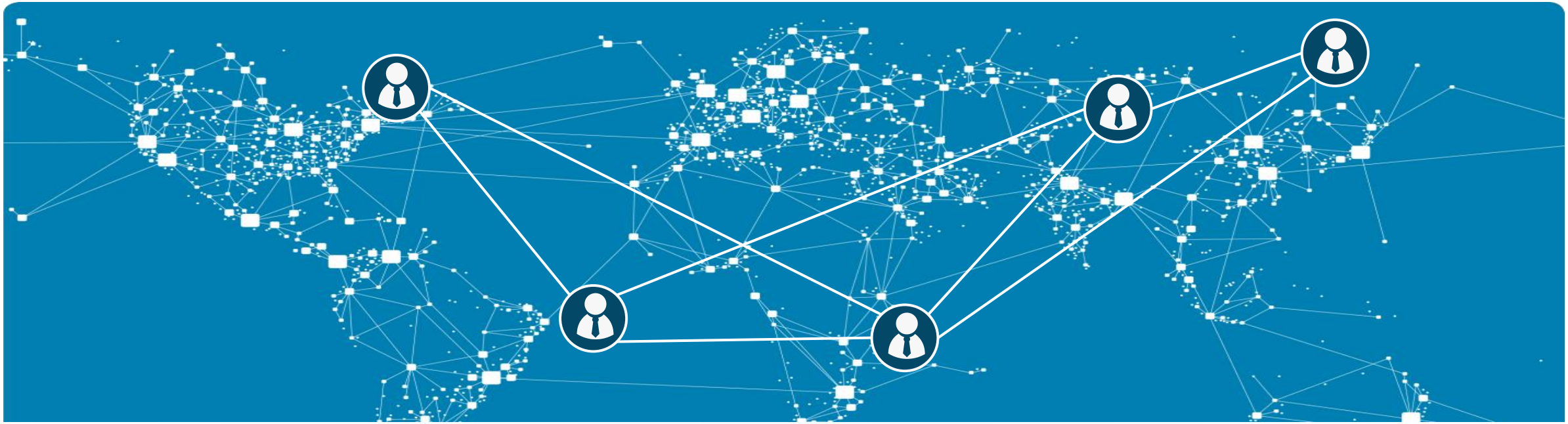
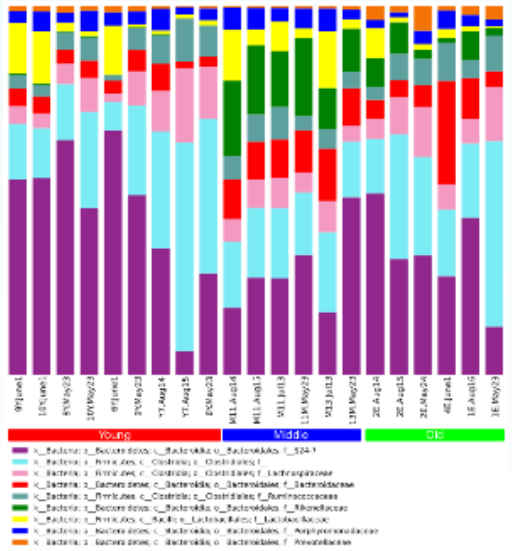
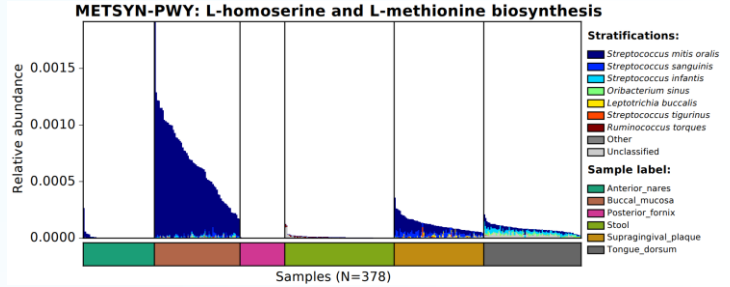
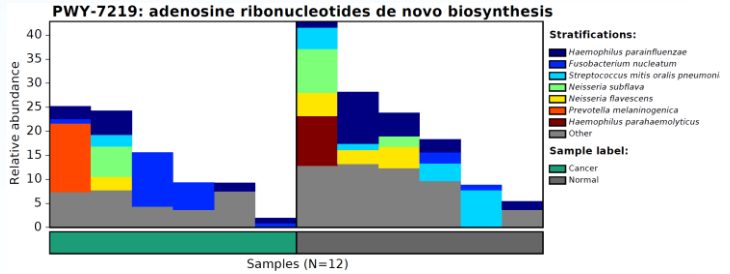
微生物组—宏基因组分析专题研讨会第14期



23HUMAnN2物种和功能组成







易生信

2022年3月26日

宏基因组基于读长(Read-based)的分析流程

一. 软件安装和数据库部署

二. KneadData质控

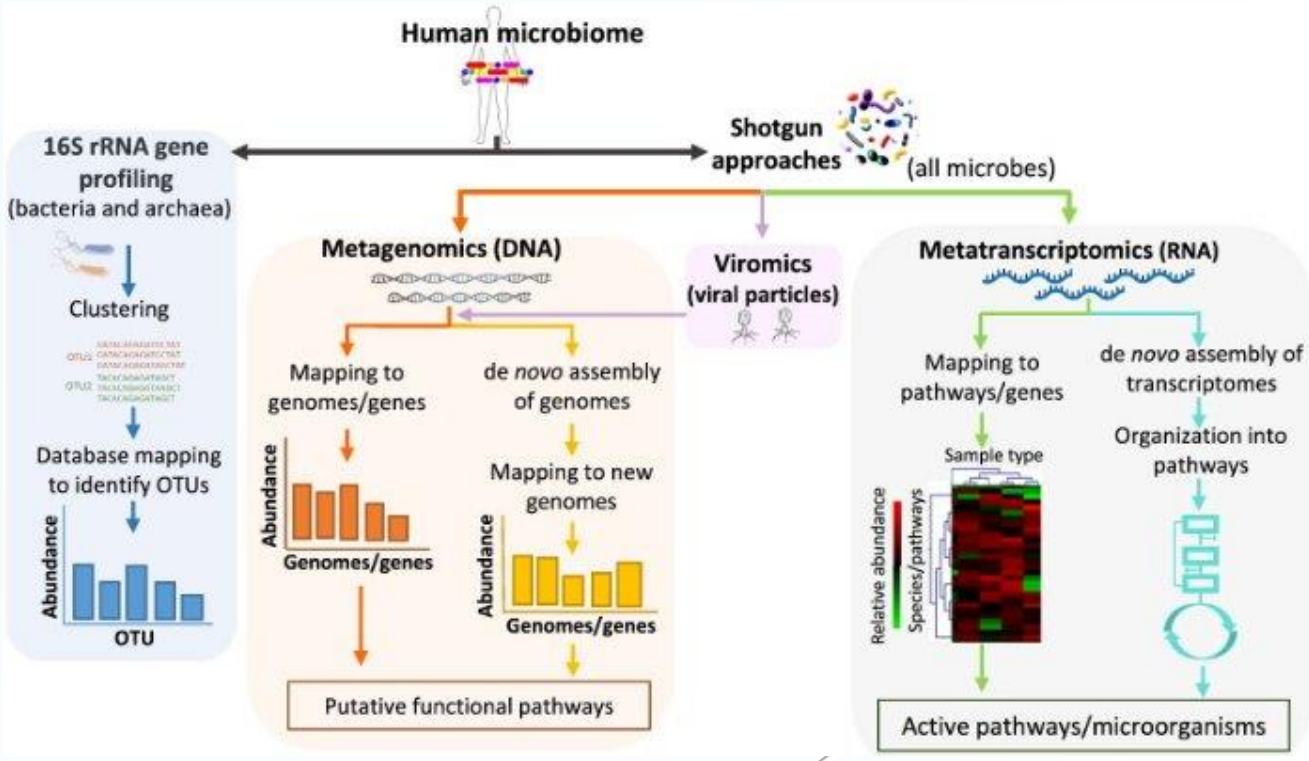
三. **MetaPhlAn2物种组成**

四. HUMAnN2功能组成

五. GraPhlAn可视化物种

六. LEfSe分析物种差异

七. STAMP功能组成分析



MetaPhlAn2 (metagenomic phylogenetic analysis，宏基因组系统发育分析，2015)

易生信

- MetaPhlAn2是分析微生物群落(细菌、古菌、真核生物和病毒)组成的工具，只需一条完命令即可获得微生物的物种丰度信息。同时提供脚本可进一步统计和可视化。
- 主页：http://segatalab.cibio.unitn.it/tools/metaphlan2/

Metagenomic microbial community profiling using unique clade-specific genes

N Segata, L Waldron, A Ballarini, V Narasimhan… - Nature …, 2012 - nature.com

Metagenomic shotgun sequencing data can identify microbes populating a microbial community and their proportions, but existing taxonomic profiling methods are inefficient for increasingly large data sets. We present an approach that uses clade-specific marker genes to unambiguously assign reads to microbial clades more accurately and> 50× faster than current approaches. We validated our metagenomic phylogenetic analysis tool, MetaPhlAn, on terabases of short reads and provide the largest metagenomic profiling to date of the …

★ Cited by 989 Related articles All 21 versions

Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature Methods* **9**, 811, doi:10.1038/nmeth.2066 (2012).

MetaPhlAn2 for enhanced metagenomic taxonomic profiling

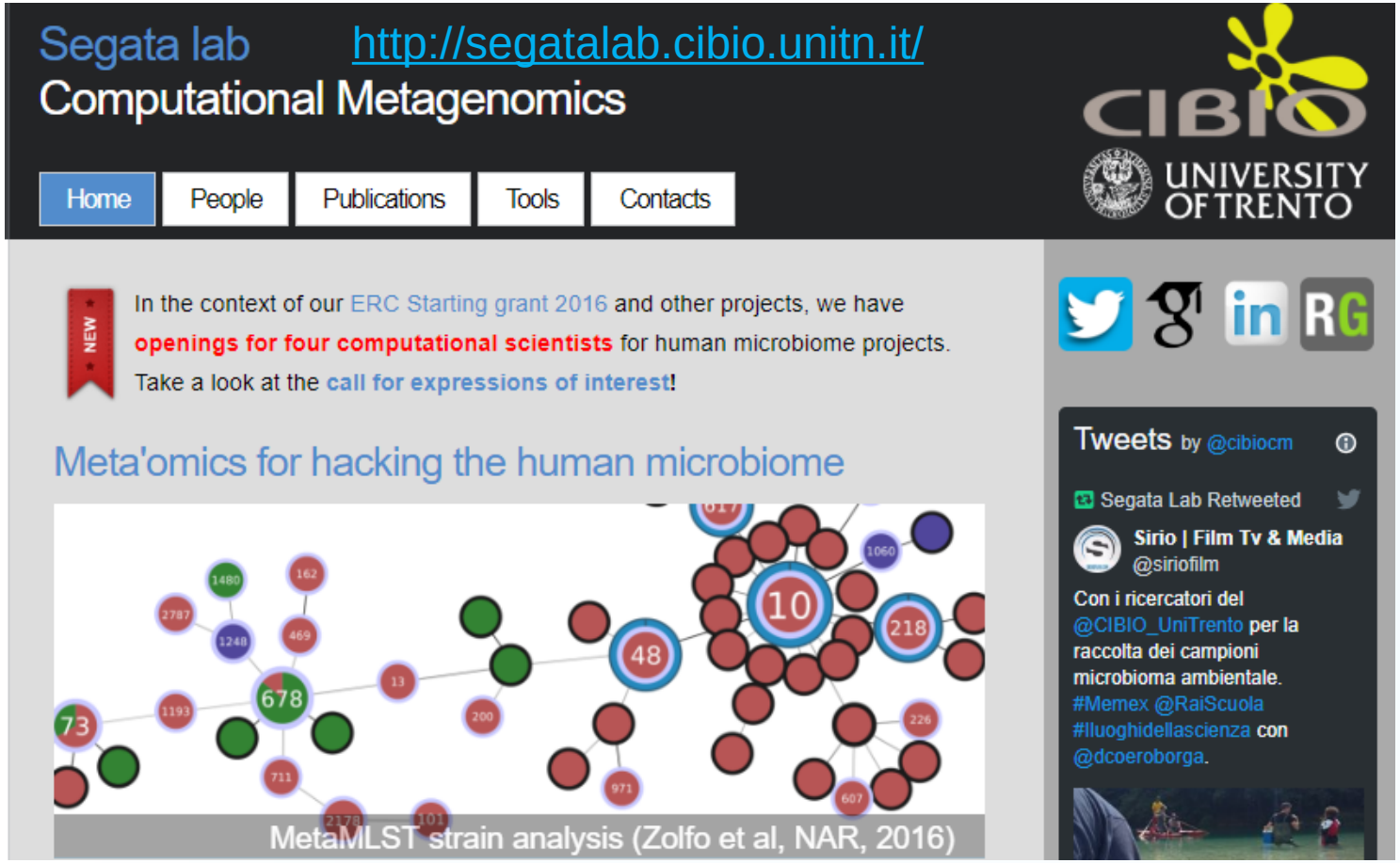
DT Truong, EA Franzosa, TL Tickle, M Scholz… - Nature …, 2015 - nature.com

MetaPhlAn (metagenomic phylogenetic analysis) 1 is a method for characterizing the taxonomic profiles of whole-metagenome shotgun (WMS) samples that has been used successfully in large-scale microbial community studies 2, 3. This work complements the original species-level profiling method with a system for eukaryotic and viral quantitation, strain-level identification and strain tracking. These and other extensions make the MetaPhlAn2 computational package (http://segatalab. cibio. unitn. it/tools/metaphlan2/and

★ Cited by 673 Related articles All 7 versions

Truong, D. T. *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods* **12**, 902, doi:10.1038/nmeth.3589 (2015).

意大利特伦托大学Nicola Segata组——宏基因组软件(10-12年哈佛公共卫生学院Huttenhower组博后)



Cell：宏基因组分箱15万人体微生物基因组

开发维护和参与的软件

- curatedMD (2017) 发表人类微生物组物种和功能组成整理数据库R包
- MetaMLST (2016)宏基因组多位点序列分型
- StrainPhlAN (2016)菌株水平群体基因组分析
- MetAML (2016)3000个宏基因组数据微生物与表型关联预测
- PanPhlAn (2016)预测菌株水平基因组成和转录活性
- MetaPhlAn2 (2015) 宏基因组物种组成
- GraPhlAn (2015) 物种或进化树圈图美化
- ShortBRED (2015)蛋白归类基因家族和宏基因组功能定量
- MicroPITA (2014)宏基因样本挑选
- MetaRef (2014)微生物类特异基因数据库
- PhyloPhlAn (2013)新微生物基因组分类和进化关系鉴定
- HUMAnN (2012) 宏基因组功能组成定量
- LEfSe (2011)生物标志物挖掘

Nicola Segata

FOLLOW

Department CIBIO, University of Trento
Verified email at unitn.it - Homepage

Human Microbiome Computational Biology Metagenomics Microbial Genomics
Machine Learning

TITLE	CITED BY	YEAR
Structure, function and diversity of the healthy human microbiome C Huttenhower, D Gevers, R Knight, S Abubucker, JH Badger, ... nature 486 (7402), 207	6261	2012
Metagenomic biomarker discovery and explanation N Segata, J Izard, L Waldron, D Gevers, L Miropolsky, WS Garrett, ... Genome biology 12 (6), R60	4961	2011
Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2 E Bolyen, JR Rideout, MR Dillon, NA Bokulich, CC Abnet, GA Al-Ghalith, ... Nature biotechnology 37 (8), 852-857	1837 *	2019
A framework for human microbiome research BA Methé, KE Nelson, M Pop, HH Creasy, MG Giglio, C Huttenhower, ... nature 486 (7402), 215	1652	2012
Metagenomic microbial community profiling using unique clade-specific marker genes N Segata, L Waldron, A Ballarini, V Narasimhan, O Jousson, ... Nature methods 9 (8), 811-814	1191	2012
Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis JU Scher, A Sczesnak, RS Longman, N Segata, C Ubeda, C Bielski, ...	1127	2013

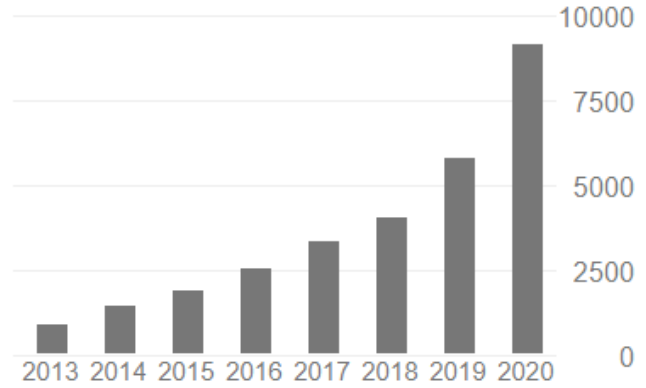
LEfSe

QIIME 2

MetaPhlAn2

Cited by

	All	Since 2016
Citations	29814	24837
h-index	53	51
i10-index	96	90



Co-authors VIEW ALL


Curtis Huttenhower
Department of Biostatistics, Harv...


Levi Waldron
City University of New York Scho...


Edoardo Pasolli
Department of Agricultural Scien...

MetaPhlAn2的数据量和特征

MetaPhlAn2整理了超过17000个参考基因组，包括13500个细菌和古菌，3500个病毒和110种真核生物，汇编整理了100万+类群特异的标记基因，可以实现：

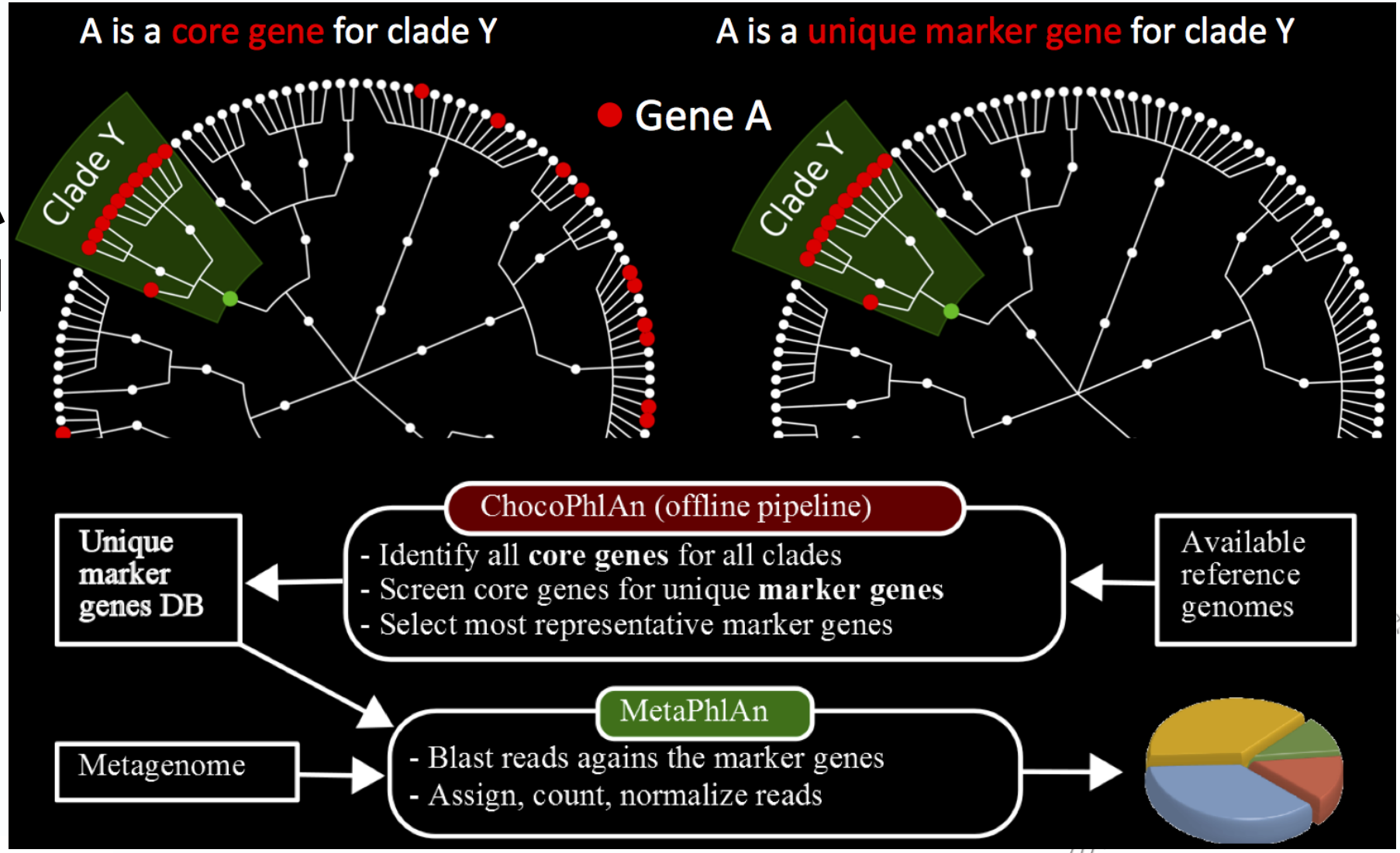
- 精确的分类群分配
- 准确估计物种的相对丰度
- 种水平精度
- 株鉴定与追踪
- 超快的分析速度

MetaPhlAn Marker的选择(核心算法)

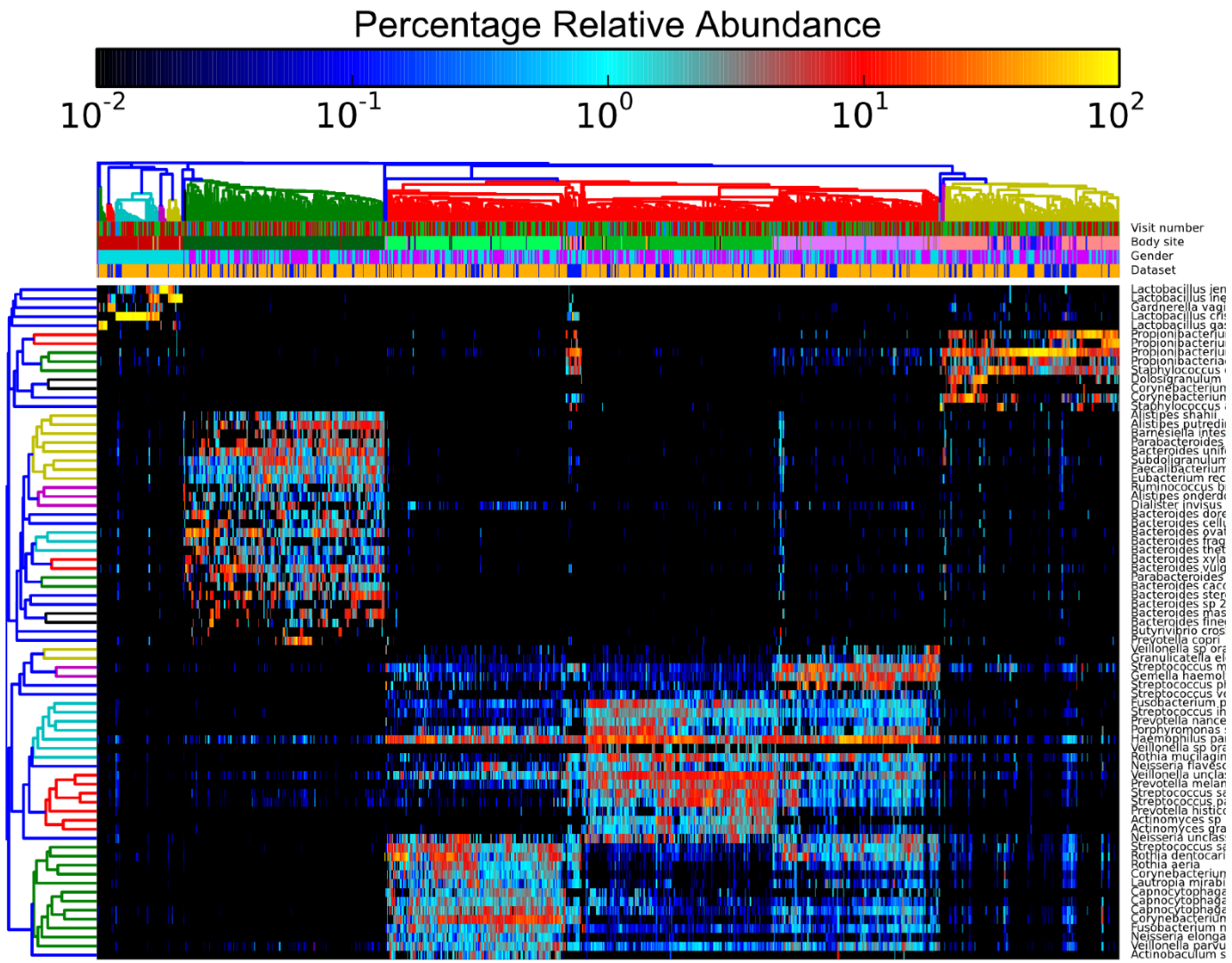
类核心基因

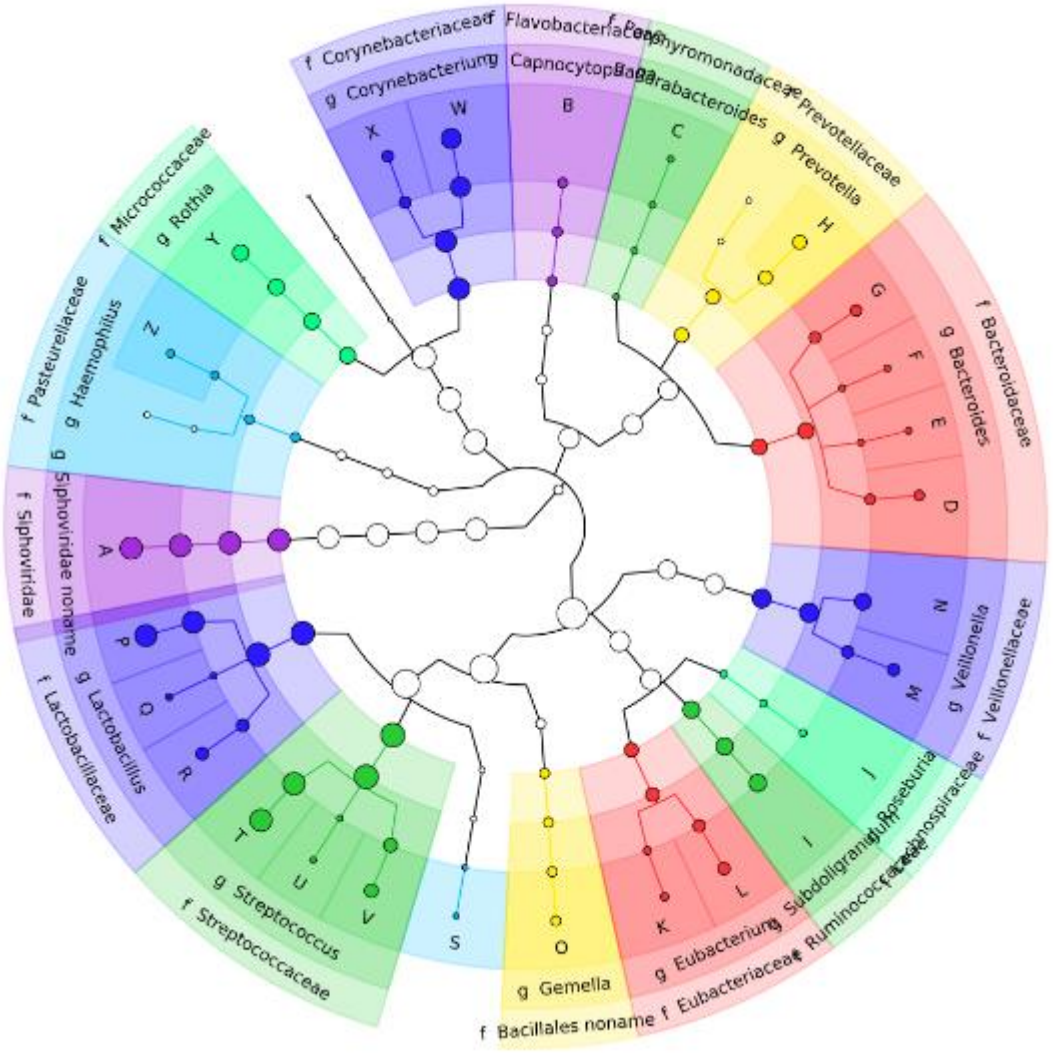
类特异基因

泛基因组数据库



MetaPhlAn2结果常用展示方式





MetaPhlAn2的使用

```
metaphlan2.py metagenome.fastq --bowtie2out metagenome.bowtie2.bz2 \
    --nproc 9 --input_type fastq > profiled_metagenome.txt
```

输入fastq单个文件，输出物种组成表，可选bowtie2比对中间文件

- 使用方法详见：MetaPhlAn2一条命令获得宏基因组物种组成
- 软件安装代码见1soft_db.sh或2pipeline.sh中附录metaphlan2部分
- 除非只关注物种组成，否则MetaPhlAn很少单独使用
- HUMAnN2整合了MetaPhlAn2软件，可实现一条命令完成物种、功能、以及功能对应物种组成三个文件，多角度挖掘宏基因组数据。

开发中的MetaPhlAn3

- 标记基因与HUMAnN3**中的**ChocoPhlAn 一致，近10万个基因组
- --unknown_estimation支持未知微生物的估计
- --index latest 自动安装最新版的数据库
- --add_viruses 支持病毒组
- 结果包括NCBI的分类学ID
- 默认移除低质量结果
- 安装 conda install -c bioconda python=3.7 metaphlan

主页：http://huttenhower.sph.harvard.edu/metaphlan3
代码：https://github.com/biobakery/MetaPhlAn/tree/3.0
教程： https://github.com/biobakery/biobakery/wiki/metaphlan3

宏基因组基于读长(Reads-based)的分析流程

一. 软件安装和数据库部署

二. KneadData质控

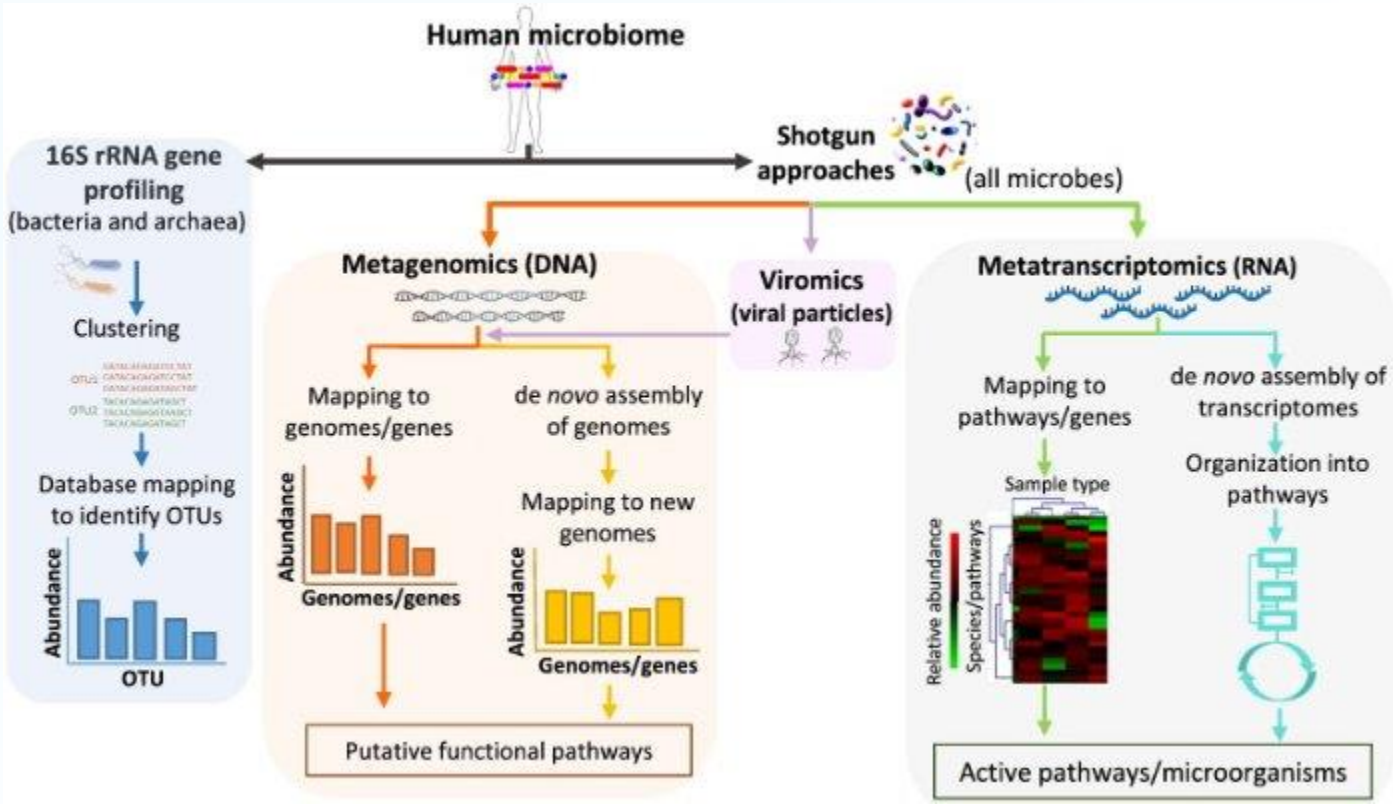
三. MetaPhlAn2物种组成

四. **HUMAnN2功能组成**

五. GraPhlAn可视化物种

六. LEfSe分析物种差异

七. 功能组成统计分析



HUMAnN2 http://www.huttenhower.org/humann2

易生信

- HUMAnN2: The HMP Unified Metabolic Analysis Network 2, HUMAnN是基于**宏基因组、宏转录组数据分析**微生物通路丰度的有效工具。这一过程称为功能谱，目的是描述群体成员的代谢潜能。可以**回答微生物群体成员可能干什么，或在干什么的问题。**

[HTML] Metabolic reconstruction for metagenomic data and its application to the human microbiome

S Abubucker, N Segata, J Goll, AM Schubert... - PLoS Comput ..

Microbial communities carry out the majority of the biochemical a they play integral roles in processes including metabolism and im human microbiome. Shotgun sequencing of such communities' m information complementary to organismal abundances from taxo resulting data typically comprise short reads from hundreds of dif best challenging to assemble comparably to single-organism gen

☆ 99 Cited by 856 Related articles All 31 versions »

Species-level functional profiling of metagenomes and metatranscriptomes

EA Franzosa, LJ McIver, G Rahnavard, LR Thompson... - Nature ..., 2018 - nature.com

Functional profiles of microbial communities are typically generated using comprehensive metagenomic or metatranscriptomic sequence read searches, which are time-consuming, prone to spurious mapping, and often limited to community-level quantification. We developed HUMAnN2, a tiered search strategy that enables fast, accurate, and species-resolved functional profiling of host-associated and environmental communities. HUMAnN2 identifies a community's known species, aligns reads to their pangenomes, performs ...

☆ 99 Cited by 491 Related articles All 11 versions

Abubucker, S. *et al.* Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLOS Computational Biology* **8**, e1002358, doi:10.1371/journal.pcbi.1002358 (2012).

Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nature Methods* **15**, 962-968, doi:10.1038/s41592-018-0176-y (2018).

Nature Methods：HUMAnN2实现宏基因组和宏转录组种水平功能组成分析

HUMAnN2：人类微生物组统一代谢网络分析2　　宏基因组有参流程 (HUMAnN2)

HUMAnN2的特点

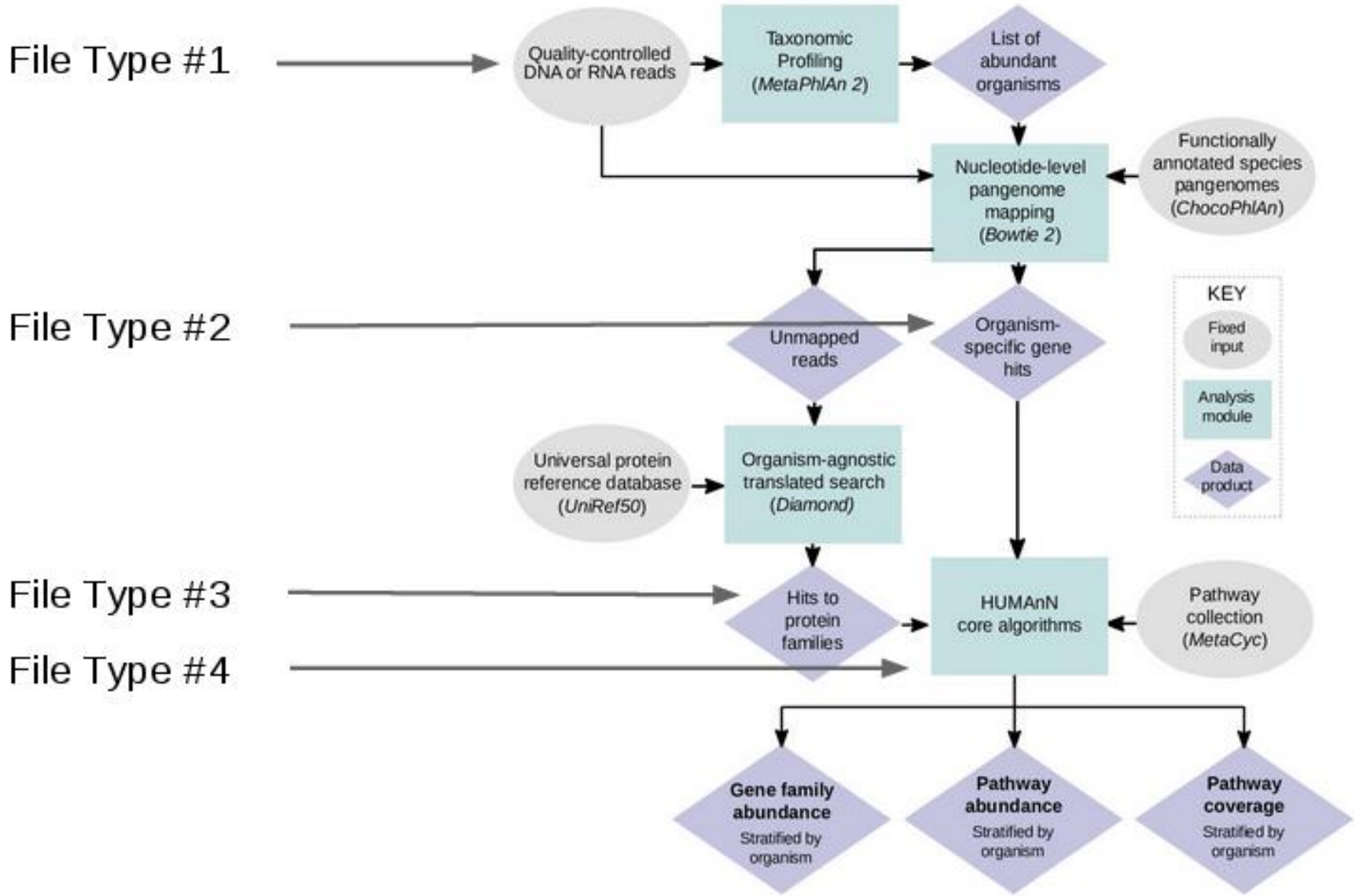
- 可对已知和未知生物分析群体功能谱
  MetaPhlAn2和ChocoPhlAn泛基因组数据库
- 可获得基因组、基因和通路层面的结果
  UniRef基因家族，MetaCyc基因通路，MinPath定义最小通路集
- 简单的使用界面(单行命令实现全部工作流程)
  用户只需提供质控后的宏基因组或宏转录组数据
- 加速序列比对
  采用Bowtie2加速核酸水平比对
  采用Diamond加速核酸翻译蛋白水平比对

HUMAnN2的依赖关系

- MetaPhlAn2 (version <= 2.6.0)
- Bowtie2 (version >= 2.2) (automatically installed)
- Diamond (0.9.0 > version >= 0.8.22) (automatically installed)
- Python (version >= 2.7)
- MinPath (automatically installed)
- Usearch (version >= 7.0) (only required if using usearch for translated search)
- SAMtools (only required if bam input files are provided)
- Biom-format (only required if input or output files are in biom format)

HUMAnN2工作原理

- File Type 1 (a quality-controlled metagenome or metatranscriptome)
  - fastq (fastq.gz)
  - fasta (fasta.gz)
- File Type 2 (alignment results type 1)
  - sam
  - Bam
- File Type 3 (alignment results type 2)
  - blast-like tsv
- File Type 4 (gene table)
  - tsv
  - biom



软件安装和数据库下载

- **HUMAnN2有参宏基因组物种和功能定量流程**

  conda install humann2

- 查看可用数据库并设置下载位置

  humann2_databases # 显示可用数据库

- 输助比对数据库

  humann2_databases --download utility_mapping full ~/db/humann2

- 微生物物种核心基因 5.37G

  humann2_databases --download chocophlan full ~/db/humann2

- 功能基因diamond索引 10.3G

  humann2_databases --download uniref uniref90_diamond ~/db/humann2

http://huttenhower.sph.harvard.edu/humann2_data/uniprot/uniref_annotated/

目前数据为2016版，2019版为HUMAnN3的数据库，数据量增大一倍

配置数据库配置

- 查看参数和数据库位置是否正确

  humann2_config --print

- 常用修改线程数、核酸、蛋白库和多种功能注释数据库位置

  humann2_config --update run_modes threads 8

  humann2_config --update database_folders utility_mapping

  ~/db/humann2/utility_mapping

  humann2_config --update database_folders nucleotide ~/db/humann2/chocophlan

  humann2_config --update database_folders protein ~/db/humann2

开发中的HUMAnN3

- 支持pip, conda和github源码3种安装方式
- 数据库更新2016版为201901版
- 微生物泛基因组数据库从1.7万增加至10万个， 5.37G增加为15.3G
- 蛋白功能注释数据库， 5.87G增加为19.31G，增大3倍多
- 安装 conda install -c biobakery humann

数据库较大，有时下载缓慢，
国家微生物组科学数据中心-工具资源下载 http://nmdc.cn/datadownload
百度云备份链接详见：
https://github.com/YongxinLiu/MicrobiomeStatPlot/blob/master/Data/BigDataDownlaodList.md

主页： http://huttenhower.sph.harvard.edu/humann3
代码： https://github.com/biobakery/humann
教程： https://github.com/biobakery/biobakery/wiki/humann3

2.1 合并双端文件

- 有参宏基因组不考虑双端，将双端文件直接合并为一个文件

```
# 创建目录存放合并后的序列
mkdir -p temp/concat
# for循环调用cat合并每一个样品
for i in `tail -n+2 result/metadata.txt | cut -f 1`;do \
    cat temp/qc/${i}*_?.fastq > temp/concat/${i}.fq; done
# 查看样品数量和大小
ls -sh temp/concat/*.fq
```

大文件加速(只用单端和/或截取，牺牲精度换速度)

- 方法1. 仅链接左端为输入文件(提速50%，节省1倍文件空间)

```
for i in `tail -n+2 result/metadata.txt|cut -f1`;do
  ln -sf `pwd`/temp/qc/${i}_1.fastq temp/concat/${i}.fq
done
```

- 方法2. 控制标准样比对时间。测序数据量通常为6~50G，同一样本分析时间可达10h~100h，严重浪费时间而浪费硬盘空间。可用head对单端分析截取20M序列，即3G，则为80M行

```
for i in `tail -n+2 result/metadata.txt|cut -f1`;do
   head -n80000000 temp/qc/${i}_1.fastq  > temp/concat/${i}.fq
done
```

2.2 HUMAnN2计算物种和功能组成

```
mkdir -p temp/humann2
```

- 如果数据库位置正确，只需输入文件和输出目录，经rush管理批量任务队列

```
tail -n+2 result/metadata.txt|cut -f1|rush -j 2 \
  'humann2 --input temp/concat/{1}.fq  \
  --output temp/humann2/'
# 核心步骤，测序数据2 X 8 = 16线程，用时1h，真实数据可能要几小时至几天
```

2.3 物种组成分析

```
mkdir -p result/metaphlan2
```

- # 样品结果合并

```
merge_metaphlan_tables.py
temp/humann2/*_humann2_temp/*_metaphlan_bugs_list.tsv | \
  sed 's/_metaphlan_bugs_list//g' > result/metaphlan2/taxonomy.tsv
```

- # 转换为spf格式方便stamp分析

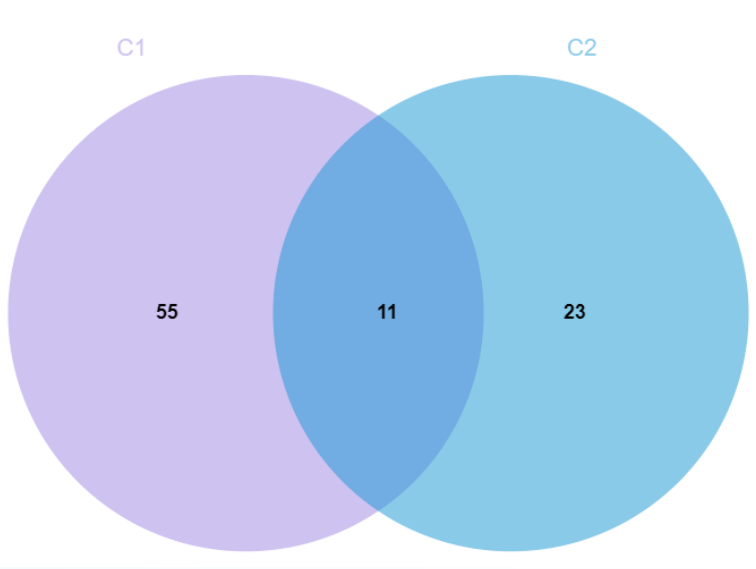
```
metaphlan_to_stamp.pl result/metaphlan2/taxonomy.tsv \
    > result/metaphlan2/taxonomy.spf
```

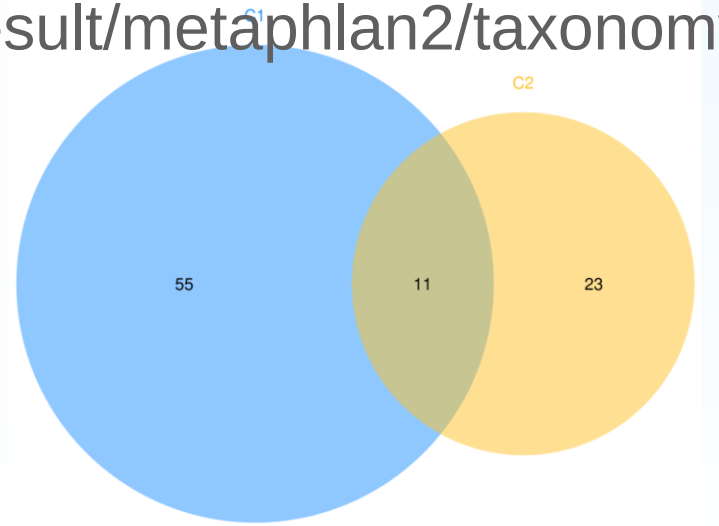
- metaphlan_to_stamp.pl 脚本来自 microbiome_helper 项目，整合至 EasyMicrobiome

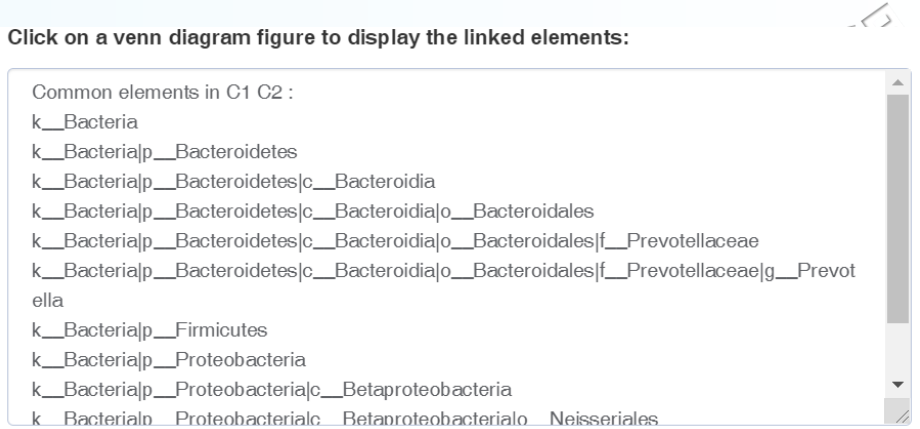
筛选样本中某一丰度的分类

- # (可选)筛选>0.5%的分类，绘制维恩图

```
awk 'BEGIN{OFS=FS="\t"}{if(FNR==1)
{for(i=2;i<=NF;i++) a[i]=$i;} \
   else {for(i=2;i<=NF;i++) if($i>0.5) print $1, a[i];}}' \
   result/metaphlan2/taxonomy.tsv \
   > result/metaphlan2/taxonomy_high.tsv
wc -l result/metaphlan2/taxonomy_high.tsv
```







http://www.ehbio.com/test/venn/#/

Tong Chen, Haiyan Zhang, Yu Liu, Yong-Xin Liu & Luqi Huang. (**2021**). EVenn: Easy to create repeatable and editable Venn diagrams and Venn networks online. ***Journal of Genetics and Genomics***, doi: 10.1016/j.jgg.2021.07.007

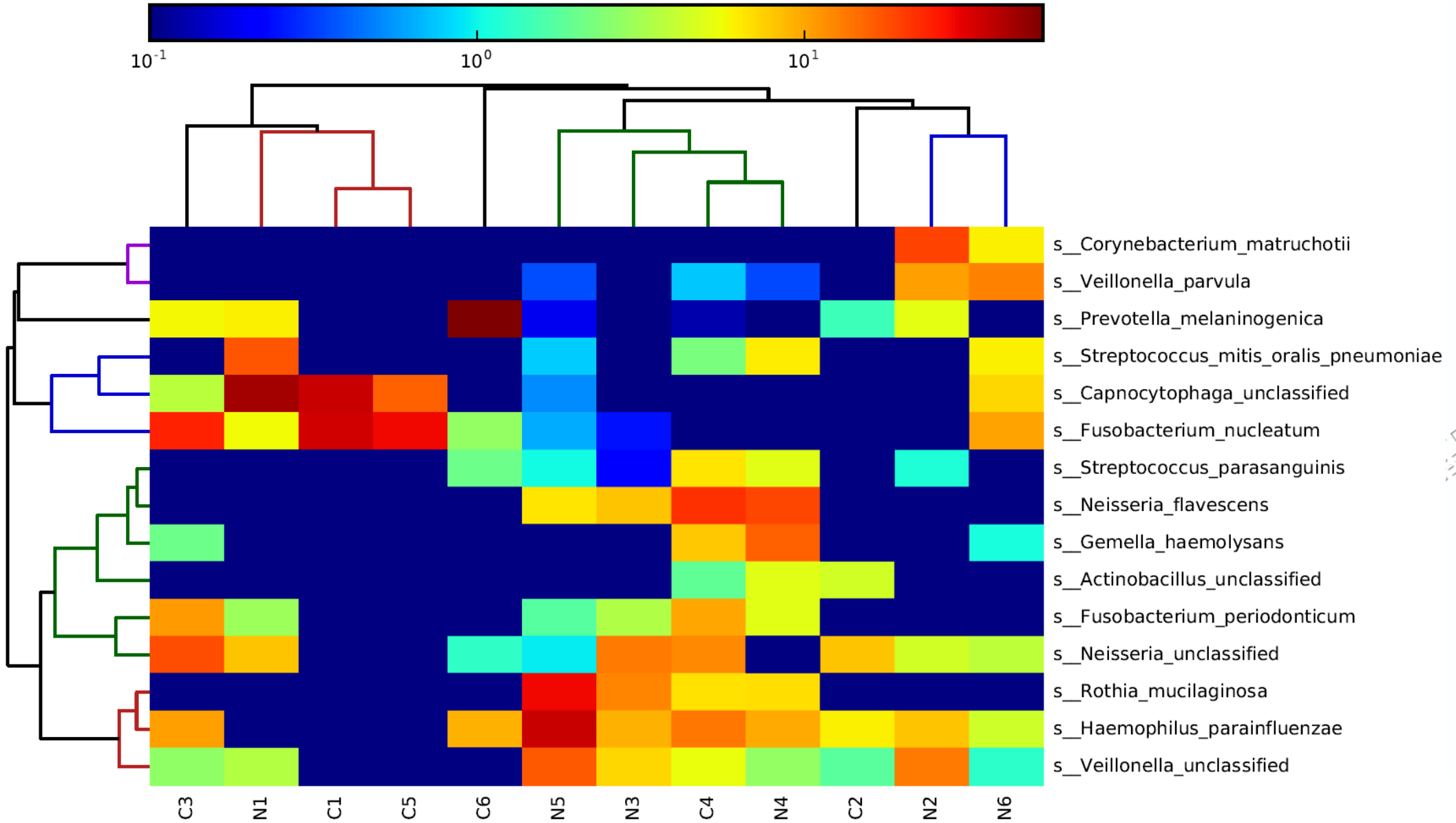
物种组成热图

- MetaPhlAn2提供了很多脚本，常用的有聚类热图

```
metaphlan_hclust_heatmap.py --in result/metaphlan2/taxonomy.tsv \
  --out result/metaphlan2/heatmap.pdf \
  -c bbcry --top 15 --minv 0.1 -s log -x 0.4 -y 0.2
```

- # c设置颜色方案，top设置物种数量，minv最小相对丰度，s标准化方法，log为取10为底对数，文件名结尾可先pdf/png/svg三种图片格式。更多说明详见 metaphlan_hclust_heatmap.py –h
- 科研结果有时要服从人类可读性，选择合适数据的结果展示，读者容易懂你是关键

前15个物种丰度对数聚类热图

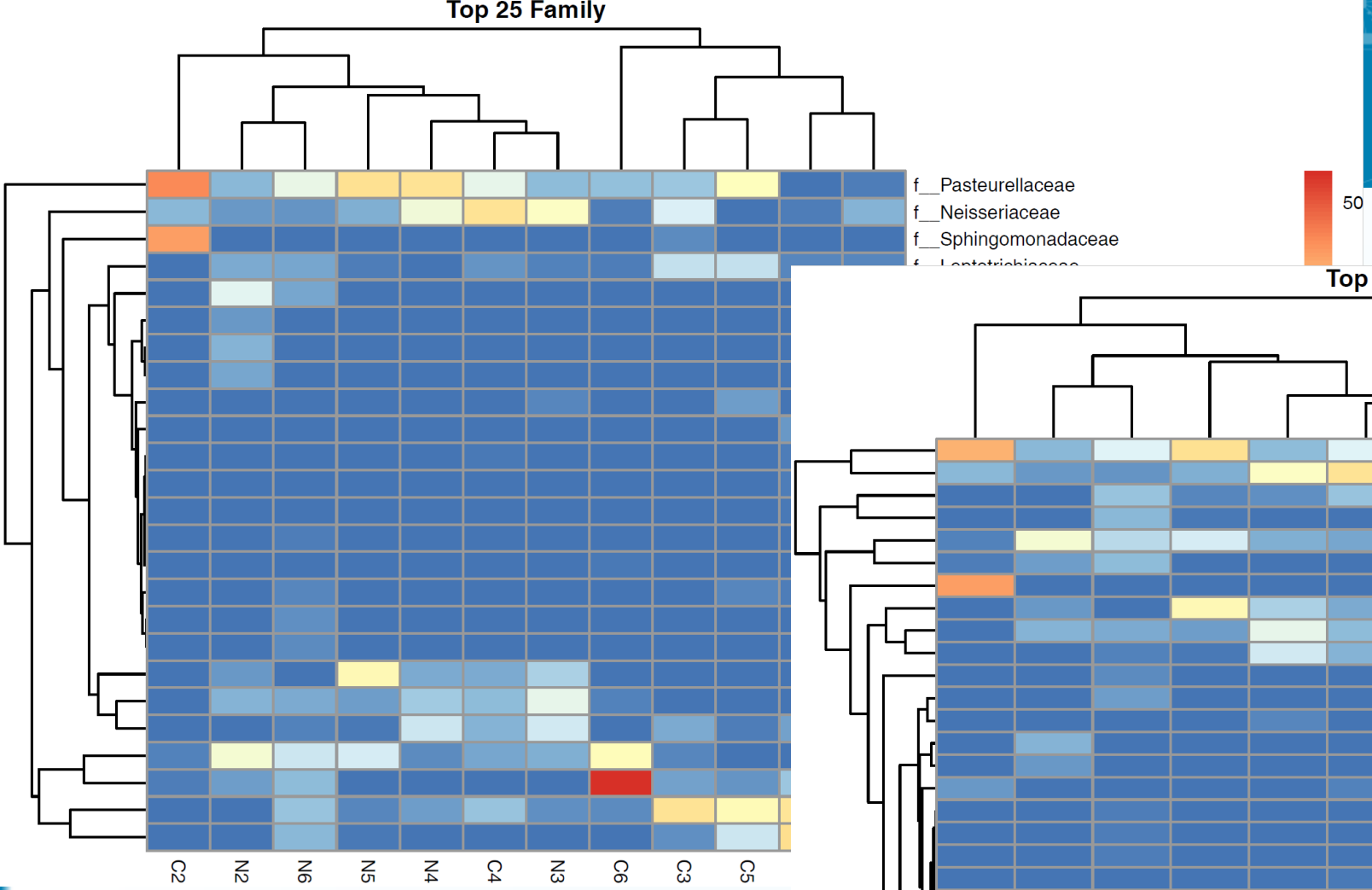


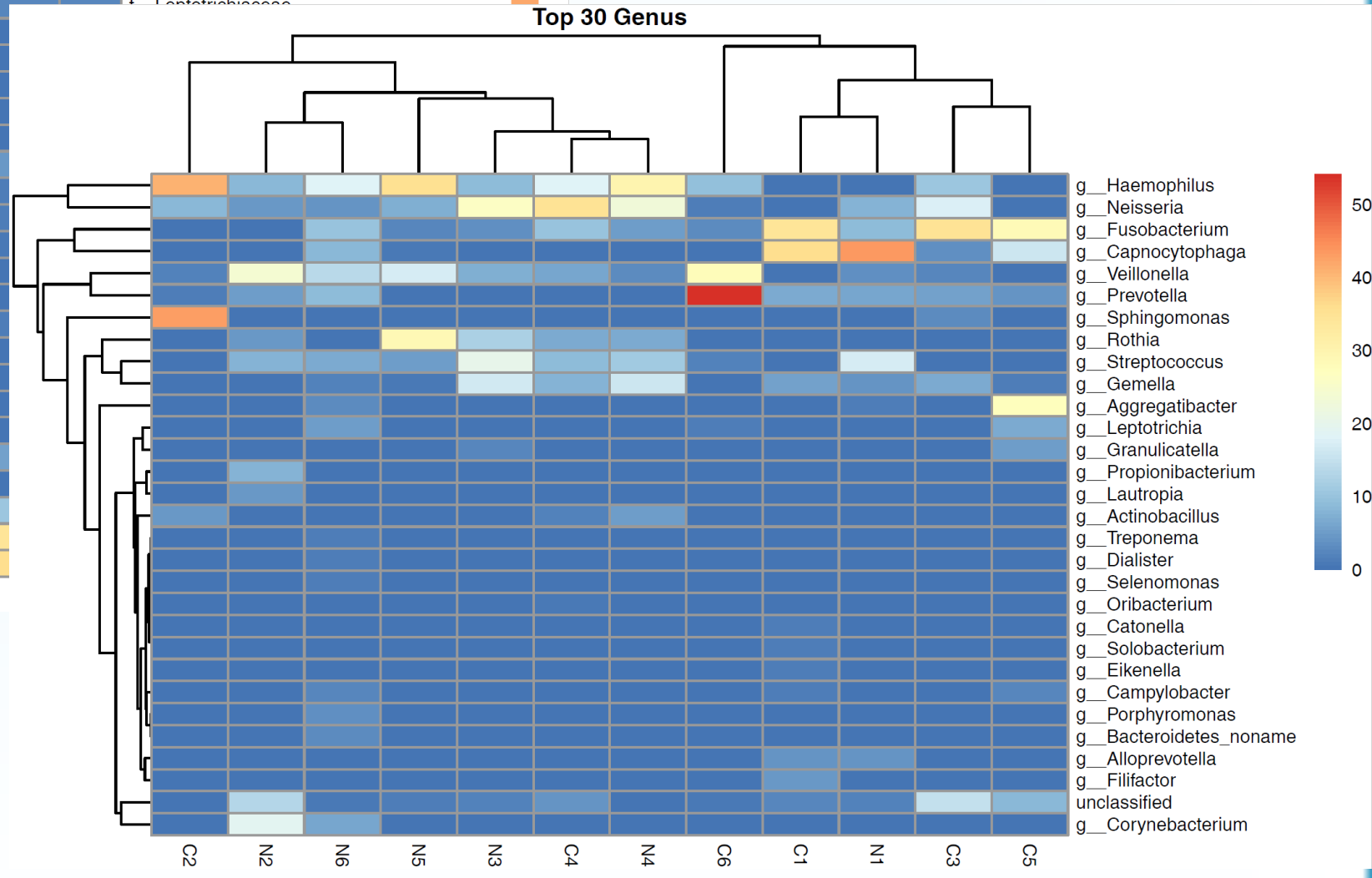
评估样本物种组成和分组聚类——更灵活绘制热图

- 方法1. metaphlan_hclust_heatmap.py 脚本服务器绘制热图，依赖关系和环境变量复杂
- 方法2. Excel筛选 metaphlan2/taxonomy.tsv 并在线绘制热图
- 方法3. R数据筛选taxonomy.spf并用pheatmap绘制热图，可指定分类级别、物种数量

```
Rscript $sd/metaphlan_hclust_heatmap.R \
    -i result/metaphlan2/taxonomy.spf \
    -t Order \
    -n 25 \
    -o result/metaphlan2/heatmap_Order
```

R脚本可以windows下运行，详见 3StatPlot.sh


Top 25 Family
f__Pasteurellaceae
f__Neisseriaceae
f__Sphingomonadaceae
50
C2
N2
N6
N5
N4
C4
N3
C6
C3
C5


Top 30 Genus
g__Haemophilus
g__Neisseria
g__Fusobacterium
g__Capnocytophaga
g__Veillonella
g__Prevotella
g__Sphingomonas
g__Rothia
g__Streptococcus
g__Gemella
g__Aggregatibacter
g__Leptotrichia
g__Granulicatella
g__Propionibacterium
g__Lautropia
g__Actinobacillus
g__Treponema
g__Dialister
g__Selenomonas
g__Oribacterium
g__Catonella
g__Solobacterium
g__Eikenella
g__Campylobacter
g__Porphyromonas
g__Bacteroidetes_noname
g__Alloprevotella
g__Filifactor
unclassified
g__Corynebacterium
50
40
30
20
10
0
C2
N2
N6
N5
N3
C4
N4
C6
C1
N1
C3
C5

2.4.1 功能组成合并、标准化和分层

- # 合并所有样品

```
humann2_join_tables --input temp/humann2 --file_name pathabundance \
  --output result/humann2/pathabundance.tsv
sed -i 's/_Abundance//g' result/humann2/pathabundance.tsv
```

- # 标准化为相对丰度relab或百万分数cpm

```
humann2_renorm_table --input result/humann2/pathabundance.tsv --units relab \
  --output result/humann2/pathabundance_relab.tsv
```

- # 分层结果

```
humann2_split_stratified_table --input result/humann2/pathabundance_relab.tsv \
  --output result/humann/
```

2.4.2 添加分组和差异比较

手动或用Shell有表头下面添加分组行

# Pathway	C1	C2	C3	C4	C5	C6	N1	N2	N3	N4	N5	N6
Group	Cancer	Cancer	Cancer	Cancer	Cancer	Cancer	Normal	Normal	Normal	Normal	Normal	Normal
ANAGLYCOLYSIS-PWY: glycolysis III (from glucose)				12.7763613320	3.3803273747	3.4161087641	15.4265476544	12.7441068874	8.6332745611	0	3.5400735715	18.6514
ANAGLYCOLYSIS-PWY: glycolysis III (from glucose)|g__Fusobacterium.s__Fusobacterium_nucleatum							5.0005085292	0	0	0	4.1457115522	0
ANAGLYCOLYSIS-PWY: glycolysis III (from glucose)|g__Haemophilus.s__Haemophilus_parainfluenzae							0	0	0	3.5541892578	0	2.0826450377

KW差异比较：输入input、分组focal、分组类型type、分组行结果、FDR和输出output；结果包括特征对应各组均值和统计值

```
humann2_associate --input result/humann2/pathabundance.pcl  \
        --focal-metadatum Group --focal-type categorical \
        --last-metadatum Group --fdr 0.05 \
        --output result/humann2/associate.txt
```

样本少无显著差异，改用HMP中每组各十个样本的数据集

HMP数据——差异比较

使用HMP示例数据，样本名下面有分组信息，需要手动制作

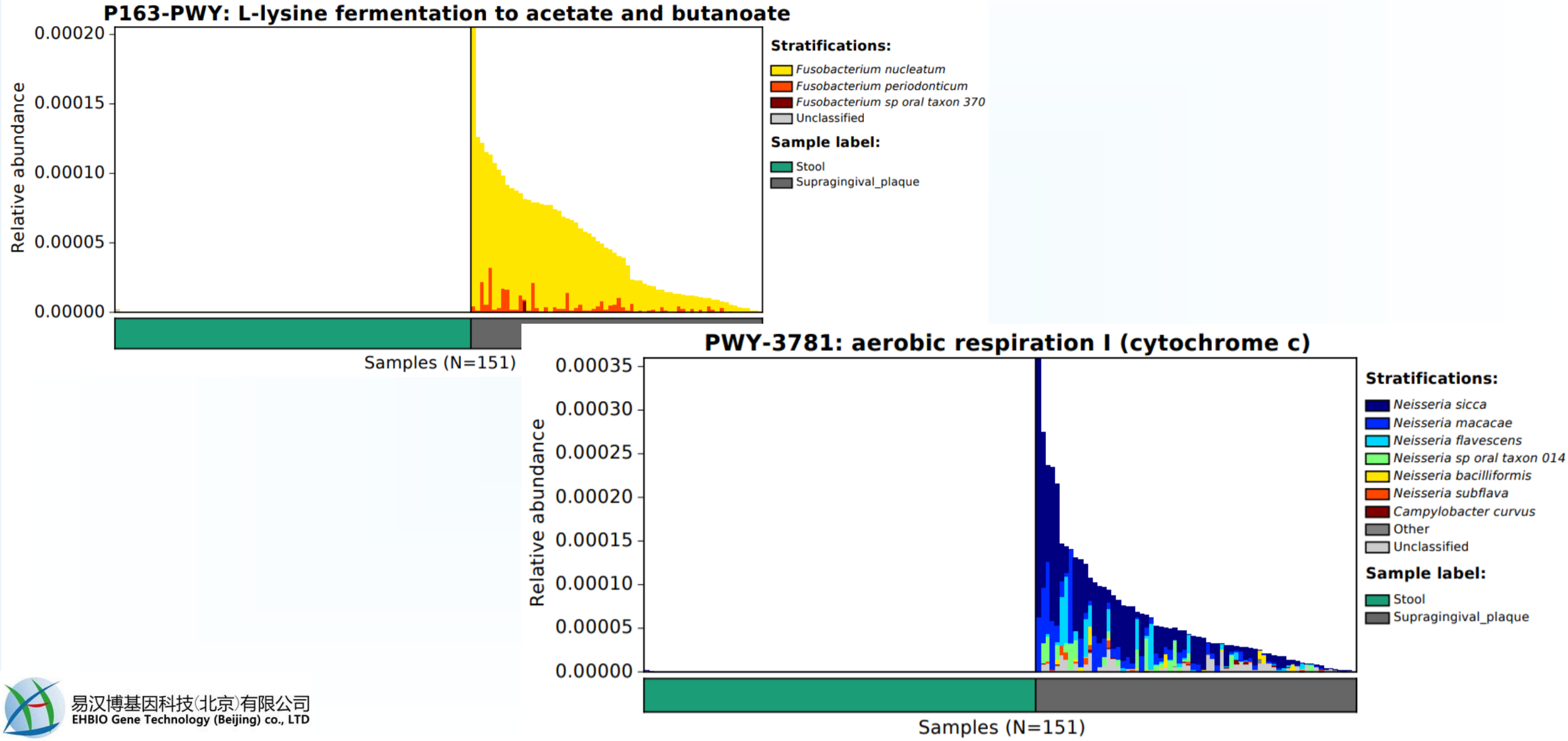
FEATURE \ SAMPLE	SRS011084	SRS011086	SRS011090
Group	Stool	Stool	Stool
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis	0.000498359	0.000628096	0.000304951
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g__Acidovorax.s__Acidovorax_ebreus	0	0	0

humann2_associate --input hmp_pathabund.pcl --focal-metadatum Group --focal-type categorical --last-metadatum Group --fdr 0.05 --output associate.txt

# Feature	Level means (|ed)	P-value	Q-value
P163-PWY: L-lysine fermentation to acetate and butanoate	Stool:3.133e-08|Supragingival_plaque:5.385e-05	1.253e-30	4.261e-28
PWY-3781: aerobic respiration I (cytochrome c)	Stool:1.464e-07|Supragingival_plaque:0.0002634	4.076e-30	4.62e-28
PWY66-409: superpathway of purine nucleotide salvage	Stool:0|Supragingival_plaque:0.000106	2.986e-30	4.62e-28
PWY1F-823: leucopelargonidin and leucocyanidin biosynthesis	Stool:4.877e-09|Supragingival_plaque:4.695e-05	1.402e-29	1.192e-27

humann2_barplot **--sort sum metadata** --input hmp_pathabund.pcl --focal-feature PWY-3781 --focal-metadatum Group --last-metadatum Group --output barplot_PWY-3781.pdf

--sort sum metadata先按分组，再按丰度



2.4.3 转换为KEGG注释

```
# 转换基因家族为KO(uniref90_ko)，可选eggNOG(uniref90_eggnog)或酶(uniref90_level4ec)
for i in `tail -n+2 result/metadata.txt|cut -f1`;do
  humann2_regroup_table \
    -i temp/humann2/${i}_genefamilies.tsv \
    -g uniref90_ko \
    -o temp/humann2/${i}_ko.tsv
done
# 合并，并修正样本名
humann2_join_tables \
  --input temp/humann2/ \
  --file_name ko \
  --output result/humann2/ko.tsv
sed -i '1s/_Abundance-RPKs//g' result/humann2/ko.tsv
```

# Gene Family	KO1	WT1
K00029	0	23.999
K00029|g__Pseudomonas.s__Pseudomon	0	23.999
K00031	8.81	7.018
K00031|g__Agrobacterium.s__Agrobacter	0	1.802
K00031|unclassified	8.81	5.215
K00032	0	8.48
K00032|g__Pseudomonas.s__Pseudomon	0	8.48
K00033	0	12.989
K00033|g__Bacillus.s__Bacillus_megateriu	0	3.454
K00033|g__Pseudomonas.s__Pseudomon	0	9.535
K00035	0	1.285
K00035|g__Agrobacterium.s__Agrobacter	0	1.285
K00036	0	46.161
K00036|g__Agrobacterium.s__Agrobacter	0	0.754
K00036|g__Bacillus.s__Bacillus_megateriu	0	5.084
K00036|g__Pseudomonas.s__Pseudomon	0	40.324

总结

- HUMAnN2调用MetaPhlAn2基于bowtie2比对上万个物种的数据库，可快速、准确获得细菌、真菌、古菌、病毒、真核生物等的物种组成
- merge_metaphlan_tables.py合并、metaphlan_hclust_heatmap.py绘制丰度热图，metaphlan_to_stamp.pl生成STAMP输入文件
- HUMAnN2采用diamond比对UniRef数据库获得功能组成；注意数据库位置设置；了解其依赖关系以便解决依赖关系错误的问题
- 结果包括功能通路丰度组成和**功能通路具体来源的物种**，提供join, norm, stratified等脚本实现合并、标准化和分层
- associate, barplot脚本实现组间KW检验和通路组成可视化
- humann2_regroup_table脚本转换基因家族为KO/eggNOG/EC等注释

参考资源

- 宏基因组公众号文章目录　　生信宝典公众号文章目录
- 科学出版社《微生物组数据分析》——50+篇
- Bio-protocol《微生物组实验手册》——153篇
- Protein Cell：扩增子和宏基因组数据分析实用指南
- CMJ：人类微生物组研究设计、样本采集和生物信息分析指南
- 加拿大生信网 https://bioinformatics.ca/ 宏基因组课程中文版
- 美国高通量开源课程 https://github.com/ngs-docs
- Curtis Huttenhower http://huttenhower.sph.harvard.edu/
- Nicola Segata http://segatalab.cibio.unitn.it/

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