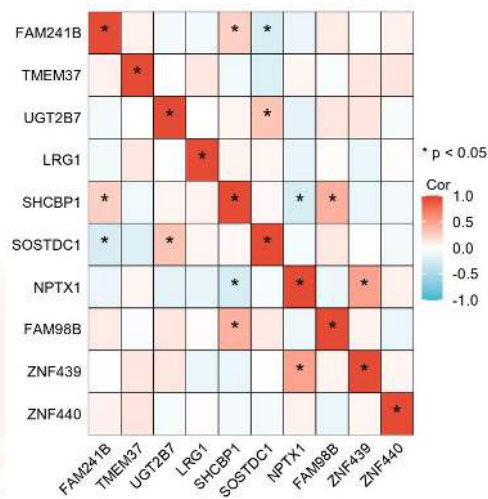


## 交互网络 - 相关性热图-指定 xy[云]



网址: <https://www.xiantao.love>



更新时间: 2023.03.03

## 目录

基本概念 .....	3
应用场景 .....	3
分析流程 .....	3
结果解读 .....	6
数据格式 .....	8
参数说明 .....	9
ID 列表 .....	9
统计 .....	11
映射 .....	12
布局 .....	14
热图 .....	15
标注 .....	17
标题文本 .....	18
图注 .....	18
坐标轴 .....	19
风格 .....	19
图片 .....	20
结果说明 .....	21
主要结果 .....	21
补充结果 .....	22
方法学 .....	24
如何引用 .....	25
常见问题 .....	26

## 基本概念

- 热图：热图是一个以颜色变化来显示数据情况的矩阵
- 相关性热图：通过热图的方式来展示变量之间的相关性

## 应用场景

相关性热图-指定 xy[云]: 基于云端数据 将数据分成两个不同方向上变量（即将所有变量分成两个不同的组别）进行两两间（两分组间）相关性分析，再以热图的形式展示其结果

## 分析流程

云端数据 → 相关性分析 → 可视化

- 云端数据：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：选择了不同的平台，搜索出来的分子可能是不一样的

- 指定 xy:

- ◆ 选择云端数据

- ◆ 选择主要参数[ID 列表] 中 X 方向分子与 Y 方向分子

- **X 方向分子**：表示云端数据中除了第 1 列(样本/样本 id)外的每一列(每一个变量/分子/基因)，至少需要提供 1 个与数据匹配的分子
- **Y 方向分子**：表示云端数据中除了第 1 列(样本/样本 id)外的每一列(每一个变量/分子/基因)，至少需要提供 1 个与数据匹配的分子

- 第 1 列为分类类型，表示样本(样本/样本 id)
- 第 2 列及以后为数值类型数据，表示每个变量/样本值

➤ 相关性分析：

- ID 列表的 X，Y 两个方向的分子列表进行两两间相关性分析，结果如下：



The screenshot shows a web interface for selecting molecules. It has a dropdown menu labeled 'ID列表'. Below it, there are two sections: 'X方向分子' (X-direction molecules) and 'Y方向分子' (Y-direction molecules). Both sections contain a list of three molecules: FAM241B, TMEM37, and UGT2B7. The 'X方向分子' section is highlighted with a red border, and the 'Y方向分子' section is highlighted with a blue border.

◆ 相关性系数表

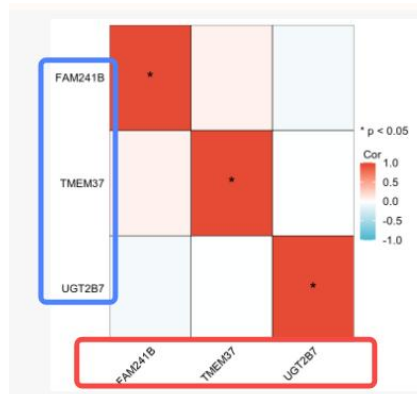
表1: 相关系数表格			
	FAM241B	TMEM37	UGT2B7
FAM241B		0.0753	-0.066
TMEM37	0.0753		-0.00483
UGT2B7	-0.066	-0.00483	

◆ 相关性检验表

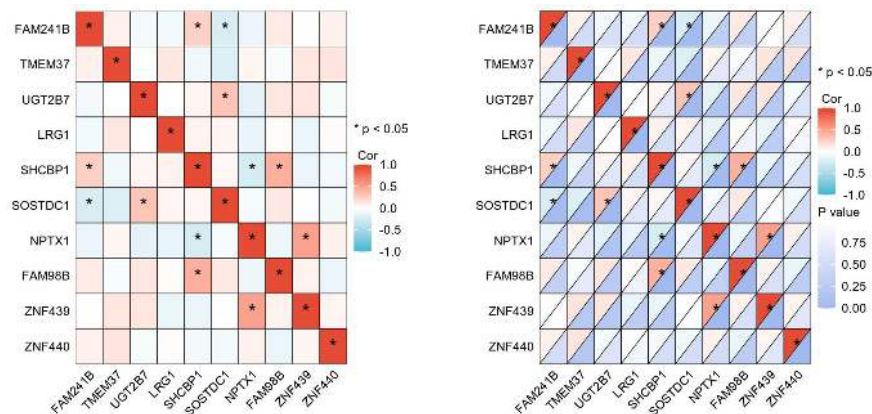
表2: 相关性检验表格

	FAM241B	TMEM37	UGT2B7
FAM241B		0.503	0.558
TMEM37	0.503		0.966
UGT2B7	0.558	0.966	

➤ 将分析所得结果进行可视化



## 结果解读



左图：（完整热图）

➤ 行、列都代表分子（代表不同方向/不同组别分别对应的分子）

■ 行对应主要参数[ID 列表] 中 X 方向的分子

■ 列对应主要参数[ID 列表] 中 Y 方向的分子

➤ 每一个小方格表示分子之间的相关系数，颜色越深，变量间越相关

右图：（完整热图-三角对角线）

➤ 行、列都代表分子

■ 行对应主要参数[ID 列表] 中 X 方向的分子

■ 列对应主要参数[ID 列表] 中 Y 方向的分子

➤ 每一个小方格分为两个部分（三角形），上部分表示分子间的相关系数，颜色越深，分子间越相关；下部分表示 p 值，颜色越深 p 值越小

补充：

➤ \* 表示 Pvalue < 0.05

➤ Correlation 代表相关性系数(Correlation)：|Correlation| 越大，分子间相关性越高，反之相关性越低

- $\text{Correlation} < 0$ , 分子间呈负相关关系
  - $\text{Correlation} = 0$ , 分子间没有相关关系
  - $\text{Correlation} > 0$ , 分子间呈正相关关系
- Pvalue 代表分子间相关系数对应的 Pvalue 值, Pvalue 值越小, 分子间相关系数越显著



## 数据格式

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：选择了不同的平台，搜索出来的分子可能是不一样的

(该样本数据：如下：)

云端数据 ⓘ 食管腺癌 / TCGA / TCGA-ESAD / RNAseq / STAR / FPKM @过滤:去除正常 @处理:log2(value+1)





## 参数说明

(说明：标注了颜色的为常用参数。)

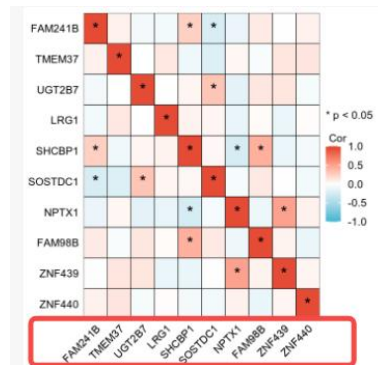
## ID 列表



这部分分子可以来自「[单基因差异分析](#)」或者「[单基因相关性筛选](#)」两个模块筛选后再进行选择，建议是结合两者一起来看，如果想要热图结果好看一些，建议是从「单基因相关性筛选」模块中挑相关性高的分子进行可视化（因为相关趋势更加明显）

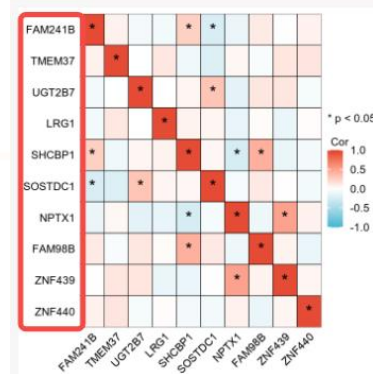
- **X 方向分子**：这部分输入的是[热图部分每一列](#)的分子列表，表示匹配到云端数据中[除了第 1 列](#)(样本/样本 id)外的每一列(每一个变量/分子/基因)

- 一列为一个 ID，最多支持 20 个（最多支持 20 个分子进行相关性分析，也就是热图的行/列不能超过 20）



- Y 方向分子：这部分输入的是热图部分每一行的分子列表，表示匹配到云端数据中除了第 1 列(样本/样本 id)外的每一列(每一个变量/分子/基因)

- 一行为一个 ID，最多支持 20 个（最多支持 20 个分子进行相关性分析，也就是热图的行/列不能超过 20）



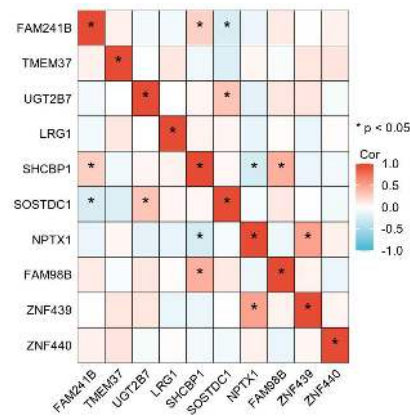
## 统计

统计

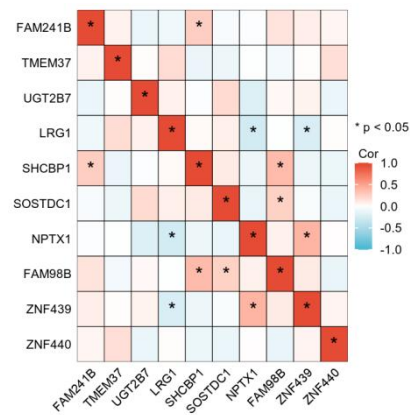
统计方法 Spearman

➤ 统计方法：可以选择分子间进行相关性分析的方法

- spearman: Spearman(默认)为非参数检验方法，数据可以不需要满足正态性



- pearson: Pearson 为参数检验方法，数据需要满足双正态



## 映射

映射

上半颜色映射

相关系数

下半颜色映射

p值

大小映射

相关性系数

- 上半颜色映射：对应整体颜色的映射，当热图选择的是三角对角线类型时，则对应三角的上半颜色映射，如下：（左侧为上半颜色映射分子相关系数，下半颜色映射分子 p 值；右侧为上半颜色映射分子 p 值，下半颜色映射分子相关系数）

热图

类型

三角对角线

上半(全)颜色

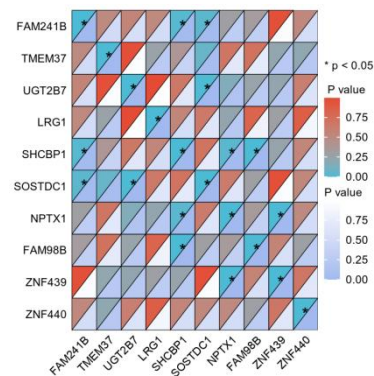
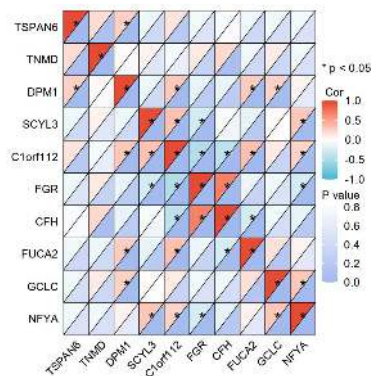
☒

☒

下半颜色

☒

☐



- 下半颜色映射：当热图选择的是三角对角线类型时，则对应三角的下半颜色映射（如上：上半颜色映射）

- 大小映射：可以对热图进行大小映射，只有在**非三角对角线类型**的时候会有效果，对应映射方块大小，默认为不映射，还可以选择相关系数绝对值，如下：

映射

上半颜色

映射

相关系数

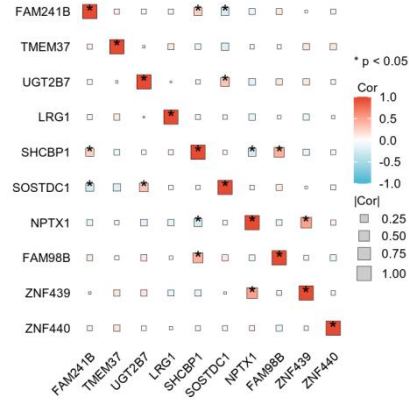
下半颜色

映射

p值

大小映射

相关系数绝对值



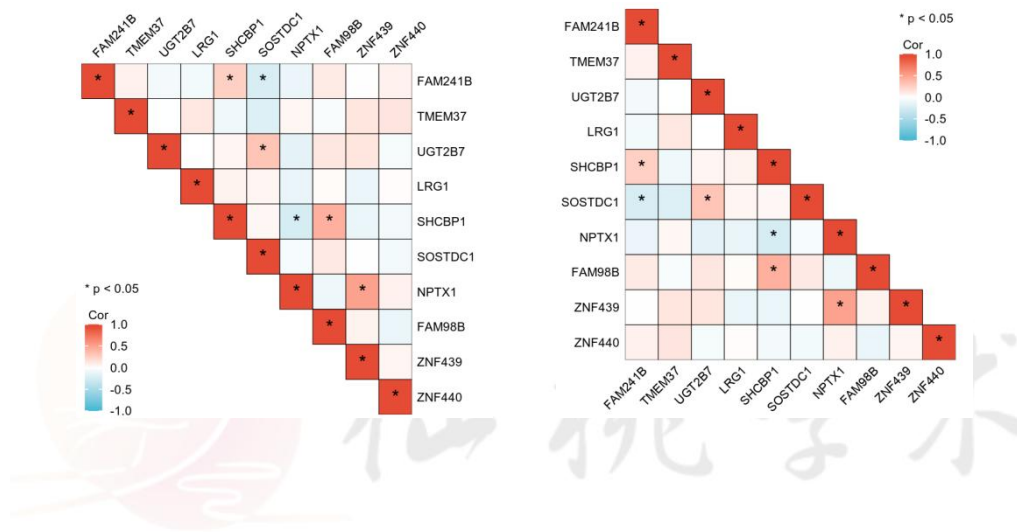
## 布局

布局

类型

完整热图

- 布局：可以选择热图的类型，默认为完整热图，还可以选择上半热图、下半热图，如下（左侧为上半，右侧为下半）



## 热图

热图

类型 方块

上半(全)颜色

下半颜色

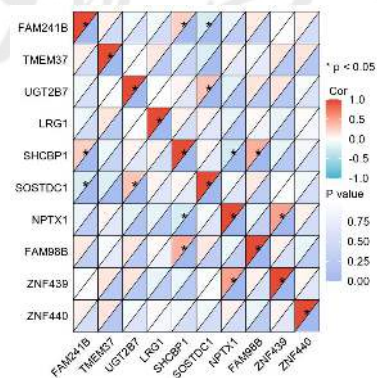
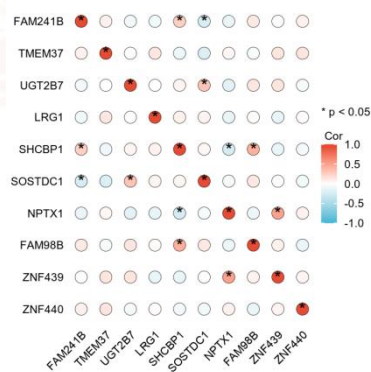
描边颜色

描边粗细 0.25pt

大小比例 1

不透明度 1

- 方块：可以选择热图中每个小块（一行一列）的类型，默认为方块，还可以选择圆形，三角对角线，如下：



- 上半（全）颜色：选择三角类型时可以修改对应方块上半部分颜色，其他类型则修改对应整个方块颜色

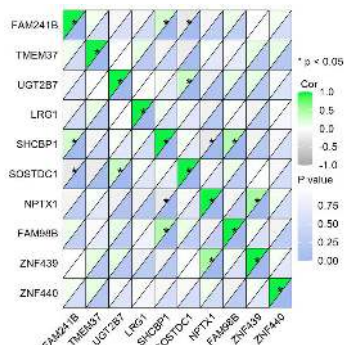
热图

类型 三角对角线

上半(全)颜色

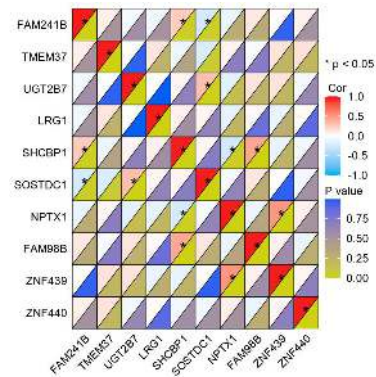
下半颜色

描边颜色





- 下半颜色：当选择三角类型时可以修改对应方块下半部分颜色，其他类型无法修改



- 描边颜色：可以修改热图对应方块的描边颜色
- 描边粗细：可以修改热图对应方块的描边粗细
- 大小比例：可以修改热图对应方块的大小比例，默认为 1
- 不透明度：可以修改热图对应方块的不透明度，默认为 1，表示完全不透明，0 表示完全透明



## 标注

标注

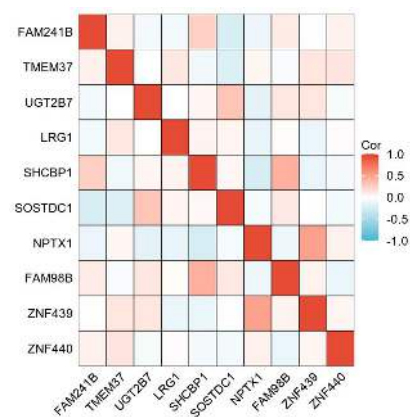
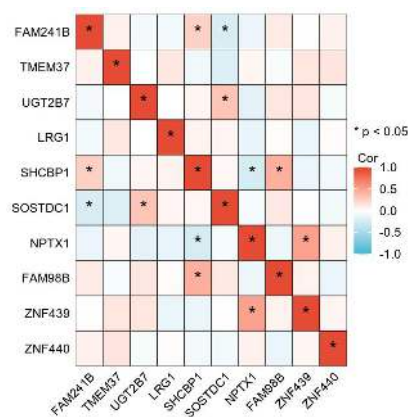
标注映射 星号(\*p<0.05)

颜色

标注大小 6pt

- 标注映射：可以选择是否在热图矩阵对应的每一个小矩形上进行标注映射，默认为星号(\*p<0.05)进行标注，还可以选择：如下：（左侧为(\*p<0.05)映射，右侧为不映射）

- 星号(\*p<0.05|\*\*p<0.01)
- 星号(\*p<0.05|\*\*p<0.01|\*\*\*p<0.001)
- p 值(2 位小数)
- 相关系数(2 位小数)
- 无



- 颜色：可以选择当进行标注映射时，标注的颜色
- 标注大小：可以选择并修改标注的大小，默认为 6pt

## 标题文本

标题

大标题 大标题内容

- 大标题：大标题文本

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

图注

是否展示 ☒

图注标题 图注标题内容

图注位置 默认

- 是否展示：可以选择是否展示图注信息，默认展示
- 图注标题：可以修改图注标题内容，默认没有
- 图注位置：默认为图片的右侧，还可以选择上、下

## 坐标轴



- x 轴标注旋转：可选择并修改 x 轴对应刻度文本的旋转角度

## 风格



- 网格：可以选择是否展示网格，默认不展示
- 文字大小：控制整体文字大小，默认为 6pt

## 图片

图片	▼
宽度 (cm)	6
高度 (cm)	6
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



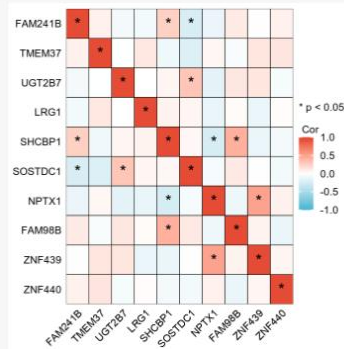
## 结果说明

## 主要结果

### 相关性热图-指定XY-云

相关性热图: 用热图展示相关性情况

统计方法: spearman



相关性热图.pdf

相关性热图.tiff

相关性热图.pptx

分析数据.xlsx

## 补充结果

### 相关性分析

统计方法: spearman

表1: 相关系数表格

	FAM241B	TMEM37	UGT2B7	LRG1	SHCBP1	SOSTDC1	NPTX1	FAM98B	ZNF439	ZNF440
FAM241B		0.0753	-0.066	-0.0733	0.254	-0.236	-0.118	0.11	0.00578	0.0717
TMEM37	0.0753		-0.00483	0.127	-0.0981	-0.205	0.0456	-0.0407	0.138	0.14
UGT2B7	-0.066	-0.00483		-0.000565	0.048	0.311	-0.159	0.128	0.136	-0.0573
LRG1	-0.0733	0.127	-0.000565		0.0677	0.0507	-0.136	0.0245	-0.12	0.0174
SHCBP1	0.254	-0.0981	0.048	0.0677		0.0422	-0.242	0.431	-0.129	-0.0681
SOSTDC1	-0.236	-0.205	0.311	0.0507	0.0422		-0.0524	0.118	0.00467	-0.0705
NPTX1	-0.118	0.0456	-0.159	-0.136	-0.242	-0.0524		-0.106	0.507	0.0742
FAM98B	0.11	-0.0407	0.128	0.0245	0.431	0.118	-0.106		0.064	-0.12
ZNF439	0.00578	0.138	0.136	-0.12	-0.129	0.00467	0.507	0.064		0.0469
ZNF440	0.0717	0.14	-0.0573	0.0174	-0.0681	-0.0705	0.0742	-0.12	0.0469	

这里提供相关性分析表：可以查看分子之间的相关系数

- 相关系数为正数，说明两个分子之间可能存在正相关关系；相关系数为负数，说明两个分子可能存在负相关关系
  - 相关系数绝对值在 0.8-1.0 之间，说明两个分子之间强相关
  - 相关系数绝对值在 0.5-0.8 之间，说明两个分子之间中等程度相关
  - 相关系数绝对值在 0.3-0.5 之间，说明两个分子之间相关程度一般
  - 相关系数绝对值在 0.0-0.3 之间，说明两个分子之间弱相关或者不相关

表2: 相关性检验表格

	FAM241B	TMEM37	UGT2B7	LRG1	SHCBP1	SOSTDC1	NPTX1	FAM98B	ZNF439	ZNF440
FAM241B		0.503	0.558	0.515	0.0226	0.0342	0.295	0.325	0.959	0.525
TMEM37	0.503		0.966	0.259	0.383	0.0664	0.686	0.717	0.22	0.212
UGT2B7	0.558	0.966		0.996	0.67	0.00472	0.156	0.254	0.225	0.611
LRG1	0.515	0.259	0.996		0.547	0.653	0.228	0.828	0.286	0.878
SHCBP1	0.0226	0.383	0.67	0.547		0.708	0.0294	6.98e-05	0.25	0.546
SOSTDC1	0.0342	0.0664	0.00472	0.653	0.708		0.642	0.296	0.967	0.532
NPTX1	0.295	0.686	0.156	0.228	0.0294	0.642		0.346	1.38e-06	0.511
FAM98B	0.325	0.717	0.254	0.828	6.98e-05	0.296	0.346		0.569	0.287
ZNF439	0.959	0.22	0.225	0.286	0.25	0.967	1.38e-06	0.569		0.678
ZNF440	0.525	0.212	0.611	0.878	0.546	0.532	0.511	0.287	0.678	

[相关性.xlsx](#)

相关性系数表格:

- 表中包含了各个变量间的相关系数(r)值, 相关系数一般是 -1到1 之间, 正负号表示正相关和负相关, 系数绝对值大小表示相关性大小
- 一般关系强度是:  $|r| > 0.95$ : 显著性相关;  $|r| \geq 0.8$ : 高度相关;  $0.5 \leq |r| < 0.8$ : 中度相关;  $0.3 \leq |r| < 0.5$ : 低度相关;  $|r| < 0.3$ : 弱相关

相关性p值表格:

- 表中包含了各个变量间的相关性的检验p值

这里提供相关性分析表: 可以查看各个分子间的相关性的检验 p 值

➤ p 值表示检验 p 值



## 方法学

统计分析和可视化均在 R 4.2.1 版本中进行

涉及的 R 包：ggplot2 包（用于可视化）

处理过程：

- (1) 对主要参数[ID 列表] 中匹配到数据中的分子进行两两相关性分析，分析结果用热图进行可视化





## 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](http://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 因为是非参数的方法，可以不需要满足。可以先选择非参数的 Spearman 相关进行尝试。

### 2. 图的内容被压缩了，如何处理？

答：由于文字不会被压缩，如果热图部分很长，就可能会导致热图部分重叠。解决方案可以是：

- ① 增加图片高度；
- ② 减少分子列表中的分子。

### 3. 相关系数多少为好？

答：这个没有很统一的标准，可以参考以下：

- 相关系数强弱：
  - 绝对值在 0.8 以上：强相关
  - 绝对值在 0.5-0.8：中等程度相关
  - 绝对值在 0.3-0.5：相关程度一般
  - 绝对值在 0.3 以下：弱或者不相关