课程概览

1 数据库基本概念

2 NCBI & EMBL-EBI & 在线分析工具

3 TCGA GTEx & GEPIA 1/2/2021

4 单细胞/空间转录组数据库

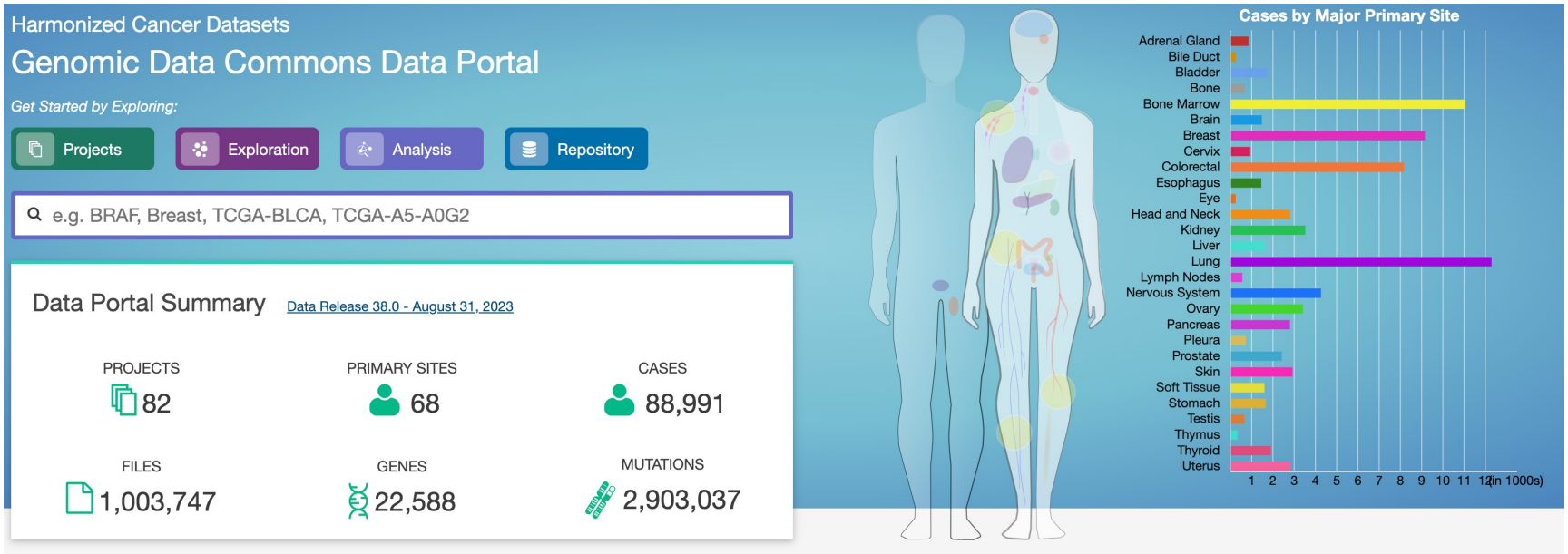
TCGA

- **肿瘤基因组图谱 (The Cancer Genome Atlas,TCGA)计划由美国National Cancer Institute(NCI)和National Human Genome Research Institute（NHGRI）于2006年联合启动的项目。**

https://www.cancer.gov/ccg/research/genome-sequencing/tcga

TCGA 成果与影响

1.建立了丰富的基因组学数据资源，支撑了计算生物学领域

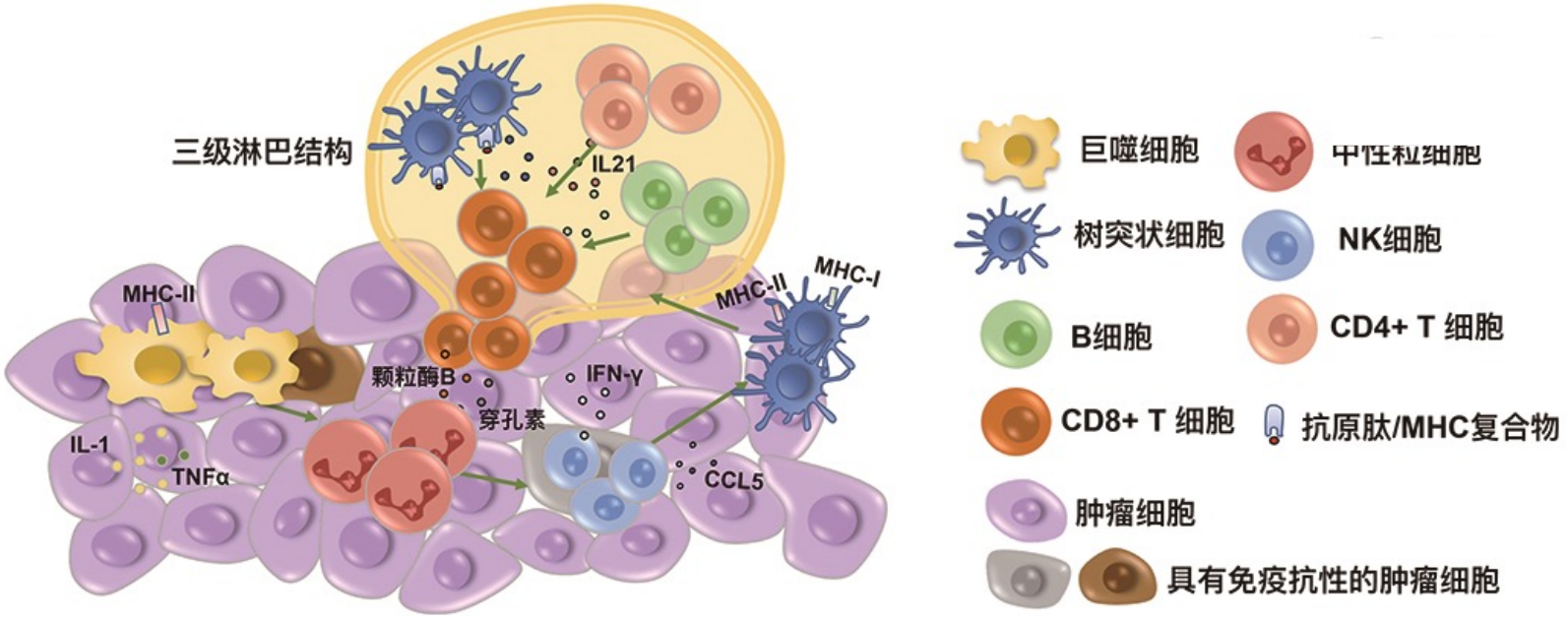


- **8W+患者样本数据， 33种癌症类型，5种测序类型数据(WXS, RNA-Seq, Genotyping Array, Targeted Sequencing, WGS)以及对应的临床信息(药物响应，生存期等)**

- **TCGA 产生的海量数据刺激了计算生物学领域的巨大发展。研究人员针对开发很多生物信息学数据挖掘工具，例如识别体细胞和种系突变、 预测具有预后意义的基因、 构建基因调控网络、 以及癌症图像的自动分析等。**

TCGA 成果与影响

2. 加深了癌症的分子机制的探索

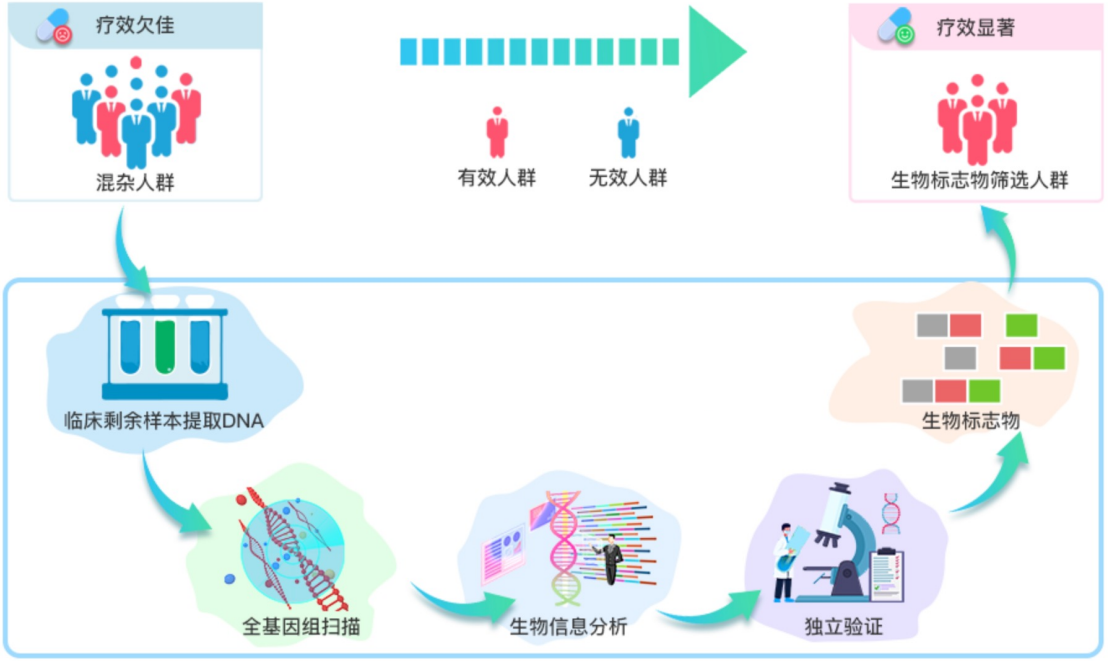


- 理解DNA各种类型的改变, 例如置换, 嵌合, 融合, 拷贝数改变以及其他复杂的结构变异

- 肿瘤异质性来源: 癌细胞和免疫细胞组成的微环境理解
- 不同癌种之间共有的分子演化规律

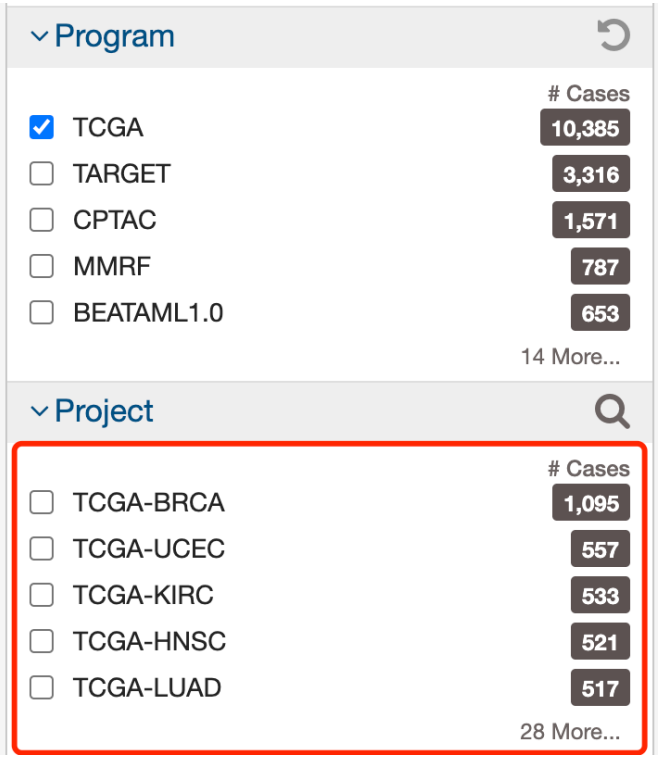
TCGA 成果与影响

3.推进医疗相关科学进步



- 推动了新兴测序技术的巨大发展，提高数据质量并降低测序成本。比如反相蛋白质阵列、福尔马林固定石蜡包埋的样品分析物提取等测序建库分子技术。
- 通过分子和临床数据对疾病进行更精确的诊断,分类和预后，推动精准医疗
- 通过海量数据挖掘癌症治疗潜在靶点分子以及对应分子机制。

TCGA RNA-seq数据介绍

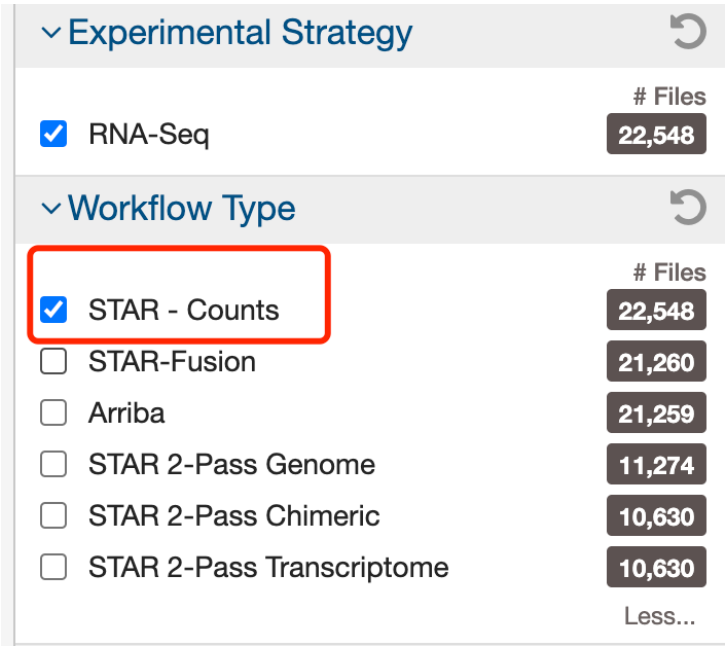


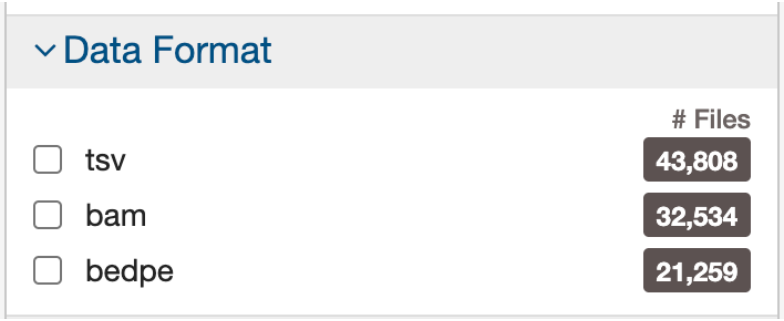
ACC 腺样囊性癌
BLCA 膀胱癌
BRCA 乳腺癌
CESC 宫颈鳞状细胞癌
CHOL 胆癌
COAD 结肠癌
DLBC 淋巴癌
ESCA 食管癌
GBMLGG 脑癌
HNSC 头颈部鳞癌
KICH 嫌色细胞癌
KIRC 肾透明细胞癌
KIRP 乳头状肾细胞癌
LAML 骨髓癌
LIHC 肝癌
LUAD 肺腺癌
LUSC 肺鳞状细胞癌
MESO 胸膜癌
OV 卵巢癌
PAAD 胰腺癌
PCPG 肾上腺癌
PRAD 前列腺癌
READ 直肠癌
SARC 软组织癌
SKCM 皮肤癌
STAD 胃癌
TGCT 睾丸癌
THCA 甲状腺癌
THYM 胸腺癌
UCEC 子宫内膜癌
UCS 子宫癌
UVM 眼癌

- **RNA-seq 数据中 TCGA 规范定义了33种常见癌种命名方式**
- **每种癌种对应均有上百样本数**

https://portal.gdc.cancer.gov/repository

TCGA RNA-seq数据介绍

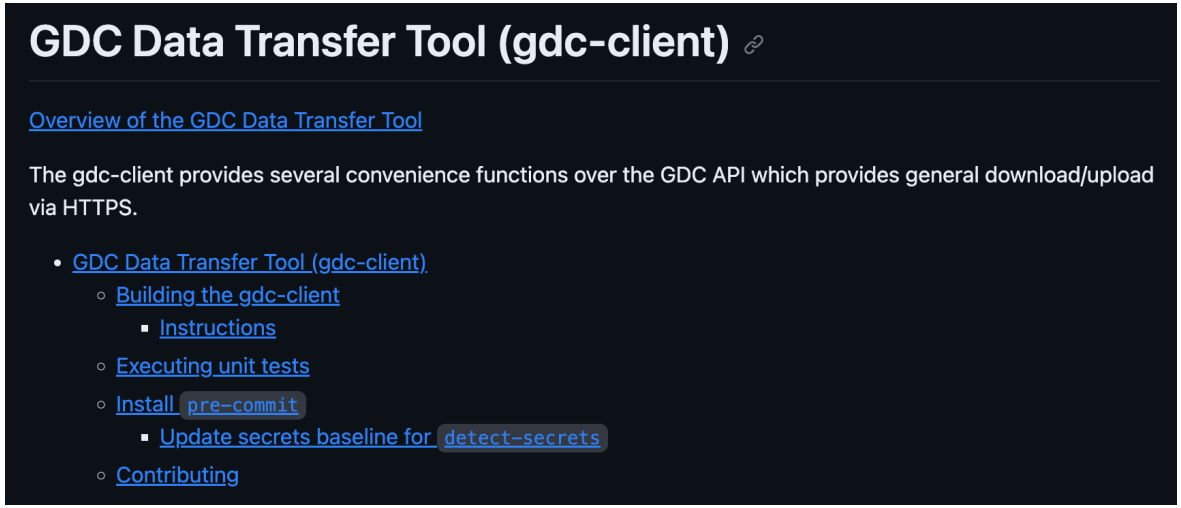




- 支持 多种常见分析软件的结果(STAR等) 同时提供多种文件格式(bam,bed等)
- 基本的基因表达矩阵支持公开下载，其他上游数据需要额外申请

TCGA RNA-seq数据下载方式

1. gdc-client(官网可执行程序下载方式)



GDC Data Transfer Tool (gdc-client)

Overview of the GDC Data Transfer Tool

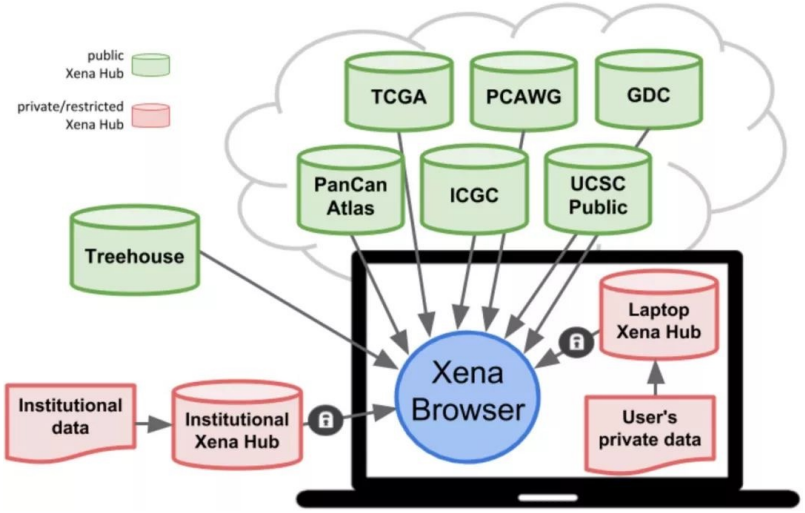
The gdc-client provides several convenience functions over the GDC API which provides general download/upload via HTTPS.

- GDC Data Transfer Tool (gdc-client)
  - Building the gdc-client
    - Instructions
  - Executing unit tests
  - Install `pre-commit`
    - Update secrets baseline for `detect-secrets`
  - Contributing

- **官方原生数据，不存在二次修改**
- **以上俩种下载方式需要 从官网对应筛选数据下载 (有多个文件需要对应)**
- **下载后的数据需要一一对照整理 (费时费力)**

TCGA RNA-seq数据下载方式

2. Xena（网页工具）新手入门推荐



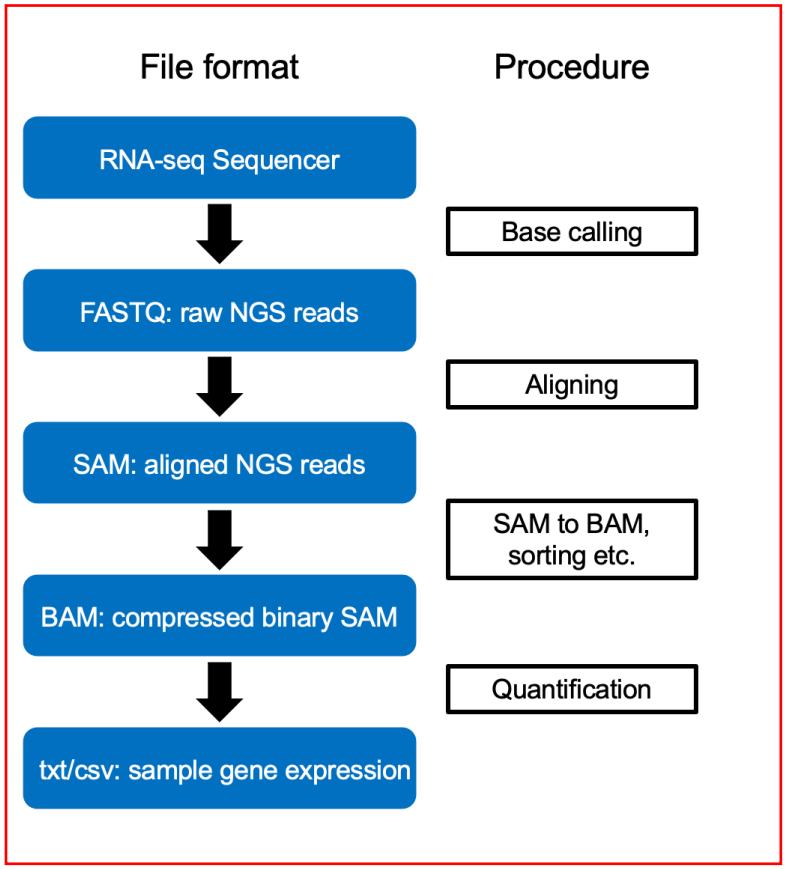
GDC TCGA Acute Myeloid Leukemia (LAML) (15 datasets)
GDC TCGA Adrenocortical Cancer (ACC) (14 datasets)
GDC TCGA Bile Duct Cancer (CHOL) (14 datasets)
GDC TCGA Bladder Cancer (BLCA) (14 datasets)
GDC TCGA Breast Cancer (BRCA) (20 datasets)
GDC TCGA Cervical Cancer (CESC) (14 datasets)
GDC TCGA Colon Cancer (COAD) (15 datasets)
GDC TCGA Endometrioid Cancer (UCEC) (15 datasets)
GDC TCGA Esophageal Cancer (ESCA) (14 datasets)
GDC TCGA Glioblastoma (GBM) (15 datasets)
GDC TCGA Head and Neck Cancer (HNSC) (14 datasets)
GDC TCGA Kidney Chromophobe (KICH) (14 datasets)
GDC TCGA Kidney Clear Cell Carcinoma (KIRC) (15 datasets)
GDC TCGA Kidney Papillary Cell Carcinoma (KIRP) (15 datasets)
GDC TCGA Large B-cell Lymphoma (DLBC) (14 datasets)
GDC TCGA Liver Cancer (LIHC) (14 datasets)

- **UCSC Xena 整理了很多常用数据库的数据，其中包括TCGA**
- **只需要在线网页选择数据点击即可下载 处理好的 TCGA数据**
- **需要注意单位换算，该网页存储方式一般取对数化log2(X+1)存储**

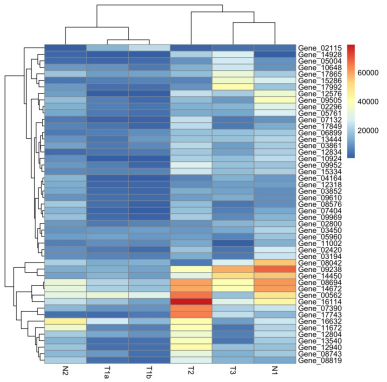
https://xenabrowser.net/datapages/

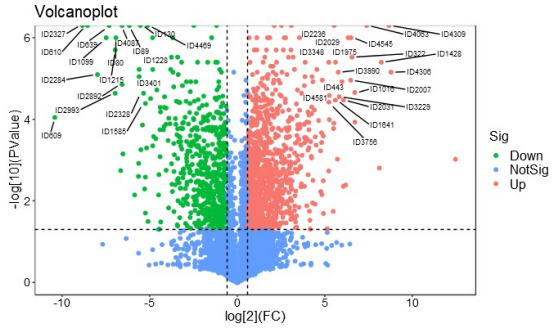
TCGA RNA-seq数据 常见处理流程和下游分析

- 自测数据处理流程



- TCGA RNA-seq数据处理流程
  - 下载获得上游**处理好**的count 矩阵, **一般不需要**重新上游处理
  - 根据需求类别进行整理和**标准化**
  - 下游差异分析,生存期分析等

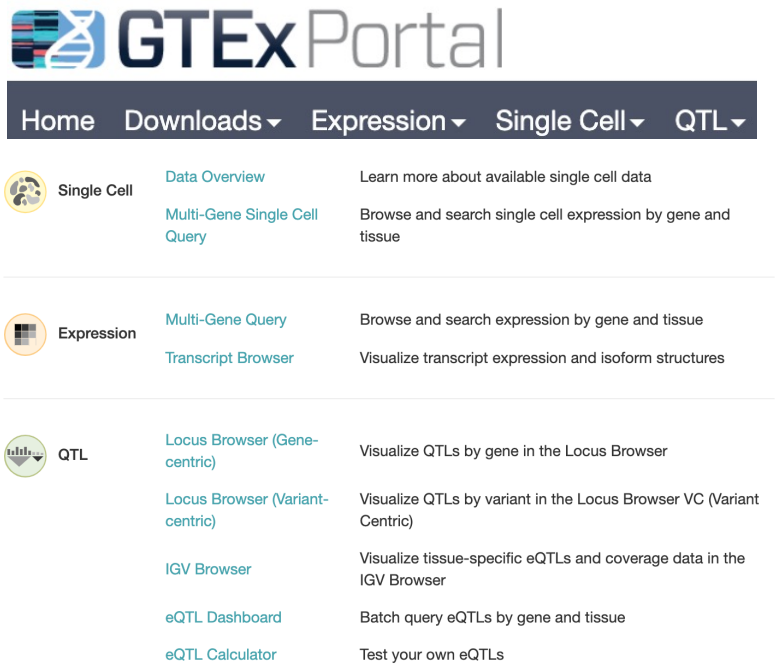




- TCGA 数据库中正常人对照的 RNA-seq数据较少? → 联合**GTEx**数据库

GTEx 数据库

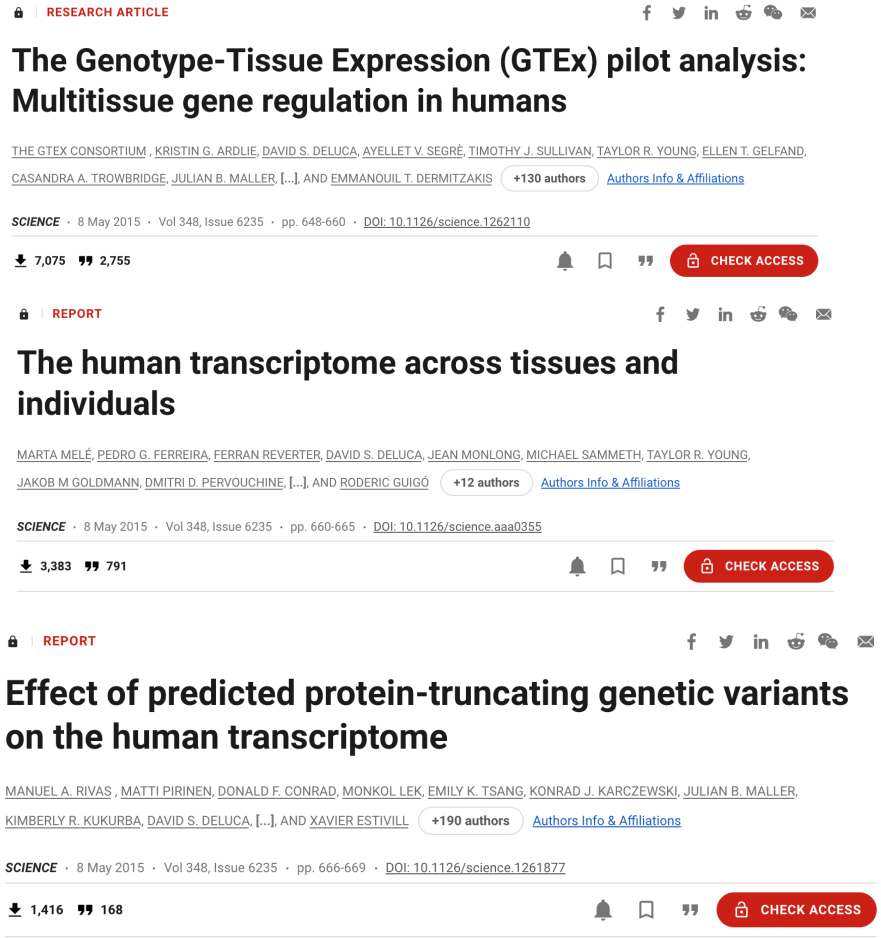
Genotype-Tissue Expression (GTEx)



https://gtexportal.org/home/

- **主要由美国NIH（国立卫生研究院）的公共基金计划资助**
- **其主要针对正常人(非疾病患者)死亡后的尸体组织进行相关基因测序 (RNA-seq, scRNA-seq, WGS等)**

GTEx 数据库 第一阶段成果



RESEARCH ARTICLE

The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans

THE GTEX CONSORTIUM , KRISTIN G. ARDLIE, DAVID S. DELUCA, AYELLET V. SEGRÈ, TIMOTHY J. SULLIVAN, TAYLOR R. YOUNG, ELLEN T. GELFAND, CASANDRA A. TROWBRIDGE, JULIAN B. MALLER, [...], AND EMMANOUIL T. DERMITZAKIS +130 authors Authors Info & Affiliations

SCIENCE · 8 May 2015 · Vol 348, Issue 6235 · pp. 648-660 · DOI: 10.1126/science.1262110

7,075 2,755 CHECK ACCESS

REPORT

The human transcriptome across tissues and individuals

MARTA MELÉ, PEDRO G. FERREIRA, FERRAN REVERTER, DAVID S. DELUCA, JEAN MONLONG, MICHAEL SAMMETH, TAYLOR R. YOUNG, JAKOB M GOLDMANN, DMITRI D. PERVOUCHINE, [...], AND RODERIC GUIGÓ +12 authors Authors Info & Affiliations

SCIENCE · 8 May 2015 · Vol 348, Issue 6235 · pp. 660-665 · DOI: 10.1126/science.aaa0355

3,383 791 CHECK ACCESS

REPORT

Effect of predicted protein-truncating genetic variants on the human transcriptome

MANUEL A. RIVAS , MATTI PIRINEN, DONALD F. CONRAD, MONKOL LEK, EMILY K. TSANG, KONRAD J. KARCZEWSKI, JULIAN B. MALLER, KIMBERLY R. KUKURBA, DAVID S. DELUCA, [...], AND XAVIER ESTIVILL +190 authors Authors Info & Affiliations

SCIENCE · 8 May 2015 · Vol 348, Issue 6235 · pp. 666-669 · DOI: 10.1126/science.1261877

1,416 168 CHECK ACCESS

- **2015**年，GTEx发布了第一阶段成果，一次性在Science杂志上发表**3**篇研究成果，该成果还被选为当年封面文章。
- 从**175**名死者身上采集到了**1641**个尸检样本，这些样本来自**54**个不同的身体部位
- 这三篇文章分别从**基因表达模式**和对基因组中影响基因表达的**特定区域**(增强子,启动子等)以及组织特异性基因的表达调控模式和截短的**蛋白变异体**对不同组织中的基因表达影响。

GTEx 数据库 第二阶段成果

Open access | Published: 12 October 2017

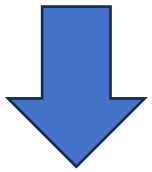
Landscape of X chromosome inactivation across human tissues

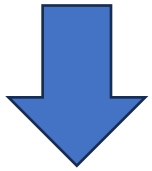
Taru Tukiainen, Alexandra-Chloé Villani, Angela Yen, Manuel A. Rivas, Jamie L. Marshall, Rahul Satija, Matt Aguirre, Laura Gauthier, Mark Fleharty, Andrew Kirby, Beryl B. Cummings, Stephane E. Castel, Konrad J. Karczewski, François Aguet, Andrea Byrnes, GTEx Consortium, Tuuli Lappalainen, Aviv Regev, Kristin G. Ardlie, Nir Hacohen & Daniel G. MacArthur

Nature **550**, 244–248 (2017) | Cite this article

- **2017**年，GTEx发布了第二阶段成果，在Nature杂志上发表**4**篇研究成果。
- 收集来自**449**名生前健康的人类捐献者的**7000多**份尸检样本，涵盖44个组织（42种不同的组织类型），包括31个实体器官组织
- 结合GTEx数据研究了与基因表达相关联的**基因变异**调节RNA编辑和X染色体失活现象的机制

基于TCGA分析的数据库：GEPIA 1





癌症样本RNA-seq
(比如脑肿瘤)

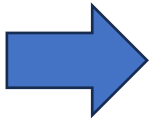
+

正常人对应组织样本RNA-seq
(比如大脑)

1. 联合分析可高通量地筛选相关肿瘤特异性相关表达基因
2. 可以在多个肿瘤疾病中寻找异同表达基因

…

- 对于不熟悉生信的生物研究者如何获取数据推动相关研究呢?



GEPIA1/2

基于TCGA分析的数据库：GEPIA 1

张泽民

邮　　箱：zemin@pku.edu.cn

职　　称：教授

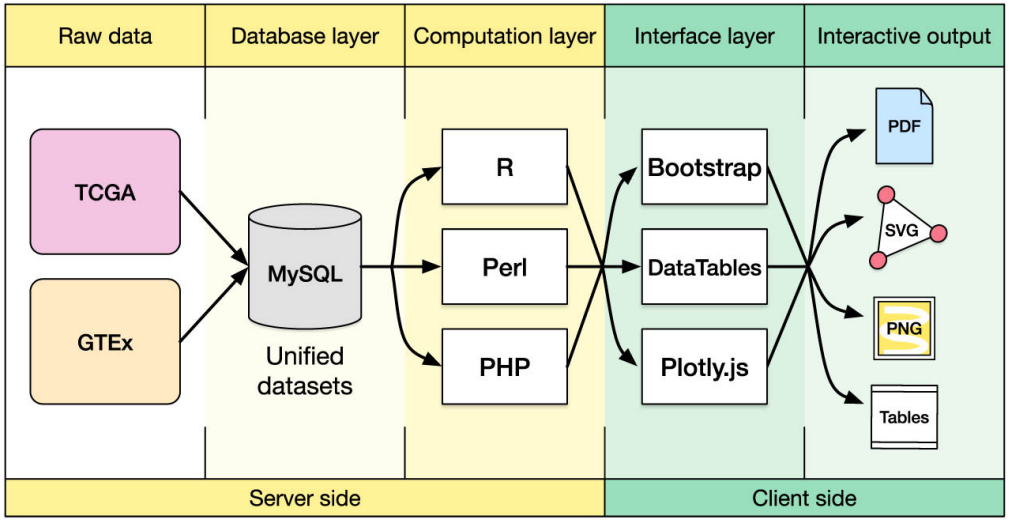
办公室地址：北京市海淀区颐和园路5号，综合科研2号楼，100871

所属实验室：张泽民实验室,生物动态光学成像中心（BIOPIC）,生命科学联合中心

个人主页：http://cancer-pku.cn/

个人介绍:

张泽民，北京大学生物医学前沿创新中心研究员、副主任，北京大学未来基因诊断高精尖创新中心研究员，中国"海外高层次人才"计划，长江特聘教授，北大-清华生命科学联合中心高级研究员。 1988年本科毕业于南开大学生物系, 1989年由 CUSBEA 项目赴美留学。1995年研究生毕业于宾州州立大学生物化学和分子生物学专业, 获得博士学位 ;1995-1998年在美国旧金山加州大学进行博士后研究工作。1998年起任职于美国 GENENTECH 公司开始生物信息学研究, 曾任博士后导师, 生物信息学部主任, 生物信息首席科学家。前期主要工作包括: 在世界上首次报道实体癌的全基因组测序; 首次在全基因组水平研究病毒插入事件在肝癌发生发展中的作用; 首次通过计算方法鉴别癌症的驱动突变。通过对癌症高通量数据的分析, 成功地发现了多种特异性抗癌靶点, 获得60多项美国专利授权。2014 年加入北京大学。2017年获The Boehringer Ingelheim Investigator Award，2017年获Bayer Investigator Award, 相关成果入选2017年度中国十大医学科技新闻、2017年度中国生命科学十大进展、细胞出版社2017和2019中国年度论文、2018和2019年度中国生物信息十大进展、2020年细胞杂志最佳论文。



GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses

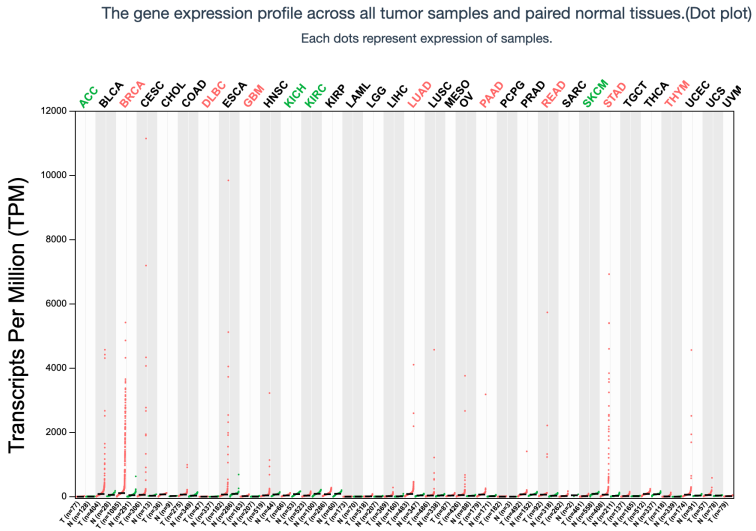
Z Tang, C Li, B Kang, G Gao, C Li… - Nucleic acids …, 2017 - academic.oup.com

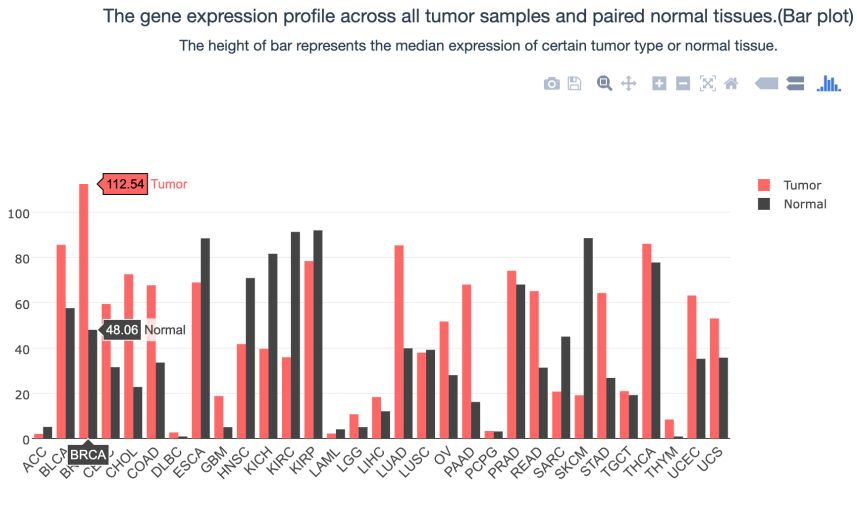
… GEPIA provides key interactive and customizable … through GEPIA greatly facilitate data mining in wide research areas, scientific discussion and the therapeutic discovery process. GEPIA …

☆ 保存　99 引用　被引用次数：7235　相关文章　所有 10 个版本

- **北大的张泽民教授课题组2017年基于上述思想 开发了一键化操作 数据库用于方便生物领域内研究者便捷地获得相关数据分析结果**

基于TCGA分析的数据库：GEPIA 1

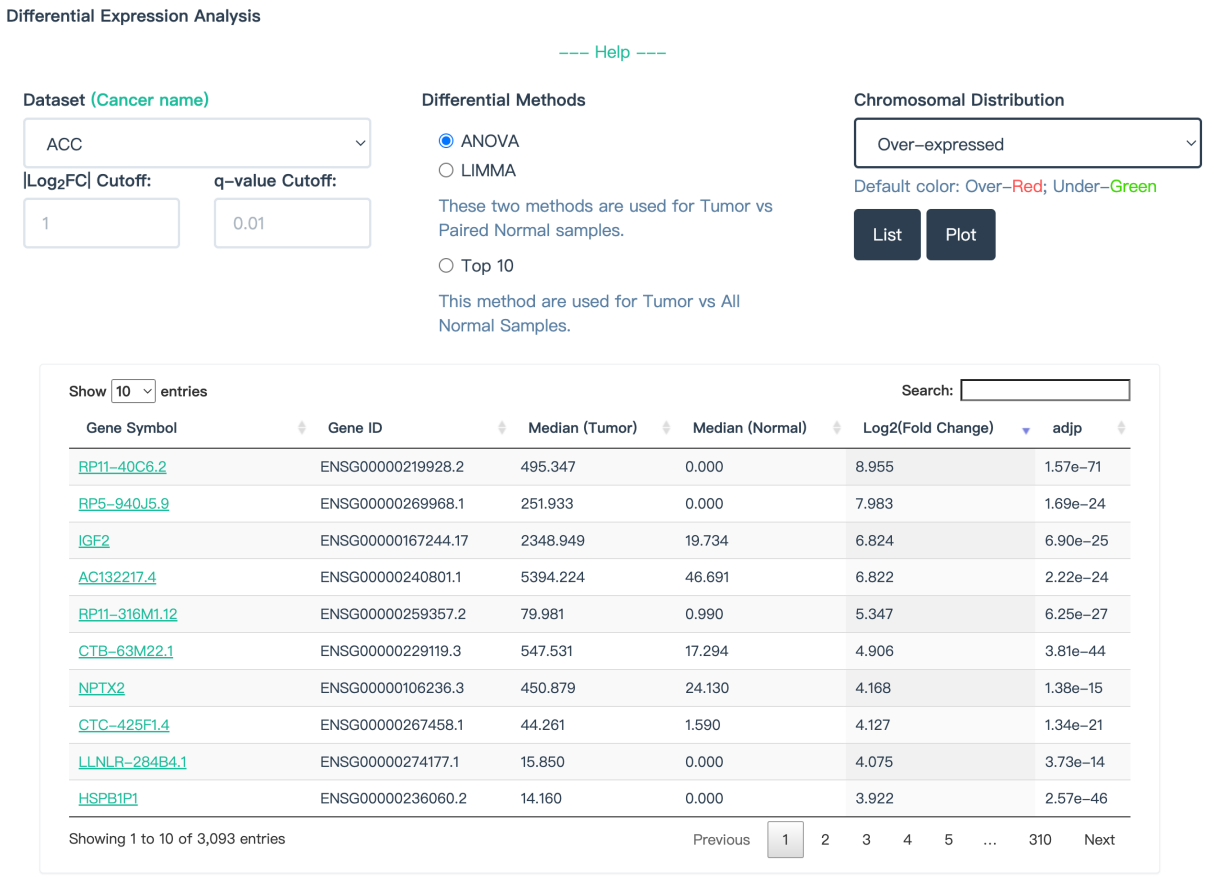




http://gepia.cancer-pku.cn/

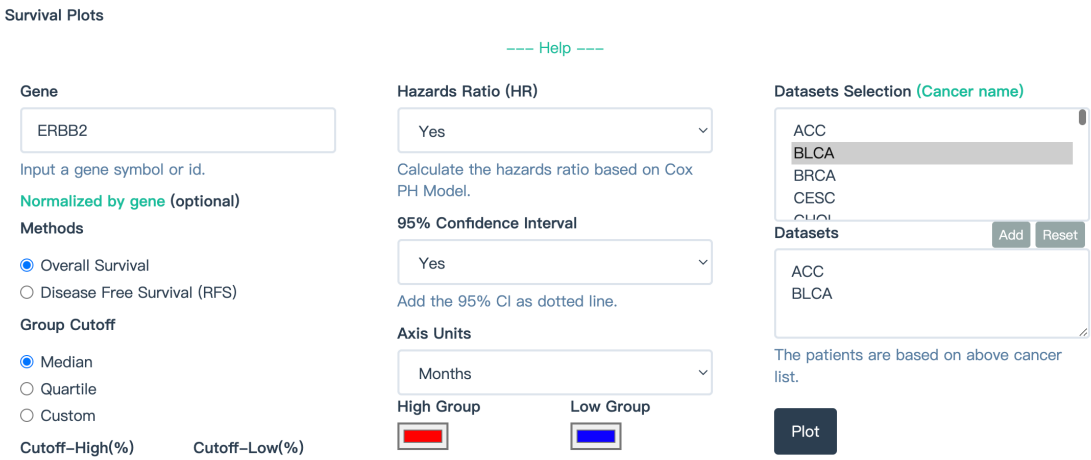
- 通过输入基因可**直接查看**对应癌症和正常样本中该基因的TPM的情况

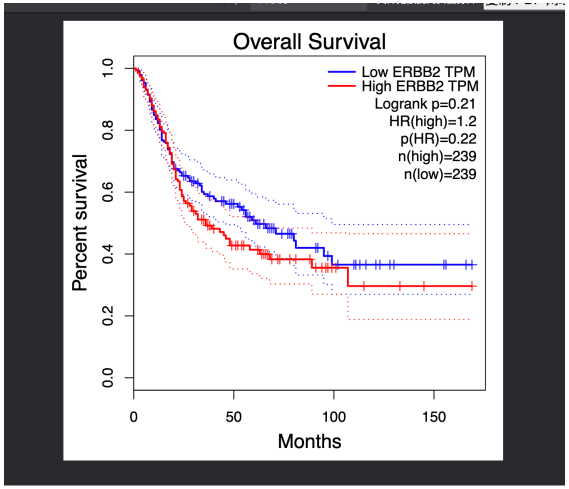
基于TCGA分析的数据库：GEPIA 1



- **通过选择相关癌种和差异分析方法可直接获得对应癌症和正常样本RNA-seq的差异分析结果 (无需手动运行deseq2/edgeR等分析软件)**

基于TCGA分析的数据库：GEPIA 1





- **通过选择相关癌种,对应基因和生存分析方法可直接获得对应癌症和正常样本生存分析结果(无需手动导入样本生存期信息并用相关软件分析)**

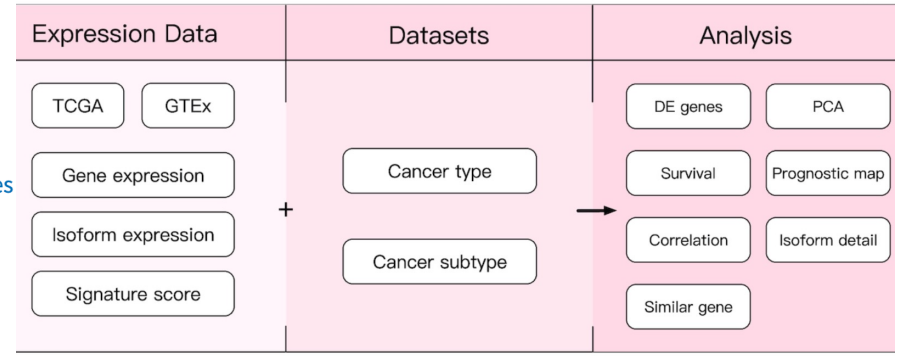
基于TCGA分析的数据库：GEPIA 2

JOURNAL ARTICLE

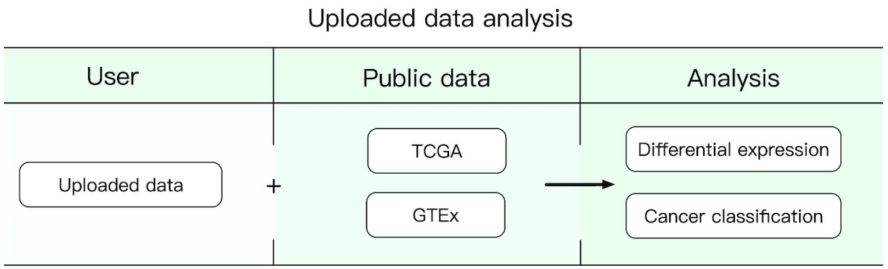
GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis

Zefang Tang, Boxi Kang, Chenwei Li, Tianxiang Chen, Zemin Zhang ✉ Author Notes

Nucleic Acids Research, Volume 47, Issue W1, 02 July 2019, Pages W556–W560,
https://doi.org/10.1093/nar/gkz430
Published: 22 May 2019 **Article history ▼**



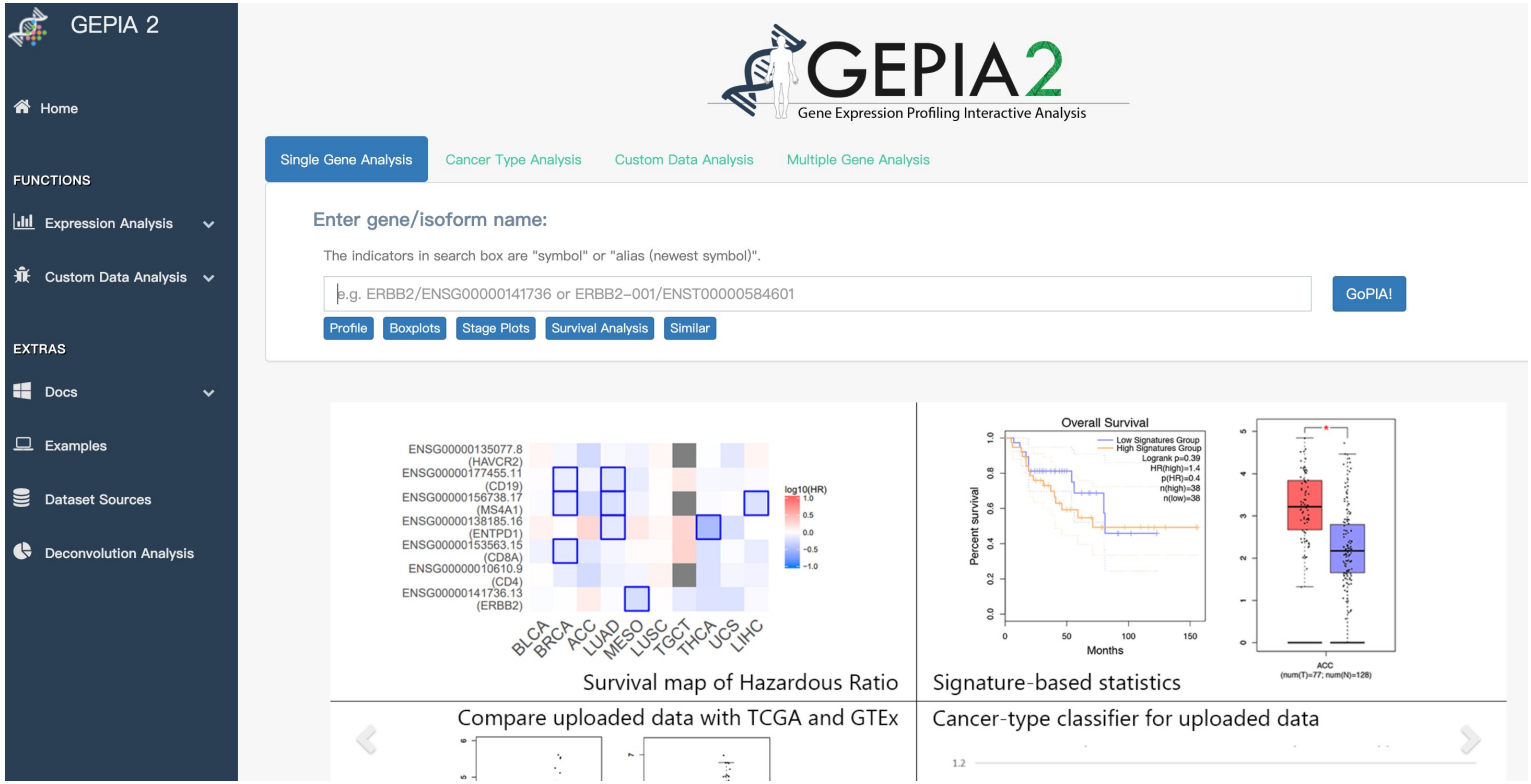
GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis
Z Tang, B Kang, C Li, T Chen, Z Zhang - Nucleic acids research, 2019 - academic.oup.com
… of different cancer subtypes, GEPIA2 allows users to fine-… GEPIA2 as we provide signature-based functionalities with a curated list of signatures for efficient investigation. Finally, …
☆ 保存 99 引用 被引用次数：2381 相关文章 所有 10 个版本



2019年，GEPIA2更新发布第二版

- 数据方面主要更新了不同**剪切体**水平(isoform)的定量结果以及**癌症的亚型**(84种)
- 提供用户**上传自己**的RNA-seq数据并和TCGA/GTEx进行比较以及对应批处理的API。

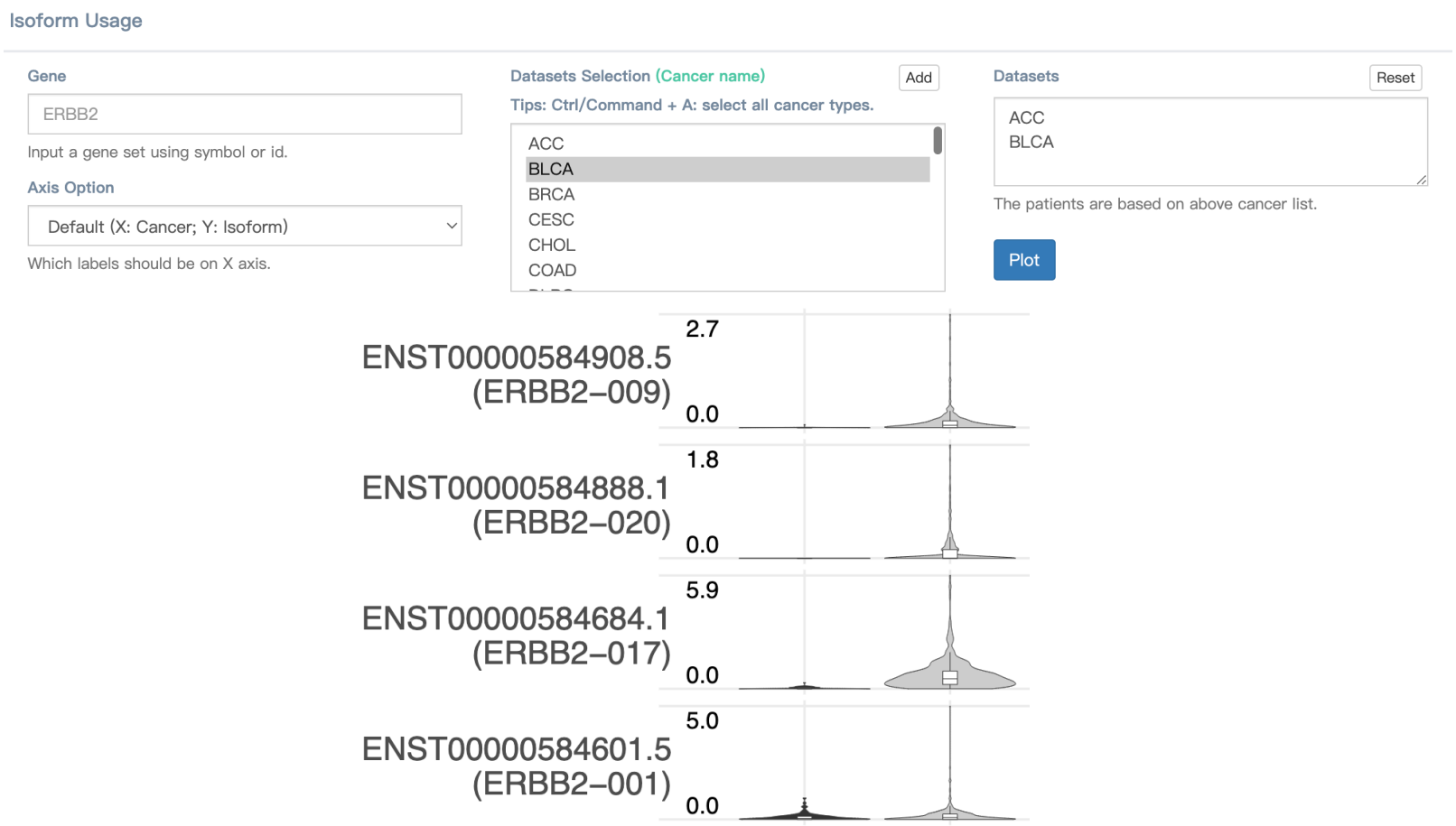
基于TCGA分析的数据库：GEPIA 2



http://gepia2.cancer-pku.cn

- **更新了前端框架并集成了GEPIA1的全部功能**
- **侧边栏中包含所有支持的功能以及数据下载和上传功能**

基于TCGA分析的数据库：GEPIA 2



- 通过输入基因可**直接查看**该基因在对应癌症的样本中该基因已知的**isoform**的表达情况

基于TCGA分析的数据库：GEPIA 2

The cancer subtype classifier takes an RNA–seq profile and makes a prediction.

Choose your model to be tested on and you will get a probability matrix by your sample for each subtype.

The uploaded gene expression profile should be TPM values with Hugo gene names.

This classifier is based on the naive bayes algorithm implemented by sklearn package in python with some modification (e.g., feature selection)

WARNING: please select the correct cancer type of your file uploaded, or you will get meaningless results.

If you want to determine the origin of the cancer metastasis, please select 'TCGA_Subtype', which has the overall subtypes of all cancer types provided by TCGA officially.

Step 1: Choose the cancer type for comparation :

BRCA

Click here to show details for cancer subtype information.

Step 2: Choose one file to upload

选择文件 未选择任何文件

Note: You need the csv format like this.

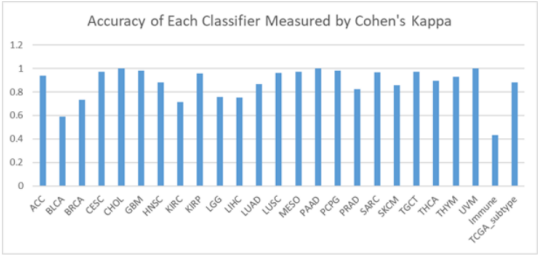
	A1BG	ADA	AKT3	ZBTB11-AS1	MED6
sample1	5.443278	1.008734	1.008734	0	0
sample2	0	8.684895	0	0	0
sample3	0	0	5.439709	0	7.926596
sample4	0	0	0	0	0
sample5	0	0	1.362182	0	0

An example can be downloaded here.

Step 3: Click to submmit your file.

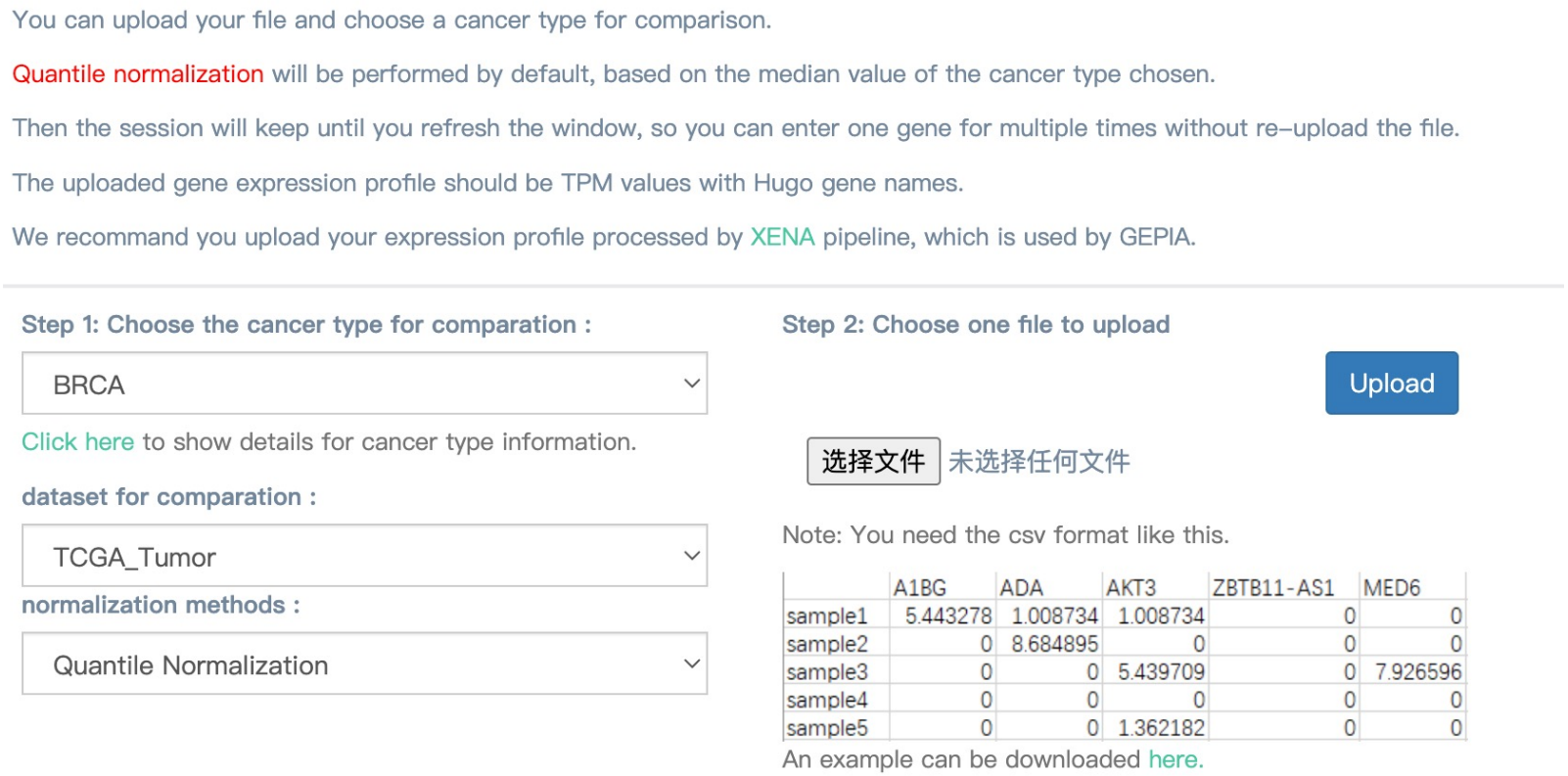
Submit

Here is the result of perfomance measured by a robust static cohen's kappa for each classifier.



- **通过上传自己的RNA-seq矩阵可获得预测的癌症分型**

基于TCGA分析的数据库：GEPIA 2



	A1BG	ADA	AKT3	ZBTB11-AS1	MED6
sample1	5.443278	1.008734	1.008734	0	0
sample2	0	8.684895	0	0	0
sample3	0	0	5.439709	0	7.926596
sample4	0	0	0	0	0
sample5	0	0	1.362182	0	0

- **通过上传自己的RNA-seq矩阵可获得自己的数据和数据库的数据基因表达比较情况 (相当于不需要用户自己做标准化)**

基于TCGA分析的数据库：GEPIA 2

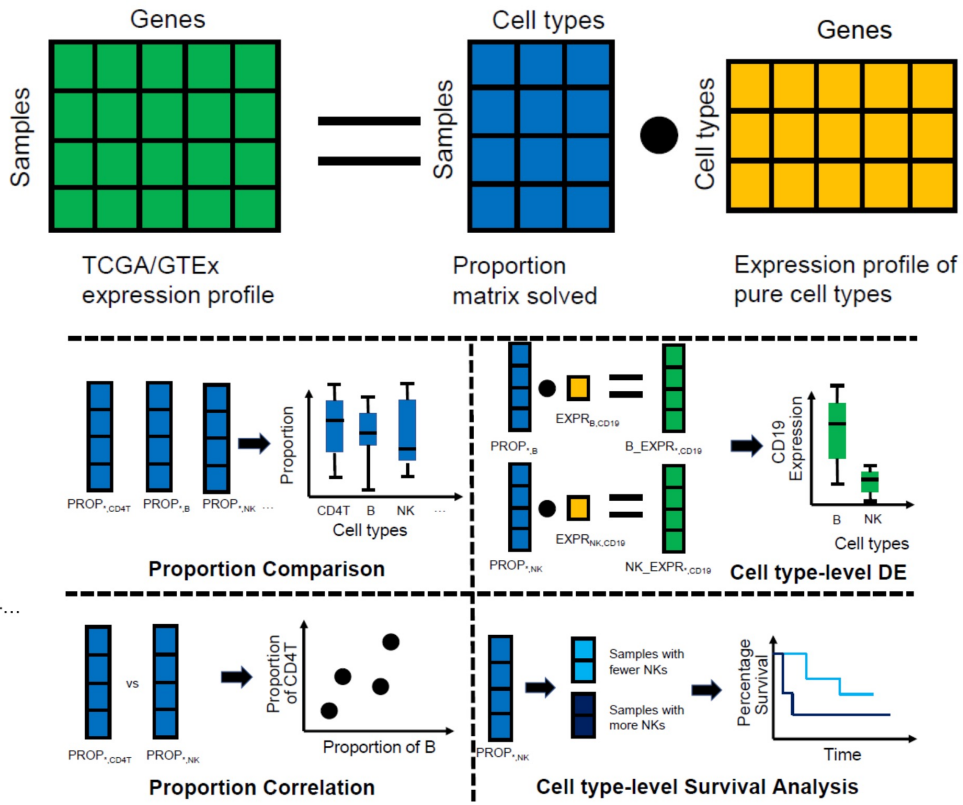
JOURNAL ARTICLE

GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA

Chenwei Li, Zefang Tang, Wenjie Zhang, Zhaochen Ye, Fenglin Liu ✉ Author Notes

Nucleic Acids Research, Volume 49, Issue W1, 2 July 2021, Pages W242–W246,
https://doi.org/10.1093/nar/gkab418
Published: 29 May 2021 **Article history ▼**

GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA
C Li, Z Tang, W Zhang, Z Ye, F Liu - Nucleic Acids Research, 2021 - academic.oup.com
… Thus, we present GEPIA2021, a standalone extension of GEPIA, … With GEPIA2021, experimental biologists could easily … GEPIA2021 is publicly accessible at http://gepia2021.cancer-…
☆ 保存 99 引用 被引用次数：214 相关文章 所有 9 个版本



详细信息参照 http://gepia2021.cancer-pku.cn/

2021年，GEPIA-2021更新发布第三版

- **主要更新基于scRNA-seq数据拆分大规模RNA-seq数据获得单细胞维度后的一系列分析(包括细胞亚型的差异分析，细胞类型的生存分析等)**

课程概览

1 数据库基本概念

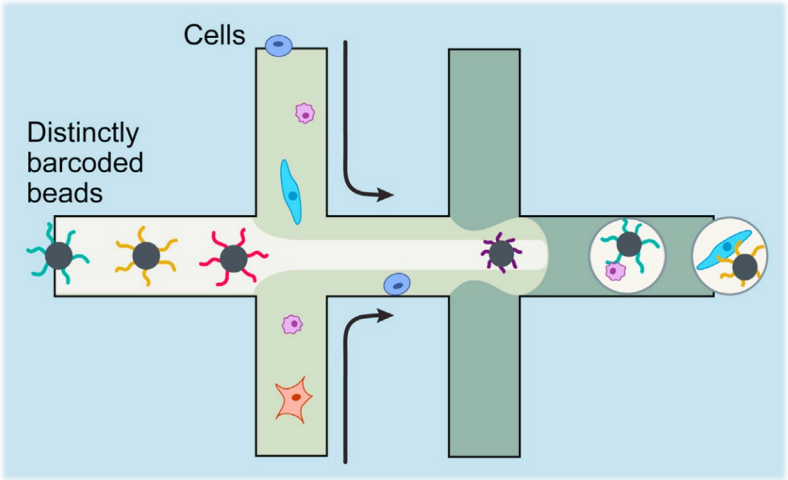
2 NCBI & EMBL-EBI & 在线分析工具

3 TCGA GTEx & GEPIA 1/2/2021

4 单细胞/空间转录组数据库

单细胞转录组(scRNA-seq) 和空间转录组(ST) 背景介绍

单细胞转录组测序技术
Single Cell RNA-seq

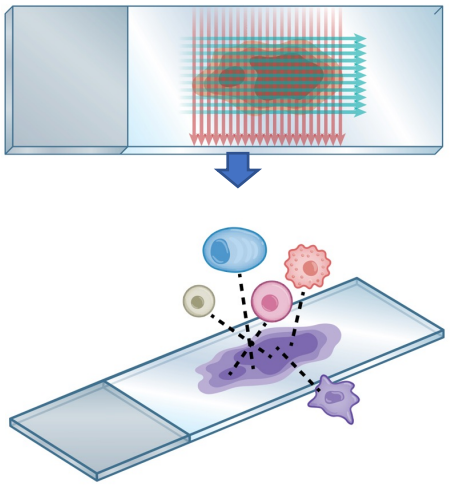


Macosko et al., ***Cell.*** 2015

2018年 *Science* 年度突破

- 检测**单个细胞**的转录组
- 发现稀有特殊细胞亚群
- 获得复杂组织细胞构成

空间转录组测序技术
Spatial Transcriptome



Ståhl P et al., ***Science.*** 2016

2020年 *Nature Methods* 年度技术

- 检测组织原位的转录组
- 发现细胞**空间位置**关系
- 识别重要空间组织区域

单细胞转录组(scRNA-seq) 数据库

TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment

D Sun, J Wang, Y Han, X Dong, J Ge, R Zheng, X Shi, B Wang, Z Li, P Ren, L Sun, Y Yan…

Nucleic acids research, 2021 • academic.oup.com

Abstract

Cancer immunotherapy targeting co-inhibitory pathways by checkpoint blockade shows remarkable efficacy in a variety of cancer types. However, only a minority of patients respond to treatment due to the stochastic heterogeneity of tumor microenvironment (TME). Recent advances in single-cell RNA-seq technologies enabled comprehensive characterization of the immune system heterogeneity in tumors but posed computational challenges on integrating and utilizing the massive published datasets to inform

展开

☆ 保存 99 引用 被引用次数：404 相关文章 所有 12 个版本

Chenfei Wang

Professor of Bioinformatics, Tongji University

在 tongji.edu.cn 的电子邮件经过验证 - 首页

Bioinformatics Artificial Intelligence Single Cell Genomics

同济大学 王晨飞教授

- **2021年 对标 TCGA 开发了 scRNA-seq技术的肿瘤疾病组学数据库 Tumor Immune Single Cell Hub(TISCH)**
- **数据包含27 种癌症类型, 76 个高质量肿瘤数据集, 总计 200 万个细胞的单细胞转录组图谱。**

单细胞转录组(scRNA-seq) 数据库

JOURNAL ARTICLE

TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment

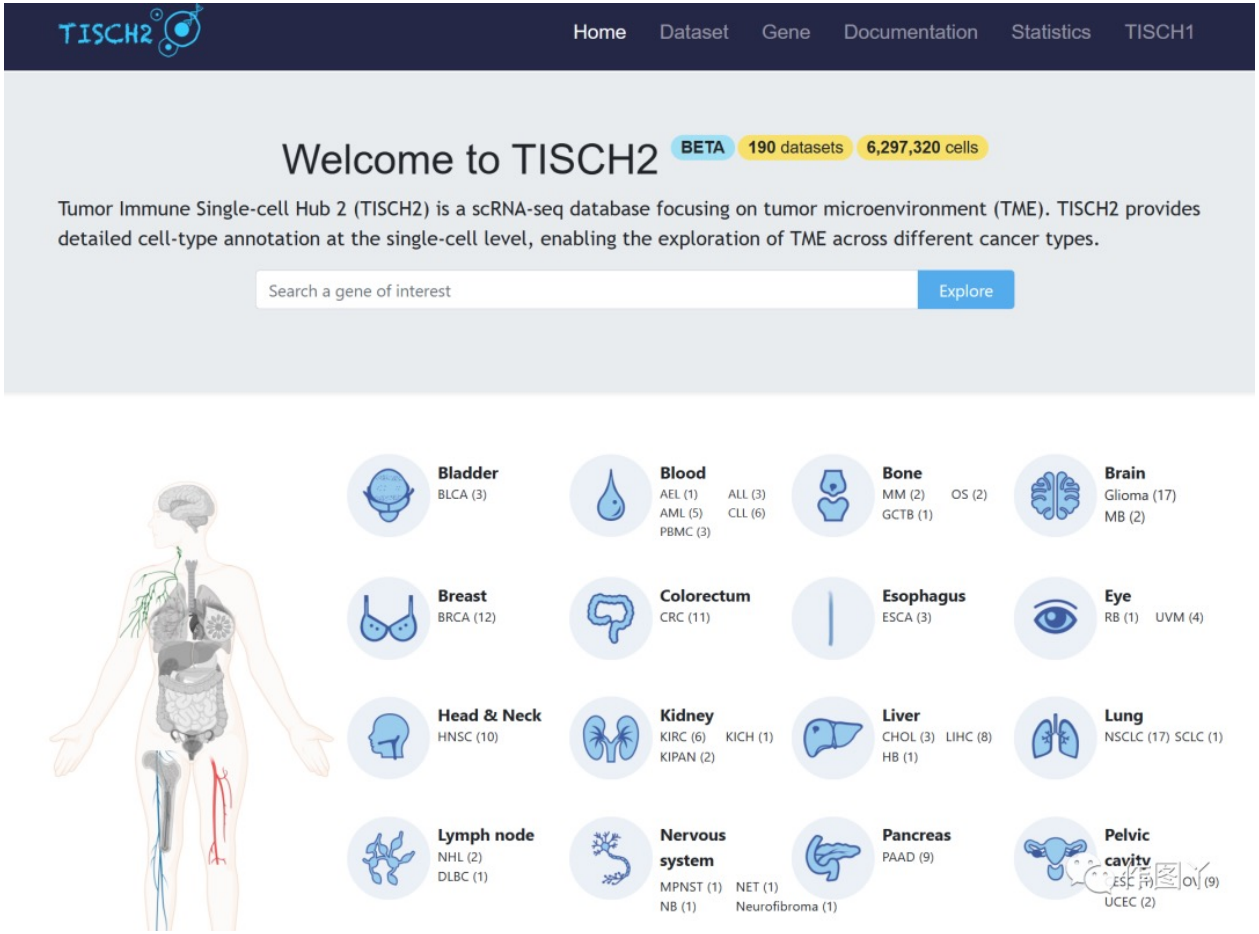
Ya Han, Yuting Wang, Xin Dong, Dongqing Sun, Zhaoyang Liu, Jiali Yue, Haiyun Wang, Taiwen Li ✉, Chenfei Wang ✉ Author Notes

Nucleic Acids Research, Volume 51, Issue D1, 6 January 2023, Pages D1425–D1431,
https://doi.org/10.1093/nar/gkac959
Published: 02 November 2022 **Article history** ▾

- **2022年12月 Tumor Immune Single Cell Hub2(TISCH2) 第二版发布**
- **数据增加至 50种癌症类型, 190个高质量肿瘤数据集, 总计 600 万个细胞的单细胞转录组图谱(细胞数为之前三倍)。**
- **增加了细胞通讯分析以及转录因子预测的可视化分析**

http://tisch.comp-genomics.org/

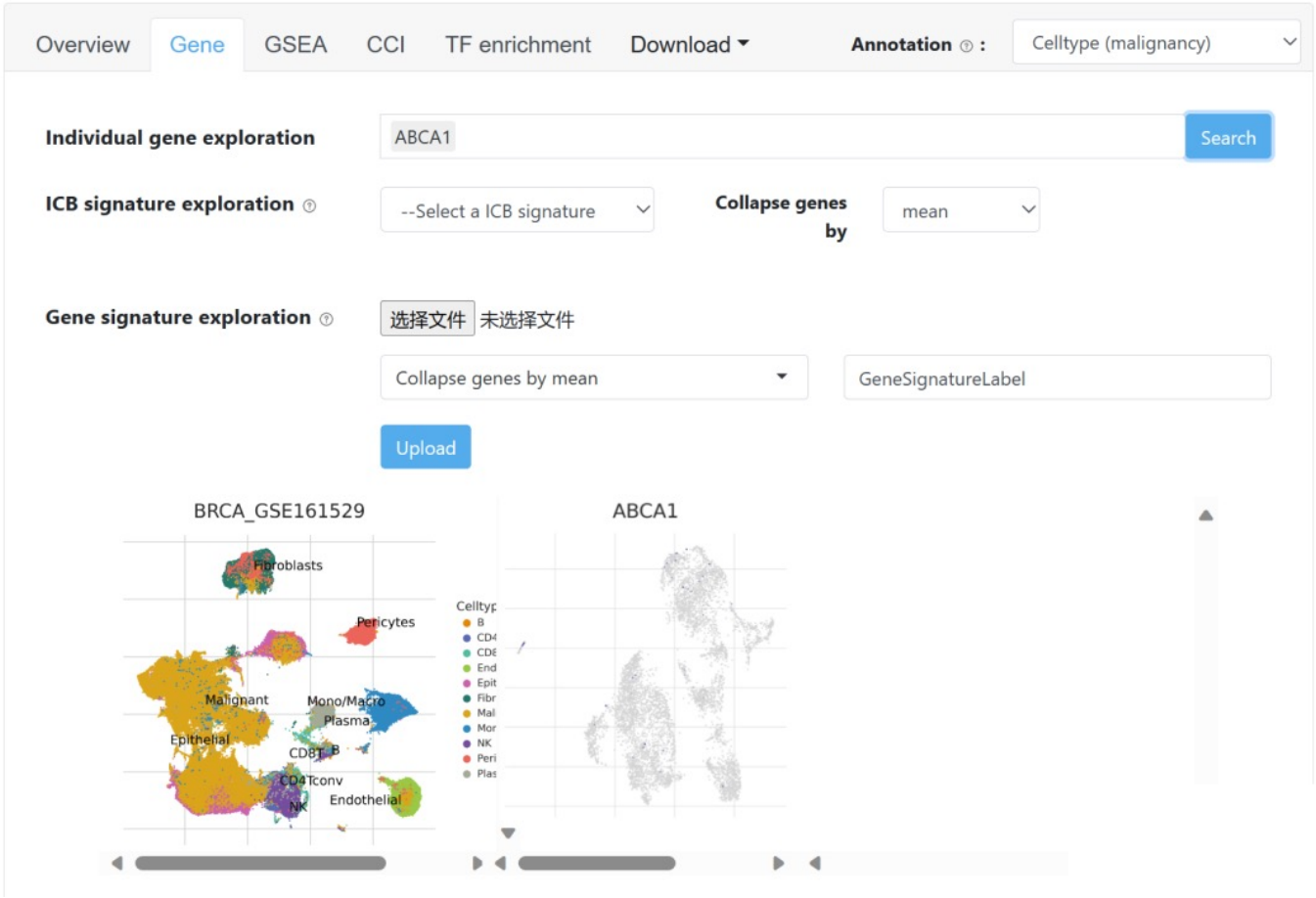
单细胞转录组(scRNA-seq) 数据库



- **数据库主页显示支持在线分析的肿瘤类型种类(命名与TCGA一致)**

http://tisch.comp-genomics.org/

单细胞转录组(scRNA-seq) 数据库

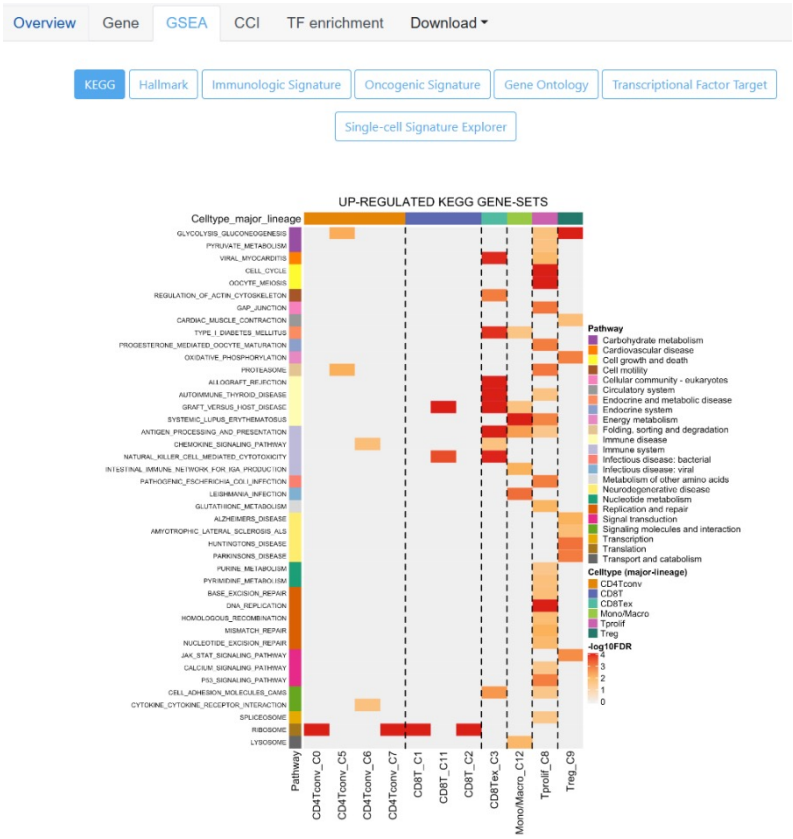


- 支持查看**自定义**基因在单细胞数据中 **分群**结果和**基因表达**情况

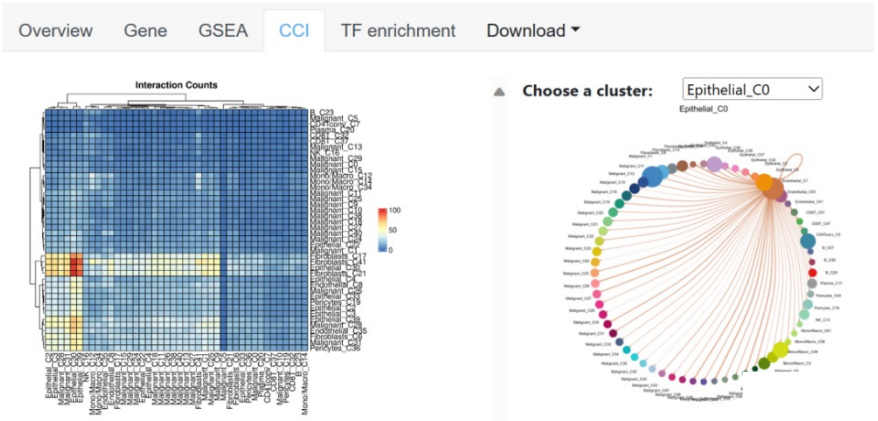
http://tisch.comp-genomics.org/

单细胞转录组(scRNA-seq) 数据库

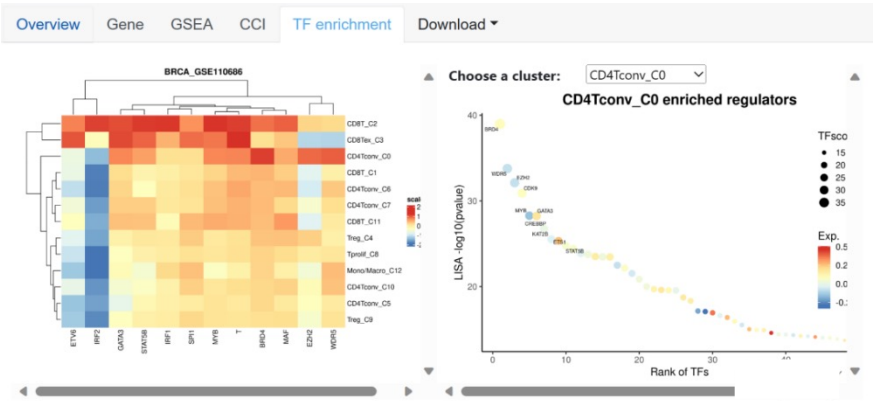
GSEA 通路分析



细胞互作分析



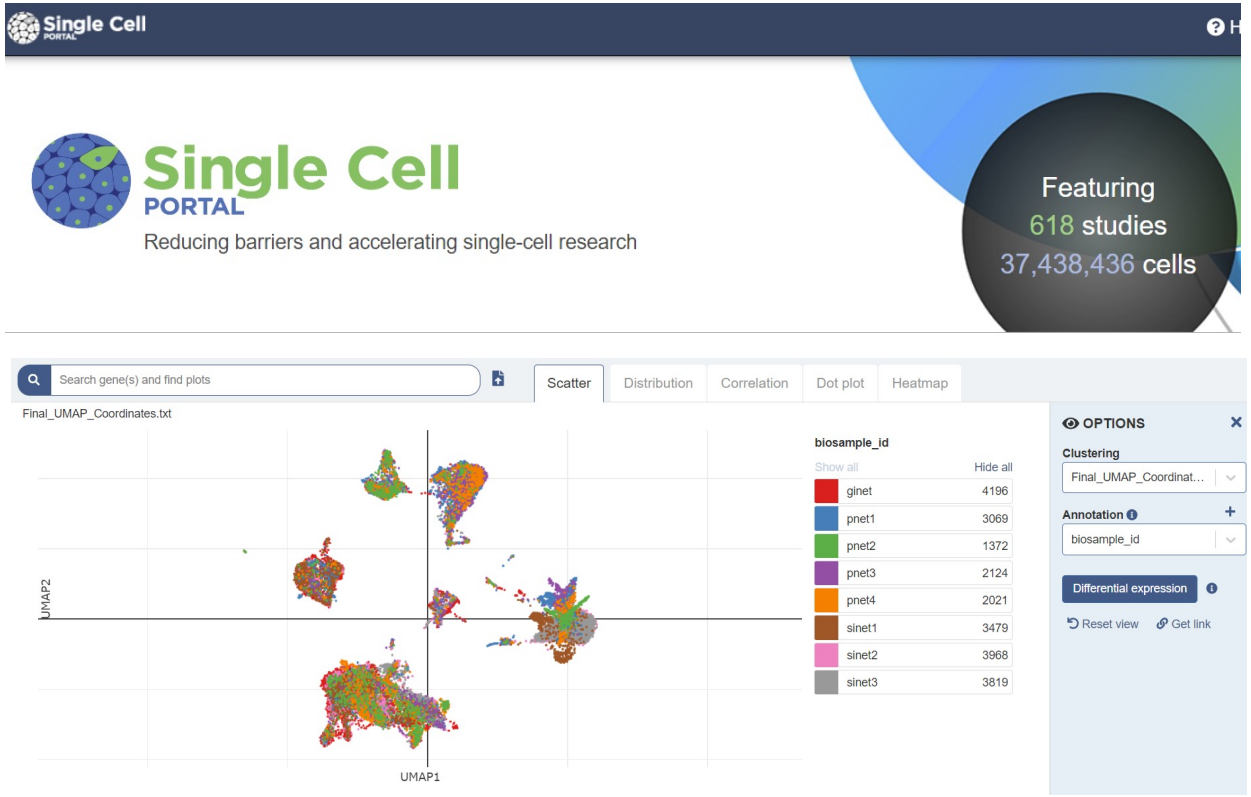
转录因子预测分析



- 支持查看在单细胞数据中**通路富集**分析，**细胞互作**分析以及**TF预测**分析

http://tisch.comp-genomics.org/

单细胞转录组(scRNA-seq) 数据库



- 来自美国博德研究所(隶属于MIT和哈佛大学)建立了更加全面scRNA-seq检索收录数据库
- 数据搜集包括618项研究中的3700W+细胞
- 支持单个研究的数据对应查看基因表达情况，缺少大规模整合

https://singlecell.broadinstitute.org/single_cell

空间转录组(ST) 数据库

nature methods

Resource

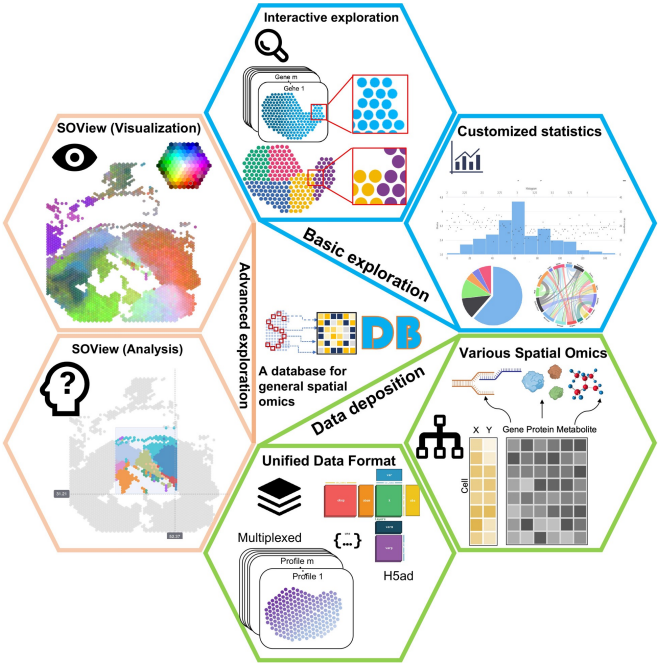
https://doi.org/10.1038/s41592-023-01773-7

SODB facilitates comprehensive exploration of spatial omics data

Received: 10 August 2022

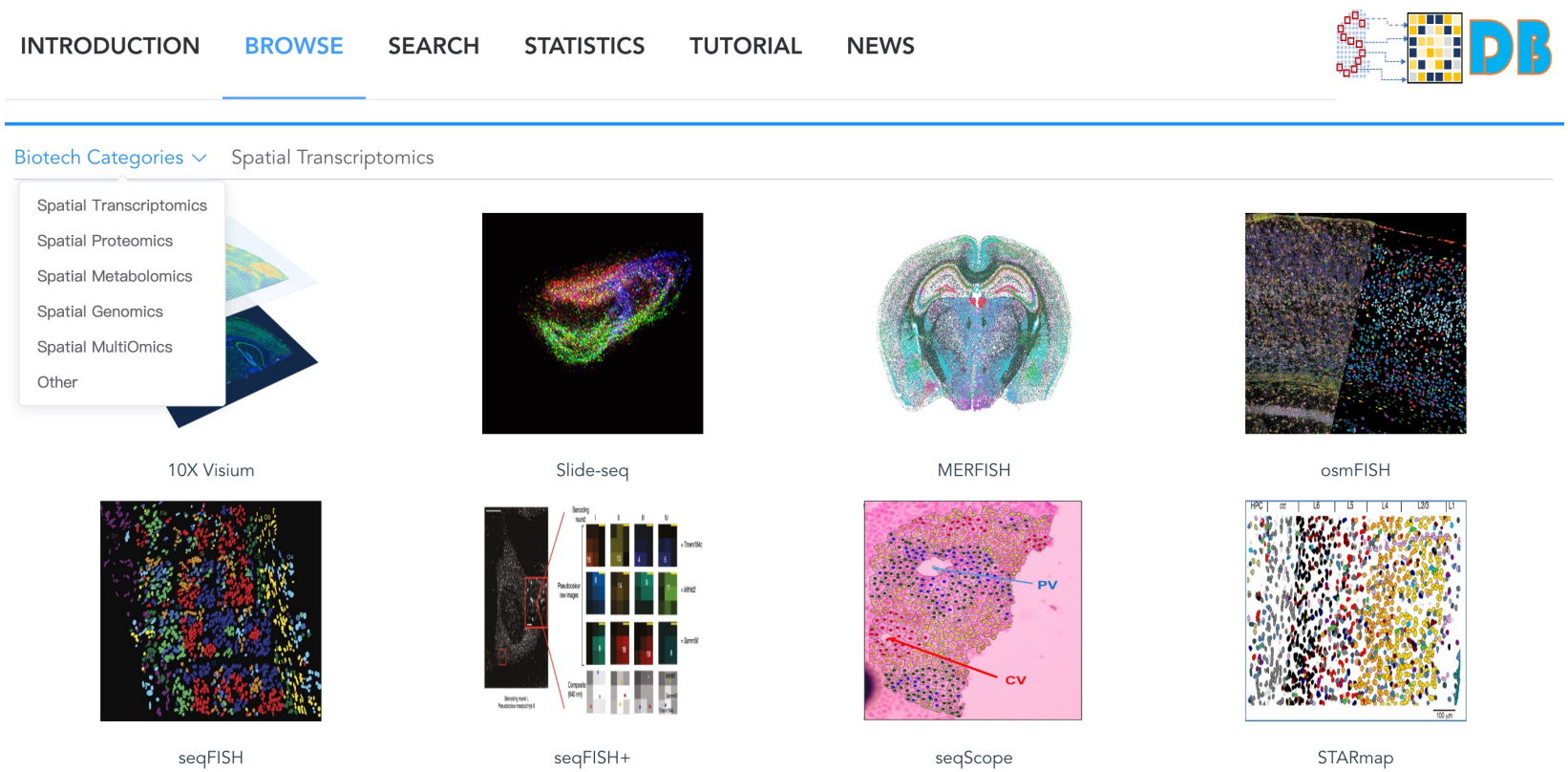
Accepted: 6 January 2023

Zhiyuan Yuan[1,2,7], Wentao Pan[2,3,7], Xuan Zhao[2], Fangyuan Zhao[4,5], Zhimeng Xu[2], Xiu Li[3], Yi Zhao[4,5], Michael Q. Zhang[6] & Jianhua Yao[2]



- 2023年1月 复旦大学类脑研究院联合德州大学达拉斯分校以及腾讯 AI lab联合开发了空间转录组第一个公开发表的大规模数据库发表于Nature Methods
- 收录了26种空间组学技术，细胞数5000W+，同时开发优化了超大数据的内存读取效率工具包-pysodb，时间效率和内存提升均在百倍以上。
- 支持多种在线数据分析和可视化模块，包括基因空间表达、细胞类型注释、基因表达比较等

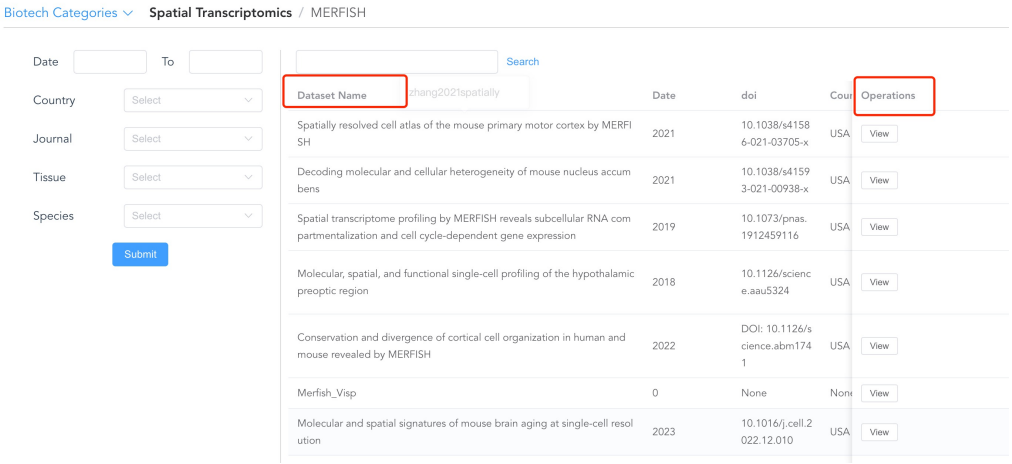
空间转录组(ST) 数据库



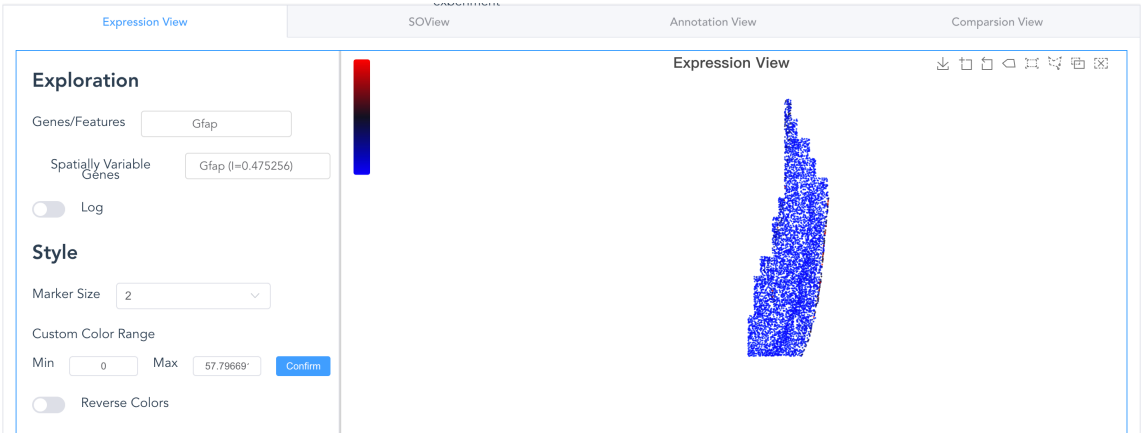
https://gene.ai.tencent.com/SpatialOmics/

- **可以通过选择技术类型以及单个研究数据进行查看(类似于**singlecell portal)
- **除了空间转录组数据，数据库同时支持其他空间代谢组，空间蛋白组等数据**

空间转录组(ST) 数据库



- **支持选择近年来不同研究的数据进行单独查看**



https://gene.ai.tencent.com/SpatialOmics/

Gfap基因 Merfish技术

- **可以通过左边选择对应的可视化参数查看不同基因的空间表达以及细胞注释**

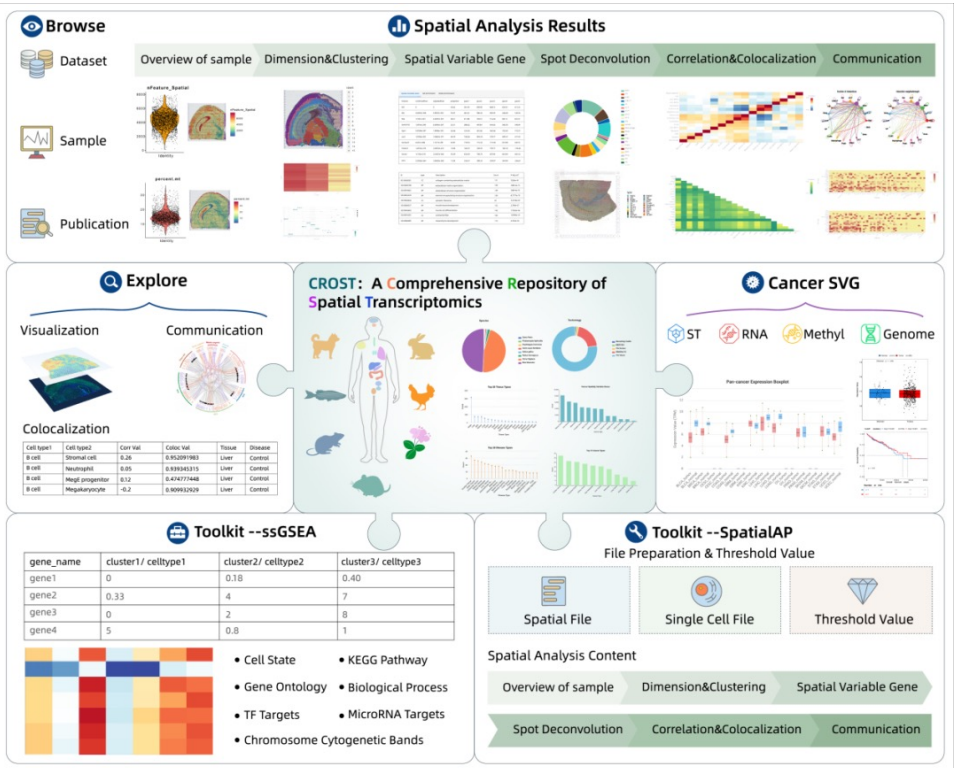
空间转录组(ST) 数据库

JOURNAL ARTICLE

CROST: a comprehensive repository of spatial transcriptomics

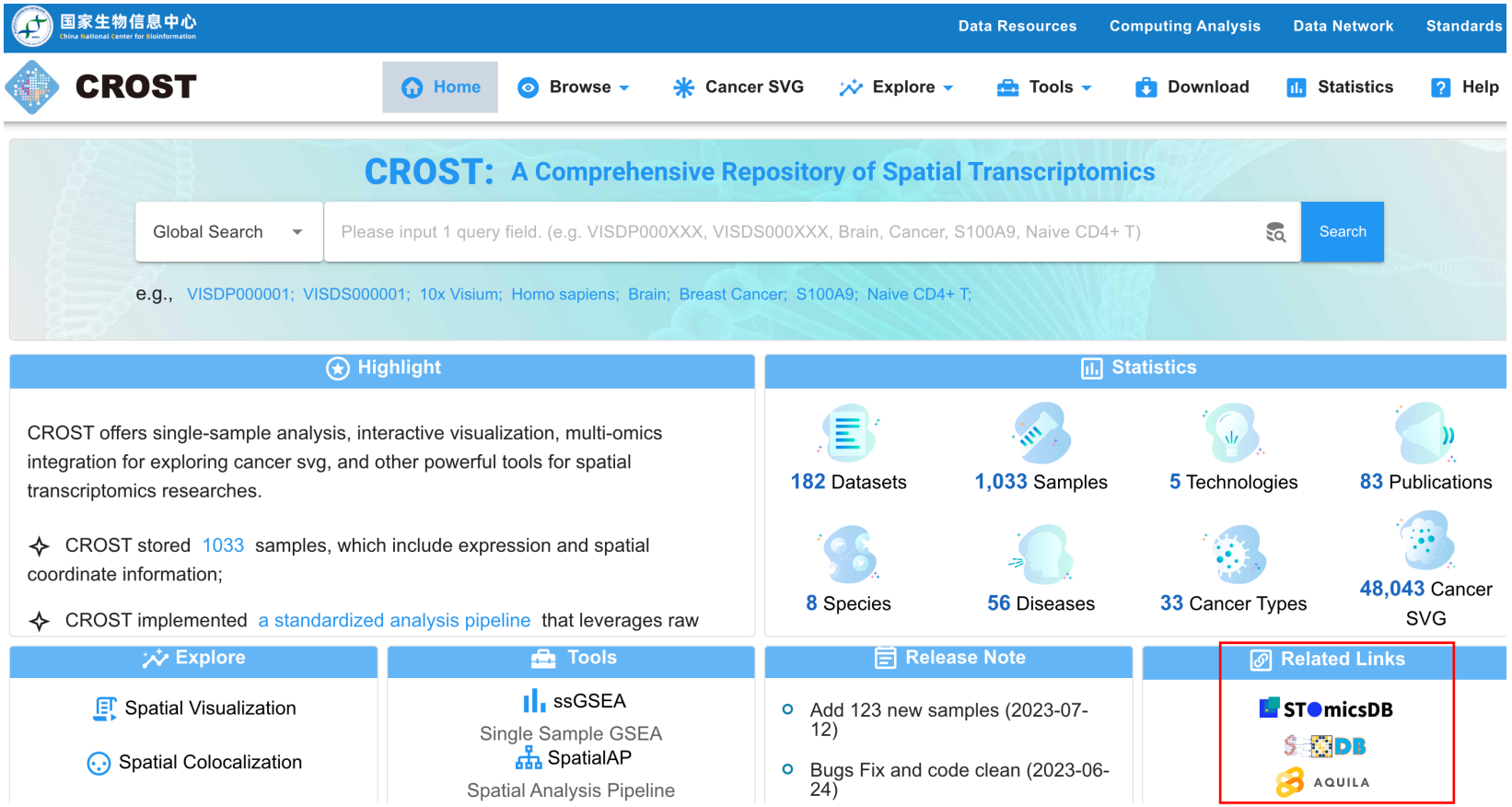
Guoliang Wang, Song Wu, Zhuang Xiong, Hongzhu Qu ✉, Xiangdong Fang ✉, Yiming Bao ✉ Author Notes

Nucleic Acids Research, gkad782, https://doi.org/10.1093/nar/gkad782
Published: 04 October 2023 **Article history ▼**



- **2023年10月 北京基因组所国家基因组科学数据中心鲍一明和方向东研究员开发了更大规模的空间转录组数据库，发表于NAR**
- **收录了8类物种，35个组织类型，56种疾病，182个高质量空间转录组数据集**

空间转录组(ST) 数据库

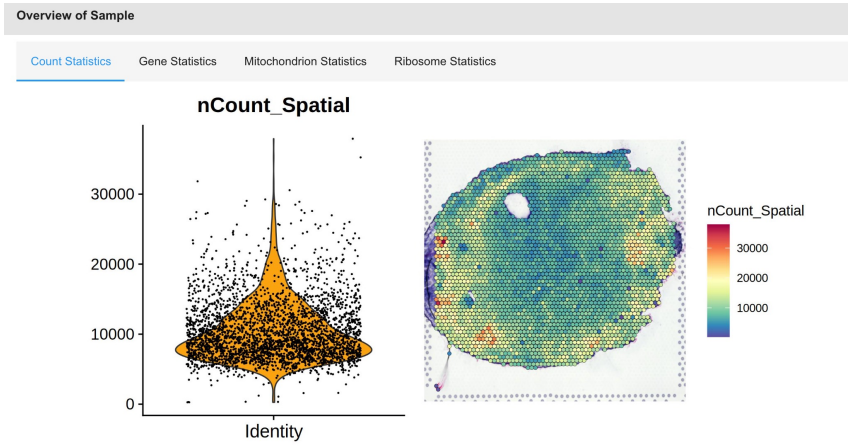


- **该数据库质量和可操作性很高，用户友好度高，同时兼有相关数据库链接**

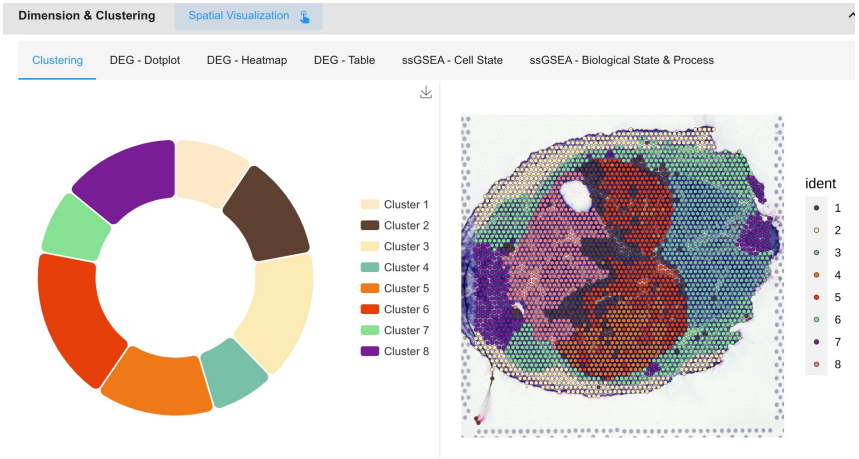
https://ngdc.cncb.ac.cn/crost/

空间转录组(ST) 数据库

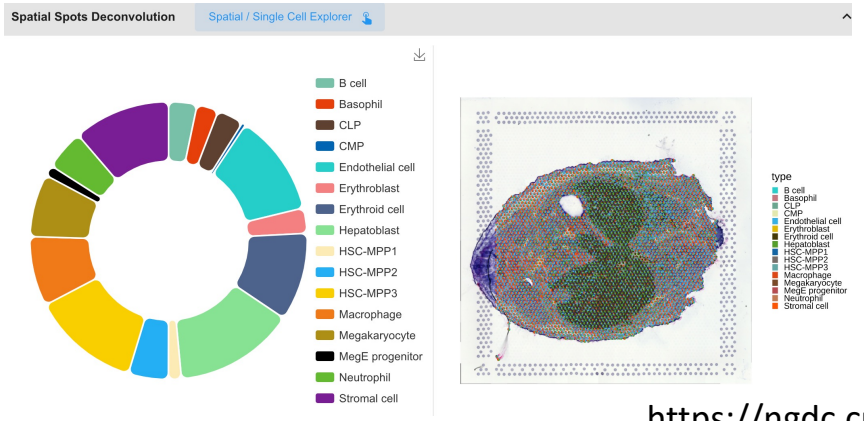
基因查看空间分布



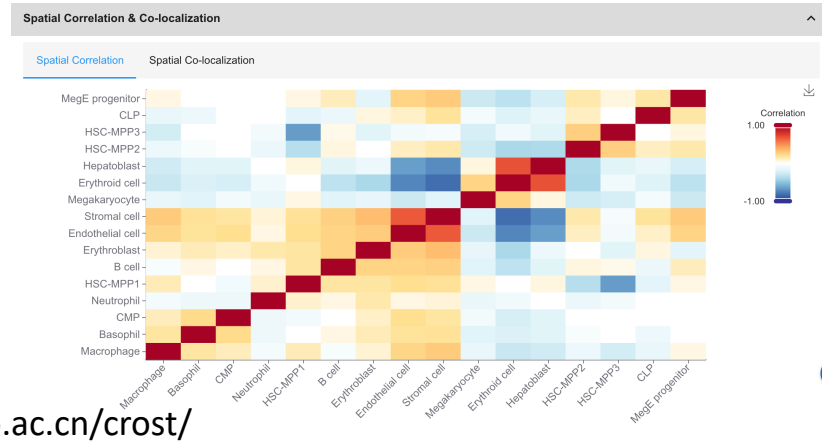
分群与差异基因分析



空间去卷积分析



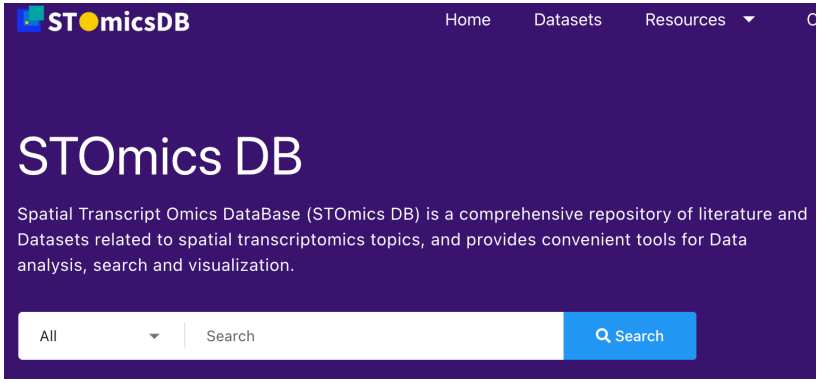
空间共定位分析



https://ngdc.cncb.ac.cn/crost/

- 支持**多种在线分析方法**，通过选择**某一个研究的数据**继而可以进行基因查看，分群和差异基因分析，空间去卷积，共定位，细胞互作等分析

空间转录组(ST) 数据库



STOmicsDB: a database of spatial transcriptomic data

Z Xu, W Wang, T Yang, J Chen, Y Huang, J Gould, W Du, F Yang, L Li, T Lai, C Hua, S Hu…

bioRxiv, 2022 • biorxiv.org

Abstract

Recent technological development in spatial transcriptomics allows researchers to measure gene expression of cells and their spatial locations at the almost single-cell level, which generates detailed biological insight into biological processes. However, specialized spatial transcriptomics databases are rare. Here, we present the Spatial TranscriptOmics DataBase (STOmicsDB), a user-friendly database with multifunctions including search of relevant publications and tools, public dataset visualization, customized specialized

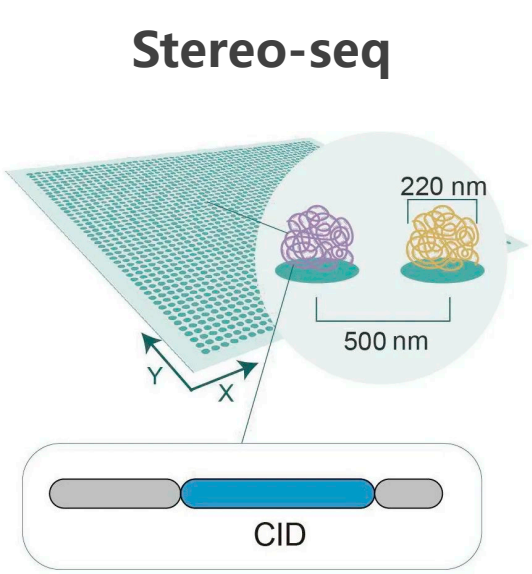
展开

☆ 保存 引用 被引用次数：12 相关文章 所有 2 个版本 »

https://db.cngb.org/stomics/

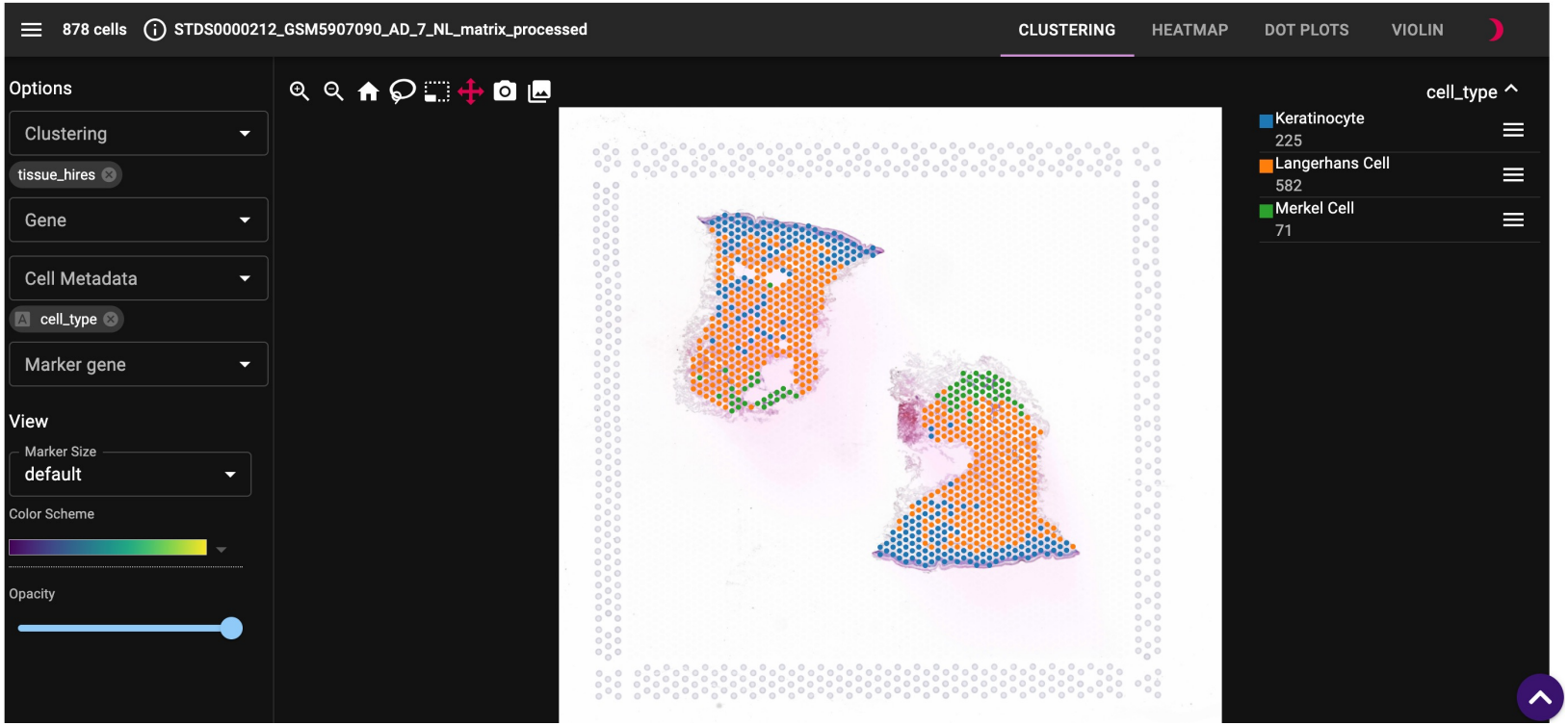
- **2022年 华大基因已经着手开发空间转录组大规模数据库，目前发表于预印本上，尚未正式发表**
- **截止目前收录了17类物种，128个组织类型，218个高质量空间转录组数据集**

空间转录组(ST) 数据库



- 近年来华大基因自主研发的空间转录组技术Stereo-seq 领域内影响很大，其多个图谱类工作均发表在CNS上(小鼠胚胎发育Cell, 蝾螈脑Science)，并且还有很多子刊
- STOmicsDB相比于其他数据库更多的亮点在于其自研技术衍生数据的展示

空间转录组(ST) 数据库



https://db.cngb.org/stomics/

- 支持**多种在线分析方法**，通过选择**某一个研究的数据**可以进行基因可视化交互
- 其他进阶诸如**差异基因分析，空间互作**等分析为结果展示，暂不支持交互展示

课程总结

1 数据库基本概念

2 NCBI & EMBL-EBI & GO

3 TCGA GTEx & GEPIA 1/2/2021

4 单细胞/空间转录组数据库