

# 肠宇宙数据库 (GutUDB): 一个全面的肠道疾病多组学数据库

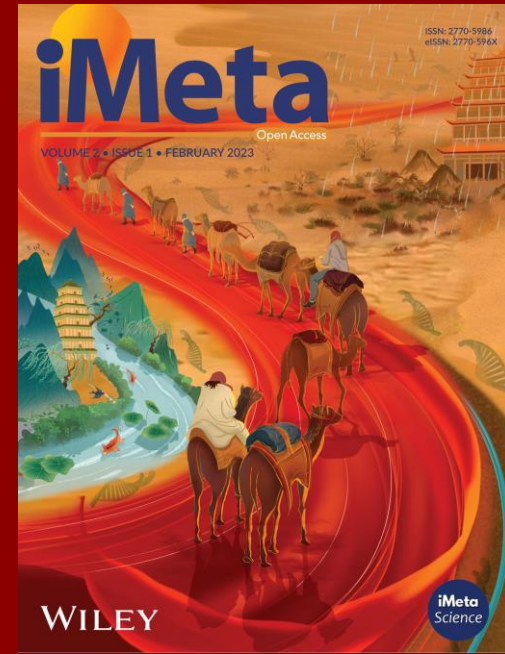
包苡<sup>1</sup>, 陈亚鑫<sup>2</sup>, 林丽珠<sup>3</sup>, 李静怡<sup>4</sup>

<sup>1</sup>广西医科大学艾滋病防治研究重点实验室

<sup>2</sup>四川大学疾病分子网络前沿科学中心

<sup>3</sup>钦州市第一人民医院麻醉科

<sup>4</sup>广东省第二人民医院病理科

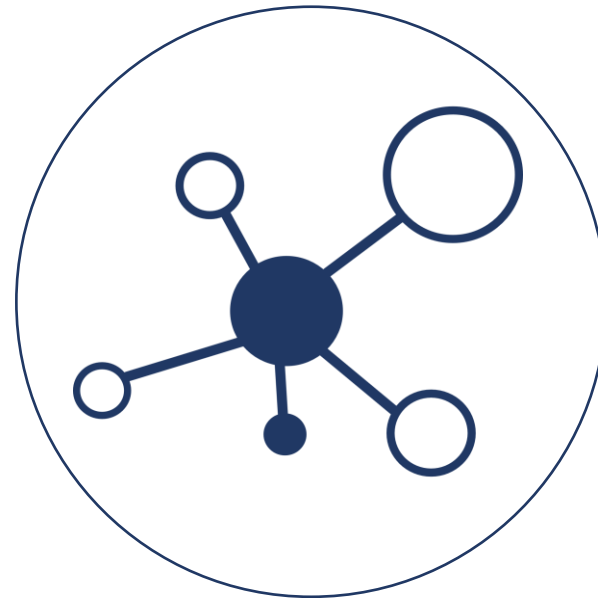


Bao Yi, YaxinChen, LizhuLin, JingyiLi, XinliLiu, GangWang, Yueqi Li, etal. 2024.  
“GutUDB: A Comprehensive Multiomics Database for Intestinal Diseases.” *iMeta* 3: e195.  
<https://doi.org/10.1002/imt2.195>



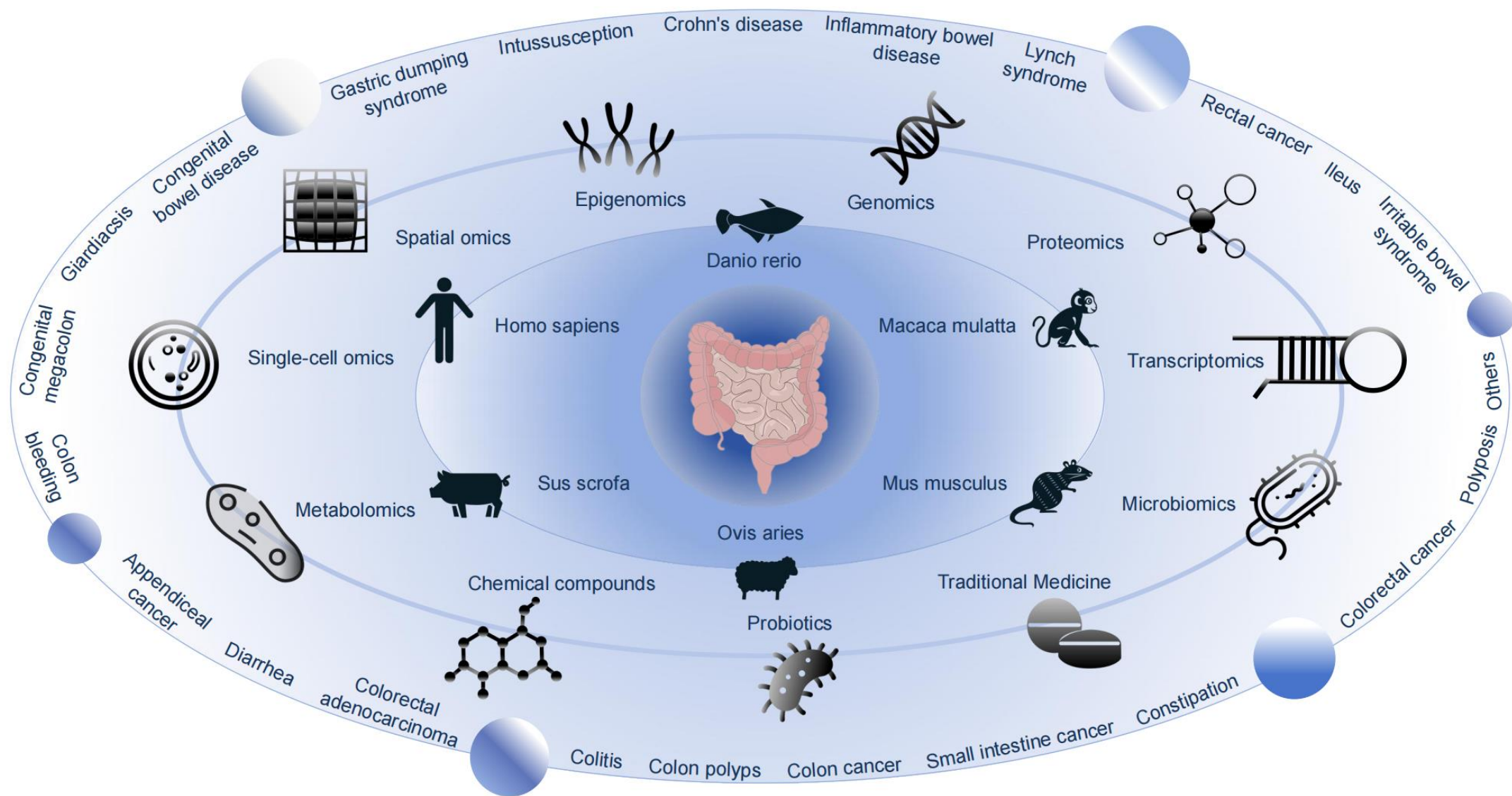
# 引言

高通量测序：  
RNA-seq  
ChIP-seq  
scRNA-seq  
MeRIP-seq





# 引言

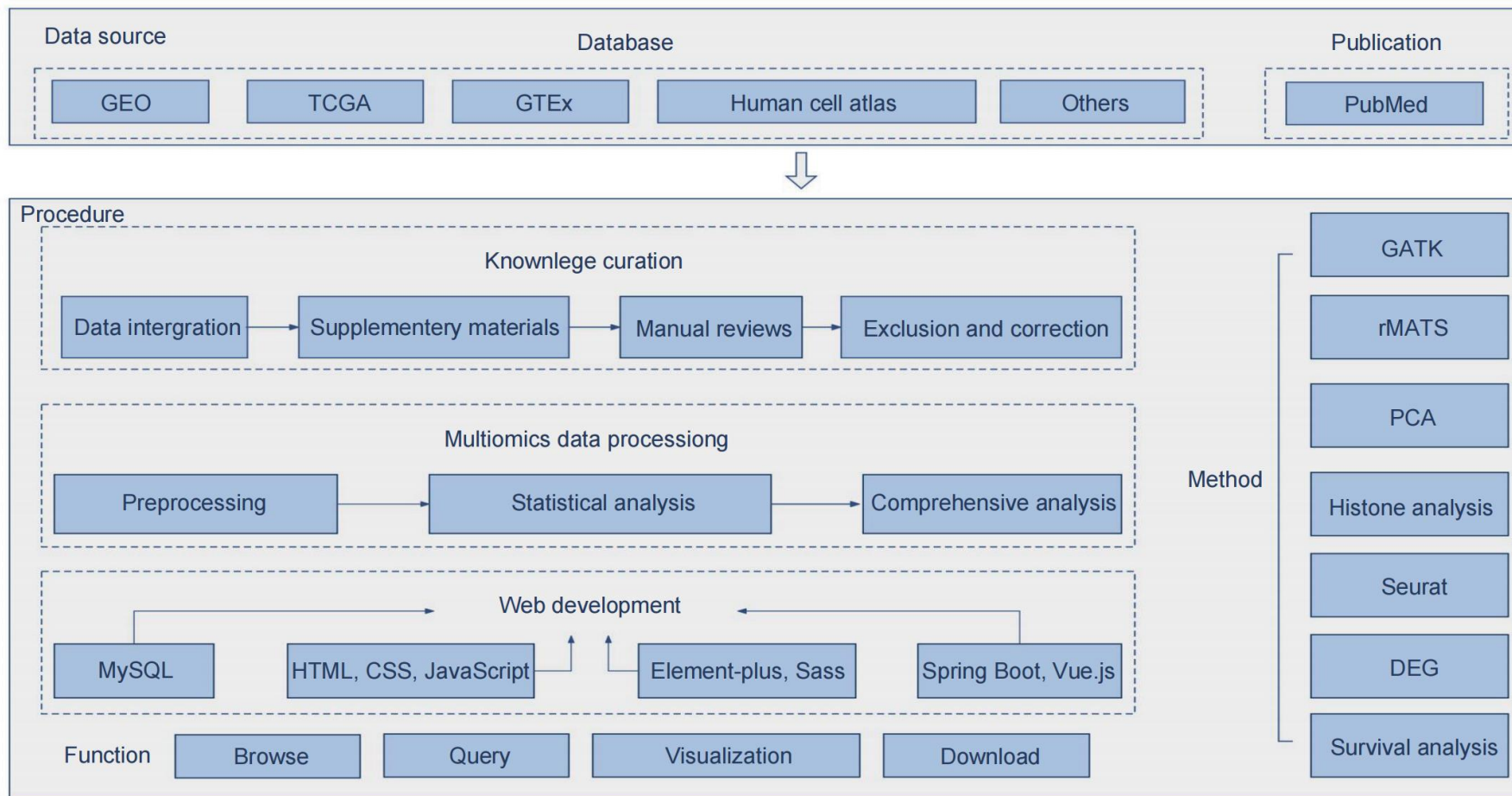


肠宇宙数据库 (GutUDB)：一个全面的肠道疾病多组学数据库



# 结果: GutUDB 概述

(A)



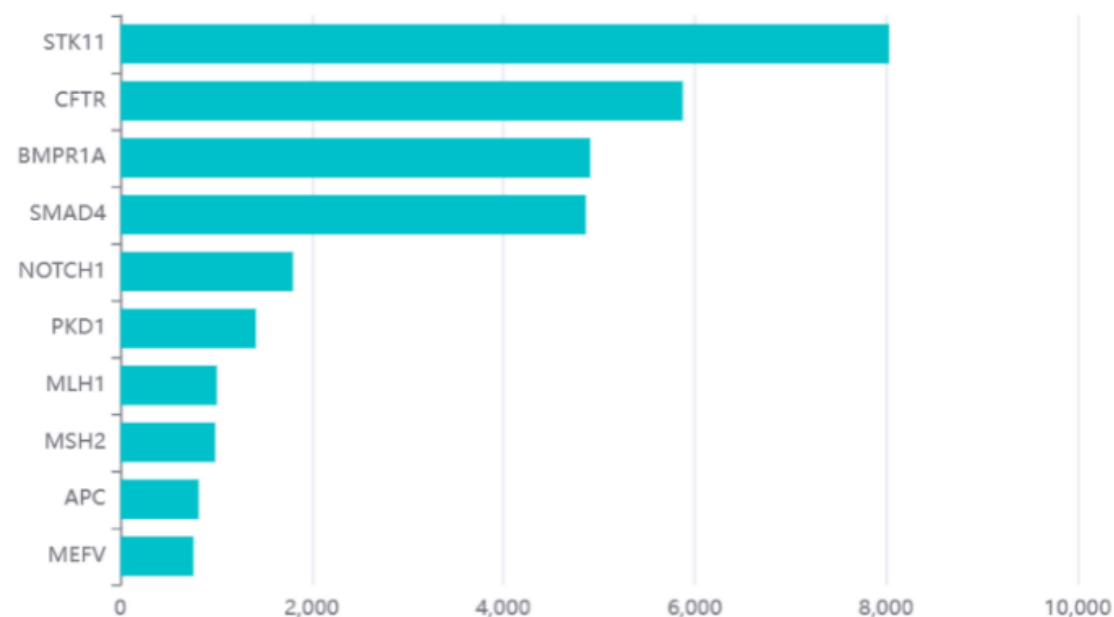
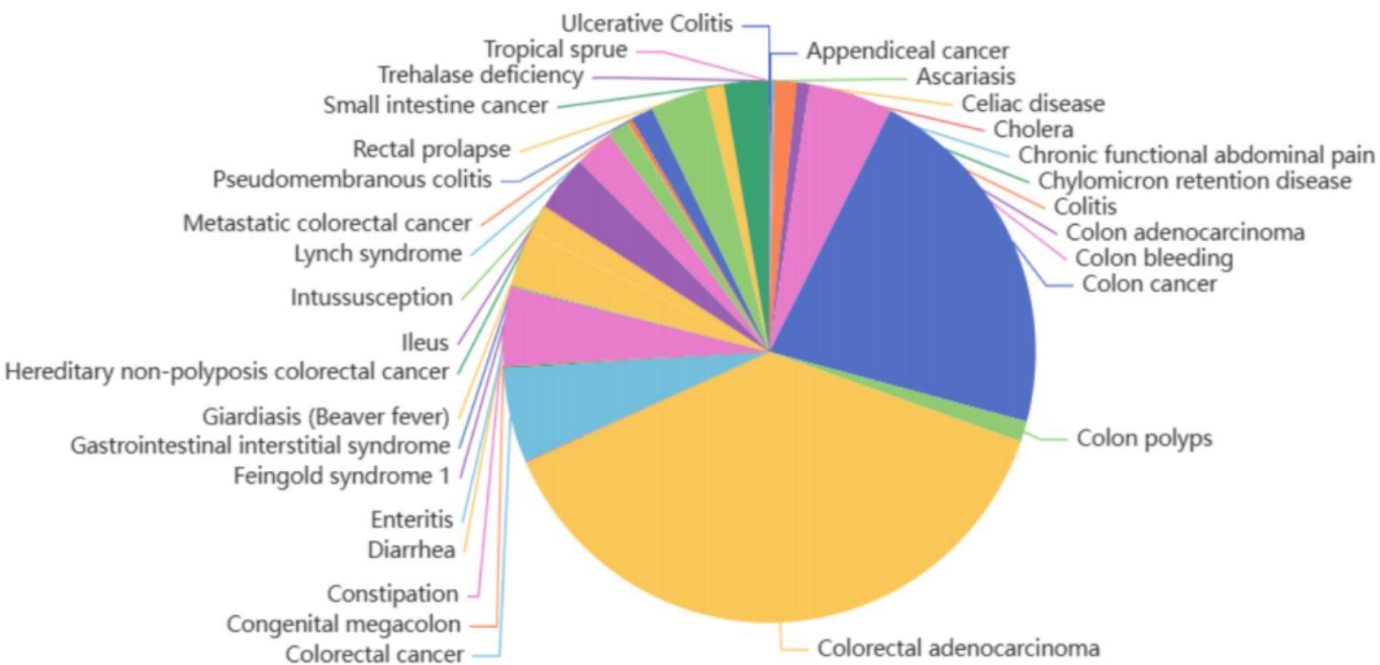
- 八大组学
- 6类物种
- 56种肠道疾病
- 11种肠道组织
- 63种肠道细胞系
- 58970个基因

图1 (A) 数据收集和处理过程, 以及数据库网站的构建过程, 包括四个核心功能 (浏览、查询、可视化和下载)。



# 结果: 不同肠道疾病之间的疾病基因关联

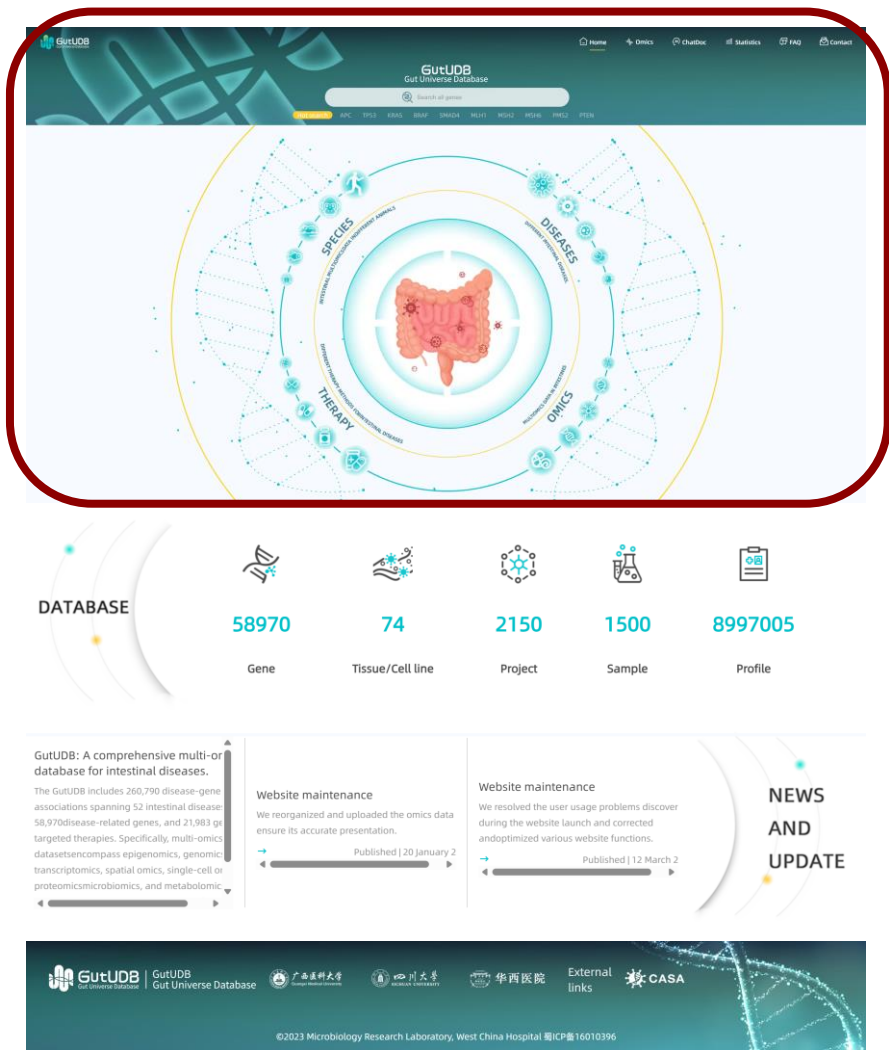
260790 个疾病基因关联



图S1 (A) 各种肠道疾病的比例。

图S1 (B) 与基因疾病相关的前10个基因。

# 结果: 主页



四大模块: 组学, 物种, 疾病, 治疗



# 结果: GutADB 概述

Gene Name: **METTL3**

Start: 21498133

🔍 RBP: Yes

🔍 RNA editing: NO

🔍 Therapeutic targets: Yes

Ensemble ID: ENSG00000165819

End: 21511342

🔍 m6A: YES

🔍 TF: Yes

🔍 Diagnostic targets:

Seqnames: chr14

Strand: -

🔍 AS: NO

Motif: NA

[NCBI](#)

## 基因详情信息

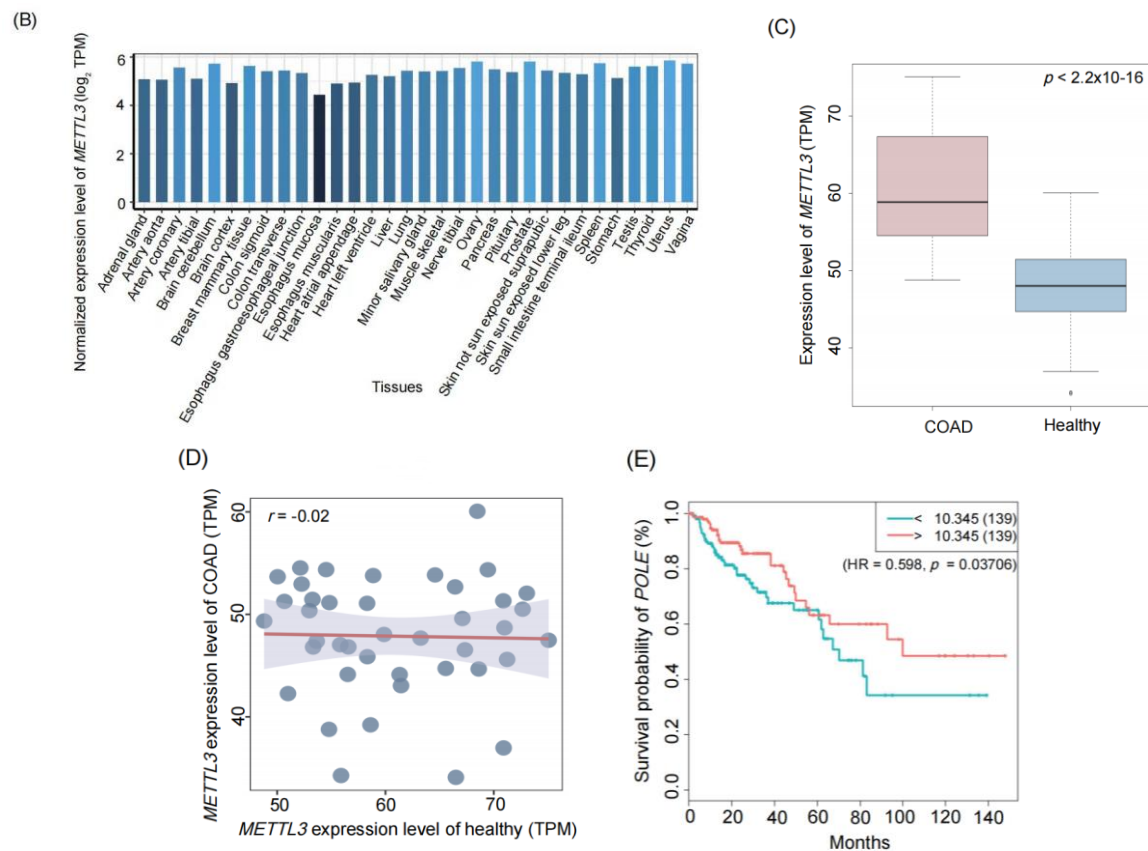


图1. (B) METTL3 在不同组织中的表达水平，数据来源于GTEx。(C) 箱线图显示结肠癌组织和健康组织的基因表达差异。(D) 散点图显示结肠癌组和健康组之间METTL3基因表达的相关性分析。(E) 结肠癌相关基因POLE 的生存分析曲线。



# 结果: GutUDB 概述

profile | omics | drugs

## Epigenomics

### DNA Methylation

Epigenomics | Genomics | Transcriptomics | Spatial Omics | Singel Cell Omics | Proteomics | Microbiomics | Metabolomics

DNA Methylation | Genomic Alteration | Non-coding RNA | Gene Expression | Associated Genera

CNA Genes | circRNA | Alternative Splicing

Mutated Genes | Inc RNA

SNP | miRNA

Structural Variant Genes

Virulence Gene

### Genomic Alteration: CNA Genes

Disease	Is Cancer Gene (source: OncoKB)	Cytoband	CNA	Freq	PMID
Colon Cancer	No	14q11.2	HOMDEL	0.048	<a href="#">31031003</a>



图S1 (D) 基因相关的八个主要组学信息。

# 结果: 治疗模块

Chemical Compounds

Traditional Medicine

Probiotics

## Chemical Compounds

#	Disease related genes	Disease	Species	Direct Evidence
1	ABCC4	Inflammatory Bowel Dis...	Human	NA
2	ABCC4	Colonic Neoplasms	Human	NA
3	ADRA2A	Colonic Neoplasms	Human	NA
4	ABCC4	Intestinal Pseudo-Obstr...	Human	NA
5	ABCC4	Pediatric Crohn's disease	Human	NA
6	ADRA2A	Enteritis	Human	NA

- 6281种化合物
- 393种传统药物
- 22种益生菌

图S1 (E) 治疗部分包含化合物、传统药物和益生菌。



# 结果: 疾病模块

## Intestinal Diseases

**Diseases name :** [All\(24388\)](#) [Colon Adenocarcinoma\(3450\)](#) [Colitis\(2711\)](#) [Intussusception\(2606\)](#) [Colorectal Neoplasms\(1734\)](#)  
[Colonic Neoplasms\(1662\)](#) [Intestinal Neoplasms\(1469\)](#) [Inflammatory Bowel Diseases\(1366\)](#) [Enteritis\(1308\)](#)  
[Duodenal Ulcer\(1175\)](#) [Crohn Disease\(1165\)](#) [Anus Neoplasms\(1164\)](#) [Ileus\(1140\)](#) [Rectal Neoplasms\(950\)](#) [Ileitis\(668\)](#)  
[Megacolon\(522\)](#) [Intestinal obstruction\(483\)](#) [Proctocolitis\(425\)](#) [Duodenal Neoplasms\(390\)](#)

**Omics level:** [All](#) [Genomic](#) [Transcriptomic](#) [Proteomics](#) [Metabolomics](#) [Single cell omics](#) [Spatial omics](#) [Epigenomics](#)  
[Microbiomics](#)

**Hot genes:** [All\(55970\)](#) [APC\(9906\)](#) [STK11\(9471\)](#) [MEN1\(5750\)](#) [SMAD4\(5137\)](#) [BMPR1A\(5048\)](#) [CFTR\(4425\)](#) [PKHD1\(2886\)](#)  
[CDH1\(2650\)](#) [NOTCH1\(2260\)](#) [MLH1\(1568\)](#) [PKD1\(1491\)](#) [KCNT1\(1446\)](#) [AP3B1\(1360\)](#) [MSH2\(1336\)](#) [TP53\(1236\)](#)

#	Disease related genes	Disease	Omics	Source
1	<a href="#">CCND3</a>	Appendiceal Cancer	Genomic	<a href="#">36493333</a>
2	<a href="#">CDKN2A</a>	Appendiceal Cancer	Genomic	<a href="#">36493333</a>
3	<a href="#">TERT</a>	Appendiceal Cancer	Genomic	<a href="#">36493333</a>

图S1 (F) 疾病部分的详细页面展示。



# 结果: 物种模块

## | Species

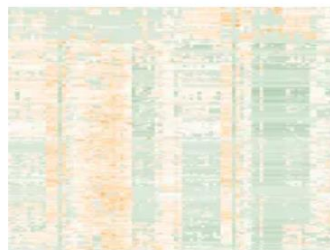
#	Species	Disease	Disease related genes	Direct Evidence	Inference
1	Human	Ileus	<a href="#">ALPI</a>	NA	Methotrex
2	Human	Colitis	<a href="#">ADRA2A</a>	NA	Arsenic Tri
3	Human	Inflammato...	<a href="#">ADRA2A</a>	NA	Dexameth
4	Human	Ileus	<a href="#">ANGPTL4</a>	NA	Cisplatin P
5	Human	Duodenal U...	<a href="#">ADRA2A</a>	NA	Capsaicin
6	Human	Intestinal O...	<a href="#">ABCC4</a>	NA	Carbamaz

图S1 (G) 物种部分的详细页面展示。

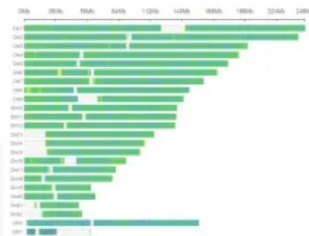


# 结果：组学

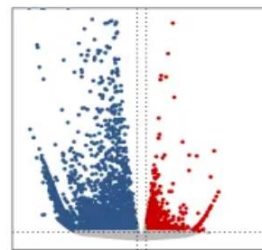
## Omics



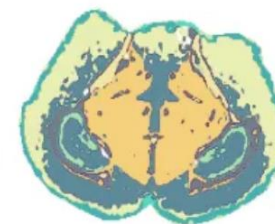
EPIGENOMICS



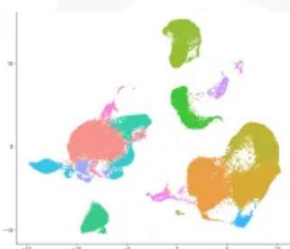
GENOMICS



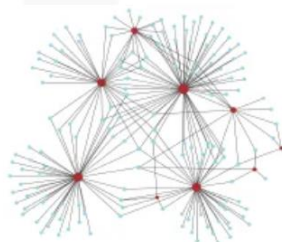
TRANSCRIPTOMICS



SPATIAL OMICS



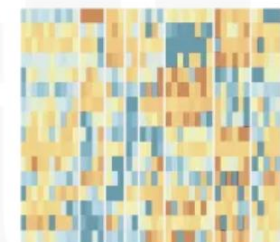
SINGLE CELL OMICS



PROTEOMICS



MICROBIOMICS



METABOLOMICS

点击可查看每个组学的内容



# 结果：浏览肠道疾病相关的空间组学和单细胞组学数据



空间组学页面详情

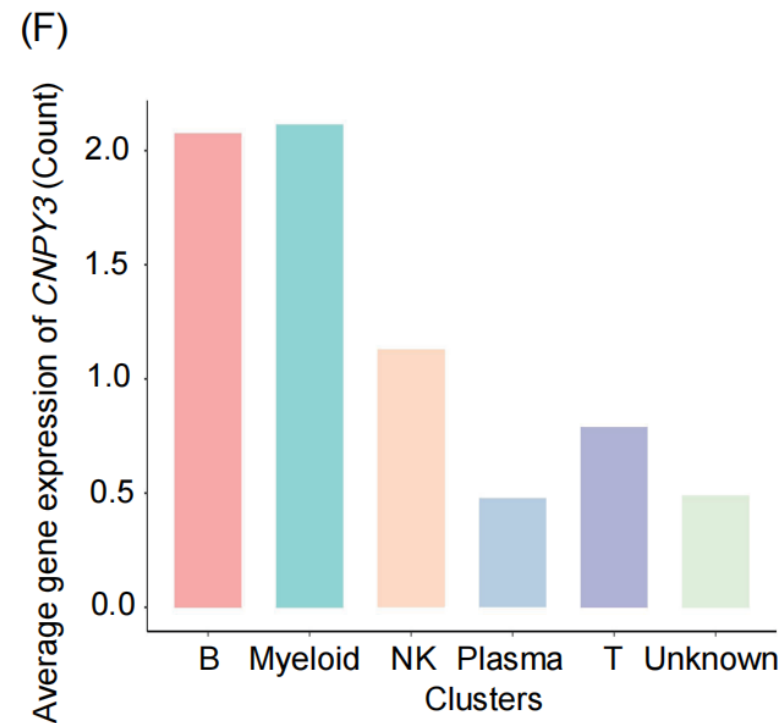


图1 (F) scRNA-seq数据中CNPY3在不同细胞中的表达水平图谱。



# 结果：与肠道疾病相关的多组学图谱的交互式可视化

(G)

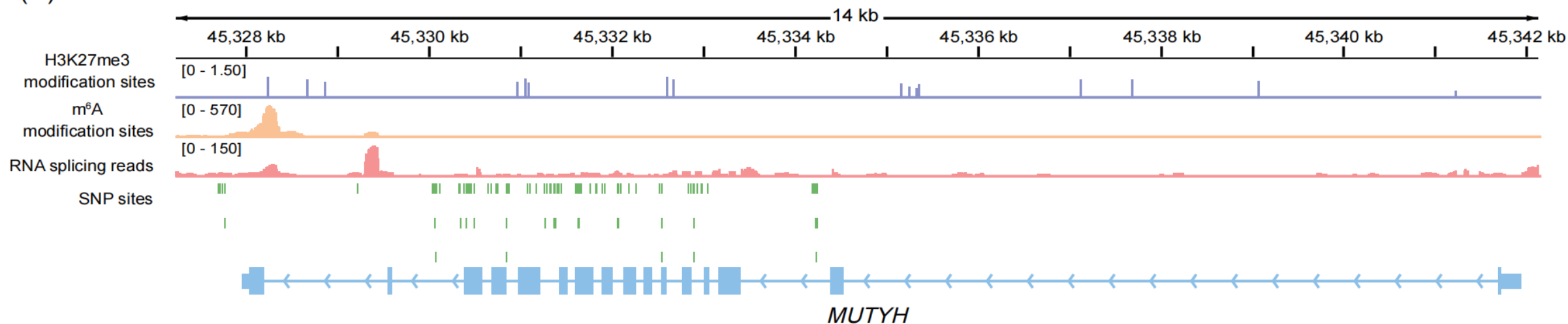


图1 (G) *MUTYH*在结直肠癌中的H3k27me3、m6A位点以及 RNA-seq和SNP数据的读数覆盖率的轨迹。



# 结果：与肠道疾病相关的多组学图谱的交互式可视化

## Summary

Project: **GSE127938**

Disease: Normal

Species: Homo sapiens

Tissue/Cell line: Pylorus, Duodenum, Jejunum, Ileum, Colon

CASE: Pylorus organoids, Duodenum organoids, Jejunum organoids, Ileum organoids, Colon organoids

Control: Corpus organoids

PMID: 525840

[NCBI](#)

profile

Sample: **colon** duodenum ileum jejunum pylorus

Gene expression data

#	Gene Name	Ensemble ID	Mean (Case)	Mean (Control)	Log2(Fold Change)	p.value
1	<a href="#">A1BG</a>	ENSG00000121410	0.0467	0.0074	2.8625	0.0736
2	<a href="#">A1BG-AS1</a>	ENSG00000268895	0.1485	0.1236	0.8733	0.7962
3	<a href="#">A1CF</a>	ENSG00000148584	0.2387	0.0186	0.1082	0.1942
4	<a href="#">A2M</a>	ENSG00000175899	0.0953	0.0383	0.4870	0.5991
5	<a href="#">A2M-AS1</a>	ENSG00000245105	0.3033	0.1406	0.5495	0.2810

< 1 2 3 4 5 6 ... 11343 > 5 1 page


- 137个RNA-seq数据集
- 790万个转录图谱

图1 (H) 基因组学中数据集详细信息的页面展示。



# 结果：Chatdoc

## ChatDoc



### ChatDoc

TOP Top 10 key words:

- Temporal and geographical distributi...
- Spatial distribution of data
- Statistical processing of data
- Reliable sources of data
- Openness of data
- Timeliness of datasets
- Biological significance of the data
- Sustainability of data
- Functional classification of disease-a...
- Prognostic markers of disease

#### Single Cell Omics

What is Single Cell Omics?

Single-cell Omics is a field of biology that aims to study and analyse the molecular properties and functions of individual cells within an organism, including the genomes, transcriptomes, proteomes, and metabolomes of individual cells. The field has evolved thanks to high-resolution technologies that allow scientists to delve into cellular diversity, heterogeneity, and interactions at the single-cell level. Below are the main components of Single Cell Omics: (1) Single-cell analysis techniques: Single-cell genomics relies on highly accurate experimental techniques such as single-cell RNA sequencing (scRNA-seq), single-cell mass spectrometry (scMS), single-cell protein mass spectrometry (scPMS), and single-cell metabolomics. These techniques allow the measurement and analysis of biomolecules in individual cells. (2) Studies of cellular heterogeneity: Single-cell histomics can reveal differences and heterogeneity between individual cells within cell populations. This helps to understand the diversity of different cell types, states and functions. (3) Cell development and differentiation: the study of changes in gene expression in individual cells during development and differentiation can help us understand cell fate decisions and tissue development. (4) Disease research: Single-cell genomics can be used to study cellular changes in disease onset and progression, helping to identify potential therapeutic targets and biomarkers. (5) Immunological studies: Single Cell Omics helps to gain insight into the function, interaction and response of various immune cells in the immune system. (6) Drug Screening and Individualised Medicine: Based on the information from single-cell genomics, drug screening experiments can be carried out to find treatments suitable for individual patients.

#### Metabolomics

What is Metabolomics?

Metabolomics is a subfield of systems biology that studies the composition, structure, quantification and dynamics of all metabolites or small-molecule metabolites in an organism. Metabolomics aims to gain a comprehensive understanding of the metabolic network in organisms to reveal the physiological state of organisms, biochemical processes and the interrelationships between biomolecules. The main components of metabolomics include: (1) Metabolites: Metabolomics studies metabolites, which are the products of metabolic reactions in organisms,

Enter content to start the chat (Ctrl+Enter line wrap) Send clear

用户可以输入问题关键字来获取答案

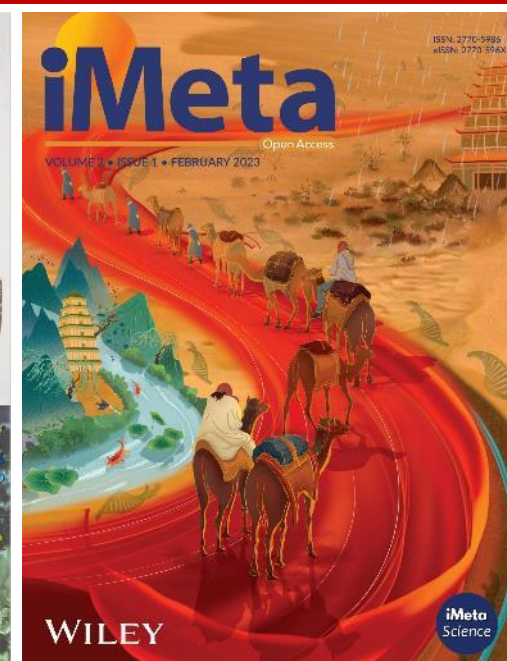


# 总结

- ❑ GutUDB是一个全面的肠道疾病多组学数据库，目前数据库已纳入260,790个疾病-基因关联，涉及基因组学、表观组学、转录组学、蛋白组学等八大组学，涵盖了56种肠道疾病和11个肠道部位，覆盖了人类、家鼠、绵羊、猪等多种物种类型。此外，数据库还包含了各种潜在的临床治疗方法。
- ❑ GutUDB集成四个核心功能：浏览、查询、可视化和下载，旨在方便用户访问和使用。基于该数据库，研究人员可进行各种组学数据分析，探索某种肠道疾病的相关基因。
- ❑ 数据库网址：<https://intestine.splicedb.net>

Bao Yi, YaxinChen, LizhuLin, JingyiLi, XinliLiu, GangWang, Yueqi Li, etal. 2024.

“GutUDB:AComprehensiveMultiomicsDatabaseforIntestinalDiseases.” *iMeta* 3: e195. <https://doi.org/10.1002/imt2.195>



“**iMeta**”由威立、肠菌分会和华人科学家出版的开放获取期刊，主编由中科院微生物所刘双江和荷兰格罗宁根大学傅静远教授共同担任。目的是发表原创研究、方法和综述以促进宏基因组学、微生物组和生物信息学发展。目标是发表前10%(IF>20)的高影响力论文。期刊特色包括视频投稿、可重复分析、图片打磨、青年编委、中英双语、50万用户的社交媒体宣传等。2022年2月发行，相继被**ESCI**、**Google Scholar**、**DOAJ**、**Scopus**等数据库收录，发文161篇，被引2316次([Dimension](#), 2024/2/19)!



主页: <http://www.imeta.science>

出版社: <https://wileyonlinelibrary.com/journal/imeta>



投稿: <https://wiley.atyponrex.com/journal/IMT2>



[office@imeta.science](mailto:office@imeta.science)



宣传片



[iMeta](#)

