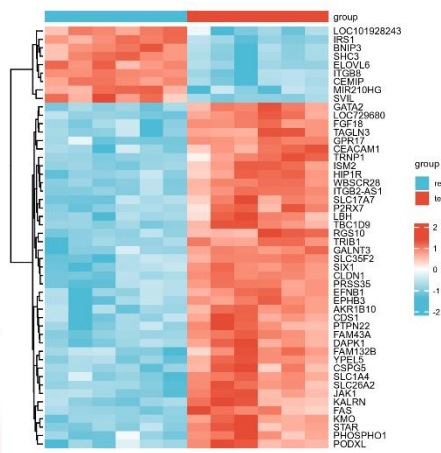


数据集工具 - [数据集] 热图



网址: <https://www.xiantao love>



更新时间: 2023.03.13

目录

基本概念	3
应用场景	3
分析流程	5
主要结果	6
云端数据	8
参数说明	9
数据设置	9
标注	10
数据处理	11
聚类(顺序)/分割	11
主图	13
上注释	13
文字	14
图注	14
图片	15
结果说明	16
主要结果	16
方法学	17
如何引用	18
常见问题	19

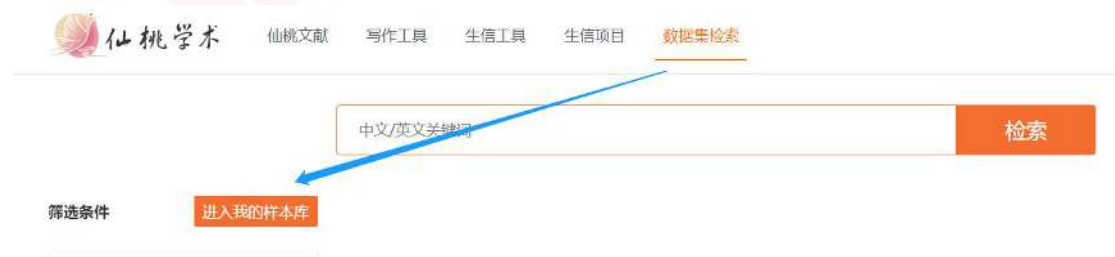
基本概念

- 数据集差异分析：从数据集检索模块中，针对特定 GEO 数据集的数据，进行芯片差异分析的过程，类似 GEO2R。
- **热图**：主要用于可视化矩阵数据的结果，每个小正方形表示行（基因）列（样本）所对应的数值情况。

应用场景

本模块为 数据集检索 - 差异分析 后结果的可视化展示。用于展示数据集差异分析后样本中一些(差异表达)基因的表达情况和样本信息。

注意：模块需要**先进行 数据集检索 - 差异分析 后**，此处的云端数据才会有结果记录，然后才能进行可视化的操作。





数据集/样本库

刷新

<input type="checkbox"/>	分组	备注	数据集	平台	样本编号	Title
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214917	UET-13TR-EWS/FLI1 0hr
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214918	UET-13TR-EWS/FLI1 24hr
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214919	UET-13TR-EWS/FLI1 48hr
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214920	UET-13TR-EWS/FLI1 72hr
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214921	UET-13TR-EWS/FLI1 24hr tet+
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214922	UET-13TR-EWS/FLI1 48hr tet+
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214923	UET-13TR-EWS/FLI1 72hr tet+
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214924	UET-13TR-EWS/ERG 0hr
<input type="checkbox"/>		GSE8665	GSE8665	GPL570	GSM214926	UET-13TR-EWS/ERG 48hr

差异分析

缺失值处理

标准化处理

探针处理

提交分析 (15/20)

免费版/基础版/高级版每日次数不同

加入参考组

加入实验组

未分组17个

取消分组 删除所选

选择参数进行分析

分析记录

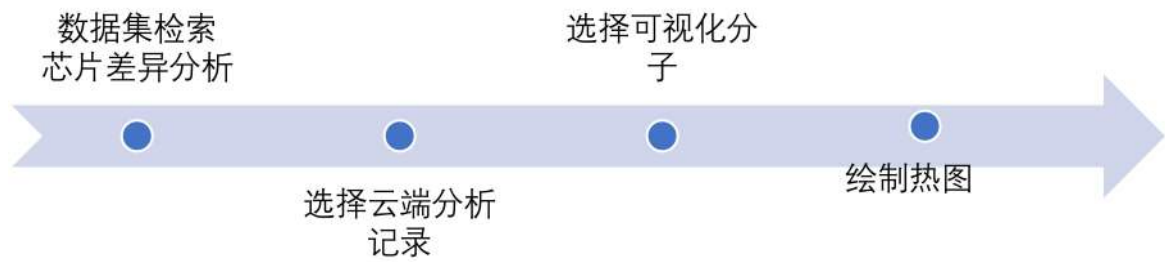
① 历史记录中超过30天的记录会自动清理

刷新

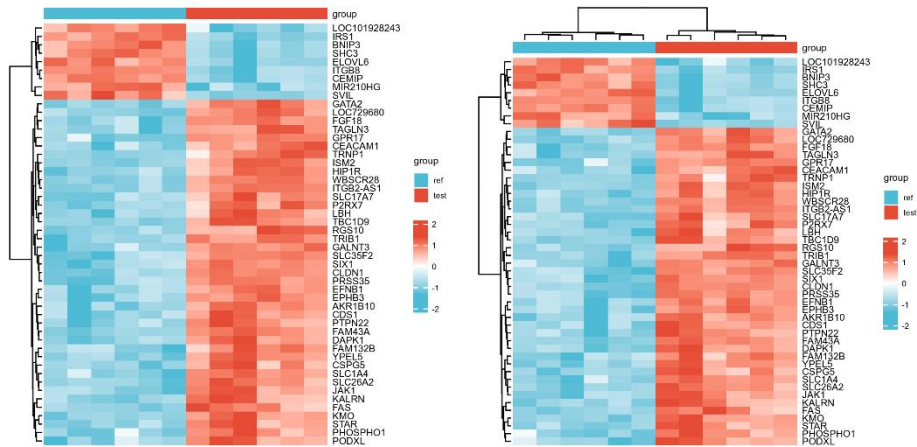
ID	名称	模块	状态	类型	时间	操作
1		芯片-差异分析	完成	表格	2023-03-12 20:35:56	更名 删除 下载 查看
2		芯片-差异分析	完成	表格	2023-03-12 20:14:16	更名 删除 下载 查看
3		芯片-差异分析	完成	表格	2023-03-12 20:13:12	更名 删除 下载 查看
4		芯片-差异分析	完成	表格	2023-03-12 20:10:10	更名 删除 下载 查看

时间最新在最上方记录

分析流程



主要结果



- **主要部分**：每一个小方块代表第 j 列样本对应第 i 行基因的表达值经过行 $zscore$ 转换后的值对应的颜色大小。其中红色代表相对表达高，蓝色代表相对表达低。

- $zscore$ 转换是绘制热图中常用的一种对数据进行转换的方法（每个基因的表达值减去其在所有样本中的表达均值后，再除以标准差），可以减少不同分子表达值差异过大而影响整个热图的可视化效果，并且 $zscore$ 转换保留了单个分子在样本间的差异情况（如果有一个分子表达值在样本间很大，另外一个在样本间很小，可以想象到前一个的热图基本都是红色，另外一个基本都是蓝色，单个基因样本间的差异就很小，很难看出单个基因的效果）。

- 如果不需要进行 $zscore$ 转换，仅使用原始的值，可以在参数【数据处理】中选择【不归一化】。

- **主要部分的左侧**：包含行的聚类树状图。
 - 聚类树状图：根据每一行的具体情况进行聚类，不同的聚类方法会影响这部分的顺序。如果选择了行聚类，则最终行的顺序不同于设定的分子顺序。
- **主要部分的上方**：包含列的注释，可以添加聚类树状图。
 - 注释条：根据云端记录中样本所属的分组信息，即在差异分析阶段，自定义的参考组（默认 ref）和实验组（默认 test）。

加入参考组

加入实验组



◦ ref组6个

◦ test组6个

◦ 未分组2个

- 聚类树状图：根据**每一列**样本的具体情况进行聚类，不同的聚类方法会影响这部分的顺序。如果选择了列聚类，则最终列的顺序不同于默认分组的顺序。



云端数据

云端数据

	记录名称	来源模块	时间	补充说明
<input checked="" type="checkbox"/>		芯片-差异分析 @1.0	2023-03-12 20:35:56	数据记录可以在历史记录中找到

这里的云端数据与历史记录汇总 数据集检索工具样本库中【差异分析】的数据记录是保持一致的，可以在历史记录中找到相应的数据记录。

根据需要可视化的项目 选择好对应的云端数据记录。默认使用最近生成的分析记录。



参数说明

(说明：标注了颜色的为常用参数。)

数据设置



- 可视化分子：输入想要可视化的 ID，默认为对应云端记录<数据集检索-差异分析>中选取“adj.P.Val”前 50 个分子，可以根据需要进行输入修改。注意：输入的 ID 来自所选云端数据记录的结果，需要先在**历史记录**中找到对应的记录，下载 **差异分析** excel 结果，复制想要展示的 ID 到这个输入框中，**一行代表一个**。最多支持 1 张图同时绘制 **500 个分子**。

标注

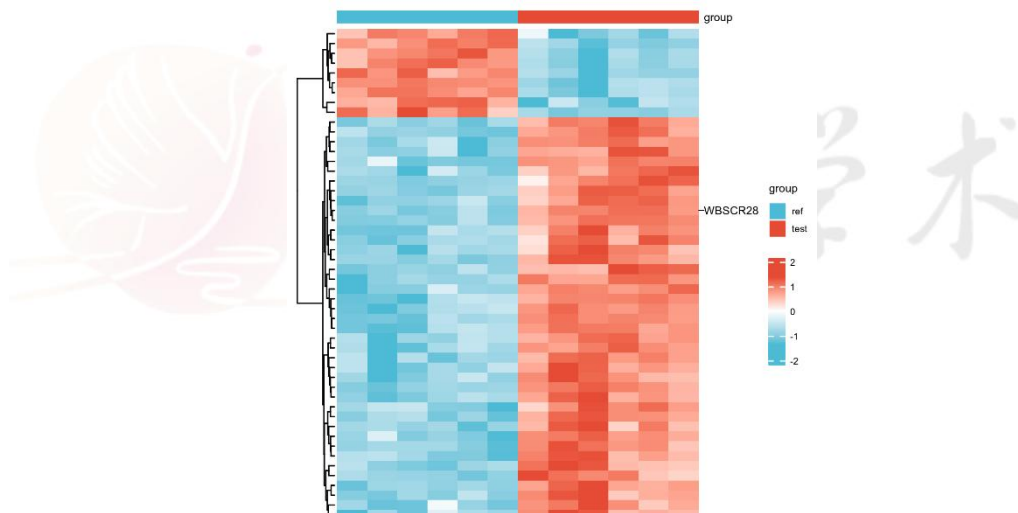
标注

分子标注

标注的字体

正体

- **分子标注**：可以填入想要标注在图中的对应行名(分子名)，一行一个分子，如果为空，默认就标注数据中所有行名(分子)。



- **标注的字体**：标注的字体类型，可选择 正体、斜体。

数据处理

数据处理

归一化

对行归一化

- 归一化：可以选择对行或列进行归一化的方式，可选择 对行归一化、对列归一化、对行归一化(处理极值: 归一化后大于 3 的值固定为 3,小于-3 固定为 -3)、对列归一化(处理极值: 归一化后大于 3 的值固定为 3,小于-3 固定为 -3)、不归一化。一般是 对行归一化。

聚类(顺序)/分割

聚类(顺序)/分割

列聚类

不聚类

行聚类

欧式距离(eu)

列分割

不分割

行分割

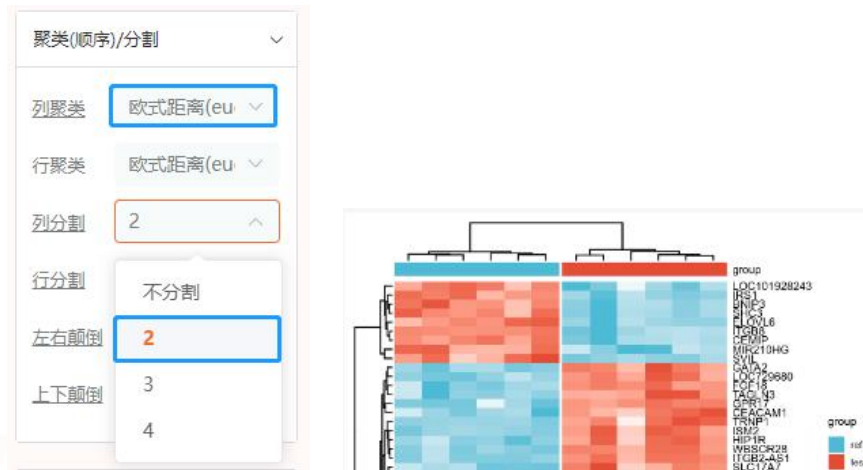
不分割

左右颠倒

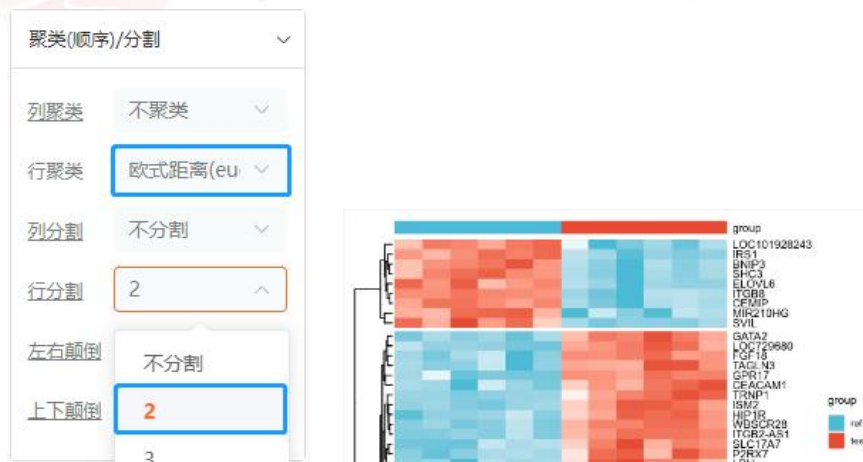
上下颠倒

- **列聚类**：可以选择是否对列进行聚类分析，可以选择不同[距离计算方法](#)，不聚类、欧式距离(euclidean)、切比雪夫距离(maximum)、曼哈顿距离(manhattan)、堪培拉距离(canberra)、闵可夫斯基距离(minkowski)。

- **行聚类**：可以选择是否对行进行聚类分析，可以选择不同距离计算方法，不聚类、欧式距离(euclidean)、切比雪夫距离(maximum)、曼哈顿距离(manhattan)、堪培拉距离(canberra)、闵可夫斯基距离(minkowski)。
- 列分割：可以对列进行分割展示，需要对列聚类才能进行列分割，如下，【列聚类】->【列分割】：



- 行分割：可以对行进行分割展示，需要对行聚类才能进行行分割，如下，【行聚类】->【行分割】：



- 左右颠倒：可以对热图中的列的顺序进行颠倒。
- 上下颠倒：可以对热图中的行的顺序进行颠倒。

主图



- 色阶：热图块的颜色，色卡的顺序是从低到高的渐变。受配色方案全局性修改。
- 色块描边：色块是否展示描边
- 外框：是否展示主图的外边框

上注释



- **注释颜色**：可以对图上方的注释内容进行颜色调节，本模块第一色卡控制参考组（默认 ref）分组，第二色卡控制实验组（默认 test）分组，最多支持修改 2 个颜色。不受配色方案全局性修改。

文字

文字

标题 标题内容

标题大小 7pt

列名大小 不显示

行名大小 6pt

- 标题：图形的标题文本
- 标题大小：标题的文字大小
- 列名大小：图中列名（下方样本）的字体大小，默认不显示
- 行名大小：图中行名（右侧基因）的字体大小

图注

图注

是否展示 ☒

图注标题 图注标题内容

图注标签 图注标签内容

图注位置 默认

文字大小 6pt

- 是否展示：是否展示图注
- 图注标题：可以添加图注标题，色阶的标题
- 图注标签：可以添加图注标签信息，如：



- 图注位置：可选择 默认、右、上。
- 字体大小：控制图注的文字大小，默认为 6pt

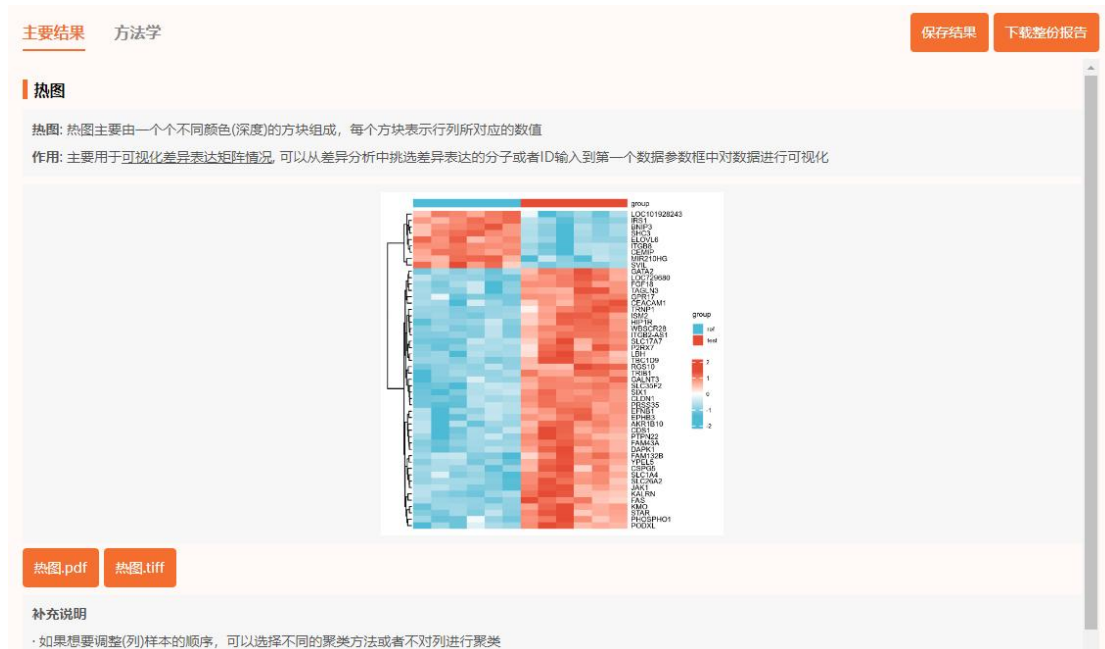
图片



- 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 格式下载, 结果报告可以下载包括 pdf 以及说明文本的内容。

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：ComplexHeatmap 包（用于可视化）

处理过程：利用 ComplexHeatmap 包进行热图的可视化。



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. 如何修改上注释的颜色?

答:

可以在【上注释】参数中修改【注释颜色】:



2. 如何修改绘图数据?

答:

本模块为 数据集检索 - 差异分析 后结果的可视化展示,默认为对应云端记录 <数据集检索-差异分析>中选取“adj.P.Val”前 50 个分子,可以根据需要进行输入修改。

输入的 ID 来自所选云端数据记录的结果,需要先在历史记录中找到对应的记录,下载 差异分析 excel 结果,复制想要展示的 ID 到这个输入框中,一行代表一个。最多支持 1 张图同时绘制 500 个分子。

我的收藏 分析工具 历史记录 拼图工具

历史记录中超过30天的记录会自动清理!

← 返回 [数据集] ... 批量删除 刷新

ID	名称	模块	状态	类型	时间	操作
1		芯片-差异分析	完成	表格		差异分析.xlsx 下载 查看
2		GSEA分析	完成	表格		样本信息.xlsx 表达谱.csv 下载 查看
3		预后筛选-云	完成	表格		箱式图.pdf 箱式图.tiff 下载 查看
4	edgeR流程	筛选分子-云	完成	表格		PCA图.pdf PCA图.tiff 下载 查看
5	DESeq2流程	筛选分子-云	完成	表格		样本-降维坐标.xlsx 火山图.pdf 火山图.tiff 下载 查看
6		单基因差异分析	完成	表格		差异表格.csv 热图.pdf 下载 查看
7		单基因差异分析	完成	表格		热图.tiff 下载 查看
8	edgeR流程	单基因差异分析	完成	表格	2023-02-28 22:22:55	下载报告 删除 下载 查看

	A	B	C	D	E	F	G	H
1	id	logFC	AveExpr	t	P.Value	adj.P.Val	B	anno
2	202718_at	5.7281172	4.80658994	6.64835371	2.5555E-05	0.00905302	2.94843022	IGFBP2
3	205440_s_at	5.60802162	3.05377225	6.42716187	3.5165E-05	0.01050627	2.64915668	NPY1R
4	206190_at	5.34344492	3.52105515	11.656637	7.8072E-08	0.00026679	7.99647107	GPR17
5	205858_at	5.21931466	4.8924156	7.08049219	1.3953E-05	0.00605454	3.51138783	NGFR
6	236218_at	5.186296	4.95628659	8.45121823	2.3782E-06	0.0020004	5.1179919	PHOSPHO1
7	238258_at	5.14528665	2.44629715	16.1199523	2.0986E-09	3.6202E-05	10.603214	WBSCR28
8	204818_at	4.55371322	4.83227605	4.77628425	0.00047146	0.04836215	0.17163268	HSD17B2
9	1554012_at	4.4825641	3.85713918	4.00715746	0.00179278	0.10242454	-1.11950474	RSPO2
10	1556168_s_at	3.98026514	4.58300691	4.95452663	0.00034976	0.04148177	0.45950646	MROH2A
11	222549_at	3.8999545	6.94312616	15.7921463	2.6485E-09	3.6202E-05	10.4515444	CLDN1
12	205051_s_at	3.86970788	5.82163002	4.72922739	0.0005105	0.05038147	0.09486319	KIT
13	213880_at	3.79042516	1.50477889	4.21404039	0.00124284	0.08307113	-0.76515703	LGR5