

# 菌群 + 对应代谢产物介导 + 机制研究

2024-02-23

LiChuang Huang



@ 立效研究院

# Contents

<b>1</b>	<b>摘要</b>	<b>1</b>
1.1	需求概要 . . . . .	1
1.2	分析结果 . . . . .	1
<b>2</b>	<b>前言</b>	<b>2</b>
<b>3</b>	<b>材料和方法</b>	<b>2</b>
3.1	材料 . . . . .	2
3.2	方法 . . . . .	2
<b>4</b>	<b>分析结果</b>	<b>2</b>
<b>5</b>	<b>结论</b>	<b>2</b>
<b>6</b>	<b>附：分析流程</b>	<b>2</b>
6.1	Microbiota 16s RNA . . . . .	2
6.1.1	Fastp QC . . . . .	2
6.1.2	元数据 . . . . .	3
6.1.3	Qiime2 分析 . . . . .	3
6.1.4	MicrobiotaProcess 分析 . . . . .	3
6.1.4.1	样本聚类 . . . . .	3
6.1.4.2	Alpha 多样性 . . . . .	4
6.1.4.3	Alpha 稀疏曲线 . . . . .	5
6.1.4.4	Beta 多样性 . . . . .	5
6.1.4.5	差异分析 . . . . .	7
6.1.5	差异菌关联到代谢物 . . . . .	11
6.1.5.1	从 gutMDisorder 数据库检索关联代谢物 . . . . .	11
6.1.5.2	尝试从已发表研究 (孟德尔随机化相关) 中寻找关联代谢物 . . . . .	12
6.1.6	代谢物的富集分析 . . . . .	12
6.1.7	从结肠炎或结肠癌已发表的代谢物研究中验证 . . . . .	15
6.1.7.1	DepressionAndYuan2021 结肠炎 (肠道菌) . . . . .	15
6.1.7.2	AlterationsInScovil2018 结肠炎 (代谢物) . . . . .	16
6.1.7.3	LossOfSymbiotSadegh2024 结肠癌 (肠道菌) . . . . .	16
6.1.7.4	IntegratedAnalChen2022 结肠癌 (肠道菌和代谢物) . . . . .	17
	<b>Reference</b>	<b>19</b>

## List of Figures

1	PCoA . . . . .	4
2	Alpha diversity . . . . .	4

3	Alpha rarefaction . . . . .	5
4	Beta diversity group test . . . . .	6
5	Ancom test group level 2 volcano . . . . .	7
6	Ancom test group level 2 Percentile abundance . . . . .	8
7	Ancom test group level 6 volcano . . . . .	9
8	Ancom test group level 6 Percentile abundance . . . . .	10
9	MetaboAnalyst kegg enrichment . . . . .	13
10	Enrichment with algorithm PageRank . . . . .	14

## List of Tables

1	Microbiota metadata . . . . .	3
2	GutMDisorder database . . . . .	11
3	MendelianRandoLiuX2022 matched data . . . . .	12
4	Compounds ID . . . . .	13
5	Data of enrichment with algorithm PageRank . . . . .	14
6	DepressionAndYuan2021 published data significant microbiota . . . . .	15
7	AlterationsInScovil2018 published data metabolites . . . . .	16
8	LossOfSymbiotSadegh2024 published data microbiota . . . . .	17
9	LossOfSymbiotSadegh2024 matched Phylum microbiota . . . . .	17
10	IntegratedAnalChen2022 published data microbiota . . . . .	18
11	IntegratedAnalChen2022 published data metabolites . . . . .	18

# 1 摘要

## 1.1 需求概要

数据分组：

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

肠道菌群测序结果 + 生信分析，得出：菌群 + 对应代谢产物介导 + 机制研究 + 再闭环回到临床。

具体：

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

## 1.2 分析结果

- 基本分析：
  - alpha、beta 多样性，A、B、C 组均无显著性差异 (6.1.4.2, 6.1.4.4)。
  - 差异菌筛选 (level 6, Species) ,筛得差异菌 (Fig. 7): d\_\_\_Bacteria;p\_\_\_Proteobacteria;c\_\_\_Alphaproteobacteria;o\_\_\_Methylobacterium 该差异菌主要存在于 A、B 组，不存在 (或少量于) 于 C (对照) 组。含量见 Fig. 8
  - ...
  - 差异菌 (level 2, Phylum) (Fig. 5), 同样的有: d\_\_\_Bacteria;p\_\_\_Proteobacteria 该差异菌主要存在于 A、B 组，不存在 (或少量于) 于 C (对照) 组。含量见 Fig. 6
- 从肠道菌到相关代谢物：
  - 使用 gutMDisorder 未发现相关代谢物。
  - 从一孟德尔随机化相关研究中<sup>1</sup>，发现了与差异菌相关的代谢物，见 Tab. 3。这些代谢物为 (详细信息见 Tab. 4): 5-methyltetrahydrofolic acid, selenium, L-cystine, Glutamic acid
  - 用 MetaboAnalystR 对相关代谢物进行富集分析，富集到两条通路 (见 Fig. 9)
  - 用 FELLA 对相关代谢物富集分析，可以发现相关联的更多通路或反应模块 (结果见 Fig. 10, Tab. 5)
- 尝试从已有的关于结肠炎或结肠癌的研究中验证上述发现：
  - 从结肠癌相关研究中匹配到<sup>2</sup> (Tab. 9): d\_\_\_Bacteria;p\_\_\_Proteobacteria (注：在 Phylum 水平上得到验证)
  - 未从其它文献中匹配到代谢物或差异肠道菌 (见 6.1.7)。
- 结肠炎向结肠癌之间的转化：
  - A 为结肠炎，B 为结肠癌前病变；A 与 B 组间无显著差异菌，因此无法从这一批数据探究可能的发展机制 (A -> B)。

## 2 前言

## 3 材料和方法

### 3.1 材料

Other data obtained from published article (e.g., supplementary tables):

- Supplementary file from article refer to<sup>1</sup>.

### 3.2 方法

Mainly used method:

- R package **FELLA** used for metabolite enrichment analysis<sup>3</sup>.
- **Fastp** used for Fastq data preprocessing<sup>4</sup>.
- Database **gutMDisorder** used for finding associations between gut microbiota and metabolites<sup>5</sup>.
- R package **MicrobiotaProcess** used for microbiome data visualization<sup>6</sup>.
- **MetaboAnalyst** used for metabolomic data analysis<sup>7</sup>.
- **Qiime2** used for gut microbiome 16s rRNA analysis<sup>8-12</sup>.
- Other R packages (eg., **dplyr** and **ggplot2**) used for statistic analysis or data visualization.

## 4 分析结果

## 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)

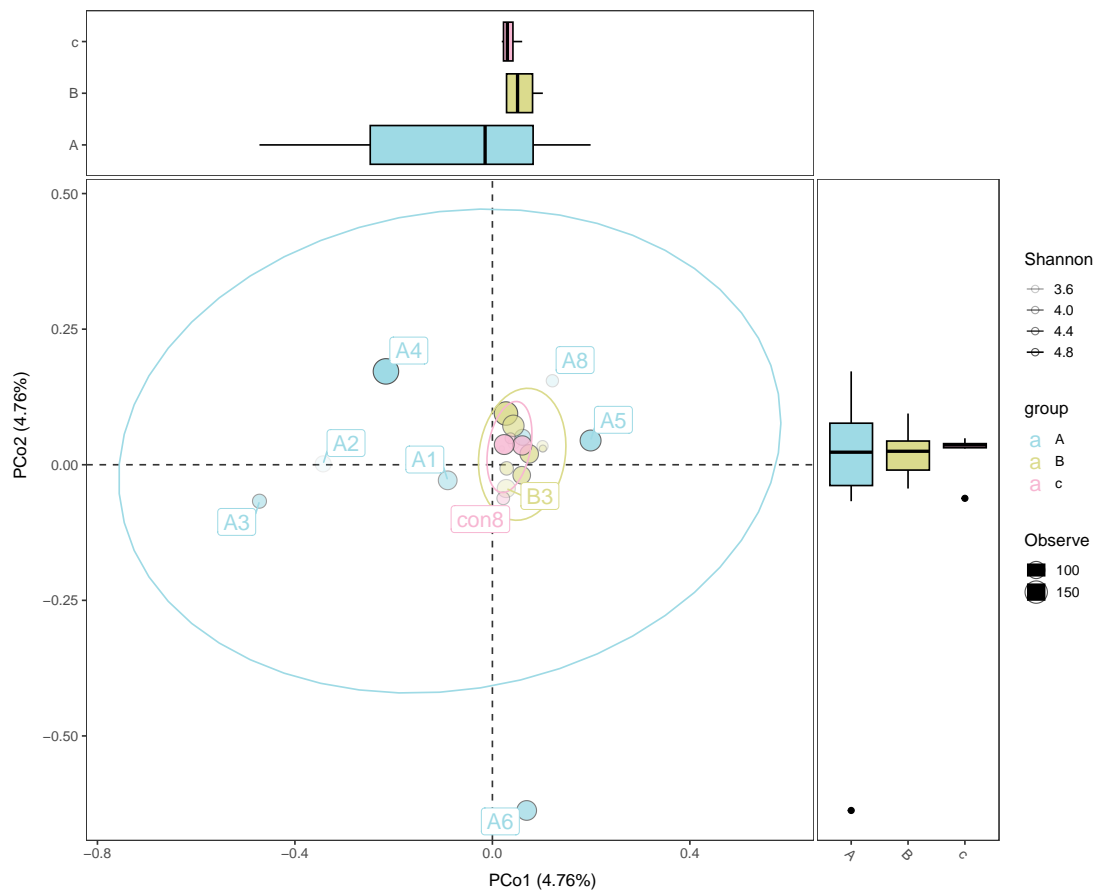


Figure 1: PCoA

#### 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)

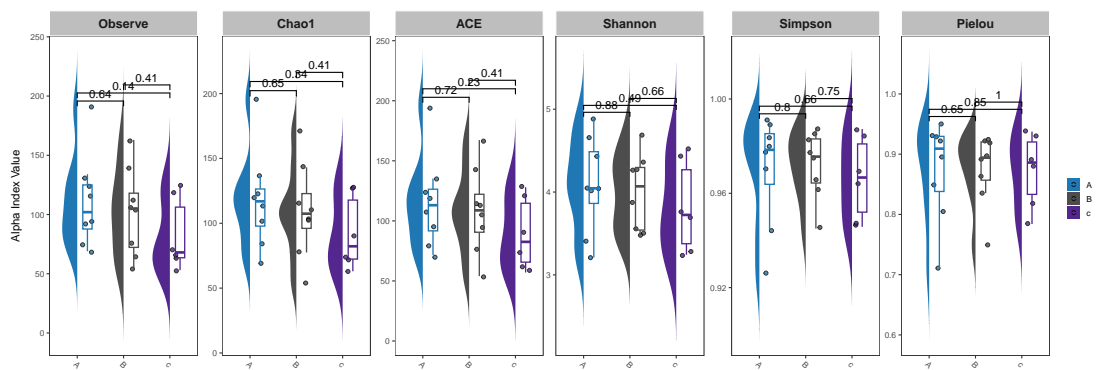


Figure 2: Alpha diversity

‘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)

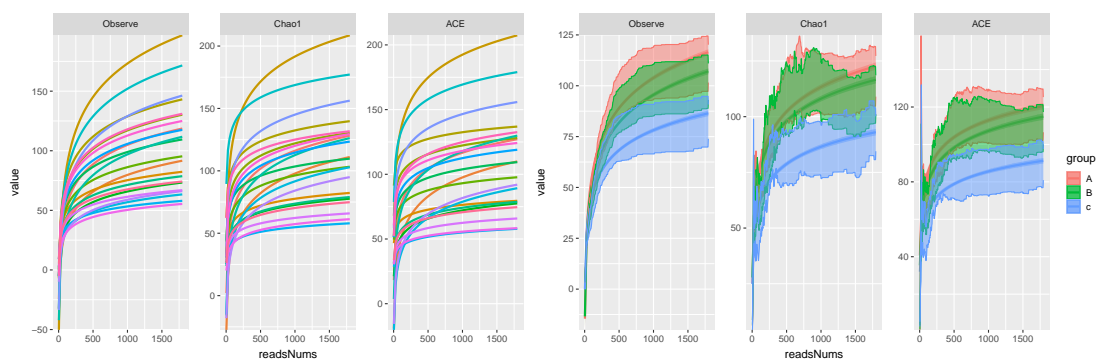


Figure 3: Alpha rarefaction

#### 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)



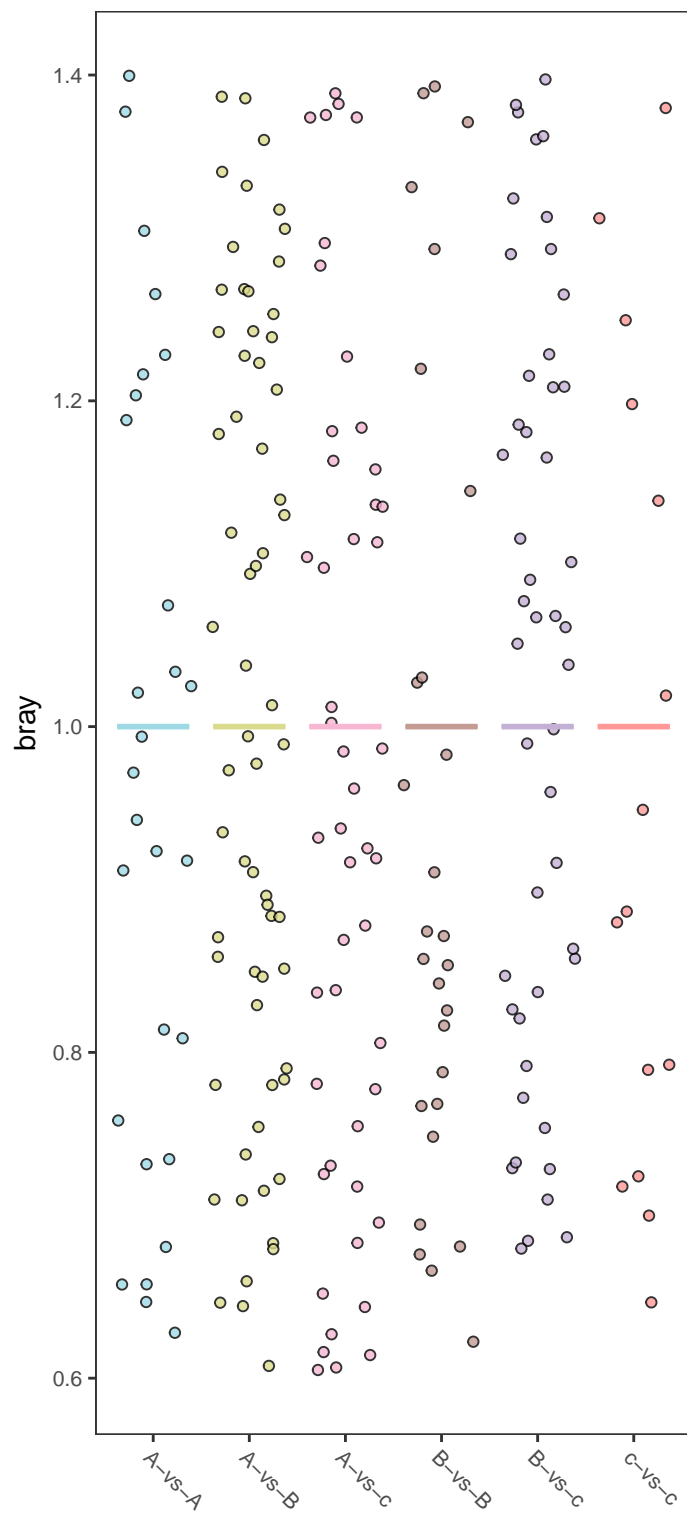


Figure 4: Beta diversity group test

‘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：

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

Figure 5 (下方图) 为图 Ancom test group level 2 volcano 概览。

(对应文件为 Figure+Table/Ancom-test-group-level-2-volcano.pdf)

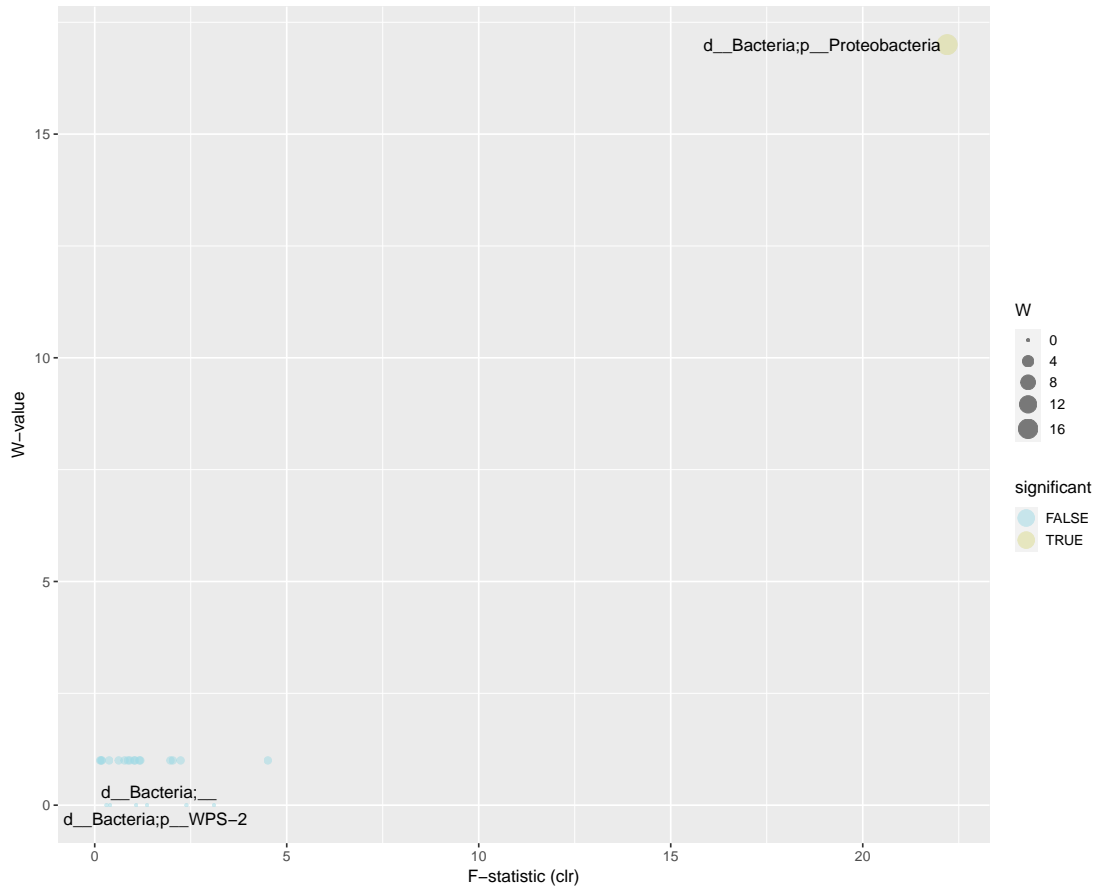


Figure 5: Ancom test group level 2 volcano

Figure 6 (下方图) 为图 Ancom test group level 2 Percentile abundance 概览。

(对应文件为 **Figure+Table/Ancom-test-group-level-2-Percentile-abundance.pdf**)

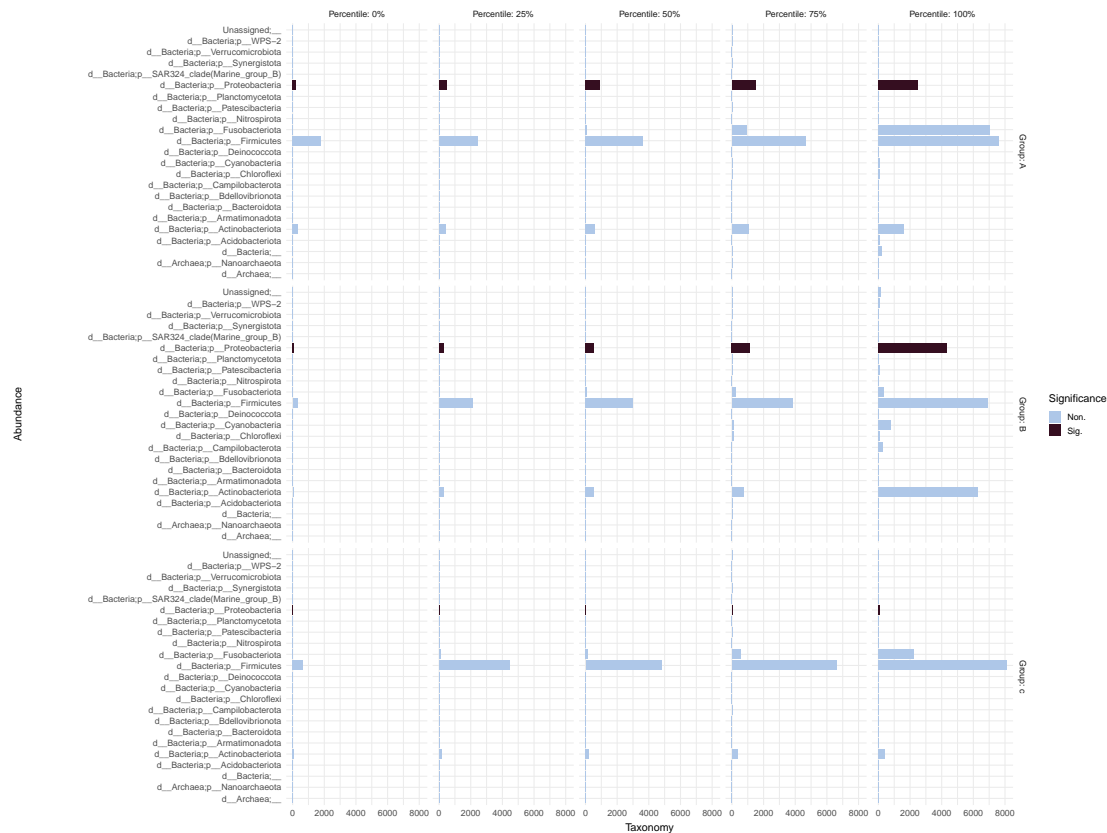


Figure 6: Ancom test group level 2 Percentile abundance

Figure 7 (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**)

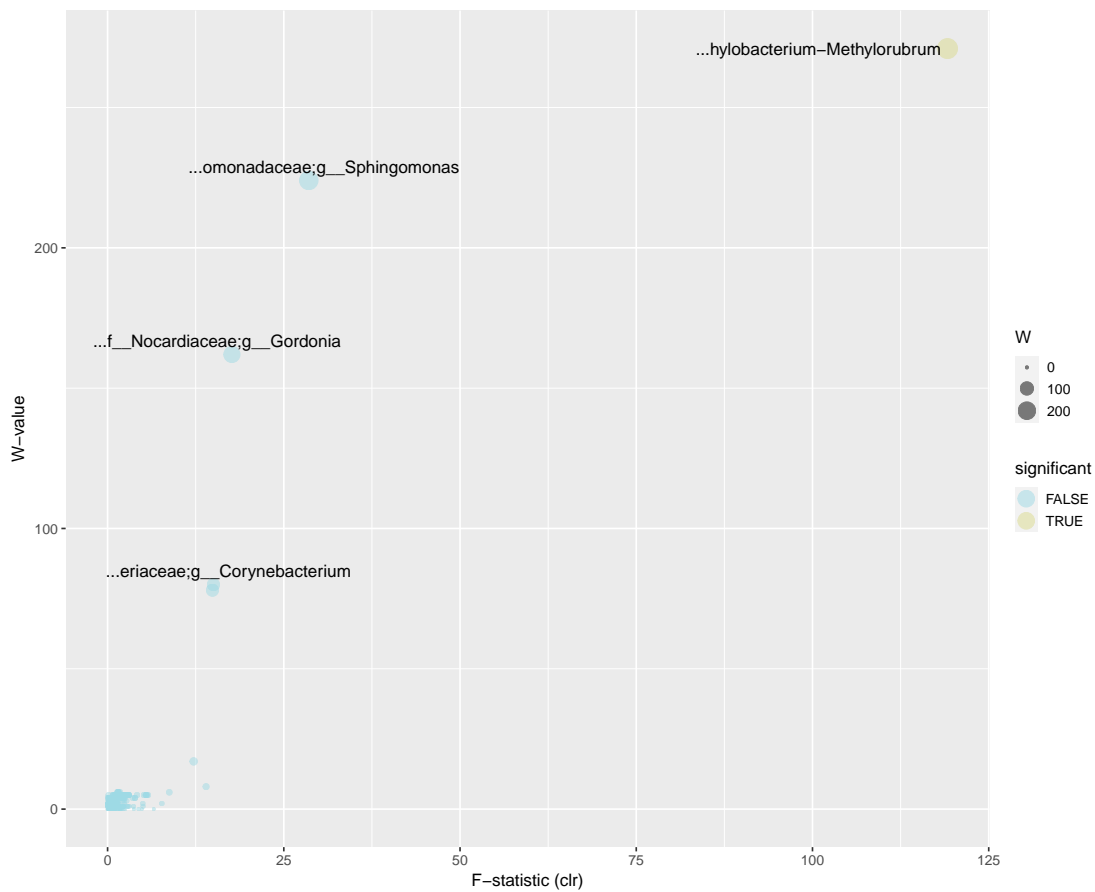


Figure 7: Ancom test group level 6 volcano

Figure 8 (下方图) 为图 Ancom test group level 6 Percentile abundance 概览。

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

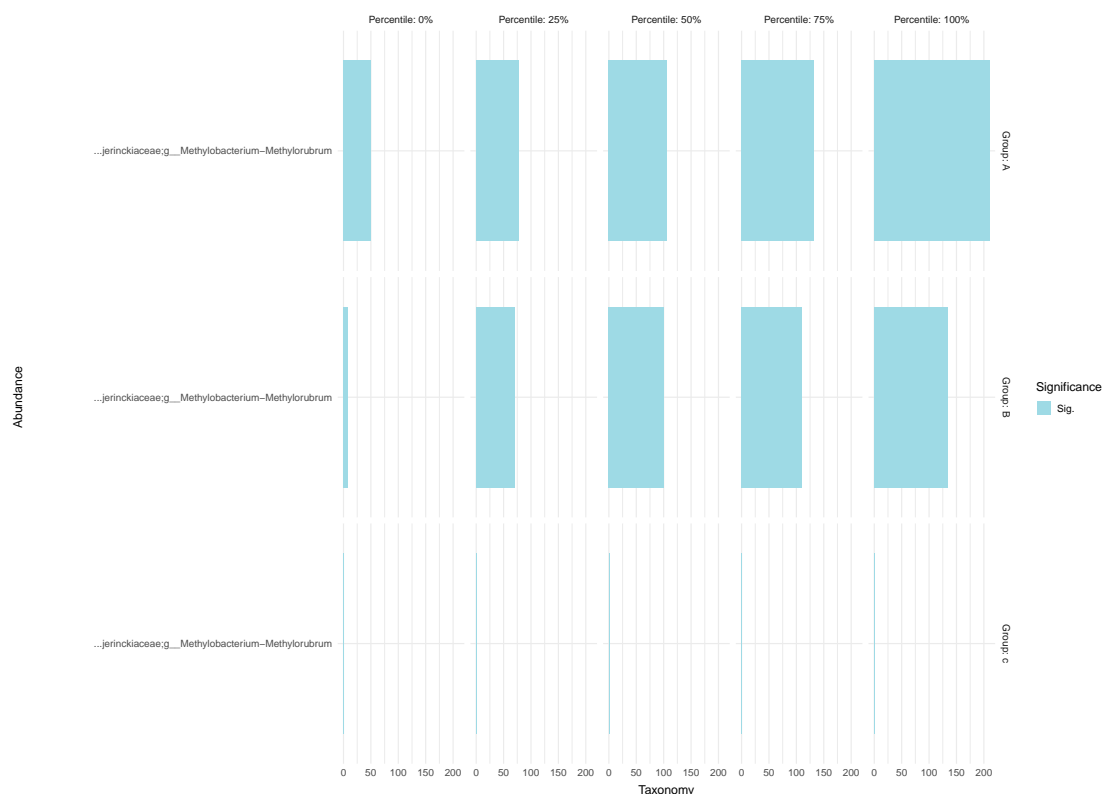


Figure 8: Ancom test group level 6 Percentile abundance

‘level 2’ 对应 Ontology 中的 Phylum。‘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 Percentile abundance’ 数据已全部提供。

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

注：文件夹 Figure+Table/Ancom-test-Percentile-abundance 共包含 5 个文件。

1. 1\_ancom\_test\_group\_level\_2.pdf
2. 2\_ancom\_test\_group\_level\_3.pdf
3. 3\_ancom\_test\_group\_level\_4.pdf
4. 4\_ancom\_test\_group\_level\_5.pdf
5. 5\_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

6.1.5 差异菌关联到代谢物

6.1.5.1 从 gutMDisorder 数据库检索关联代谢物

使用的数据库如下：

Table 2 (下方表格) 为表格 GutMDisorder database 概览。

(对应文件为 **Figure+Table/GutMDisorder-database.xlsx**)

注：表格共有 724 行 12 列，以下预览的表格可能省略部分数据；表格含有 289 个唯一 ‘Gut.Microbiota’。

Table 2: GutMDisorder database

Gut.Mi.....1	Gut.Mi.....2	Gut.Mi.....3	Classi...	Substrate	Substr.....6	Substr.....7	...
Christ...	NA	gm0883	strain	D-Glucose	5793	HMDB00...	...
Christ...	NA	gm0883	strain	Salicin	439503	HMDB00...	...
Christ...	NA	gm0883	strain	D-Xylose	135191	HMDB00...	...
Christ...	NA	gm0883	strain	L-Arab...	439195	HMDB00...	...
Christ...	NA	gm0883	strain	L-Rham...	25310	HMDB00...	...
Christ...	NA	gm0883	strain	D-Mannose	18950	HMDB00...	...
Christ...	NA	gm0883	strain	D-Glucose	5793	HMDB00...	...
Christ...	NA	gm0883	strain	Salicin	439503	HMDB00...	...
Christ...	NA	gm0883	strain	D-Xylose	135191	HMDB00...	...
Christ...	NA	gm0883	strain	L-Arab...	439195	HMDB00...	...
Christ...	NA	gm0883	strain	L-Rham...	25310	HMDB00...	...
Christ...	NA	gm0883	strain	D-Mannose	18950	HMDB00...	...
Entero...	1343173	gm0884	species	Orientin	5281675	HMDB00...	...
Clostr...	29347	gm0885	strain	Bile acid	439520		...
Clostr...	29347	gm0885	strain	Cholic...	221493	HMDB00...	...
...	...	...	...	...	...	...	...

使用差异肠道菌匹配：

**Content :**

Proteobacteria, Alphaproteobacteria, Rhizobiales, Beijerinckiaceae, Methylobacterium-Methylobacterium

未找到相关代谢物。

### 6.1.5.2 尝试从已发表研究 (孟德尔随机化相关) 中寻找关联代谢物

请参考<sup>1</sup>

匹配到 Phylum 水平的菌群关联的代谢物:

**Content :**

5-methyltetrahydrofolic acid, selenium, Cystine, Glutamic acid

Table 3 (下方表格) 为表格 MendelianRandoLiuX2022 matched data 概览。

(对应文件为 **Figure+Table/MendelianRandoLiuX2022-matched-data.csv**)

注: 表格共有 4 行 25 列, 以下预览的表格可能省略部分数据; 表格含有 1 个唯一 ‘X1’。

Table 3: MendelianRandoLiuX2022 matched data

X1	X2	beta	se	p	beta.1	se.1	p.1	p_MRP	beta.2
p_Prot...	5-meth...	-0.15312	0.0313679	1.0532...	-0.097...	0.0417018	0.0188806	0.1756	-0.166...
p_Prot...	selenium	-0.122431	0.0289698	2.3772...	-0.032...	0.0412256	0.431953	0.189	-0.192...
p_Prot...	Cystine	-0.097...	0.0221977	1.0141...	-0.045...	0.0323047	0.16131	0.4752	-0.101...
p_Prot...	Glutam...	0.175275	0.0392992	8.1952...	0.0363522	0.0368746	0.324216	0.0082	0.1929...

### 6.1.6 代谢物的富集分析

将匹配到的代谢物 (6.1.5.2) 进行代谢物富集分析。

以下是代谢物的数据库匹配:

Table 4 (下方表格) 为表格 compounds ID 概览。

(对应文件为 **Figure+Table/compounds-ID.csv**)

注: 表格共有 4 行 7 列, 以下预览的表格可能省略部分数据; 表格含有 4 个唯一 ‘Query’。

Table 4: Compounds ID

Query	Match	HMDB	PubChem	KEGG	SMILES	Comment
5-methylte...	5-Methylte...	HMDB0001396	439234	C00440	CN1C(CNC2=...	1
selenium	Selenium	HMDB0001349	NA	C01529	[Se++]	1
L-cystine	L-Cystine	HMDB0000192	67678	C00491	N[C@@H](CS...	1
Glutamic acid	Glutamic acid	HMDB0000148	33032	C00302	N[C@@H](CC...	1

Figure 9 (下方图) 为图 MetaboAnalyst kegg enrichment 概览。

(对应文件为 [Figure+Table/metabolites\\_ORA\\_dot\\_kegg\\_pathway\\_dpi72.pdf](#))

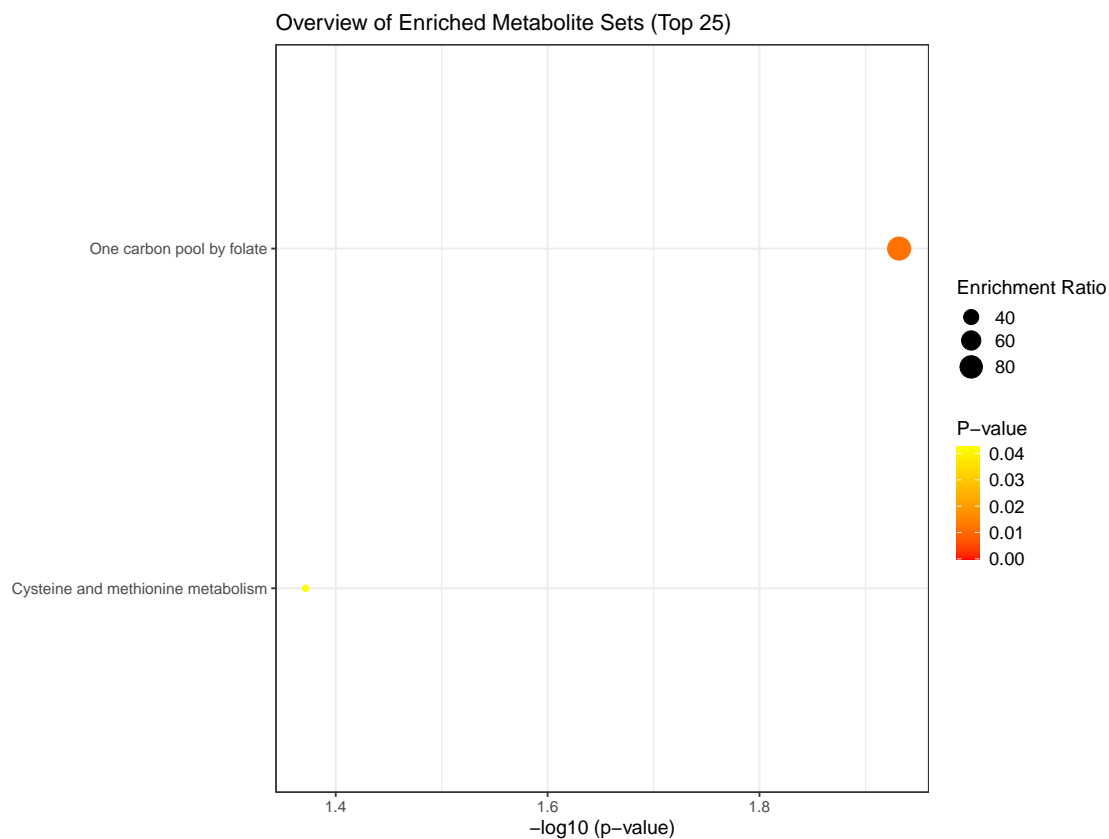


Figure 9: MetaboAnalyst kegg enrichment

Figure 10 (下方图) 为图 Enrichment with algorithm PageRank 概览。

(对应文件为 [Figure+Table/Enrichment-with-algorithm-PageRank.pdf](#))





name	com	NAME	label	input	abbrev.name	type
hsa04974	1	Protein di...	Protein di...	Others	Protein di...	Pathway
hsa04975	1	Fat digest...	Fat digest...	Others	Fat digest...	Pathway
M00017	2	Methionine...	Methionine...	Others	Methionine...	Module
M00170	2	C4-dicarbo...	C4-dicarbo...	Others	C4-dicarbo...	Module
M00171	2	C4-dicarbo...	C4-dicarbo...	Others	C4-dicarbo...	Module
...	...	...	...	...	...	...

### 6.1.7 从结肠炎或结肠癌已发表的代谢物研究中验证

#### 6.1.7.1 DepressionAndYuan2021 结肠炎 (肠道菌)

Depression and anxiety in patients with active ulcerative colitis: crosstalk of gut microbiota, metabolomics and proteomics<sup>13</sup>

以下是整理自该文献的差异肠道菌汇总：

Table 6 (下方表格) 为表格 DepressionAndYuan2021 published data significant microbiota 概览。

(对应文件为 **Figure+Table/DepressionAndYuan2021-published-data-significant-microbiota.xlsx**)

注：表格共有 91 行 4 列，以下预览的表格可能省略部分数据；表格含有 91 个唯一 ‘Taxonomy’。

Table 6: DepressionAndYuan2021 published data significant microbiota

Taxonomy	p.value	MRA.in.UC	MRA.in.HC
p__Gemmatimonadetes	0.00049553252446827	0.00171119907683608	0.000160190585722501
p__Actinobacteria	0.00834784591709094	0.0617033057472256	0.0389112516772091
p__Firmicutes	0.00478686441960355	0.589610871565727	0.499806265231797
p__Bacteroidetes	9.60337999115942e-05	0.250752809840626	0.403687121772228
p__unidentified	1.40986578996555e-07	0.000640301130981834	2.05372545798078e-06
c__Actinobacteria	0.00353055167754535	0.046725454341533	0.0237116295626934
c__Longimicrobia	7.16632518063251e-07	0.00145496100157866	0
c__Bacteroidia	6.38198131772293e-05	0.24518920382048	0.403287329883075
c__Deinococci	3.17450175630301e-06	0.00131123215732429	0
c__Bacilli	0.00330587139663399	0.0649165901033199	0.0154563377967633
c__Cytophagia	2.10636789003367e-05	0.0016664310495747	4.24436594649361e-05
c__unidentified	2.64154914758447e-08	0.0016652538833675	8.83101946931736e-05
c__Flavobacteriia	5.26714172360768e-05	0.00167703418517471	5.95580382814426e-05
o__Oceanospirillales	0.00495867783478051	0.000461817587592187	2.05372545798077e-06
o__Bifidobacteriales	0.00354779153342515	0.0443044462799416	0.0224465346805772
...	...	...	...

未从 Fig. 6 中匹配到客户数据筛选出的肠道菌。

6.1.7.2 AlterationsInScovil2018 结肠炎 (代谢物)

Alterations in Lipid, Amino Acid, and Energy Metabolism Distinguish Crohn’s Disease from Ulcerative Colitis and Control Subjects by Serum Metabolomic Profiling.<sup>14</sup>

以下是整理自该文献的代谢物汇总：

Table 7 (下方表格) 为表格 AlterationsInScovil2018 published data metabolites 概览。

(对应文件为 **Figure+Table/AlterationsInScovil2018-published-data-metabolites.xlsx**)

注：表格共有 565 行 22 列，以下预览的表格可能省略部分数据；表格含有 565 个唯一 ‘Metabolite’。

Table 7: AlterationsInScovil2018 published data metabolites

Metabo...	Pathway	Sub-Pa...	Platform	RI	Mass	MSI Id...	PUBCHEM	KEGG	HMDB
alanine	Amino ...	Alanin...	LC/MS ...	2780.3	88.0404	1	5950	C00041	HMDB00161
aspara...	Amino ...	Alanin...	LC/MS ...	2951.1	131.0462	1	6267	C00152	HMDB00168
aspartate	Amino ...	Alanin...	LC/MS Neg	640	132.0302	1	5960	C00049	HMDB00191
N-acet...	Amino ...	Alanin...	LC/MS ...	1564.8	130.051	1	88064	C02847	HMDB00766
N-acet...	Amino ...	Alanin...	LC/MS ...	785	175.0713	1	99715	NA	HMDB06028
N-acet...	Amino ...	Alanin...	LC/MS ...	3143	174.0408	1	65065	C01042	HMDB00812
creatine	Amino ...	Creati...	LC/MS ...	2920	130.0622	1	586	C00300	HMDB00064
creati...	Amino ...	Creati...	LC/MS ...	2055	114.0662	1	588	C00791	HMDB00562
guanid...	Amino ...	Creati...	LC/MS ...	2884	116.0466	1	763	C00581	HMDB00128
glutamate	Amino ...	Glutam...	LC/MS ...	1500	148.0604	1	611	C00025	HMDB00148
glutamine	Amino ...	Glutam...	LC/MS ...	1291	147.0764	1	5961	C00064	HMDB00641
N-acet...	Amino ...	Glutam...	LC/MS ...	1035	305.098	1	5255	C12270	HMDB01067
N-acet...	Amino ...	Glutam...	LC/MS ...	3106	188.0564	1	70914	C00624	HMDB01138
N-acet...	Amino ...	Glutam...	LC/MS Neg	771	187.0724	1	182230	C02716	HMDB06029
pyrogl...	Amino ...	Glutam...	LC/MS ...	1900	129.0659	2	134508	NA	NA
...	...	...	...	...	...	...	...	...	...

未从 Tab. 7 中匹配到 6.1.5.2 中的关联代谢物。

6.1.7.3 LossOfSymbiotSadegh2024 结肠癌 (肠道菌)

Loss of symbiotic and increase of virulent bacteria through microbial networks in Lynch syndrome colon carcinogenesis<sup>2</sup>

以下是整理自该文献的关联肠道菌汇总：

Table 8 (下方表格) 为表格 LossOfSymbiotSadegh2024 published data microbiota 概览。

(对应文件为 **Figure+Table/LossOfSymbiotSadegh2024-published-data-microbiota.csv**)

注：表格共有 34 行 6 列，以下预览的表格可能省略部分数据；表格含有 34 个唯一 ‘taxon’。

Table 8: LossOfSymbiotSadegh2024 published data microbiota

taxon	proportion...	position.i...	mean.auc.....	sd.auc..test.	Value.with...
k_Bacteria...	0.992	numerator	1	0	stool
k_Bacteria...	0.998	numerator	1	0	stool
k_Bacteria...	0.956	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	0.982	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	0.996	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
k_Bacteria...	1	numerator	1	0	stool
...	...	...	...	...	...

匹配到的肠道菌 (Phylum 水平):

Table 9 (下方表格) 为表格 LossOfSymbiotSadegh2024 matched Phylum microbiota 概览。

(对应文件为 **Figure+Table/LossOfSymbiotSadegh2024-matched-Phylum-microbiota.csv**)

注：表格共有 3 行 6 列，以下预览的表格可能省略部分数据；表格含有 3 个唯一 ‘taxon’。

Table 9: LossOfSymbiotSadegh2024 matched Phylum microbiota

taxon	proportion...	position.i...	mean.auc.....	sd.auc..test.	Value.with...
k_Bacteria...	1	denominator	1	0	stool
k_Bacteria...	1	denominator	1	0	stool
k_Bacteria...	1	denominator	1	0	stool

6.1.7.4 IntegratedAnalChen2022 结肠癌 (肠道菌和代谢物)

Integrated analysis of the faecal metagenome and serum metabolome reveals the role of gut microbiome-associated metabolites in the detection of colorectal cancer and adenoma<sup>15</sup>

以下是整理自该文献的肠道菌和代谢物数据 (PDF 识别结果):

Table 10 (下方表格) 为表格 IntegratedAnalChen2022 published data microbiota 概览。

(对应文件为 Figure+Table/IntegratedAnalChen2022-published-data-microbiota.csv)

注：表格共有 782 行 4 列，以下预览的表格可能省略部分数据；表格含有 3 个唯一 ‘Type’。

1. pvalue: 显著性 P。

Table 10: IntegratedAnalChen2022 published data microbiota

Type	Species	Metabolites	pvalue
Tumor_promoting_b...	Alistipes_finegoldii	X23.4_476.011mz_pos	3.87e-07
Tumor_promoting_b...	Alistipes_finegoldii	X21.4_494.68mz_pos	0.000232816
Tumor_promoting_b...	Alistipes_finegoldii	X24.5_504.692mz_pos	0.000162744
Tumor_promoting_b...	Alistipes_finegoldii	X26.1_509.03mz_pos	0.000793675
Tumor_promoting_b...	Alistipes_finegoldii	X26.3_514.705mz_pos	0.000881926
Tumor_promoting_b...	Bilophila_wadswor...	X23.4_476.011mz_pos	4.45e-05
Tumor_promoting_b...	Fusobacterium_nuc...	X21.2_512.336mz_neg	0.000581862
Tumor_promoting_b...	Fusobacterium_nuc...	X19.2_536.299mz_neg	0.000431165
Tumor_promoting_b...	Fusobacterium_sp...	X21.2_512.336mz_neg	0.000850936
Tumor_promoting_b...	Fusobacterium_sp...	X19.2_536.299mz_neg	9.61e-05
Tumor_promoting_b...	Odoribacter_splan...	X23.4_476.011mz_pos	1.93e-06
Tumor_promoting_b...	Odoribacter_splan...	X21.4_494.68mz_pos	3.38e-05
Tumor_promoting_b...	Odoribacter_splan...	X24.8_495.024mz_pos	0.000618015
Tumor_promoting_b...	Odoribacter_splan...	X22.9_499.686mz_pos	0.000395014
Tumor_promoting_b...	Odoribacter_splan...	X24.5_504.692mz_pos	0.000898428
...	...	...	...

Table 11 (下方表格) 为表格 IntegratedAnalChen2022 published data metabolites 概览。

(对应文件为 Figure+Table/IntegratedAnalChen2022-published-data-metabolites.xlsx)

注：表格共有 969 行 19 列，以下预览的表格可能省略部分数据；表格含有 905 个唯一 ‘V1’。

Table 11: IntegratedAnalChen2022 published data metabolites

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
Supple...	material	placed	on	this	supple...	material	which	has	been
Table	S2.	List	of	colore...	abnormal	associ...	metabo...		
Metabo...	feature	Metabo...	annota...	approach	F_meanN	F_meanC	F_meanA	anova_...	anova_...

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
X10.3_...	(-)-Fo...	exact	mass	match	7748.8...	5053.5...	5527.3...	0.0000...	7.36E-08
X12.4_...	(-)-Or...	exact	mass	match	45226....	31316....	27721....	0.0000...	0.0000327
X10.9_...	(-)-tr...	glucos...	mass	match	364.39...	8542.0151	9965.7...	0.0026...	0.0009...
X15_26...	(+)-2,...	Goniot...	mass	match	32595....	22986....	25975....	0.0012...	0.0013...
X20.8_...	(+/-)-...	mass	match	6812.3...	10133....	3573.2...	0.0042...	0.0002535	0.1164...
X12.5_...	(±)15-...	exact	mass	match	11931....	8063.3...	8309.9...	3.99E-10	3.99E-10
X14.9_...	(±)-Go...	exact	mass	match	24851....	16540....	15696....	0.0022...	0.0001...
X15.1_...	(±)-Pa...	exact	mass	match	78154....	37693....	61456....	0.0000...	0.0004...
X24_35...	(1S)-1...	mass	match	2067.0...	6783.2...	10613....	0.0000...	0.0001...	0.5843...
X20.7_...	(20R)-...	Rh2	exact	mass	match	19519....	12756....	12327....	4.39E-10
X11_42...	(22E)-...								
exact	mass	match	97120....						
...	...	...	...	...	...	...	...	...	...

未从上述数据中匹配到客户数据的差异肠道菌或其关联的代谢物。

## Reference

1. Liu, X. *et al.* Mendelian randomization analyses support causal relationships between blood metabolites and the gut microbiome. *Nature Genetics* **54**, (2022).
2. Sadeghi, M. *et al.* Loss of symbiotic and increase of virulent bacteria through microbial networks in lynch syndrome colon carcinogenesis. *Frontiers in oncology* **13**, (2024).
3. Picart-Armada, S., Fernandez-Albert, F., Vinaixa, M., Yanes, O. & Perera-Lluna, A. FELLA: An r package to enrich metabolomics data. *BMC Bioinformatics* **19**, 538 (2018).
4. Chen, S. Ultrafast one-pass fastq data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, (2023).
5. Cheng, L., Qi, C., Zhuang, H., Fu, T. & Zhang, X. GutMDisorder: A comprehensive database for dysbiosis of the gut microbiota in disorders and interventions. *Nucleic Acids Research* **48**, (2019).
6. Xu, S. *et al.* MicrobiotaProcess: A comprehensive r package for deep mining microbiome. *The Innovation* **4**, 100388 (2023).
7. Chong, J. *et al.* MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research* **46**, W486–W494 (2018).
8. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using qiime 2. *Nature Biotechnology* **37**, 852–857 (2019).
9. 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).

10. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from illumina amplicon data. *Nature methods* **13**, 581 (2016).
11. 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).
12. Hamday, M. & Knight, R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research* **19**, 1141–1152 (2009).
13. Yuan, X. *et al.* Depression and anxiety in patients with active ulcerative colitis: Crosstalk of gut microbiota, metabolomics and proteomics. *Gut microbes* **13**, (2021).
14. Scoville, E. A. *et al.* Alterations in lipid, amino acid, and energy metabolism distinguish crohn’s disease from ulcerative colitis and control subjects by serum metabolomic profiling. *Metabolomics : Official journal of the Metabolomic Society* **14**, (2018).
15. Chen, F. *et al.* Integrated analysis of the faecal metagenome and serum metabolome reveals the role of gut microbiome-associated metabolites in the detection of colorectal cancer and adenoma. *Gut* **71**, 1315–1325 (2022).