

开源数据库结肠炎和结肠癌的菌群差异分析

2024-03-15

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 数据来源	1
6.1.1 结肠炎数据	1
6.1.1.1 TargetedSuppreFederi2022 结肠炎	1
6.1.1.2 DepressionAndYuan2021 结肠炎	2
6.1.1.3 TransplantationSinha2022 结肠炎	2
6.1.2 结肠癌数据	3
6.1.2.1 AnIntegratedTRoelan2023 结肠癌	3
6.1.2.2 LocationAndCoSambru2023 结肠癌	4
6.2 微生物注释	4
Reference	4

List of Figures

List of Tables

1	DepressionAndYuan2021 data	2
2	TransplantationSinha2022 data	2
3	AnIntegratedTRoelan2023 data	3
4	LocationAndCoSambru2023 data	4

1 摘要

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 TransplantationSinha2022¹.
- Supplementary file from article refer to TargetedSuppreFederi2022².
- Supplementary file from article refer to DepressionAndYuan2021³.
- Supplementary file from article refer to AnIntegratedTROelan2023⁴.
- Supplementary file from article refer to LocationAndCoSambru2023⁵.

3.2 方法

Mainly used method:

- R version 4.3.2 (2023-10-31); 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 数据来源

6.1.1 结肠炎数据

6.1.1.1 TargetedSuppreFederi2022 结肠炎

‘TargetedSuppreFederi2022 data’ 数据已全部提供。

(对应文件为 **Figure+Table/TargetedSuppreFederi2022-data**)

注：文件夹 Figure+Table/TargetedSuppreFederi2022-data 共包含 4 个文件。

1. 1_Corrected France.csv
2. 2_Corrected Israel.csv
3. 3_Corrected US.csv
4. 4_Corrected Germany.csv

6.1.1.2 DepressionAndYuan2021 结肠炎

Table 1 (下方表格) 为表格 DepressionAndYuan2021 data 概览。

(对应文件为 Figure+Table/DepressionAndYuan2021-data.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 1: DepressionAndYuan2021 data

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
...

6.1.1.3 TransplantationSinha2022 结肠炎

Table 2 (下方表格) 为表格 TransplantationSinha2022 data 概览。

(对应文件为 Figure+Table/TransplantationSinha2022-data.csv)

注：表格共有 141 行 12 列，以下预览的表格可能省略部分数据；表格含有 83 个唯一 ‘W’。

Table 2: TransplantationSinha2022 data

W	detect.....2	detect.....3	detect.....4	detect.....5	Kingdom	Phylum	Class	Order	Family	...
121	FALSE	TRUE	TRUE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Rikene...	...
89	FALSE	FALSE	FALSE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Rikene...	...
87	FALSE	FALSE	FALSE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Marini...	...
94	FALSE	FALSE	FALSE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Barnes...	...

W	detect.....2	detect.....3	detect.....4	detect.....5	Kingdom	Phylum	Class	Order	Family	...
125	FALSE	TRUE	TRUE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Barnes...	...
94	FALSE	FALSE	FALSE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Barnes...	...
95	FALSE	FALSE	FALSE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Barnes...	...
135	TRUE	TRUE	TRUE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Prevot...	...
138	TRUE	TRUE	TRUE	TRUE	d__Bac...	Bacter...	Bacter...	Bacter...	Prevot...	...
103	FALSE	FALSE	TRUE	TRUE	d__Bac...	Firmic...	Bacilli	Lactob...	Entero...	...
134	TRUE	TRUE	TRUE	TRUE	d__Bac...	Firmic...	Bacilli	Erysip...	Erysip...	...
118	FALSE	TRUE	TRUE	TRUE	d__Bac...	Firmic...	Bacilli	Erysip...	Erysip...	...
90	FALSE	FALSE	FALSE	TRUE	d__Bac...	Firmic...	Bacilli	Erysip...	Erysip...	...
87	FALSE	FALSE	FALSE	TRUE	d__Bac...	Firmic...	Bacilli	Erysip...	Erysip...	...
95	FALSE	FALSE	FALSE	TRUE	d__Bac...	Firmic...	Bacilli	Erysip...	Erysip...	...
...

6.1.2 结肠癌数据

6.1.2.1 AnIntegratedTRoelan2023 结肠癌

Table 3 (下方表格) 为表格 AnIntegratedTRoelan2023 data 概览。

(对应文件为 **Figure+Table/AnIntegratedTRoelan2023-data.xlsx**)

注：表格共有 74 行 15 列，以下预览的表格可能省略部分数据；表格含有 1 个唯一 ‘sheet’。

Table 3: AnIntegratedTRoelan2023 data

sheet	Taxonomy	p_val	FDR	wilcox...	mean_N	mean_T	median_N	median_T	Direction
Supple...	D_0__B...	2.8875...	4.2447...	1965	0.0121...	0.0524...	0.0006...	0.0059...	Enrich...
Supple...	D_0__B...	6.6405...	4.8807...	657	0.0018...	0.0095...	0	0	Enrich...
Supple...	D_0__B...	7.1087...	3.4833...	21023.5	0.0190...	0.0126...	0.0128...	0.0067...	Enrich...
Supple...	D_0__B...	2.8160...	1.0348...	18604	0.0132...	0.0087...	0.0078...	0.0041...	Enrich...
Supple...	D_0__B...	1.7251...	5.0718...	11985	0.0070...	0.0043...	0.0034...	0.0019...	Enrich...
Supple...	D_0__B...	5.0924...	1.2476...	7227.5	0.0126...	0.0315...	0.0031...	0.0076...	Enrich...
Supple...	D_0__B...	9.2218...	1.9365...	223.5	0.0002...	0.0052...	0	0	Enrich...
Supple...	D_0__B...	1.5535...	2.8547...	16631.5	0.0106...	0.0076...	0.0064...	0.0041...	Enrich...
Supple...	D_0__B...	2.1475...	3.5075...	1494	0.0022...	0.0076...	0	0	Enrich...
Supple...	D_0__B...	4.3690...	6.2979...	62	0.0001...	0.0026...	0	0	Enrich...
Supple...	D_0__B...	4.7127...	6.2979...	20341	0.0280...	0.0232...	0.0228...	0.0170...	Enrich...
Supple...	D_0__B...	7.9474...	9.7356...	237.5	0.0008...	0.0058...	0	0	Enrich...
Supple...	D_0__B...	1.6820...	1.9019...	4806.5	0.0040...	0.0026...	0	0	Enrich...
Supple...	D_0__B...	1.5690...	1.5376...	4009	0.0025...	0.0039...	0.0006...	0.0013...	Enrich...
Supple...	D_0__B...	1.5362...	1.5376...	12107.5	0.0075...	0.0059...	0.0017...	0.0012...	Enrich...

sheet	Taxonomy	p_val	FDR	wilcox...	mean_N	mean_T	median_N	median_T	Direction
...

6.1.2.2 LocationAndCoSambru2023 结肠癌

Table 4 (下方表格) 为表格 LocationAndCoSambru2023 data 概览。

(对应文件为 `Figure+Table/LocationAndCoSambru2023-data.csv`)

注：表格共有 44 行 2 列，以下预览的表格可能省略部分数据；表格含有 44 个唯一 ‘tax_id’。

Table 4: LocationAndCoSambru2023 data

tax_id	taxon_name
40545	Sutterella_wadsworthensis
214856	Alistipes_finegoldii
328814	Alistipes_shahii
674529	Bacteroides_faecis
333367	[Clostridium]_asparagiforme
437898	Sutterella_parvirubra
74426	Collinsella_aerofaciens
1531	[Clostridium]_clostridioforme
239935	Akkermansia_muciniphila
901	Desulfovibrio_piger
1892897	Shigella_sp._FC569
68259	Streptomyces_purpurogeneisc...
1450439	Bacteroides_sp._UW
585543	Bacteroides_sp._D20
1581131	Actinomyces_sp._HMSC08A01
...	...

6.2 微生物注释

Reference

1. Sinha, A. *et al.* Transplantation of bacteriophages from ulcerative colitis patients shifts the gut bacteriome and exacerbates the severity of dss colitis. *Microbiome* **10**, (2022).
2. Federici, S. *et al.* Targeted suppression of human ibd-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell* **185**, 2879–2898.e24 (2022).
3. Yuan, X. *et al.* Depression and anxiety in patients with active ulcerative colitis: Crosstalk of gut microbiota, metabolomics and proteomics. *Gut microbes* **13**, (2021).

4. Roelands, J. *et al.* An integrated tumor, immune and microbiome atlas of colon cancer. *Nature medicine* **29**, 1273–1286 (2023).
5. Sambruni, G. *et al.* Location and condition based reconstruction of colon cancer microbiome from human rna sequencing data. *Genome medicine* **15**, (2023).