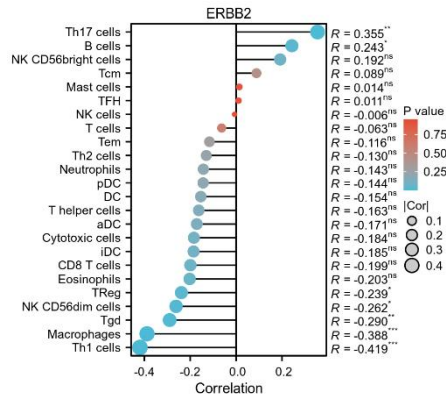


交互网络 - [免疫浸润-云] 棒棒糖图



网址: <https://www.xiantao.love>



更新时间: 2023.03.10

目录

基本概念	3
应用场景	4
分析流程	4
主要结果	5
云端数据	6
参数说明	7
特殊参数	7
算法	8
统计	8
映射	9
点	10
线	11
标注	11
标题	12
图注(Legend)	12
坐标轴	13
风格	14
图片	15
结果说明	16
主要结果	16
补充结果	17
方法学	18
如何引用	19
常见问题	20

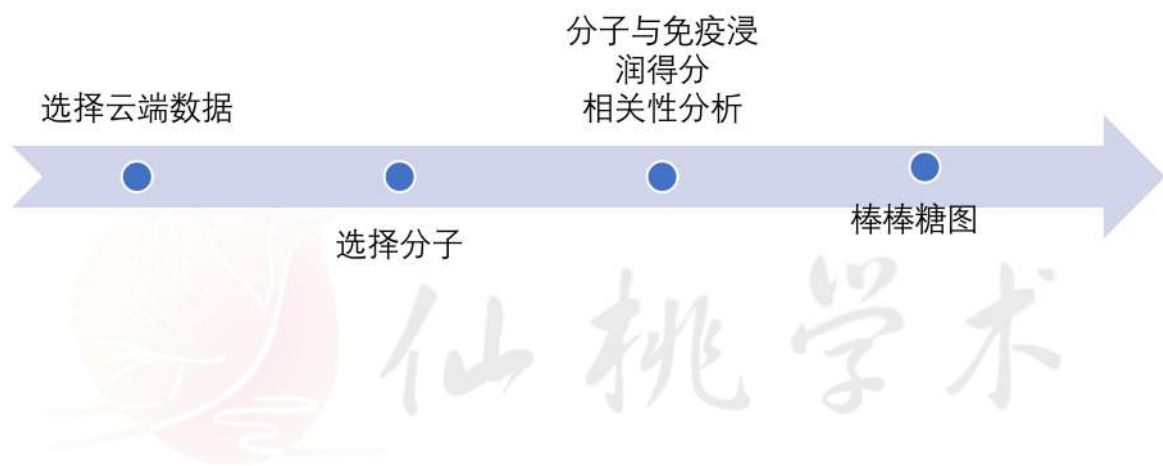
基本概念

- 免疫浸润分析：利用转录组或者其他组学的数据，通过算法计算组织中免疫细胞的分数情况，推测组织中免疫细胞的构成情况。
- 免疫浸润算法
 - **ssGSEA (single sample GSEA)**：用给定的基因集对每一个样本计算对应基因集的富集得分。在免疫浸润中的 ssGSEA 则是利用每一类免疫细胞特定的 markers（来源于说明文本中的引文）作为基因集去计算每个样本中每一类免疫细胞的富集得分，推断单个样本的免疫细胞的浸润情况。
 - **estimate**：estimate 包提供的通过内置的 markers 计算样本的免疫浸润得分、基质得分和估计得分。（主要看基质得分和免疫浸润得分，可以简单理解得分越高，内环境中基质成分和免疫浸润越高）
- 棒棒糖图：可以用于可视化 1 个分子和多个细胞分数的相关性情况。此时，圆圈的大小和棒子的高度均代表相关性程度，颜色的深浅代表 p 值的大小情况。
- 统计方法：
 - **Pearson 相关**：参数相关性检验，衡量两组之间是否存在线性关系，数据需要满足双正态性。
 - **Spearman 相关**：非参数相关性检验，通过秩次来判断两组是否存在相关性。数据可以不需要满足正态性，如果不懂具体的选择条件，可以先选择该方法。

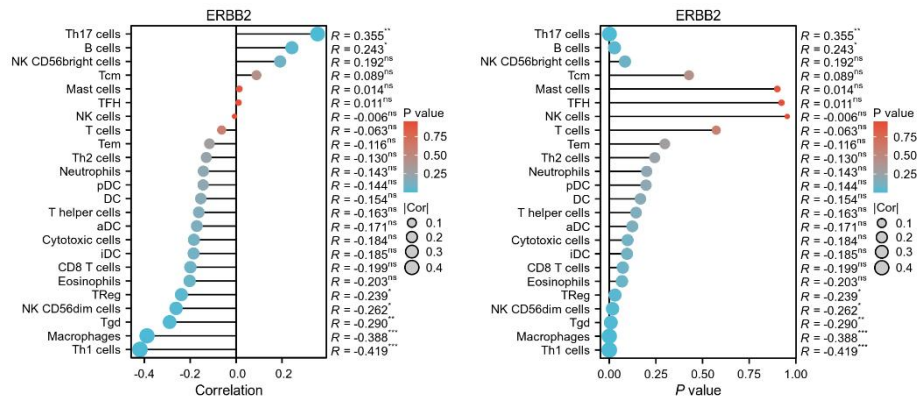
应用场景

基于公共数据（云端数据）直接分析所选分子与预先计算好的免疫细胞分数之间的相关性。一般绘制棒棒糖图进行直观比较。

分析流程



主要结果



通过棒棒糖图展示相关性结果。

- 纵坐标代表算法下的细胞种类，横坐标代表所选择的【y轴映射】内容。（可选相关系数 或 p 值，注意绘图的默认风格为 xy 轴颠倒）
 - 相关性大小，相关性为负数，则说明分子和对应的细胞负相关；相关性为正数，则说明分子和对应的细胞正相关。
 - p 值显著性，值越小，说明对应关系的可靠性越高。
- 图中点（圈）的颜色为所选择的【颜色映射】内容。（如上图，对应 p 值）
- 图中点（圈）的大小为所选择的【大小映射】内容。（如上图，对应相关系数绝对值 |Cor| ）
 - 相关性程度（|Cor|），相关性程度越高，圆圈相对越大。
- 可以认为相关性系数达到某个值以上就是相关（0.3, 0.5, 0.7 等等），或者是 p 值满足统计学意义认为是相关。

云端数据

数据参数
重置参数

云端数据

食管鳞癌 / TCGA / TCGA-ESCC / RNAseq / STAR / TPM @过滤:去除正常 @处理:log2(value+1)

↓
↓

疾病名/来源/数据集/平台/分析流程/数据格式
@数据处理方式

云端数据
选择疾病
选择数据处理方式, 默认 去除正常 (过滤样本)、log2(value+1)
×

疾病 请选择 ▼

数据过滤: 去除正常 ▼

数据格式: log2(value+1) ▼

	疾病系统	疾病名	疾病英文	来源	获取时间	数据集	平台	Wo
<input checked="" type="checkbox"/>	食管	食管鳞癌	Esophagus squamous cell carcinoma	TCGA	202208	TCGA-ESCC	RNAseq	STA
<input type="checkbox"/>	食管	食管癌	Esophageal carcinoma	TCGA	202208	TCGA-ESCA	RNAseq	STA
<input type="checkbox"/>	食管	食管癌	Esophageal carcinoma	TCGA	202208	TCGA-ESCA	RNAseq	STA
<input type="checkbox"/>	食管	食管腺癌	Esophagus adenocarcinoma	TCGA	202208	TCGA-ESAD	RNAseq	STA

共 78 条 上一页 1 2 3 4 5 6 ... 8 下一页

只有合适这个模块的云端数据才会展示

确认

本模块提供预清洗好的云端数据，不同平台的云端数据集的分子可能会有不同。注意查看当前数据参数选中的云端数据。

参数说明

(说明：标注了颜色的为常用参数。)

特殊参数



特殊参数

分子: ERBB2[ENSG00000141736.14]

OAZ1[ENSG00000104904.12]

TRMT1[ENSG00000104907.12]

STX10[ENSG00000104915.15]

RETN[ENSG00000104918.8]

FCER2[ENSG00000104921.15]

DMPK[ENSG00000104936.19]

CLEC4M[ENSG00000104938.18]

RSPH6A[ENSG00000104941.8]

主要参数

算法

标注

图片

- 分子：下拉框将列出对应所选数据集分子，可以输入关键字搜索分子，基因 symbol 或 Ensembl ID，[只能选单个分析](#)。

算法

- **算法**：选择预先计算好的每个样本的免疫细胞浸润分数的方法，可以选择 ssgsea、estimate。不同算法之间对应的内容会有一定差别，选择好算法后，下面“细胞”参数需要点击“确认”后会进行更新，具体算法的参考文献可查看返回结果的<方法学>对应的内容。
- **细胞**：算法下提供的细胞，默认是选择所有内容，具体细胞类型见<方法学>的补充说明。此参数需要点击“确认”后才会根据算法进行更新。可多选。

统计

- **统计方法**：可选择 Spearman、Pearson。统计方法的选择依据可以参考“基本概念”中统计方法的说明。

映射

映射	
y轴映射	相关系数
颜色映射	p值
大小映射	相关系数绝对值

- y 轴映射：主要影响 y 轴（默认横坐标）的取值。可选择 相关系数、p 值。
- 颜色映射：主要影响点（棒子上的圈）的颜色范围，注意映射内容的数值类型，数值型数据为渐变色，不映射时只显示一种颜色。可选择 p 值、相关系数、无。
- 大小映射：主要影响点（棒子上的圈）的相对大小。可选择 相关系数绝对值、p 值、无。

点



- **填充色**：点的填充色颜色选项，取决于 [颜色映射](#) 参数所选择的内容，展示数值型内容时，修改第一和第二色卡作为数值从小到大的渐变色，最多支持修改 2 个颜色。受配色方案全局性修改。
- **描边色**：点的描边色颜色选项，取决于 [颜色映射](#) 参数所选择的内容，展示数值型内容时，修改第一和第二色卡作为数值从小到大的渐变色，最多支持修改 2 个颜色。受配色方案全局性修改。
- **样式**：点的样式类型，可选择 [圆形、正方形、菱形、三角形、倒三角](#)。单选，[选择后将全局变化](#)。
- **大小比例**：点的大小。
- **不透明度**：点的透明度。0 为完全透明，1 为完全不透明。

线

线

颜色

类型

实线

粗细

0.75pt

不透明度

1

轴线与点之间的连线。

- 颜色：线的颜色，默认为纯黑，不受配色方案全局性影响。
- 类型：线条的类型，可选 实线、虚线。
- 粗细：线的粗细，默认为 0.75pt
- 不透明度：线的透明度。0 为完全透明，1 为完全不透明。

标注

标注

内容

相关系数-星

标注大小

6pt

- 内容：影响图中右侧的展示信息，默认为相关系数-星号。可以选 不标注、相关系数、相关系数-星号、星号、p 值科学计数法、p 值科学计数法、p 值数值(小于 0.05 自动<)、p 值数值(小于 0.001 自动<)。

- 标注大小：标注的文字大小。

标题

标题

大标题

大标题内容

x轴标题

x轴标题内容

y轴标题

y轴标题内容

- 大标题：大标题文本
- x 轴标题：x 轴标题文本
- y 轴标题：y 轴标题文本
- 补充：在要换行的中间插入\n。如果需要上标，可以用两个英文输入法下的大括号括住，比如 {{2}}；如果需要下标，可以用两个英文输入法下的中括号括住，比如 [[2]]

图注(Legend)

图注

是否展示

☒

图注位置

默认

- 展示：是否展示图注
- 图注位置：可选 默认、右、上，默认为右。

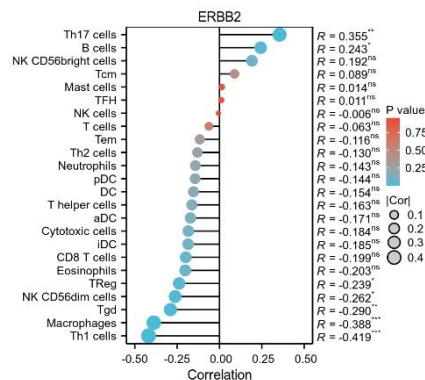
坐标轴

坐标轴

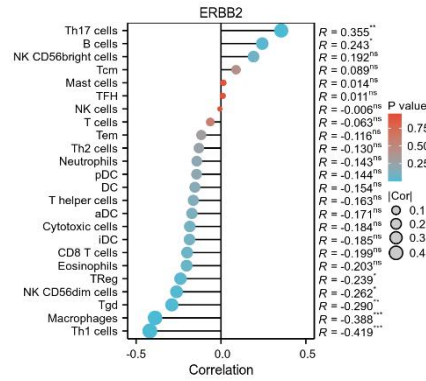
x轴标注旋
转
0

y轴范围+刻度 逗号隔开

- X 轴标注旋转：支持对 x 轴文字进行旋转。适合于 x 轴文字过长的时候。
- Y 轴范围+刻度：（注意：范围的修改如果调整过大会失效）
 - 如果只是想要修改范围，可以只输入两个范围值，比如 -0.5,0.5，自动调整刻度



- 如果同时想要修改范围+刻度，可以输入比如：-0.5,-0.5,0,0.5,0.5 。注意，此时最大和最小值会被当做范围值，不会作为刻度，如果需要刻度，需要类似于 0.5 那样同时写两次



风格

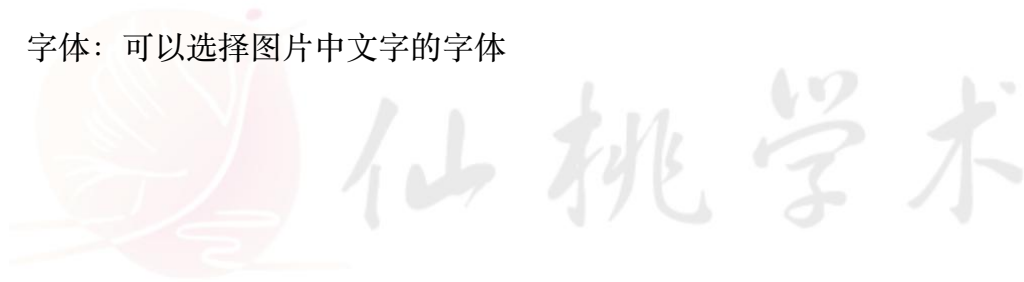


- 边框：是否添加外框
- 网格：是否添加网格
- xy 颠倒：可以颠倒 xy 轴
- 文字大小：针对图中所有文字整体的大小控制

图片

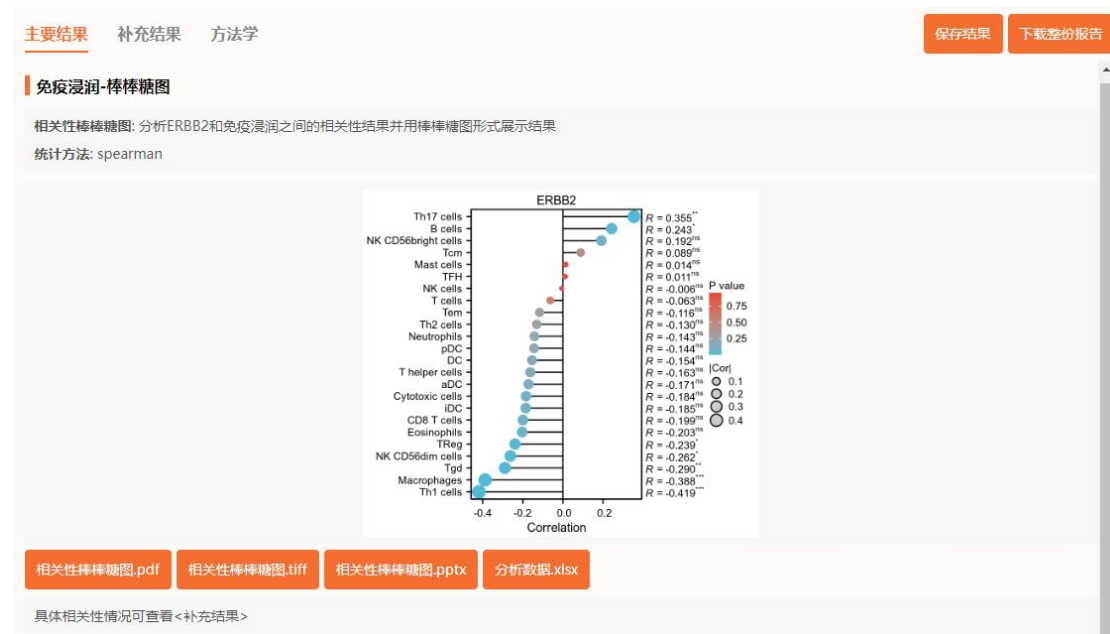
图片	▼
宽度 (cm)	8
高度 (cm)	7
字体	Arial ▼

- 宽度：图片横向长度，单位为 cm
- 高度：图片纵向长度，单位为 cm
- 字体：可以选择图片中文字的字体



结果说明

主要结果



主要结果格式为图片格式，提供 PDF、TIFF、PPTX 格式下载，结果报告可以下载包括 pdf 以及说明文本的内容。

➤ 此外，还提供公共数据中所选分子在不同样本中的表达数据及对应的免疫浸润得分，提供 EXCEL 格式下载：

	A	B	C	D	E	F
1	sample_id	ERBB2	aDC	B cells	CD8 T cells	Cytotoxic cel
2	TCGA-IC-A6RF-01A-13R-A336-31	7.08817977	0.61923636	0.44807042	0.65105878	0.42148459
3	TCGA-IG-A3I8-01A-11R-A24K-31	5.3260591	0.38598375	0.11404171	0.67006305	0.33192069
4	TCGA-IG-A3QL-01A-11R-A24K-31	5.81655619	0.35905094	0.14446447	0.65406727	0.33230351
5	TCGA-IG-A3YA-01A-11R-A24K-31	5.08085802	0.62174938	0.42042542	0.68014127	0.46116118
6	TCGA-IG-A3YB-01A-11R-A36D-31	5.45138587	0.60277327	0.18440524	0.68562326	0.4656253
7	TCGA-IG-A3YC-01A-11R-A24K-31	5.3026371	0.63814237	0.31658942	0.69536115	0.65876444
8	TCGA-IG-A4P3-01A-11R-A260-31	6.24438667	0.70767333	0.42473756	0.71399362	0.65144359
9	TCGA-IG-A50L-01A-11R-A260-31	4.93491274	0.46948508	0.12499211	0.66802117	0.37076345
10	TCGA-IG-A51D-01A-11R-A36D-31	5.84043554	0.68195733	0.30289088	0.67552269	0.57623703
11	TCGA-IG-A5B8-01A-11R-A28J-31	3.43170318	0.52258351	0.12490554	0.66550139	0.45682779
12	TCGA-IG-A5S3-01A-11R-A28J-31	5.85764018	0.47461414	0.27399567	0.66619587	0.44965659
13	TCGA-IG-A625-01A-11R-A31P-31	5.40525041	0.46444336	0.26148114	0.68779561	0.44540146

补充结果

相关性分析

提供Pearson和Spearman统计方法的结果

主变量	次变量	自由度(df)	统计量-Pearson	相关系数-Pearson	p值-Pearson	统计量-Spearman	相关系数-Spearman
ERBB2	aDC	80	-1.52964	-0.168572	0.1301	1.076e+05	-0.171311
ERBB2	B cells	80	2.3729	0.256428	0.0200	6.954e+04	0.24313
ERBB2	CD8 T cells	80	-1.49341	-0.164689	0.1393	1.102e+05	-0.19933
ERBB2	Cytotoxic cells	80	-1.43594	-0.158513	0.1549	1.088e+05	-0.18387
ERBB2	DC	80	-1.3349	-0.147611	0.1857	1.06e+05	-0.15386
ERBB2	Eosinophils	80	-1.64712	-0.181108	0.1035	1.105e+05	-0.20251
ERBB2	IDC	80	-2.02734	-0.221056	0.0460	1.089e+05	-0.18531
ERBB2	Macrophages	80	-3.41582	-0.356768	0.0010	1.276e+05	-0.38827
ERBB2	Mast cells	80	0.747594	0.0832931	0.4569	9.06e+04	0.013985
ERBB2	Neutrophils	80	-1.1801	-0.130806	0.2415	1.05e+05	-0.14265
ERBB2	NK CD56bright cells	80	1.63297	0.179603	0.1064	7.427e+04	0.191715
ERBB2	NK CD56dim cells	80	-2.21746	-0.240635	0.0294	1.159e+05	-0.26169

相关性.xlsx

此表格提供 Pearson 和 Spearman 统计方法的分析结果，即所有的相关性分析结果，提供 EXCEL 格式下载。



方法学

所有分析和可视化均在 R 4.2.1 中进行

涉及的 R 包：ggplot2 包（用于可视化）

处理过程：对数据中主变量和免疫浸润矩阵数据之间进行相关性分析，分析结果用 ggplot2 包进行棒棒糖图可视化。



如何引用

生信工具分析和可视化用的是 R 语言，可以直接写自己用 R 来进行分析和可视化即可，可以无需引用仙桃，如果想要引用仙桃，可以在致谢部分 (Acknowledge) 致谢仙桃学术 (www.xiantao love)。

方法学部分可以参考对应说明文本中的内容以及一些文献中的描述。



常见问题

1. 方法里面的 Spearman 和 Pearson 方法，应该选择哪一个？

答：

两种方法均可以选择。Pearson 会要求数据是满足正态性，Spearman 可以不需要满足正态性，可以先选择非参数的 Spearman 相关进行尝试。

2. 怎么判断相关性好坏（结果好坏）？

答：

这个没有很统一的标准，可以是相关系数在 0.3、0.5、0.7 以上（不一定说要非常相关(>0.7)才能放文章里面），也可以是 p 值，具体根据研究情况进行判断。

3. ssGSEA 算法（其他一些算法）给到的结果跟 其他方法（或者别的数据库 TIMER 等）趋势不一样，这个是什么原因？

答：

不同算法之间可能是会存在有一定的差别，况且算法只是一种推测手段，实际是什么情况还是需要通过做实验来确定的。所以，如果只是单纯想要拿一些结果来充实自己的研究，那么可以只放满足自己想要的趋势的数据。

4. 免疫浸润可以做什么实验验证？

答：

可以通过免疫组化检测对应的免疫细胞的 markers，也可以对组织做流式分析分析细胞的情况等等。具体要根据研究情况进行安排。

5. 为什么“算法选项卡”中的细胞选项的内容对不上？

答：

如果更换了算法的参数，需要重新点击“确认”按钮后，选项卡中的【细胞】选项才会变成对应算法的选项。

6. 统计学标注可以用具体 p 值吗?

答:

在“标注”选项卡中，【内容】参数，里面有显示具体 p 值的选项，也可以选择展示其他内容。

7. 结果 (ssGSEA) 里面的有一些英文简写的细胞具体是哪些细胞?

答:

可以看方法学中对应的 Immunity 的引文: Bindea, Gabriela, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 39.4 (2013): 782-795. 本模块 ssGSEA 算法基于这篇引文（补充材料）提供的 24 种细胞 markers 计算这 24 种细胞的得分情况。如果要知道具体某些细胞是什么，可以先查阅上面那篇引文。

8. 在云端数据框内看到的例数和分析时候的例数不同，这个是什么情况?

答:

云端数据的例数一般是对应组学所有的例数，分析时候可能会有剔除样本的情况，本模块使用 [去除正常+去除无临床信息](#) 的样本，具体需要看说明文本中对于数据的处理情况的说明。

9. 云端数据在哪可以查询?

答:

模块分析后，在方法学标签中，提供了公共数据（云端数据）的具体信息及下载链接。