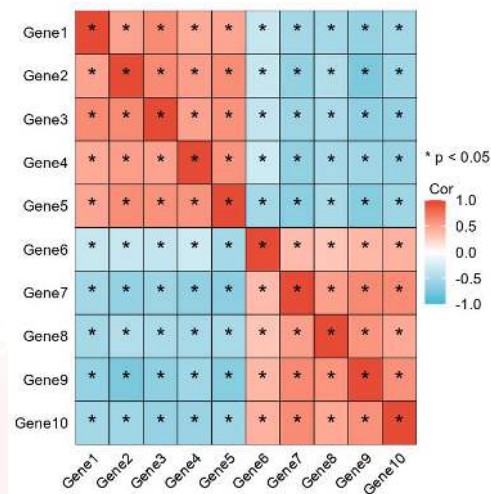


## 交互网络 - 相关性热图-原始矩阵



网址: <https://www.xiantao love>



更新时间: 2023.02.14

## 目录

基本概念 .....	3
应用场景 .....	3
分析流程 .....	3
结果解读 .....	5
数据格式（原始矩阵） .....	6
参数说明 .....	7
统计 .....	7
映射 .....	8
布局 .....	10
热图 .....	11
标注 .....	13
标题文本 .....	14
图注 .....	14
坐标轴 .....	15
风格 .....	15
图片 .....	16
结果说明 .....	17
主要结果 .....	17
补充结果 .....	18
方法学 .....	20
如何引用 .....	21
常见问题 .....	22

## 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；相关性热图这是根据变量之间的相关性来展示的热图。

## 应用场景

通过热图的形式 展示变量之间相关性分析相关的内容

## 分析流程

上传数据 ➡ 数据处理(清洗) ➡ 相关性分析 ➡ 可视化

- 数据格式：

- 数据每一列都代表一个变量/样本，都需要是数值类型的数据

	A	B	C	D	E	F	G	H	I	J
1	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
2	0.496745924	0.467723302	-0.048351761	-0.240196499	0.43226315	0.921879419	0.784514945	1.245615557	0.518298418	0.739705603
3	-0.01368477	0.06421928	0.249732944	0.537374657	-0.04316992	1.286957965	1.175662851	0.5182841	0.981862233	0.568852396
4	0.01182029	0.141322045	0.633965878	-0.263885834	0.542861073	0.490195153	1.183083571	0.987716302	1.067240227	1.033805987
5	-0.004488363	-0.058815726	0.586873137	0.222179483	0.179198701	0.44559659	0.548951482	0.965296724	1.13184169	0.354723611
6	-0.210829963	0.369931219	0.547144016	0.36844285	0.472425386	0.854233325	1.18447251	1.123654746	1.262049733	0.943975343
7	0.642966409	0.355974738	0.671350067	-0.173697559	-0.032966	0.767416042	1.249694698	0.595296227	0.60783876	0.513691745
8	-0.179486354	0.196057306	0.168928365	0.065455873	-0.014343253	0.448853135	1.244653087	0.678902879	1.150490547	1.088849734
9	0.119288873	-0.192532717	-0.155474752	0.501615558	-0.246108244	0.940513758	1.272317943	0.347395517	0.745167982	0.645567114
10	0.15636136	0.52661018	-0.077797179	0.380145601	0.315129994	1.260823014	1.181884192	1.258837779	0.556531639	1.120612391
11	0.625642134	0.319952123	0.246807153	0.013511406	0.392323231	1.198206731	0.861498273	0.492483526	1.168313728	1.199737413
12	0.223237827	0.46813969	-0.270452353	-0.271382299	-0.042795272	0.734466762	0.696293427	0.3255452	1.235287427	0.623165723
13	0.327989421	0.053296945	-0.054571989	0.113712744	-0.114236178	0.36837606	1.103604986	1.17269313	0.833974877	0.594458569
14	-0.138863375	0.651505674	0.145345091	0.532905728	0.306025015	0.945230436	0.440612899	1.112037256	1.013758465	0.42780374
15	0.049065953	-0.14631829	0.063868247	0.010649082	0.333018047	0.33428674	1.225104532	1.255972546	1.076337558	0.50712956
16	0.391187731	0.140520319	-0.000645322	0.505075229	0.51432391	0.440272447	0.4341059	1.147909249	0.70803667	0.680611736

- 数据处理

- 数据格式：

- ◆ 所有变量/列都需要纯数值类型的数据
- ◆ 不能有非数值，特殊值(特殊符号等)，并且每一个变量不能都是一个值

#### ■ 数据处理：

- ◆ 对数据中每一列非数值类型的数据进行处理

#### ➤ 相关性分析：

#### ■ 将处理(清洗)后的数据进行相关性分析

- ◆ 变量与变量之间

- ◆ 相关性系数表

表1: 相关性系数表格

	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
Gene1		0.501	0.625	0.457	0.483	-0.334	-0.537	-0.53	-0.62	-0.535
Gene2	0.501		0.63	0.525	0.613	-0.342	-0.626	-0.459	-0.75	-0.556
Gene3	0.625	0.63		0.506	0.592	-0.361	-0.577	-0.495	-0.646	-0.614
Gene4	0.457	0.525	0.506		0.581	-0.295	-0.638	-0.521	-0.557	-0.592
Gene5	0.483	0.613	0.592	0.581		-0.528	-0.659	-0.516	-0.69	-0.563
Gene6	-0.334	-0.342	-0.361	-0.295	-0.528		0.371	0.316	0.395	0.42
Gene7	-0.537	-0.626	-0.577	-0.638	-0.659	0.371		0.517	0.64	0.625
Gene8	-0.53	-0.459	-0.495	-0.521	-0.516	0.316	0.517		0.579	0.469
Gene9	-0.62	-0.75	-0.646	-0.557	-0.69	0.395	0.64	0.579		0.598
Gene10	-0.535	-0.556	-0.614	-0.592	-0.563	0.42	0.625	0.469	0.598	

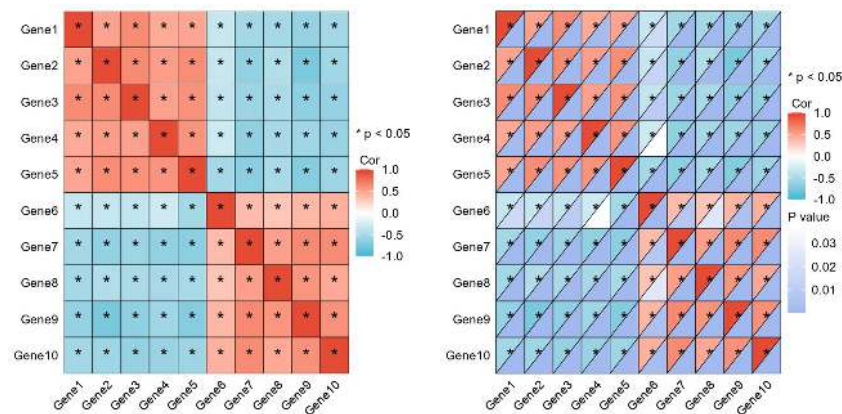
- ◆ 相关性检验表(p 值表)

表2: 相关性检验表格

	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
Gene1		0.000211	1.25e-06	0.000856	0.000376	0.0178	5.7e-05	7.55e-05	1.61e-06	6.29e-05
Gene2	0.000211		9.69e-07	8.91e-05	2.25e-06	0.0151	1.18e-06	0.0008	3.58e-10	2.81e-05
Gene3	1.25e-06	9.69e-07		0.00018	6.04e-06	0.01	1.16e-05	0.000258	3.97e-07	2.09e-06
Gene4	0.000856	8.91e-05	0.00018		9.91e-06	0.0374	6.19e-07	0.000106	2.67e-05	5.84e-06
Gene5	0.000376	2.25e-06	6.04e-06	9.91e-06		8.1e-05	1.99e-07	0.000127	3.03e-08	2.09e-05
Gene6	0.0178	0.0151	0.01	0.0374	8.1e-05		0.00795	0.0254	0.00454	0.00238
Gene7	5.7e-05	1.18e-06	1.16e-05	6.19e-07	1.99e-07	0.00795		0.00012	5.65e-07	1.21e-06
Gene8	7.55e-05	0.0008	0.000258	0.000106	0.000127	0.0254	0.00012		1.08e-05	0.000586
Gene9	1.61e-06	3.58e-10	3.97e-07	2.67e-05	3.03e-08	0.00454	5.65e-07	1.08e-05		4.42e-06
Gene10	6.29e-05	2.81e-05	2.09e-06	5.84e-06	2.09e-05	0.00238	1.21e-06	0.000586	4.42e-06	

#### ➤ 将分析后得到的相关性系数与 p 值进行后续的相关性热图可视化

## 结果解读



左图:

- 行、列都代表变量/样本
- 每一个小方格表示变量之间的相关系数，颜色越深，变量间越相关

右图:

- 行、列都代表变量/样本
- 每一个小方格分为两个部分（三角形），上部分表示变量间的相关系数，颜色越深，变量间越相关；下部分表示 p 值，颜色越深 p 值越小

补充:

- \* 表示 Pvalue < 0.05; \*\* 表示 Pvalue < 0.01
- Correlation 代表相关性系数(Correlation): |Correlation| 越大，变量间相关性越高，反之相关性越低
  - Correlation < 0，变量间呈负相关关系
  - Correlation = 0，变量间没有相关关系
  - Correlation > 0，变量间呈正相关关系
- Pvalue 代表变量间相关系数对应的 Pvalue 值，Pvalue 值越小，变量间相关系数越显著



## 数据格式（原始矩阵）

	A	B	C	D	E	F	G	H	I	J
1	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
2	0.496745924	0.467723302	-0.048351761	-0.240196499	0.43226315	0.921879419	0.784514945	1.245615557	0.518298418	0.739705603
3	-0.01368477	0.06421928	0.249732944	0.537374657	-0.04316992	1.286957965	1.175662851	0.5182841	0.981862233	0.568852396
4	0.01182029	0.141322045	0.633965878	-0.263885834	0.542861073	0.490195153	1.183083571	0.987716302	1.067240227	1.033805987
5	-0.004488363	-0.058815726	0.586873137	0.222179483	0.179198701	0.44559659	0.548951482	0.965296724	1.13184169	0.354723611
6	-0.210829963	0.369931219	0.547144016	0.36844285	0.472425386	0.854233325	1.18447251	1.123654746	1.262049733	0.943975343
7	0.642966409	0.355974738	0.671350067	-0.173697559	-0.032966	0.767416042	1.249694698	0.595296227	0.60783876	0.513691745
8	-0.179486354	0.196057306	0.168928365	0.065455873	-0.014343253	0.448853135	1.244653087	0.678902879	1.150490547	1.088849734
9	0.119288873	-0.192532717	-0.155474752	0.501615558	-0.246108244	0.940513758	1.272317943	0.347395517	0.745167982	0.645567114
10	0.15636136	0.52661018	-0.077797179	0.380145601	0.315129994	1.260823014	1.181884192	1.258837779	0.556531639	1.120612391
11	0.625642134	0.319952123	0.246807153	0.013511406	0.392323231	1.198206731	0.861498273	0.492483526	1.168313728	1.199737413
12	0.223237827	0.46813969	-0.270452353	-0.271382299	-0.042795272	0.734466762	0.696293427	0.3255452	1.235287427	0.623165723
13	0.327989421	0.053296945	-0.054571989	0.113712744	-0.114236178	0.36837606	1.103604986	1.17269313	0.833974877	0.594458569
14	-0.138863375	0.651505674	0.145345091	0.532905728	0.306025015	0.945230436	0.440612899	1.112037256	1.013758465	0.42780374
15	0.049065953	-0.14631829	0.063868247	0.010649082	0.333018047	0.33428674	1.225104532	1.255972546	1.076337558	0.50712956
16	0.391187731	0.140520319	-0.000645322	0.505075229	0.51432391	0.440272447	0.4341059	1.147909249	0.70803667	0.680611736

### 数据要求：

- 数据至少 2 列以上，每列至少 5 行（5 个观测值）最多支持 50 列和 5000 行数据
- 每一列列名对应热图的行、列名。图中各变量的顺序与上传数据中各变量的顺序保持一致，若 需要调整图中各分类的顺序，需要在上传数据内进行调整，然后再上传数据。
- 每一行可以代表在不同样本对应的列变量的情况。例如表达谱数据。
- 列名不能重复
- 每个变量不能有非数值，特殊值(特殊符号等)，并且一个变量不能全部为一样的值

## 参数说明

(说明: 标注了颜色的为常用参数。)

## 统计



统计

统计方法 Spearman

- 统计方法: 可以选择变量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：热图-->类型-->三角对角线（第1个为上半颜色映射变量相关系数，下半颜色映射变量p值；第2个为上半颜色映射变量p值，下半颜色映射变量相关系数）

热图

类型

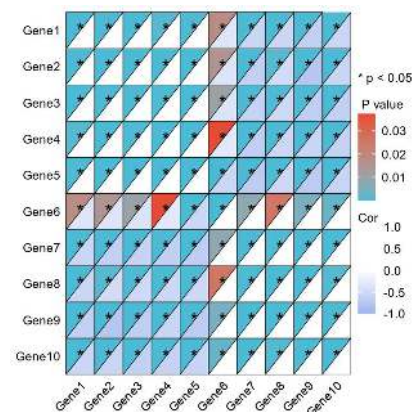
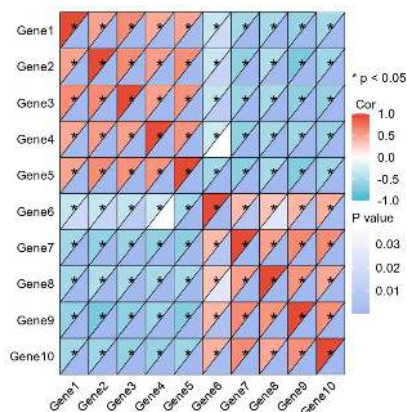
三角对角线

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

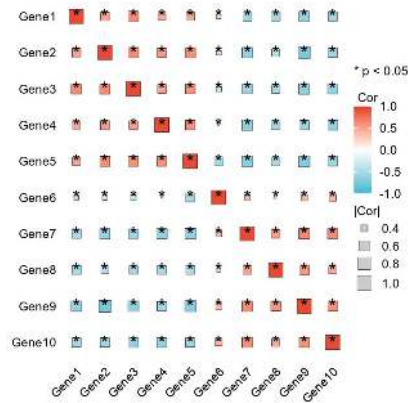
相关系数

下半颜色  
映射

p值

大小映射

相关系数绝对值



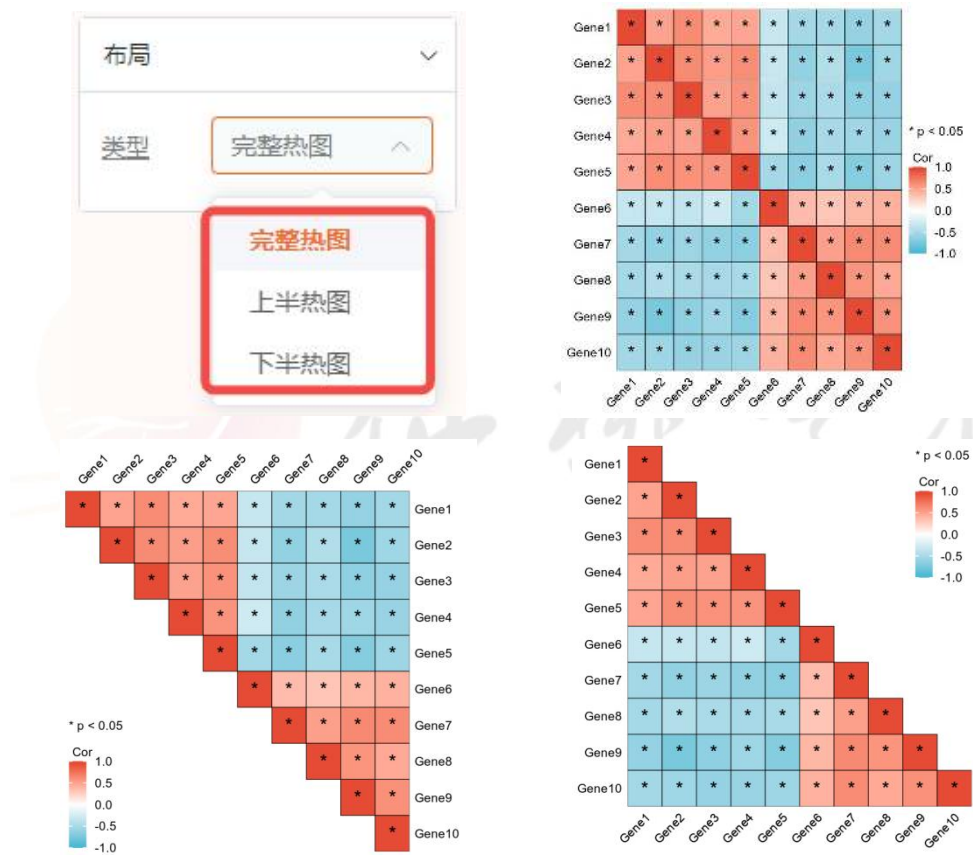
## 布局

布局

类型

完整热图

- 类型：可以选择热图矩阵的类型，默认为完整热图，还可以选择上半热图（只画上半部分热图）、下半热图（只画下半热图），如下：



## 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：

热图

类型 方块

上半(全)颜色

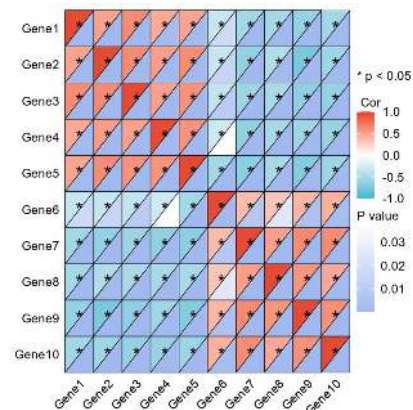
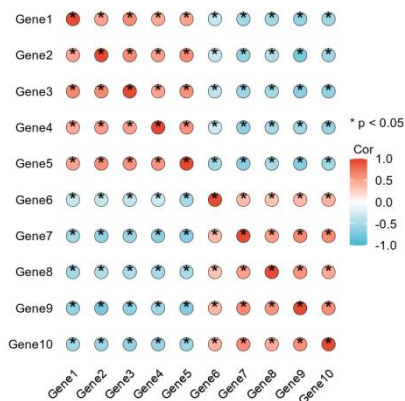
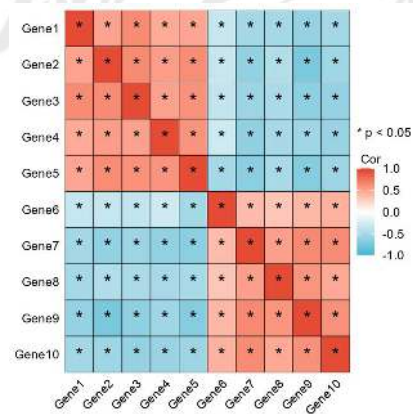
下半颜色

描边颜色

方块

圆形

三角对角线



- 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
- 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

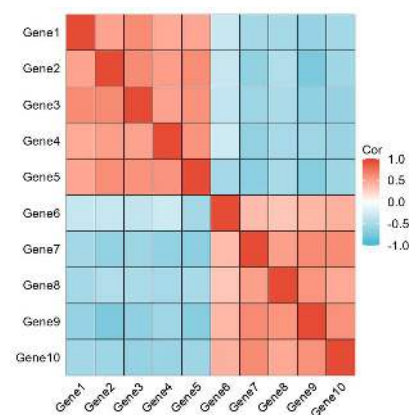
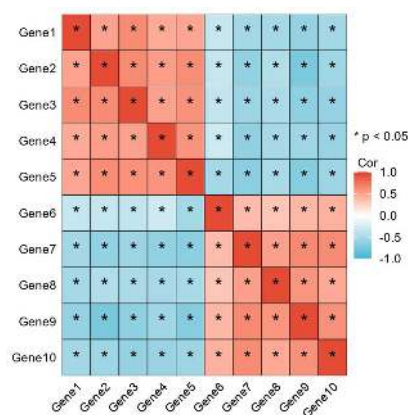
标注映射 星号(\*p<0.05)

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：（第1个为(\*p<0.05)映射，第2个为不映射）

- 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



## 补充结果

### 相关性分析

统计方法: spearman

表1: 相关系数表格

	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
Gene1		0.501	0.625	0.457	0.483	-0.334	-0.537	-0.53	-0.62	-0.535
Gene2	0.501		0.63	0.525	0.613	-0.342	-0.626	-0.459	-0.75	-0.556
Gene3	0.625	0.63		0.506	0.592	-0.361	-0.577	-0.495	-0.646	-0.614
Gene4	0.457	0.525	0.506		0.581	-0.295	-0.638	-0.521	-0.557	-0.592
Gene5	0.483	0.613	0.592	0.581		-0.528	-0.659	-0.516	-0.69	-0.563
Gene6	-0.334	-0.342	-0.361	-0.295	-0.528		0.371	0.316	0.395	0.42
Gene7	-0.537	-0.626	-0.577	-0.638	-0.659	0.371		0.517	0.64	0.625
Gene8	-0.53	-0.459	-0.495	-0.521	-0.516	0.316	0.517		0.579	0.469
Gene9	-0.62	-0.75	-0.646	-0.557	-0.69	0.395	0.64	0.579		0.598
Gene10	-0.535	-0.556	-0.614	-0.592	-0.563	0.42	0.625	0.469	0.598	

这里提供相关性分析表：可以查看变量之间的相关系数

- 表中包含了各个变量间的相关系数(r)值
- 相关系数为正数，说明两个变量（主要变量与其他变量）之间可能存在正相关关系；相关系数为负数，说明两个变量可能存在负相关关系
  - 相关系数绝对值在 0.8-1.0 之间，说明两个变量之间强相关
  - 相关系数绝对值在 0.5-0.8 之间，说明两个变量之间中等程度相关
  - 相关系数绝对值在 0.3-0.5 之间，说明两个变量之间相关程度一般
  - 相关系数绝对值在 0.0-0.3 之间，说明两个变量之间弱相关或者不相关



表2: 相关性检验表格

	Gene1	Gene2	Gene3	Gene4	Gene5	Gene6	Gene7	Gene8	Gene9	Gene10
Gene1		0.000211	1.25e-06	0.000856	0.000376	0.0178	5.7e-05	7.55e-05	1.61e-06	6.29e-05
Gene2	0.000211		9.69e-07	8.91e-05	2.25e-06	0.0151	1.18e-06	0.0008	3.58e-10	2.81e-05
Gene3	1.25e-06	9.69e-07		0.00018	6.04e-06	0.01	1.16e-05	0.000258	3.97e-07	2.09e-06
Gene4	0.000856	8.91e-05	0.00018		9.91e-06	0.0374	6.19e-07	0.000106	2.67e-05	5.84e-06
Gene5	0.000376	2.25e-06	6.04e-06	9.91e-06		8.1e-05	1.99e-07	0.000127	3.03e-08	2.09e-05
Gene6	0.0178	0.0151	0.01	0.0374	8.1e-05		0.00795	0.0254	0.00454	0.00238
Gene7	5.7e-05	1.18e-06	1.16e-05	6.19e-07	1.99e-07	0.00795		0.00012	5.65e-07	1.21e-06
Gene8	7.55e-05	0.0008	0.000258	0.000106	0.000127	0.0254	0.00012		1.08e-05	0.000586
Gene9	1.61e-06	3.58e-10	3.97e-07	2.67e-05	3.03e-08	0.00454	5.65e-07	1.08e-05		4.42e-06
Gene10	6.29e-05	2.81e-05	2.09e-06	5.84e-06	2.09e-05	0.00238	1.21e-06	0.000586	4.42e-06	

这里提供相关性分析表： 可以查看变量之间相关系数对应的统计学 p 值

➤ p 值表示统计学 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) 对数据中变量进行两两相关性分析
- (2) 使用 ggplot2 包将所得的结果进行可视化（热图绘制）



## 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 减少变量列表中的变量。

### 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