

菌群 + 对应代谢产物介导 + 机制研究

2024-02-20

LiChuang Huang



@ 立效研究院

Contents

1	摘要	1
2	前言	1
3	材料和方法	1
3.1	材料	1
3.2	方法	1
4	分析结果	1
5	结论	1
6	附：分析流程	1
6.1	Microbiota 16s RNA	1
6.1.1	Fastp QC	1
6.1.2	元数据	2
6.1.3	Qiime2 分析	2
6.1.4	MicrobiotaProcess 分析	3
6.1.4.1	样本聚类	3
6.1.4.2	Alpha 多样性	3
6.1.4.3	Alpha 稀疏曲线	4
6.1.4.4	Beta 多样性	4
6.1.4.5	差异分析	6
	Reference	8

List of Figures

1	PCoA	3
2	Alpha diversity	4
3	Alpha rarefaction	4
4	Beta diversity group test	5
5	Ancom test group level 6 volcano	7

List of Tables

1	Microbiota metadata	2
---	-------------------------------	---

1 摘要

- con: Control
- A: colitis
- B: colon precancerous lesions

想要做肠道菌群测序结果 + 生信分析，然后找菌群 + 对应代谢产物介导 + 机制研究 + 再闭环回到临床。在沟通时客户提到想要做溃疡性结肠炎和结肠癌的肠道菌群之间的区别和关联，进而研究其对应的机制，研究结肠炎向结肠癌发展的关键机制，为临床早期筛查提供理论支持

2 前言

3 材料和方法

3.1 材料

3.2 方法

Mainly used method:

- Fastp used for Fastq data preprocessing¹.
- R package MicrobiotaProcess used for microbiome data visualization².
- Qiime2 used for gut microbiome 16s rRNA analysis³⁻⁷.
- Other R packages (eg., dplyr and ggplot2) used for statistic analysis or data visualization.

4 分析结果

- A、B 组 Alpha 和 Beta 多样性无显著差异 (见 6.1.4.2 和 6.1.4.4)。
- A、B 组差异分析，未找到差异菌。

5 结论

6 附：分析流程

6.1 Microbiota 16s RNA

6.1.1 Fastp QC

原始数据质控：

‘Fastp QC’ 数据已全部提供。

(对应文件为 ./fastp_report/)

注：文件夹./fastp_report/共包含 23 个文件。

1. A1.338F_806R..html
2. A2.338F_806R..html
3. A3.338F_806R..html
4. A4.338F_806R..html
5. A5.338F_806R..html
6. ...

6.1.2 元数据

Table 1 (下方表格) 为表格 microbiota metadata 概览。

(对应文件为 Figure+Table/microbiota-metadata.csv)

注：表格共有 22 行 7 列，以下预览的表格可能省略部分数据；表格含有 22 个唯一 ‘SampleName’。

1. group: 分组名称

Table 1: Microbiota metadata

SampleName	group	dirs	reports	Run	forward-ab...	reverse-ab...
A1	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A2	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A3	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A4	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A5	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A6	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A7	A	./material...	./material...	rawData	/home/echo...	/home/echo...
A8	A	./material...	./material...	rawData	/home/echo...	/home/echo...
B1	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B2	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B3	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B4	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B5	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B6	B	./material...	./material...	rawData	/home/echo...	/home/echo...
B7	B	./material...	./material...	rawData	/home/echo...	/home/echo...
...

6.1.3 Qiime2 分析

Microbiota 数据经 Qiime2 分析后，由 MicrobiotaProcess 下游分析和可视化。

6.1.4 MicrobiotaProcess 分析

6.1.4.1 样本聚类

Figure 1 (下方图) 为图 PCoA 概览。

(对应文件为 Figure+Table/PCoA.pdf)

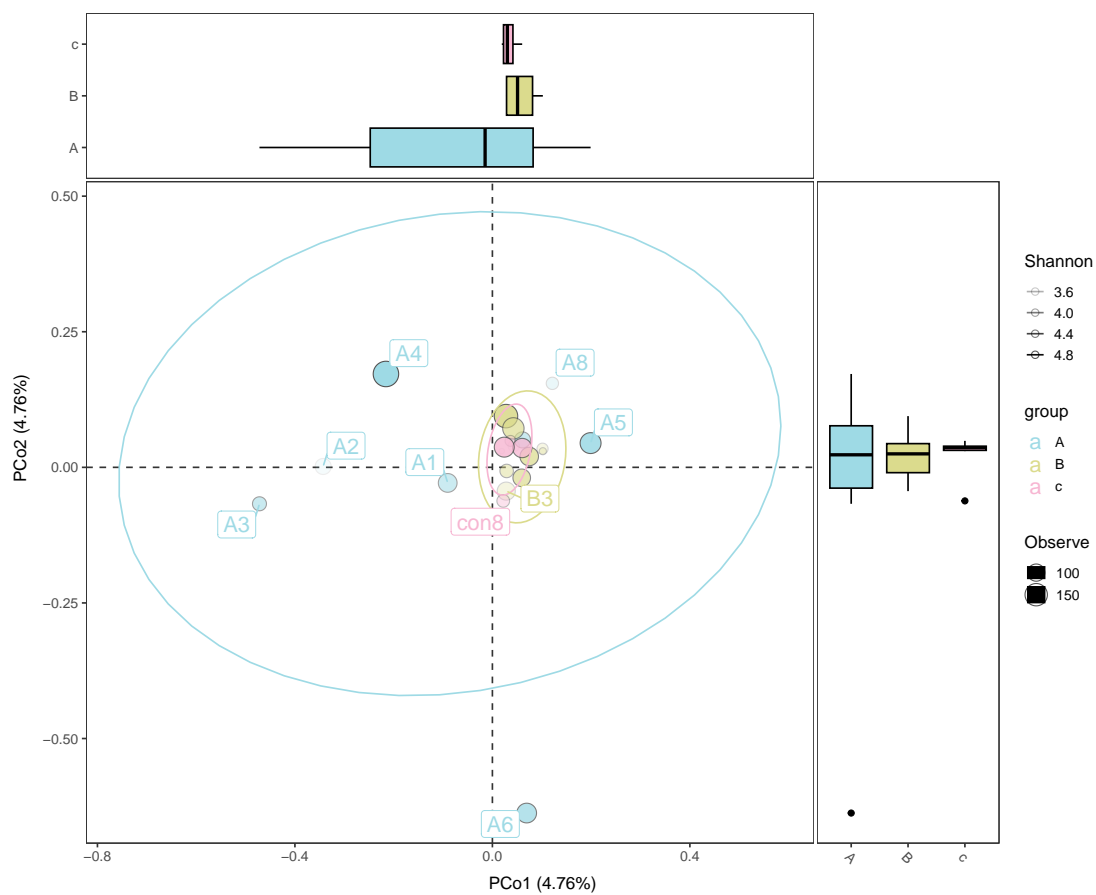


Figure 1: PCoA

6.1.4.2 Alpha 多样性

三组 alpha 多样性没有显著差异。

Figure 2 (下方图) 为图 Alpha diversity 概览。

(对应文件为 Figure+Table/Alpha-diversity.pdf)

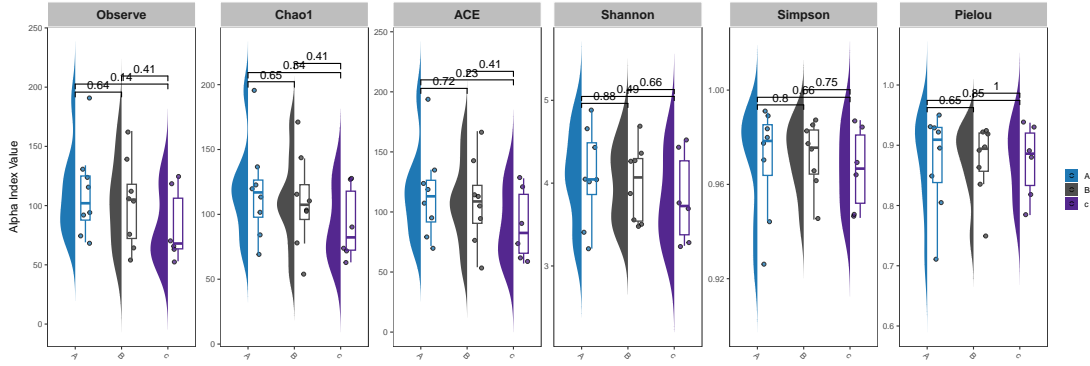


Figure 2: Alpha diversity

‘Taxonomy abundance’ 数据已全部提供。

(对应文件为 Figure+Table/Taxonomy-abundance)

注：文件夹 Figure+Table/Taxonomy-abundance 共包含 6 个文件。

1. 1_Phylum.pdf
2. 2_Class.pdf
3. 3_Order.pdf
4. 4_Family.pdf
5. 5_Genus.pdf
6. ...

6.1.4.3 Alpha 稀疏曲线

Figure 3 (下方图) 为图 Alpha rarefaction 概览。

(对应文件为 Figure+Table/Alpha-rarefaction.pdf)

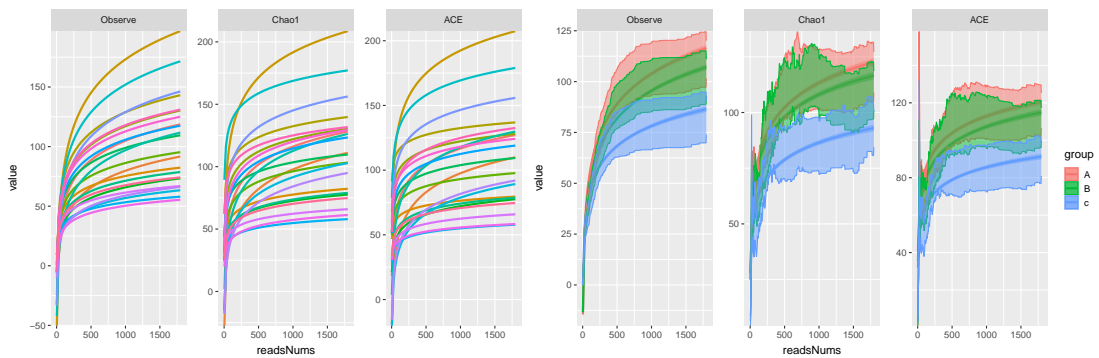


Figure 3: Alpha rarefaction

6.1.4.4 Beta 多样性

Beta 多样性无显著差异。

Figure 4 (下方图) 为图 Beta diversity group test 概览。

(对应文件为 Figure+Table/Beta-diversity-group-test.pdf)

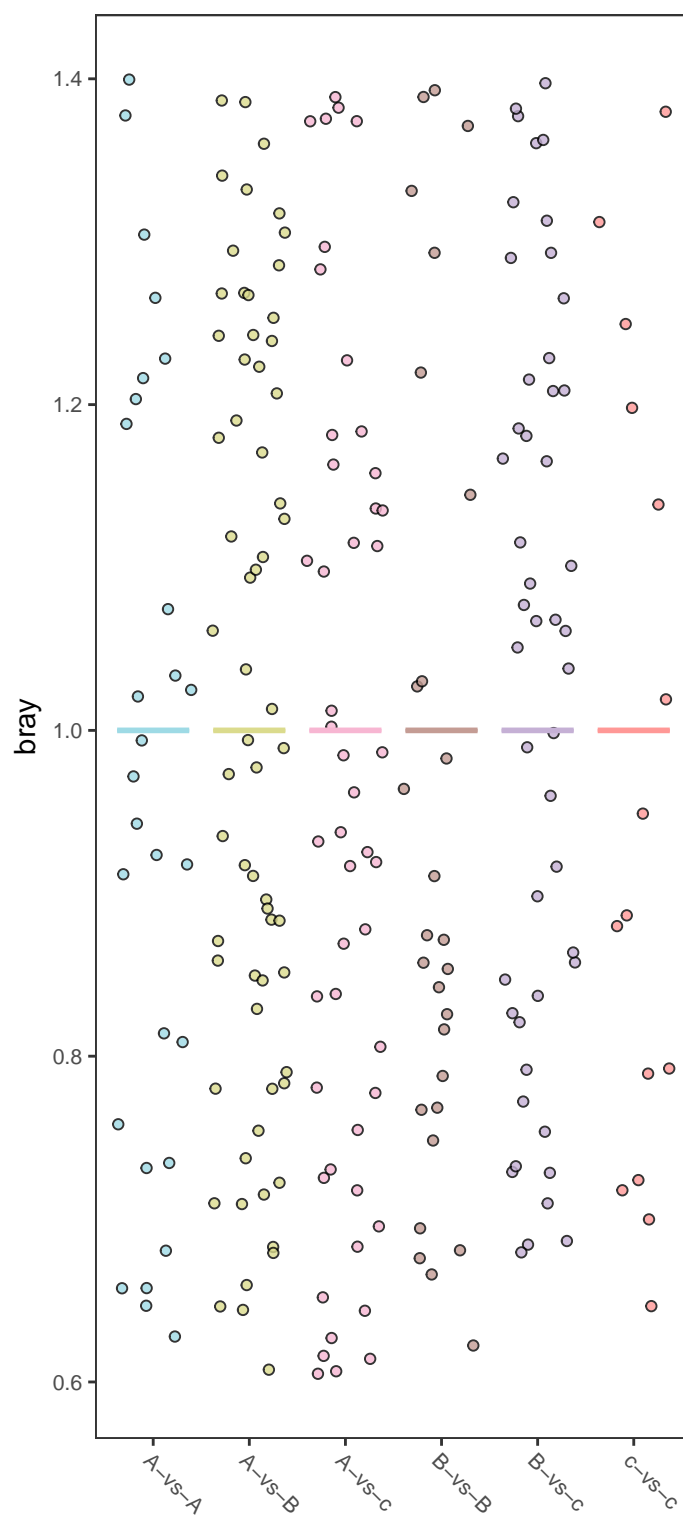


Figure 4: Beta diversity group test

‘Taxonomy hierarchy’ 数据已全部提供。

(对应文件为 Figure+Table/Taxonomy-hierarchy)

注：文件夹 Figure+Table/Taxonomy-hierarchy 共包含 6 个文件。

1. 1_Phylum.pdf
2. 2_Class.pdf
3. 3_Order.pdf
4. 4_Family.pdf
5. 5_Genus.pdf
6. ...

6.1.4.5 差异分析

MicrobiotaProcess 的差异分析 (MicrobiotaProcess::mp_diff_analysis) 未发现差异菌，因此这里主要用的 Qiime2 的差异分析结果 (ancom test)。

注：关于 ancom test 的结果的解释，可以参考：

- <https://forum.qiime2.org/t/how-to-interpret-ancom-results/1958>
- <https://forum.qiime2.org/t/specify-w-cutoff-for-anacom/1844>

Figure 5 (下方图) 为图 Ancom test group level 6 volcano 概览。

(对应文件为 Figure+Table/Ancom-test-group-level-6-volcano.pdf)

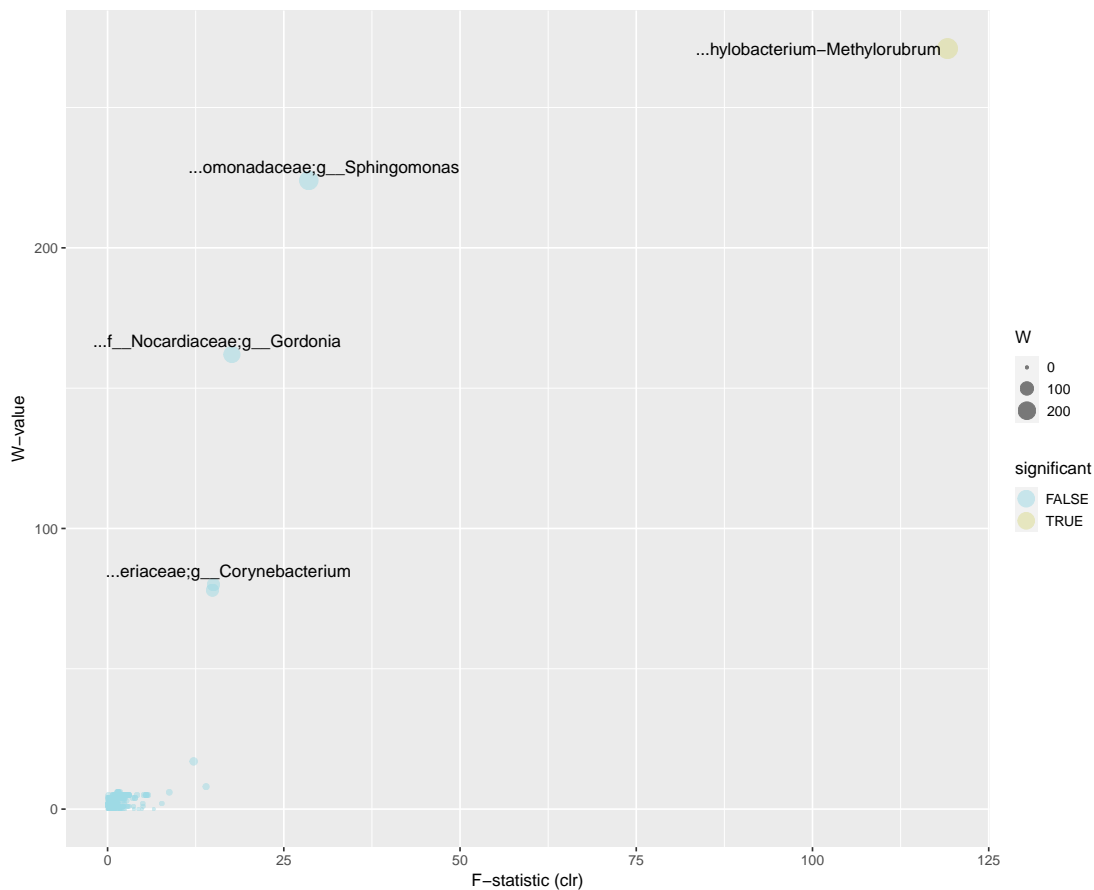


Figure 5: Ancom test group level 6 volcano

‘level 6’ 对应 Ontology 中的 Species。

其余结果的可视化见：

‘Ancom test visualization’ 数据已全部提供。

(对应文件为 **Figure+Table/Ancom-test-visualization**)

注：文件夹 Figure+Table/Ancom-test-visualization 共包含 3 个文件。

1. 1_ancom_test_group_level_4.pdf
2. 2_ancom_test_group_level_5.pdf
3. 3_ancom_test_group_level_6.pdf

‘Ancom test results’ 数据已全部提供。

(对应文件为 **Figure+Table/Ancom-test-results**)

注：文件夹 Figure+Table/Ancom-test-results 共包含 3 个文件。

1. 1_ancom_test_group_level_4.csv
2. 2_ancom_test_group_level_5.csv
3. 3_ancom_test_group_level_6.csv

Reference

1. Chen, S. Ultrafast one-pass fastq data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, (2023).
2. Xu, S. *et al.* MicrobiotaProcess: A comprehensive r package for deep mining microbiome. *The Innovation* **4**, 100388 (2023).
3. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using qiime 2. *Nature Biotechnology* **37**, 852–857 (2019).
4. McDonald, D. *et al.* The biological observation matrix (biom) format or: How i learned to stop worrying and love the ome-ome. *GigaScience* **1**, 7 (2012).
5. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from illumina amplicon data. *Nature methods* **13**, 581 (2016).
6. Hamday, M., Walker J., J., Harris, J. K., Gold J., N. & Knight, R. Error-correcting barcoded primers allow hundreds of samples to be pyrosequenced in multiplex. *Nature Methods* **5**, 235–237 (2008).
7. Hamday, M. & Knight, R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research* **19**, 1141–1152 (2009).