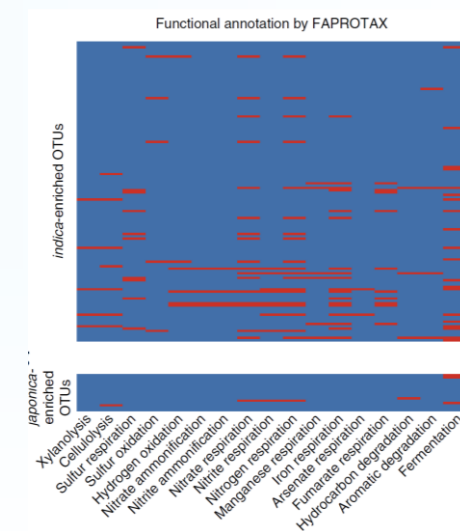
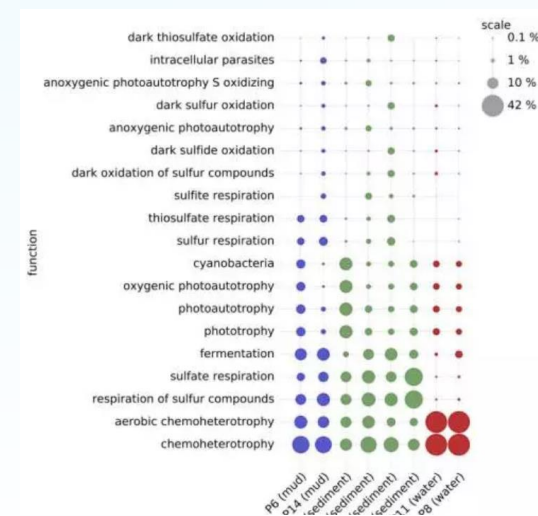


# 2022微生物组—扩增子16S分析研讨会第14期



## 31元素循环FAPROTAX

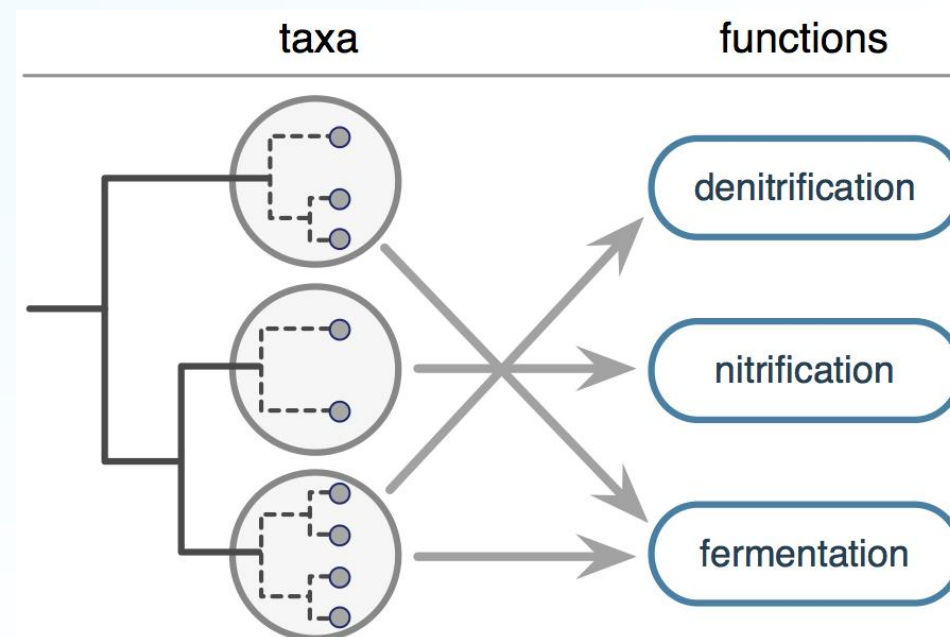
易生信  
2022年1月9日



易生信，毕生缘；培训版权所有。

# 原核物种功能注释 Functional Annotation of Prokaryotic Taxa (FAPROTAX)

- FAPROTAX是根据文献手动构建了物种分类与功能注释的数据库；编写了联系OTU分类表与数据库的python脚本；我们只要将基于16S的OTU分类表，使用python脚本，就可以获得微生物群落的功能注释预测结果。
- FAPROTAX较适用于对环境样本（如海洋、湖泊等）的生物地球化学循环过程（特别是碳、氢、氮、磷、硫等元素循环）进行功能注释预测。因其基于已发表验证的可培养菌文献，其预测准确度较好，但覆盖度较低。



# FAPROTAX结果什么样

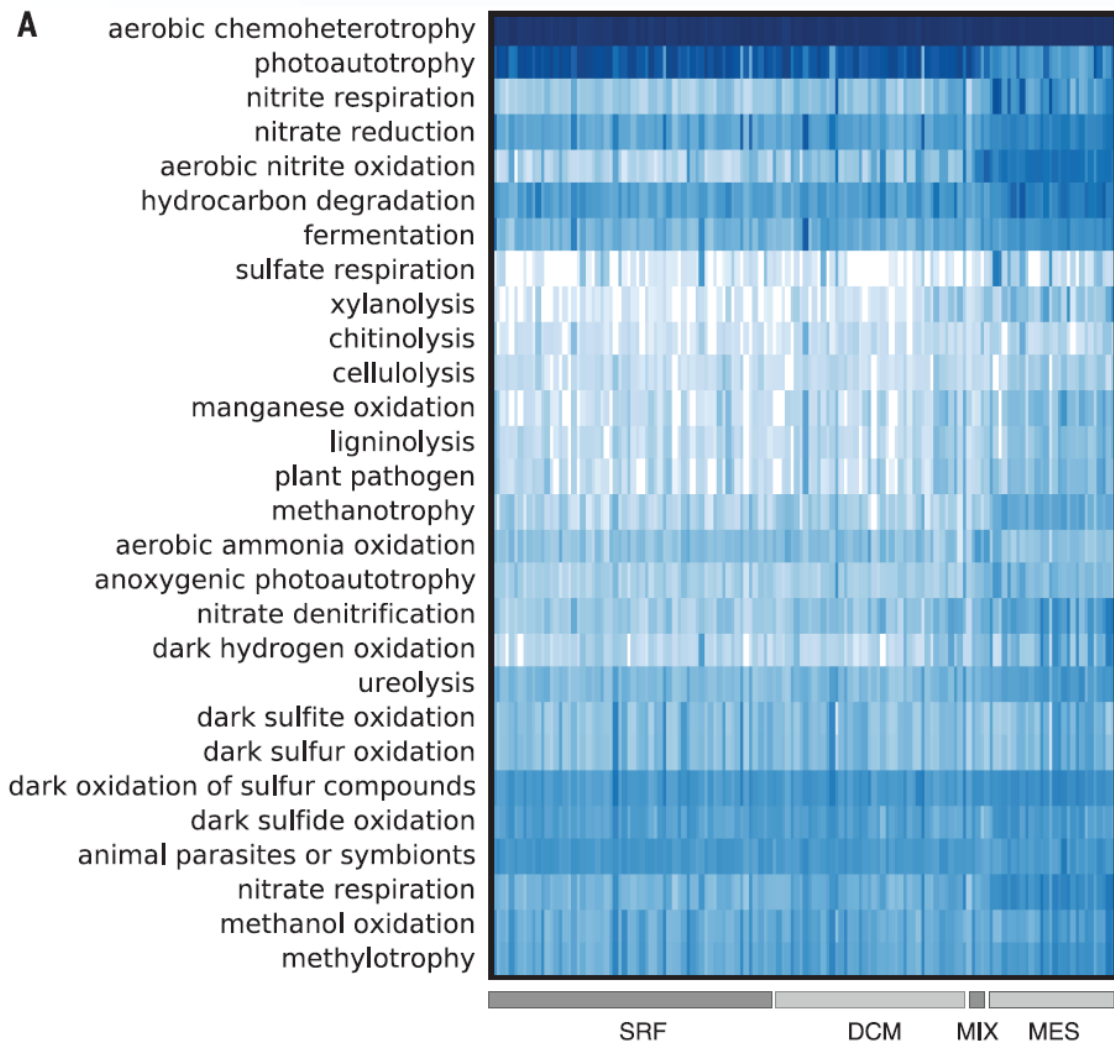
group	K01	K02	K03	K04	K05	K06	OE1	OE2	OE3	OE4
methanotrophy	0	0	0	0	0	0	0	0	0	0
acetoclastic_methanogenesis			0	0	0	0	0	0	0	0
methanogenesis_by_disproportionation_of_methyl_groups							0	0	0	0
methanogenesis_using_formate			0	0	0	0	0	0	0	0
methanogenesis_by_CO2_reduction_with_H2				0	0	0	0	0	0	0
methanogenesis_by_reduction_of_methyl_compounds_with_H2							0	0	0	0
hydrogenotrophic_methanogenesis			0	0	0	0	0	0	0	0
methanogenesis	0	0	0	0	0	0	0	0	0	0
methanol_oxidation		77	105	49	54	134	156	58	123	
methylotrophy	77	105	49	54	134	156	58	123	67	
aerobic_ammonia_oxidation			6	5	0	4	0	1	6	
aerobic_nitrite_oxidation			28	27	10	17	13	20	43	
nitrification	34	32	10	21	13	21	49	26	35	
sulfate_respiration		0	0	0	0	0	0	0	0	

faprotax.txt

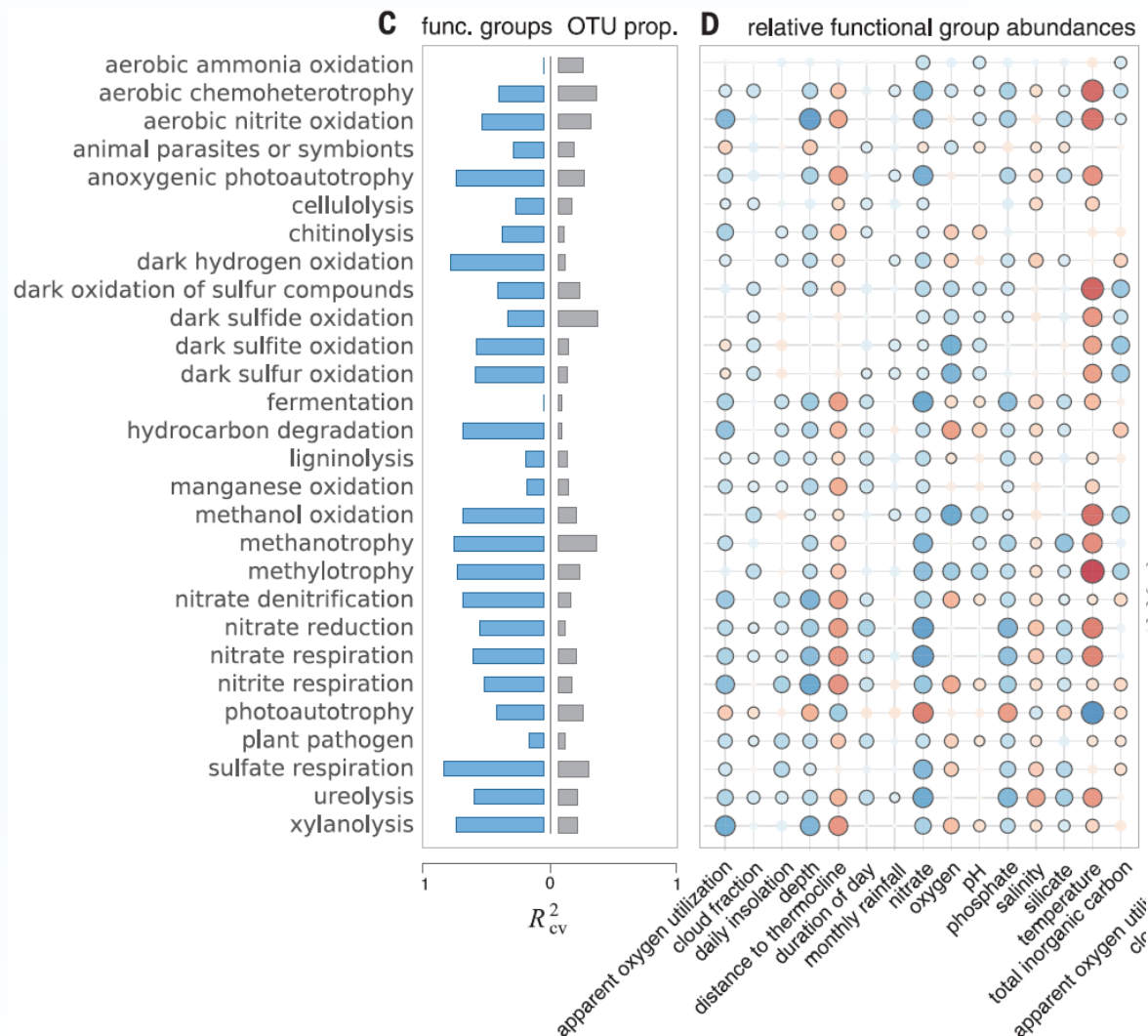
与OTU表、KO表类似，为功能描述与样本对应的矩阵



# FAPROTAX结果有什么应用



样本分组对应功能丰度热图 (Science图2A)



环境因子与功能相关性泡泡图 (Science图1D)

[Louca, S., Parfrey, L.W., Doebeli, M. \(2016\)](#)

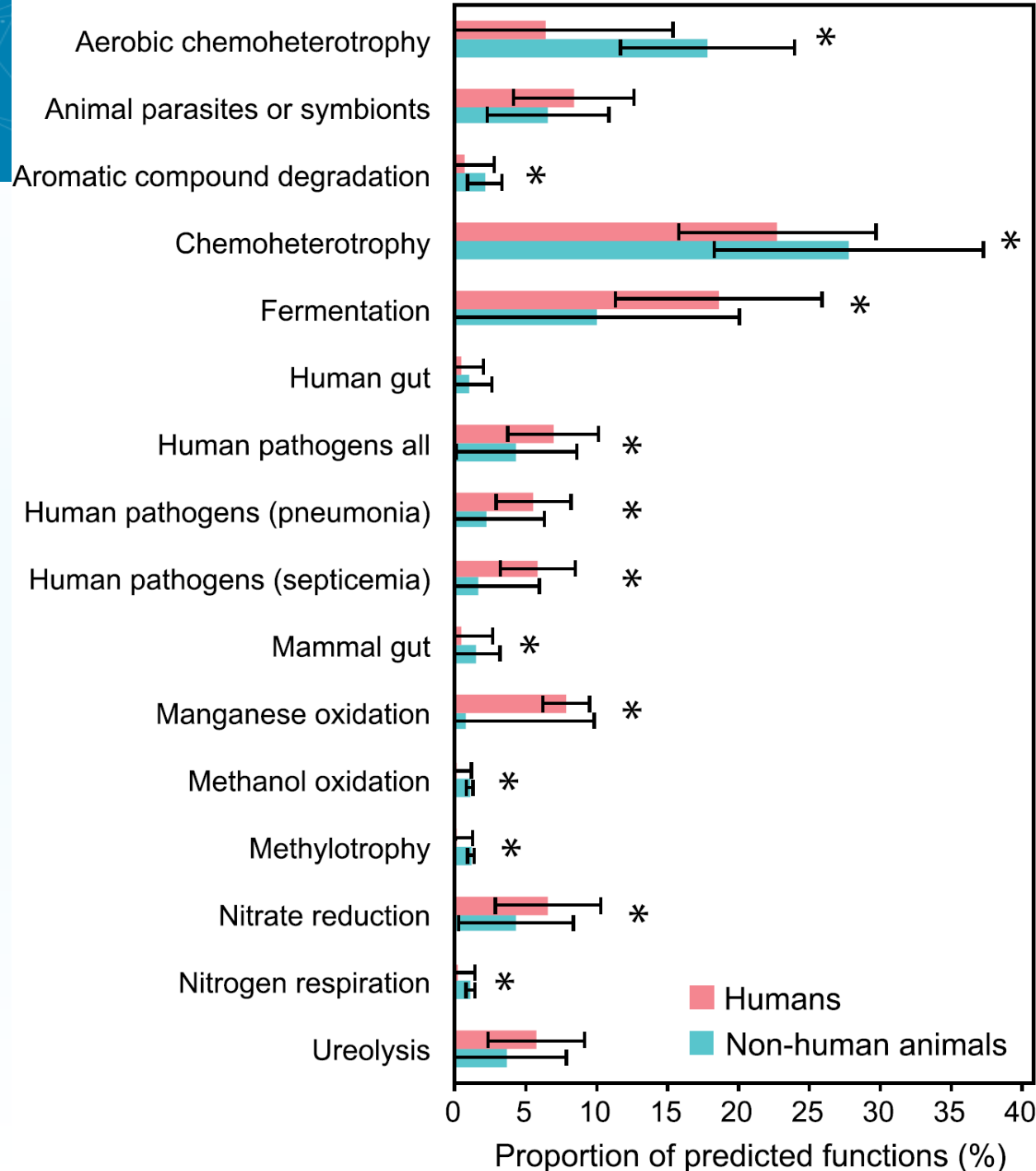
[Decoupling function and taxonomy in the global ocean microbiome. Science 353:1272-1277](#)



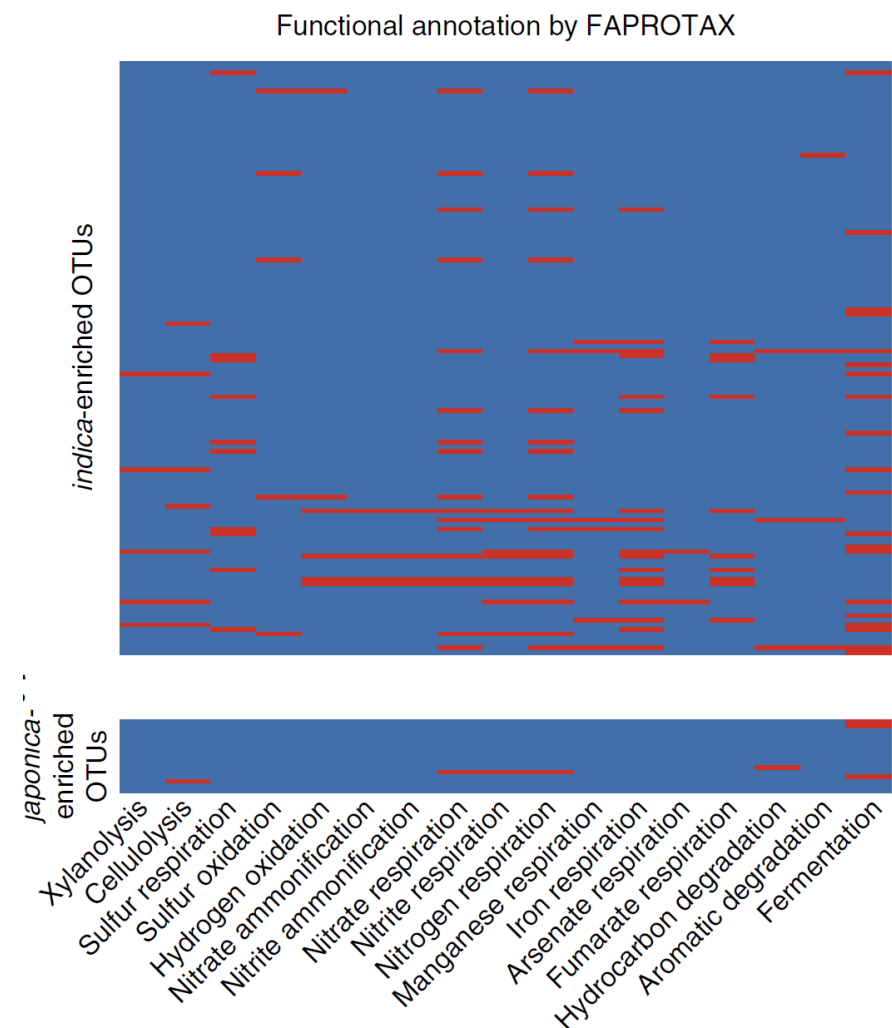
**基于FAPROTAX数据库预测菌群的功能。**  
应用Bonferroni校正后，星号表示哺乳动物和人类样品中的 $P < 0.05$ 。误差棒表示动物（ $n = 512$ ）和人类（ $n = 77$ ）样本的SD。

### PNAS：人类皮肤微生物的独特性和哺乳动物的系统发育共生现象

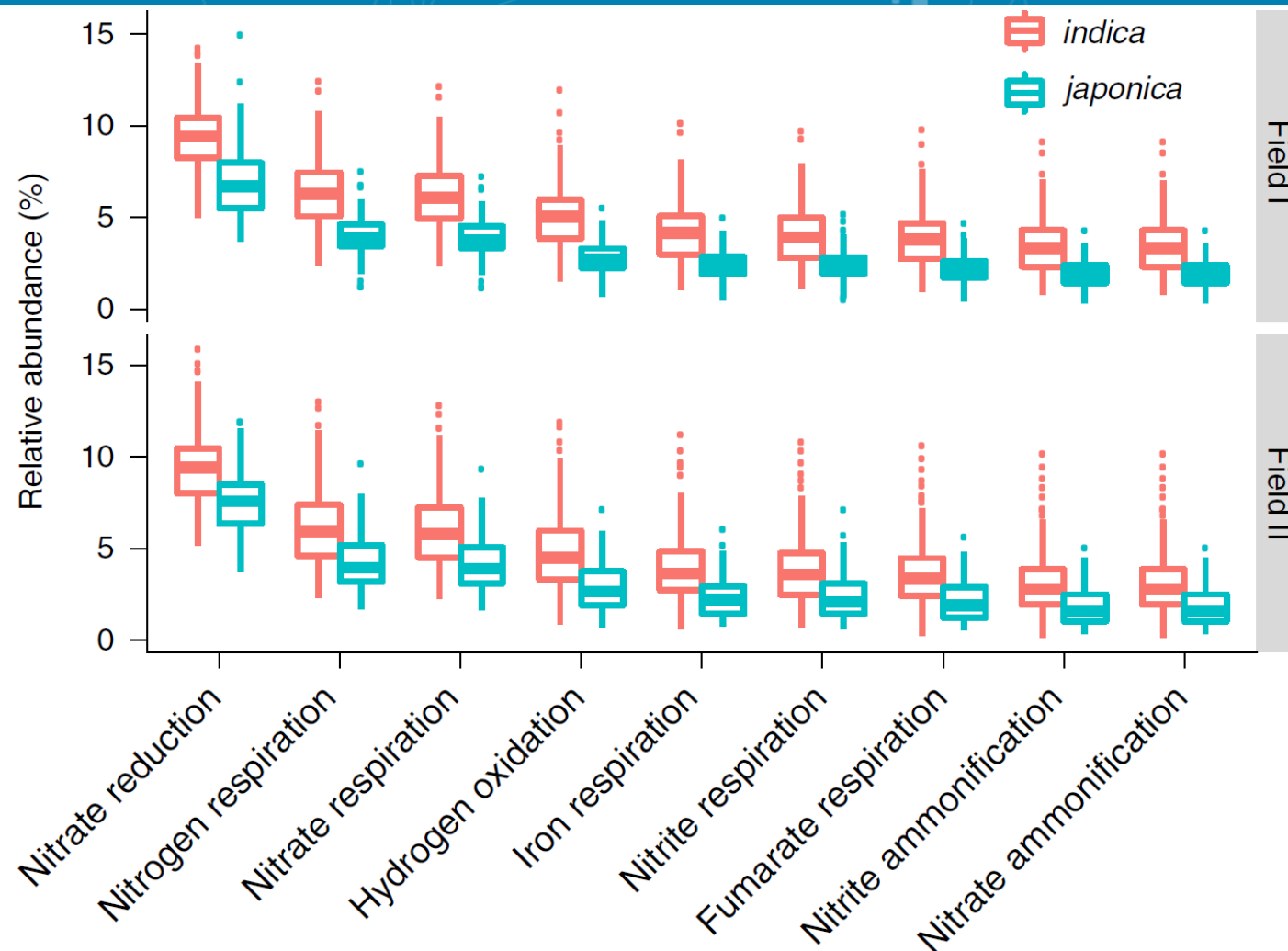
Ashley A. Ross, Kirsten M. Müller, J. Scott Weese & Josh D. Neufeld. (2018). Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phyllosymbiosis within the class Mammalia. ***Proceedings of the National Academy of Sciences of the United States of America*** 115, E5786-E5795, doi: <https://doi.org/10.1073/pnas.1801302115>



# 挖掘差异菌与氮功能相关



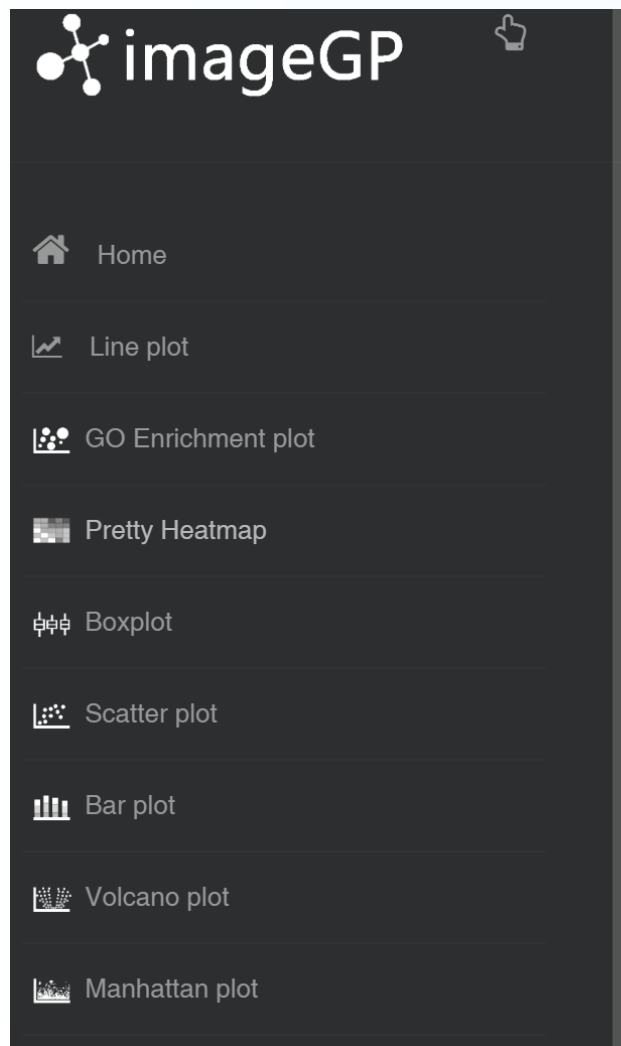
差异菌与报导的氮循环功能相关



粳稻中富集的9个最高丰度的通路中，6个与氮相关



# 方法1. 推荐ImageGP——FAPROTAX一键分析



Make plot be easy and exquisite

woodcorpse@163.com

## FAPROTAX



如果操作中遇到问题或想添加新功能，请扫码拉您入群

Map prokaryotic clades to established metabolic or other ecologically relevant functions. Please refer to [FAPROTAX](#) for input formats, results explanation and citation information.

OTU abundance table

OTUID	KO1	KO2	KO3	KO4	KO5	KO6	OE1	OE2	OE3	OE4	OE5	OE6	WT1
WT2	WT3	WT4	WT5	WT6									
OTU_91	169.0	62.0	55.0	94.0	0.0	39.0	74.0	83.0	120.0	123.0	69.0	24.0	
9.0	12.0	1.0	55.0	16.0	2.0								
OTU_684	10.0	3.0	7.0	0.0	3.0	18.0	9.0	5.0	3.0	7.0	1.0	7.0	
2.0	2.0	11.0	2.0	2.0	6.0								
OTU_7	200.0	113.0	971.0	298.0	1016.0	241.0	122.0	1084.0	727.0	494.0	172.0	202.0	631.0
715.0	450.0	174.0	456.0	706.0									
OTU_60	1447.0	1016.0	1161.0	1323.0	1907.0	1576.0	455.0	410.0	615.0	720.0	656.0	614.0	
848.0	842.0	748.0	864.0	762.0	814.0								
OTU_4	1328.0	341.0	420.0	613.0	1003.0	419.0	472.0	526.0	381.0	573.0	583.0	390.0	884.0
1120.0	448.0	468.0	871.0	758.0									



# 方法2：本地分析(限Linux系统，选学)

## 1. 软件安装(仅需运行1次)

下载软件包和解压

测试及依赖关系(如numpy、biom);

## 2. 准备带有物种注释信息的biom格式OTU表;

文本格式转换为biom格式

Biom格式添加物种信息

## 3. OTU表转换为功能表

collapse\_table.py转换带有物种信息的OTU表为功能表





# 1. 软件安装(推荐喜欢折腾和命令行的朋友，选学)

- <http://www.zoology.ubc.ca/louca/FAPROTAX> 下载软件1.2.4版， 2020/11/27更新数据库，脚本运行环境更新为python3  
wget -c  
`https://pages.uoregon.edu/slouca/LoucaLab/archive/FAPROTAX/SECTION_Download/MODULE_Downloads/CLASS_Latest%20release/UNIT_FAPROTAX_1.2.4/FAPROTAX_1.2.4.zip`  
unzip FAPROTAX\_1.2.4.zip
- 新建一个python3环境，或进入python3环境，如qiime2  
conda activate qiime2-2020.6
- 测试是否可运行，弹出帮助即正常工作  
python FAPROTAX\_1.2.4/collapse\_table.py
- 如果报错，并一般提示缺少numpy，可使用conda安装依赖包  
conda install numpy  
conda install biom



## 2. 制作输入OTU表

- # txt转换为biom json格式

```
biom convert -i otutab_rare.txt -o otutab_rare.biom --table-type="OTU table" --to-json
```

- # 添加物种注释

```
biom add-metadata -i otutab_rare.biom --observation-metadata-fp  
taxonomy2.txt -o otutab_rare_tax.biom --sc-separated taxonomy --observation-  
header OTUID,taxonomy
```

- # 指定输入文件、物种注释、输出文件、注释列名、属性列名



### 3. FAPROTAX功能预测

- # python运行collapse\_table.py脚本、-i输入带有物种注释的OTU表、-g指定数据库位置，物种注释列名，输出过程信息，强制覆盖结果，输出文件

```
python FAPROTAX_1.2.4/collapse_table.py -i otutab_rare_tax.biom \  
-g FAPROTAX_1.2.4/FAPROTAX.txt \  
--collapse_by_metadata 'taxonomy' -v --force \  
-o faprotax.txt -r faprotax_report.txt  
# Assigned 979 records to groups, 1924 records were leftovers
```

- # 下载faprotax.txt，配合实验设计可进行统计分析faprotax\_report.txt，faprotax\_report.txt查看每个类别中具体来源哪些OTUs



## 4. 制作OTU对应功能注释有无矩阵

- 对ASV(OTU)注释行，及前一行标题进行筛选

```
grep 'ASV_' -B 1 faprotax_report.txt | grep -v -P '^--$' > faprotax_report.clean
```

- Perl脚本将数据整理为三类表格

```
faprotax_report_sum.pl -i faprotax_report.clean -o faprotax_report
```

ASV_1227	ureolysis		
ASV_1232	aerobic_chemoheterotrophy		
ASV_1232	chemoheterotrophy		
ASV_1235	predatory_or_exoparasitic		
ASV_1236	aerobic_chemoheterotrophy		
ASV_1236	aromatic_compound_degradation		
ASV_1236	chemoheterotrophy		
ASV_1241	ureolysis		
ASV_1243	aerobic_chemoheterotrophy		
ASV_1243	chemoheterotrophy		
ASV_1244	chemoheterotrophy		
ASV_1244	fermentation		

OTU功能注释列表

faprotax\_report.otu\_func

aerobic_ammonia_oxidation	ASV_1646
aerobic_ammonia_oxidation	ASV_2391
aerobic_anoxygenic_phototro	ASV_293
aerobic_chemoheterotrophy	ASV_1000
aerobic_chemoheterotrophy	ASV_1001
aerobic_chemoheterotrophy	ASV_1007
aerobic_chemoheterotrophy	ASV_1009
aerobic_chemoheterotrophy	ASV_1016
aerobic_chemoheterotrophy	ASV_1017
aerobic_chemoheterotrophy	ASV_1018
aerobic_chemoheterotrophy	ASV_1019
aerobic_chemoheterotrophy	ASV_1020

功能包含OTU列表

faprotax\_report.func\_otu

OTUID	aerobic_chemoheterotrophy	aromatic_compound_degradation	cellulolysis	chemoheterotrophy
ASV_1000	1	0	0	1
ASV_1001	1	0	0	1
ASV_1007	1	0	0	1
ASV_1011	0	0	0	0
ASV_1016	1	0	0	1
ASV_1017	1	0	0	1
ASV_1018	1	0	0	1
ASV_1019	1	0	1	1
ASV_1020	1	1	0	1
ASV_1023	1	0	0	1

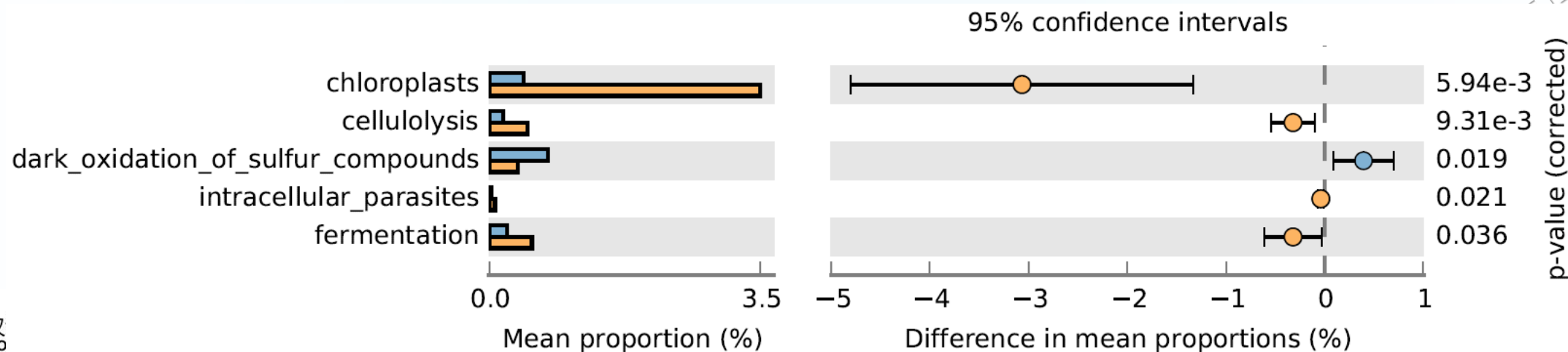
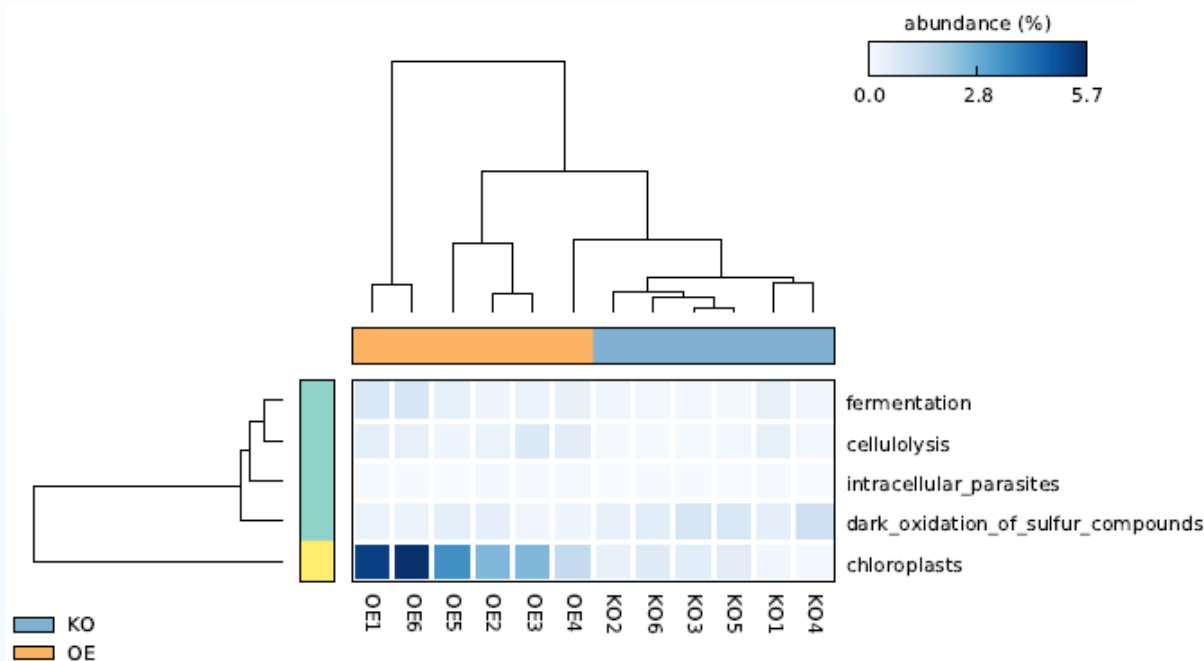
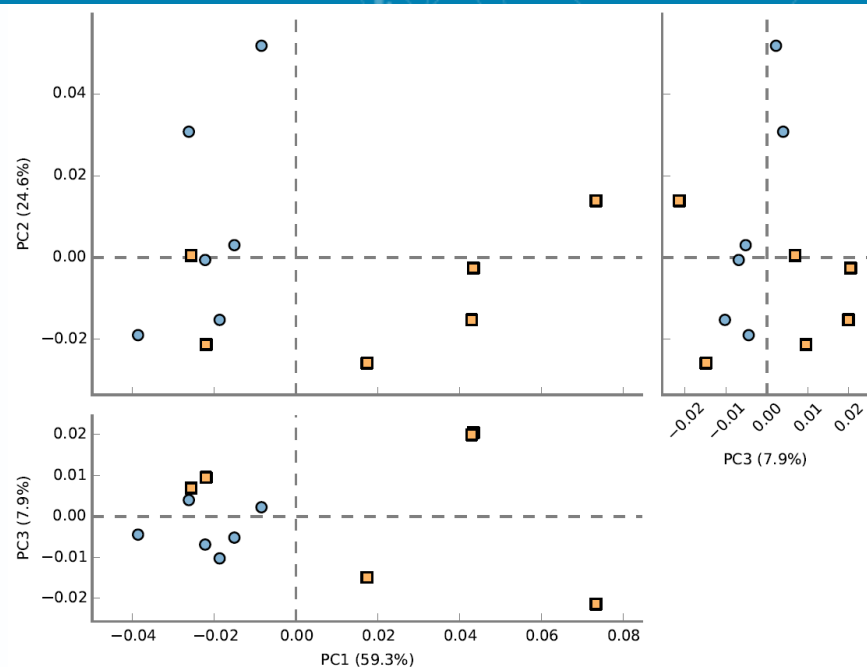
OTU功能有无矩阵

faprotax\_report.mat

筛选部分进行绘制热图



# FAPROTAX结果展示(STAMP)



- FAPROTAX于2016年发表于Science，5年引用600+次，在关注元素循环功能方面注释较好、环境、农业领域有较多应用；2020年11月27更新1.2.4版，注释比例提高，推荐使用SILVA132数据库更配；
- FAPROTAX无网络版本，需要本地脚本、数据库及依赖关系，准备带有物种注释的OTU表作为输入；
- 推荐使用ImageGP在线版，一键完成分析，目前默认使用1.1版；
- 软件不限物种注释数据库，但经验上RDP注释的结果匹配率最高；
- 软件中间文件的单个OTU功能注释结果，可用于图型展示和讨论。

- 0概述：根据16S预测微生物群落功能最全攻略，对以下4种常用的功能预测工具进行简介、实例说明和点评，必读；
- 1KO通路PICRUST：本地和在线分析，及统计分析可视化指导；
- 2元素循环FAPROTAX：元素循环相关菌代谢预测，只有本地版python脚本；识别菌种名称；
- 3表型bugbase：表型预测、革兰氏、氧气利用等；有在线和本地版，基于Greengene数据库；
- 4KO通路Tax4Fun：R包基于作者整理的Sliva123数据，准确率不可知，输出结果为百分比；最新版Tax4Fun2
- PICRUST2：OTU/ASV等16S序列随意预测宏基因组，参考数据库增大10倍





扫码关注生信宝典，学习更多生信知识



扫码关注宏基因组，获取专业学习资料

# 易生信，没有难学的生信知识

