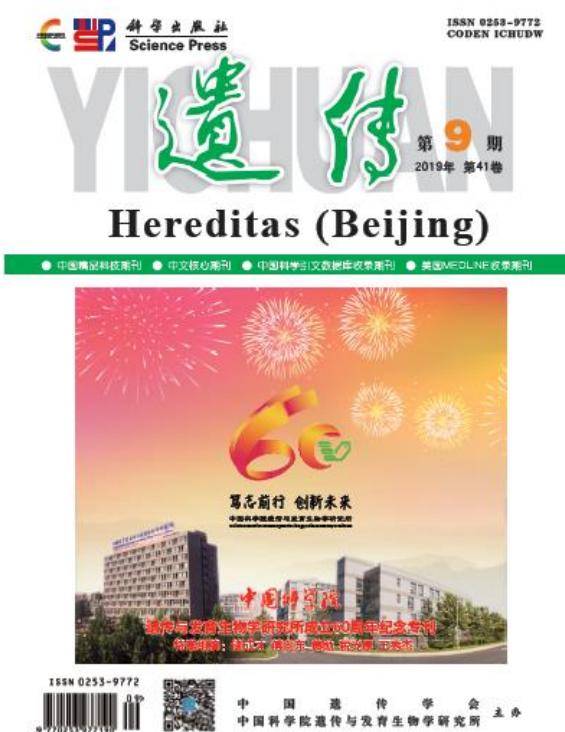
系统发育树

Yong-Xin Liu, Yuan Qin, Tong Chen, Meiping Lu, Xubo Qian, Xiaoxuan Guo & Yang Bai. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* 41, 1-16, doi: <https://doi.org/10.1007/s13238-020-00724-8> (2020).



# 微生物组领域研究方法综述中文版

- 负责撰写《遗传》杂志所庆60周年专刊特邀综述一篇



当前位置: 三年内

选择:  下载引用

 显示/隐藏图片

微生物组数据分析方法与应用 

刘永鑫,秦媛,郭晓璇,白洋

遗传 2019, 41 (9): 845-862. DOI: 10.16288/j.yczz.19-222

录用日期: 2019-09-02

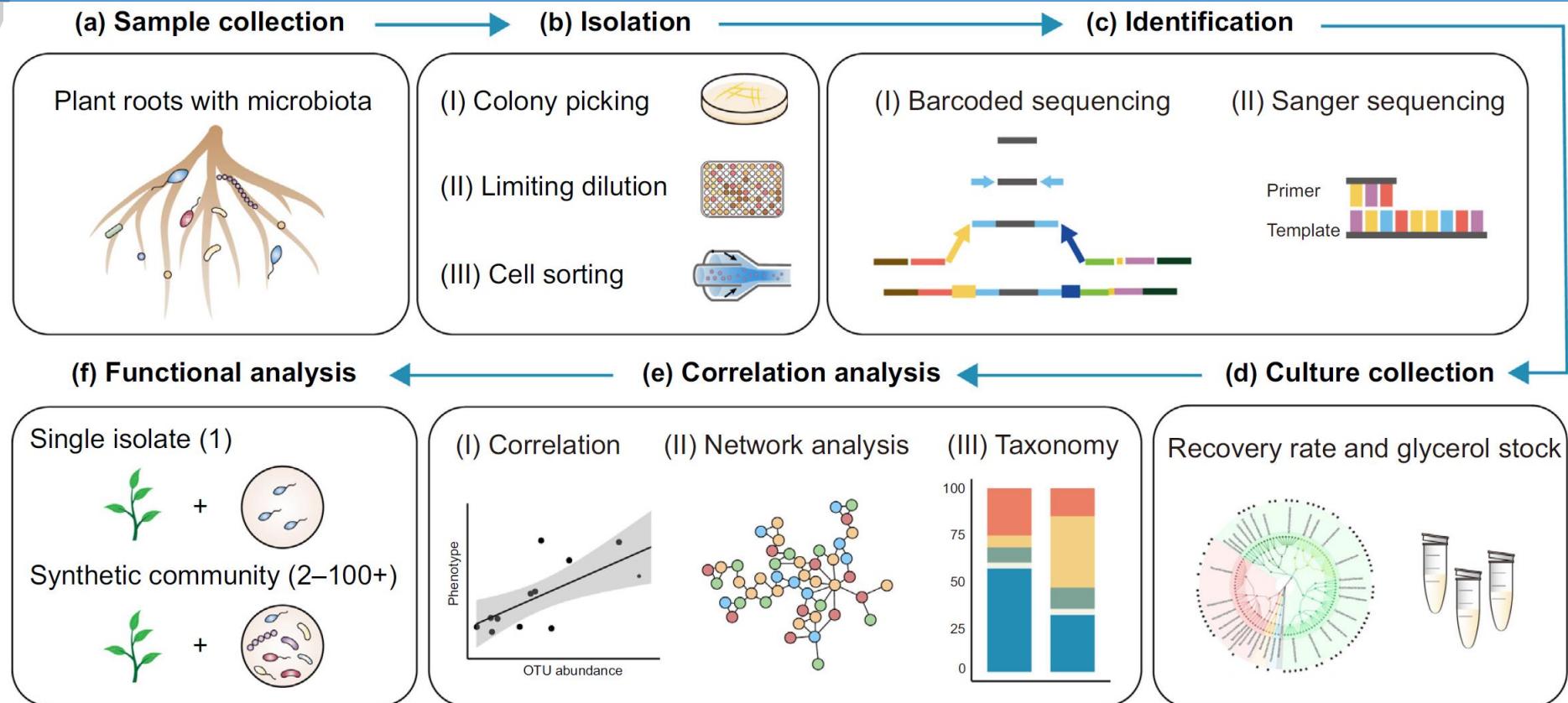
 摘要 (3706)  HTML (117)  PDF (557KB) (3843)



遗传: 微生物组数据分析方法与应用



# 微生物组功能研究方法综述



应邀为微生物组领域权威杂志 ***Current Opinion in Microbiology*** 撰写综述  
推动研究从描述性阶段向功能研究发展



# JoVE微生物组专刊征稿

Search 12,213 video articles [Advanced](#) [Log In](#)

Free access to all JoVE Education videos until June 15. Click [here](#).

REMOTE LEARNING NEW COVID-19 NEW BIOPHARMA FOR LIBRARIANS PUBLISH VIDEO JOURNAL SCIENCE EDUCATION

ACCELERATING SCIENTIFIC RESEARCH & EDUCATION

## We change the way science is done.

NOW SHOWING

Optimization for Sequencing and Analysis of Degraded FFPE-RNA Samples



CEPAMS

[JoVE微生物组专刊征稿，写方法拍视频教程发SCI](#)

# JoVE微生物组专刊征稿



## Analysis and visualization methods for amplicon and metagenomic data of microbiome

Advances in high-throughput sequencing have fostered developments in microbiome research, and massive microbiome datasets have been generated. Microbi...

### GUEST EDITORS

Taxonomy  
Differences comparison



**Yong-Xin Liu**



**Tong Chen**



**Maosheng  
Zheng**



<https://www.jove.com/methods-collections/426/analysis-visualization-methods-for-amplicon-metagenomic-data>

[JoVE微生物组专刊征稿，写方法拍视频教程发SCI](#)



# 《微生物组数据分析与可视化实战》专著

- 众筹编写《微生物组数据分析与可视化实战》——成为宏基因组学百科全书的创始人(目录)**
- 编者序：初衷、计划、要求、优势、目标和展望**
- 内容格式要求：背景知识、实例解读、实战代码；帮助同行看得懂文章，会归纳结果和写文章，实现自主分析；**
- 计划：1年发行中文专著和建立在线百科全书平台，优秀稿件推荐发表中、英文方法期刊；2年建立我国自主、有影响力的微生物组分析平台；3年走出去，发表国际有影响力的方法文章和出版外文图书**

编号	姓名	单位	研究方向	职称/年	创作/参与章节			
1	刘永鑫	中科院遗传发育所	微生物组数据分析	工程师	扩增子/宏基因组分析流程, Beta多样性, R语言应用			
2	文涛	南京农业大学	微生物组数据分析	博4	可视化章节正文, R包amplicon			
3	钱旭波	浙江大学	儿童风湿与肠道菌	主任医师	基本概念、统计、英文版发行			
4	吴一磊	中科院微生物所	环境微生物16S测序	科研助理	ggplot2可视化？R统计分析基础			
5	徐俊	北京大学	肠道微生物与肠道菌	助理研究员	真菌组18S/ITS			
6	蔡伟伟	北京交通大学	环境微生物	讲师	微生物网络分析			
7	查富蓉	华中农业大学/凌恩	微生物组数据分析	硕士	宏基因组抗性基因/整合子分析, R语言应用			
8	常帆	陕西省微生物研究所	生物废弃物综合利用	助理研究员	数据库和数据上传			
9	陈亮	中科院微生物研究所	三代测序与肠道微生物	助理研究员	network分析与真菌组			
10	邓子祺	西班牙植物生物技术研究所	系统发育工具开发	博1	利用ETE构建、绘制和分析系统发育树			
11	付先恒	中国科学院大学	微生物生态与土壤养分循环	博3	ggplot2可视化及ASV表后续统计分析			
12	葛德燕	中国科学院动物研究所	动物进化与系统学	副研究员	野生动物系统发育重建与胃肠道宏基因组			
13	何芳章	安徽医科大学	肠道宏基因组与代谢	讲师	普氏分析，相关分析的可视化，随堂练习			
14	李鸿毅	浙江大学/某测序公司	土壤微生物生态/扩增子	博士	lefse结合原代码进行各参数解读			
15	李辉	中山大学中山医学院	医学微生物学	博2	mothur; Community type			
16	李杰	常熟理工	三代测序和冠状病毒	讲师	三代测序Nanopore分析冠状病毒			
17	李金优	浙江大学	肠道衰老与肠道微生物	实验员	图片排版和美化, R统计和绘图			
18	李瑞琳	中国科学院计算机网络信息中心	高性能计算与生物信息学	助理研究员	宏基因组数据挖掘算法与分析流程			
19	李苏梅	深圳市人民医院	临床药理病理	副研究员	功能注释数据库			
20	李文耀	新加坡国立大学	环境微生物抗药性	博后	文章套路总结-抗生素抗性			
21	李延明	堪萨斯大学医学中心	机器学习, 大数据分析	助理教授	机器学习的常用算法			
22	李雨泽	西北农林科技大学	农田微生物生态	硕2	co-occurrence网络分析			
23	林禾雨	墨尔本大学	环境微生物多组学	博3	分箱专题			
24	刘华	中国科学技术大学生命医学/微生物学	博3	非限制性排序, ggplot2扩展包				
25	刘云	吉林大学	宏基因组binning算法	讲师	分箱专题			
26	吕波	湖南师范大学	胁迫对动物肠道菌群	硕1	R可视化、宏转录组			
27	毛杰	南京农业大学	WES数据分析	硕士	R统计与绘图, ggplot2可视化			
28	孙江伟	瑞典卡罗林斯卡医学院	宏基因组学分析	博3	统计基础			
29	谭乔文	同济大学	饮用水微生物组	硕3	分类树构建/机器学习的常用算法			
30	王敬敬	中国科学院天津工业生物所	微生物肥料, 根际微生物	副研究员	Gephi绘制网络图, CANOCO进行R			
31	王晓雯	凌波微课			测序技术在微生物组	中级	工程文章写作套路与实验设计	
32	王子叶	复旦大学			宏基因组学binning	博3	分箱专题	
33	魏勇军	郑州大学			微生物组学、合成生物学	副教授	Dix-seq: 16S rRNA扩增子整合分析	
34	巫您	华大基因			法医基因组学		研发工程宏基因组模块? ggplot2等绘图基	
35	吴季秋	华西生物医药大数据中心			噬菌体, 代谢, 肠道		数据分析人体肠道病毒组, 婴幼儿肠道动态	
36	吴俊	华东师范大学			宏基因组数据分析	副研究员	统计基础	
37	肖劲洲	上海海洋大学			宏病毒组	博6	宏病毒组, mothur Miseq SOP	
38	徐锐	广东省生态环境厅			土壤微生物生态	助理研究员	一步到位的R精简版网络分析方法/	
39	许亚东	西北农林科技大学			凋落物分解微生态	博3	特征与环境因子相关	
40	余珂	北京大学			宏基因组学binning	助理教授	分箱专题: 从数据组装、自动化bi	
41	袁峥嵘	北京林业大学			微生物组数据分析	副教授	物种多样性分析	
42	张麟	杭州师范大学			动物生理生态与进化	助理研究员	统计基础	
43	赵丹阳	中国药科大学			肠肝轴、酒精肝与非酒精肝	硕2	qiime2? STAMP与扩展柱状图?	
44	赵圣国	中国农业科学院北京动物营养与饲料研究所			功能基因（微生物）多样性数据分			
45	赵向阳	南京农业大学			微生物组	硕3	数据爬取及分析	
46	赵泽龙	大连理工大学			微生物组数据分析	博6	beta多样性	
47	郑茂盛	华北电力大学			污染处理	教授	LEfSe分析文章准备、本地在线分	
48	郑鹏飞	中国科学院深海科学与工程研究所			深海微生物组学	博后	网络分析, 相关分析	
49	郑伟	西北农林科技大学			果园微生物与土壤健康	副教授	tax4fun教程, circles圈图教程,	
50	周欣	中国科学院微生物研究所			根际微生物组学与病害	博士后	进化树分析及iTOL可视化、FUNGuild	
51	朱成凯	复旦大学			遗传学/肠道微生物	博2	使用过很多宏基因组软件, 可以写	
52	吴佩渝	华东师范大学			休学		本科毕业数据分析的计算机环境配置, 以Li	
53	张云增	扬州大学			微生物组		特聘教授宏基因组、宏转录组	
54	李苏梅	深圳市人民医院			临床药理病理	副研究员	Alpha多样性	
55	席娇	西北农林科技大学			微生物生态与植物抗性	博2	微生物生态与植物抗性研究	
56	石峻峰	中国科学院海洋研究所			原生动物多样性	博4	海洋原生动物多样性	
57	付先恒	中国科学院大学			微生物生态及土壤养分循环	博3		
58	李青霞	中国科学院海洋研究所			海洋原生动物分子生物学	博3		
59	马腾飞	南京农业大学			固氮	博4	植物微生物组研究方法	
60	王超	中科院遗传发育所			植物微生物组	博后	宿主与微生物关联	
61	刘宁华	山东大学			海洋微生物	博1	海洋微生物方向（基因组, 分类鉴	

211.Alpha多样性箱线图(样章, 11图2视频)



# 科学传播的问题和解决方案

- 文章分析描述缺少细节无法重现、文章多如汪洋大海、英文资料不少但缺少系统整理和中文资料、缺少实例和经验分享等问题？
- 解决方案：
- 倡导文章分享可重复代码，建立微生物组百科全书中、英文版，提高可信度、国际影响力和信誉；
- 成果新闻、新媒体综合解读宣传，让同行了解，提高全民科学素养；
- 翻译或分享较新的前沿资料，进一步总结综述并同步翻译传播；
- 利用“互联网+”模式，分析实例、经验总结通过博客让同行检索，公众号推送让别人知道，在线直播和演讲的方式让大家有机学方便学习前沿知识、技术和经验。

# 总结



中国微生物组研究	现状	对策
方法优化和开发	软件多 易生性差 国内原创少	评估+排名+点评 优化适合非专业人士 与数学、计算机联合开发
合作分析	缺少沟通 急功近得 心态定位偏差	掌握必要的对方知识 实验和分析都是长期科研 换位思考、尊重他人贡献
科学传播	中文资料少 成果被埋没 报告讲座少 问题无处问	共享经验和笔记 解读、宣传和后续更新维护 视频直播+永久访问 建立活跃的同行社区和交流群



# 解决本领域中文资料匮乏的困境 ——宏基因组公众号



- 写并分享笔记，发布于宏基因组公众号
- 3年——1900篇原创文章，300万字
- 扩增子系列图表解读、分析流程和绘图教程21篇
- QIIME2中文教程32篇
- 宏基因组分析流程30篇
- 解读翻译文章综述500多篇
- 9万+同行关注，1400万+阅读量





谢谢各位老师！

