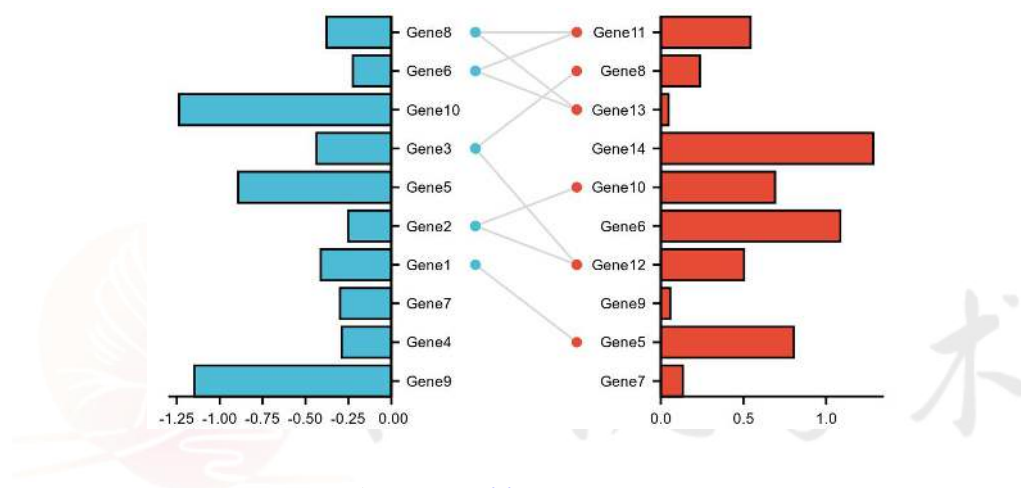


## 基础绘图 - 连接柱状图



网址: <https://www.xiantao.love>



更新时间: 2023.08.17

## 目录

基本概念 .....	3
应用场景 .....	3
分析过程 .....	3
结果解读 .....	6
数据格式 .....	7
参数说明 .....	9
柱 .....	9
点 .....	10
连接线 .....	11
标注 .....	12
标题文本 .....	13
坐标轴 .....	14
风格 .....	15
画图比例 .....	16
图片 .....	17
结果说明 .....	18
主要结果 .....	18
方法学 .....	19
如何引用 .....	20
常见问题 .....	21

## 基本概念

- 柱状图：用柱子的高度或者柱子的相对高度来表示数据的大小情况
- 连接柱状图：通过连线的方式，来展示两个柱状图的匹配情况

## 应用场景

连接柱状图主要用来展示数据分布情况。常应用于数据量对比与样本匹配

## 分析过程

上传数据 → 数据处理(清洗) → 可视化

- 数据格式：（具体数据格式要求可以看后面过程的“数据格式”部分）

- 第 1 个表

- ◆ 数据第 1 列需要提供分类类型，对应双值柱状图 y 轴
- ◆ 数据第 2 列需要提供数值类型，对应双值柱状图 x 轴
- ◆ 数据第 3 列及以后可以是数值类型也可以是分类类型

#	A	B	C	D	E
1	sample1	value1			
2	Gene9	-1.147657009			
3	Gene4	-0.289461574			
4	Gene7	-0.799215118			
5	Gene1	-0.411510833			
6	Gene2	-0.752223448			
7	Gene5	-0.891921177			
8	Gene3	-0.435683299			
9	Gene10	-1.737538472			
10	Gene6	-0.724267885			
11	Gene8	-0.377395646			
12					
13					
14					
15					
16					
17					
18					

- 第 2 个表：数据格式与第 1 个表一致

#	A	B	C	D	E
1	sample2	value2			
2	Gene7	0.133336361			
3	Gene5	0.80418951			
4	Gene9	0.057106774			
5	Gene12	0.503607972			
6	Gene6	1.085769362			
7	Gene10	0.69095384			
8	Gene14	1.284599354			
9	Gene13	0.046726172			
10	Gene8	0.235706556			
11	Gene11	0.542888255			
12					
13					
14					
15					
16					
17					
18					

- 第 3 个表：假如上传有第 3 个表的时候：

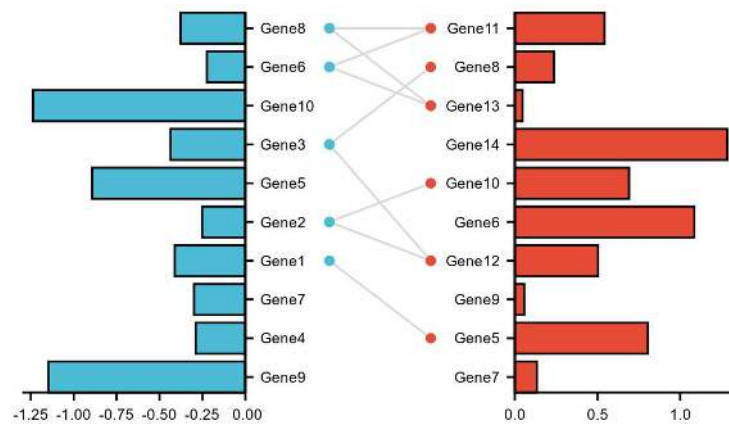
- ◆ 第 1 列需要提供分类类型，并且需要与第 1 个表的第 1 列相对应，如果对应不上将不会进行相关的可视化
- ◆ 第 2 列需要提供分类类型，并且需要与第 2 个表的第 1 列相对应，如果对应不上将不会进行相关的可视化

#	A	B	C	D	E
1	sample1	sample2			
2	Gene8	Gene11			
3	Gene2	Gene10			
4	Gene2	Gene12			
5	Gene6	Gene11			
6	Gene6	Gene15			
7	Gene1	Gene5			
8	Gene3	Gene8			
9	Gene3	Gene12			
10	Gene8	Gene13			
11	Gene6	Gene13			
12					
13					
14					
15					
16					
17					
18					

- 数据处理：对第 1、2 个表分别对分类类型数据，数值类型的数据及其他列数据进行相应处理
  - 数值类型数据只能是纯数值类型数据，不能包含非数值与不规则的值
  - 非数值类型(分类类型)数据不能含有特殊、无法识别字符串
  - .....
  - 如果有第 3 个表，将第 3 个表的第 1/2 列分别与第 1/2 个表的第 1 列进行匹配，如果匹配上了将会进行相应的连接线可视化
- 可视化：将清洗后的数据进行 ggplot2 包可视化



## 结果解读



- 横坐标（柱状图部分）表示变量对应不同的值（对应上传数据第 1/2 个表第 2 列数据（固定）数值类型的数据）
- 纵坐标（柱状图部分）表示变量（对应上传数据第 1/2 个表第 1 列分类类型的数据）
- 连接线没有具体含义，就是两个柱状图的匹配情况

## 数据格式

	A	B	C	D	E
1	sample1	value1			
2	Gene9	-1.147657009			
3	Gene4	-0.289461574			
4	Gene7	-0.299215118			
5	Gene1	-0.411510833			
6	Gene2	-0.252223448			
7	Gene5	-0.891921127			
8	Gene3	-0.435683299			
9	Gene10	-1.237538422			
10	Gene6	-0.224267885			
11	Gene8	-0.377395646			
12					
13					
14					
15					
16					
17					
18					

数据要求：

➤ 第 1 个表

- 数据至少 2 列，每列至少 1 个观测（即至少 1 行数据），最多支持 10 列和 100 行数据
- ◆ 数据第 1 列需要提供分类类型，对应双值柱状图 y 轴
- ◆ 数据第 2 列需要提供数值类型，对应双值柱状图 x 轴
- ◆ 数据第 3 列及以后可以是数值类型也可以是分类类型

	A	B	C	D	E
1	sample2	value2			
2	Gene7	0.133336361			
3	Gene5	0.80418951			
4	Gene9	0.057106774			
5	Gene12	0.503607972			
6	Gene6	1.085769362			
7	Gene10	0.69095384			
8	Gene14	1.284599354			
9	Gene13	0.046726172			
10	Gene8	0.235706556			
11	Gene11	0.542888255			
12					
13					
14					
15					
16					
17					
18					

Navigation: left **right** link +

- 第 2 个表: 数据格式与第 1 个表一致

	A	B	C	D	E
1	sample1	sample2			
2	Gene8	Gene11			
3	Gene2	Gene10			
4	Gene2	Gene12			
5	Gene6	Gene11			
6	Gene6	Gene15			
7	Gene1	Gene5			
8	Gene3	Gene8			
9	Gene3	Gene12			
10	Gene8	Gene13			
11	Gene6	Gene13			
12					
13					
14					
15					
16					
17					
18					

Navigation: left right **link** +

- 第 3 个表: 假如上传有第 3 个表的时候:

- 第 1 列需要提供分类类型, 并且需要与第 1 个表的第 1 列相对应, 如果对应不上将不会进行相关的可视化
- 第 2 列需要提供分类类型, 并且需要与第 2 个表的第 1 列相对应, 如果对应不上将不会进行相关的可视化

- 数据每一列列名不能重复



## 参数说明

(说明：标注了颜色的为常用参数。)

### 柱

柱

填充色

描边色

描边粗细

0.75pt

不透明度

1

宽度

0.8

- 填充色：可以修改绘制左右两侧柱状图的填充颜色
- 描边色：可以修改绘制左右两侧柱状图的描边颜色
- 描边粗细：可以选择左右两侧柱状图的每根柱子外部边框的粗细
- 不透明度：可以修改左右两侧柱状图的不透明度
- 宽度：可以修改左右两侧柱状图没跟柱子的宽度

## 点

点

填充色

描边色

样式

圆形

大小比例

1

不透明度

1

- 填充色：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以点的形式展示），这时可以修改点的填充色
- 描边色：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以点的形式展示），这时可以修改点的描边色
- 样式：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以点的形式展示），这时可以修改点的样式，默认为圆形，还可以选择正方形、菱形、三角形、倒三角形
- 当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以点的形式展示），这时可以修改点的大小比例
- 当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以点的形式展示），这时可以修改点的不透明度，1 表示完全不透明，0 表示完全透明

## 连接线

连接线

颜色

☐

类型

实线

粗细

0.75pt

- 颜色：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以线的方式将两个柱状图进行连接并展示），这时可以修改连接线的颜色
- 类型：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以线的方式将两个柱状图进行连接并展示），这时可以修改连接线的类型，默认为实线，还可以选择虚线的形式
- 粗细：当上传有第三个表格的并且与第 1/2 个表匹配上的时候，将会进行绘制匹配到的样本（以线的方式将两个柱状图进行连接并展示），这时可以修改连接线的粗细

## 标注

标注

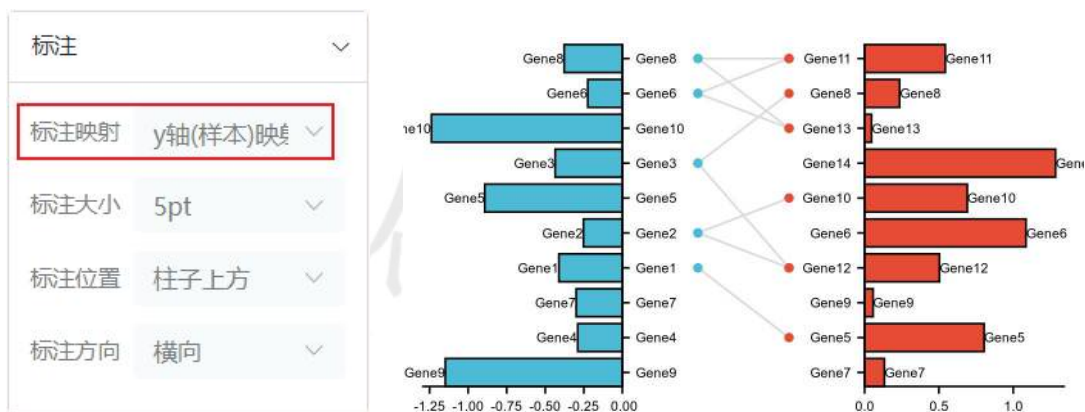
标注映射 不映射

标注大小 6pt

标注位置 柱子上方

标注方向 横向

- 标注映射：可以选择是否对左右两侧柱状图进行标注操作，如下：



- 标注大小：可以选择并修改进行标注映射时标注文本的字体大小
- 标注位置：可以选择进行标注映射时标注的位置，默认在柱子上方，还可以选择柱子中间和柱子底部，如下：



- 标注方向：可以选择进行标注映射时标注的方向，默认为横向，还可以选择纵向

## 标题文本



- 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]]

## 坐标轴

坐标轴
▼

是否显示x轴
☒

是否显示y轴
☒

左侧x轴范围  
+刻度

逗号隔开

右侧x轴范围  
+刻度

逗号隔开

- 是否显示 x 轴：可以选择是否显示左右两侧柱状图的 x 轴
- 是否显示 y 轴：可以选择是否显示左右两侧柱状图的 y 轴
- 左侧 x 轴范围+刻度：可以控制左侧柱状图 x 轴范围和刻度，可只提供 2 个值来控制范围。形如 0.1, 0.1, 0.2, 0.3 (最小值和最大值不能不能可视化数据范围 20%，如果调整过大可能会无作用)
- 右侧 x 轴范围+刻度：可以控制右侧柱状图 x 轴范围和刻度，可只提供 2 个值来控制范围。形如 0.1, 0.1, 0.2, 0.3 (最小值和最大值不能不能可视化数据范围 20%，如果调整过大可能会无作用)

## 风格



- 边框：可以选择是否进行添加图形边框的操作
- 网格：可以选择是否进行添加图形网格线的操作
- 文字大小：控制整体文字大小，默认为 6pt



## 画图比例

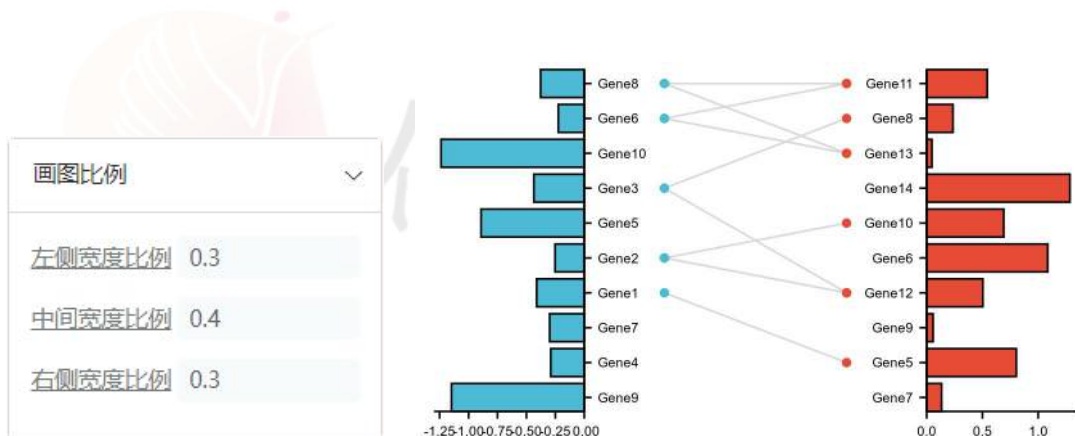
画图比例

左侧宽度比例 0.4

中间宽度比例 0.2

右侧宽度比例 0.4

- 左侧宽度比例：可以控制左侧柱状图宽度比例，0-1 之间
- 中间宽度比例：可以控制中间连线图宽度比例，0-1 之间
- 右侧宽度比例：可以控制右侧柱状图宽度比例，0-1 之间，如下：





## 图片

图片	▼
宽度 (cm)	9
高度 (cm)	5
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

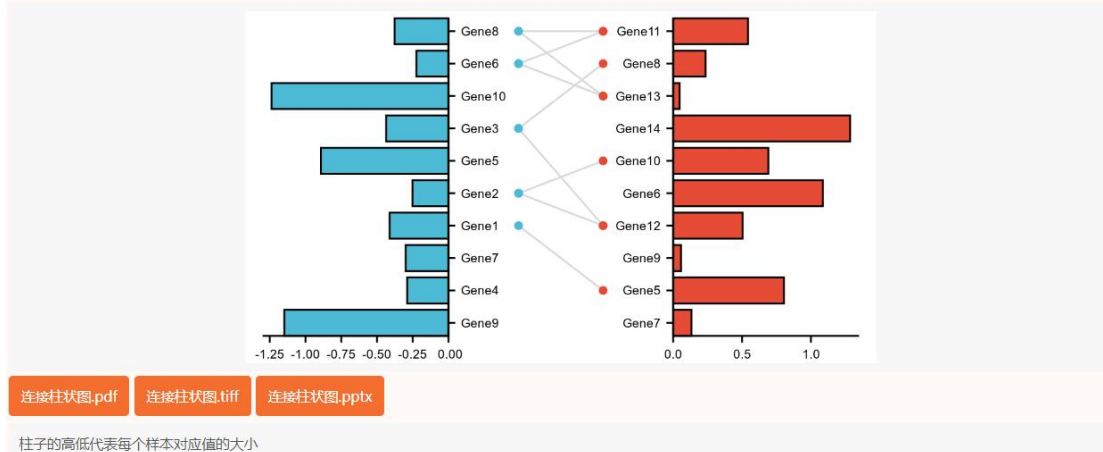


## 结果说明

## 主要结果

### 连接柱状图

连接柱状图: 用柱子绘制不同分组/变量中各样本的情况, 用连接线来绘制不同分组/变量间的关系



## 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) 使用 ggplot2 包对数据进行可视化。



## 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

