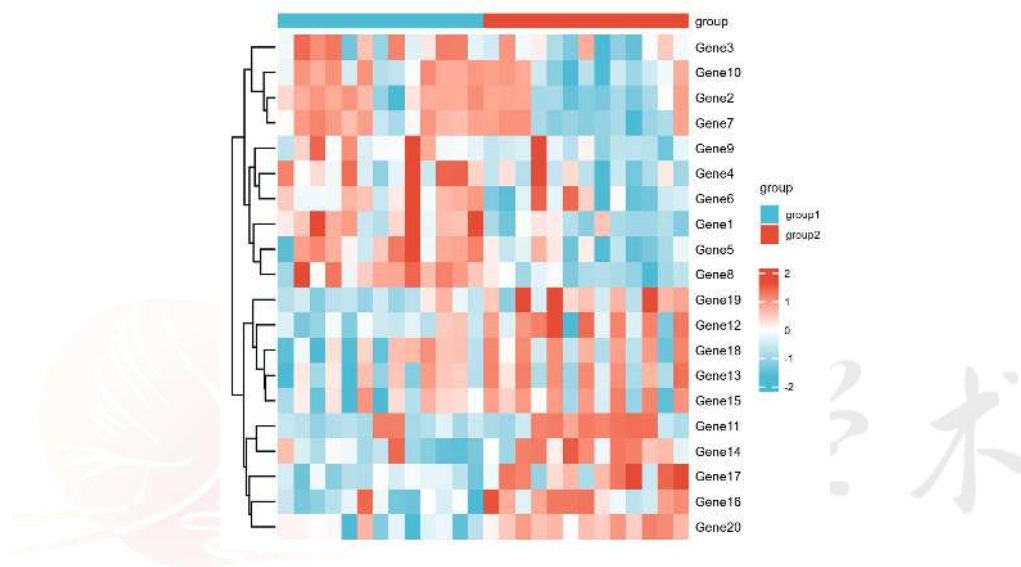


表达差异 - 简易数值热图



网址: <https://www.xiantao.love>



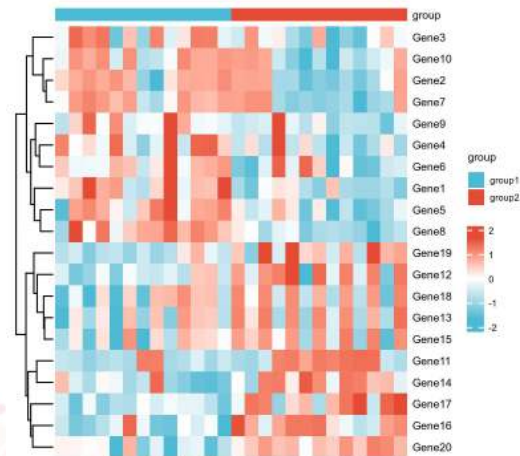
更新时间: 20230225

目录

基本概念	3
应用场景	3
结果解读	4
数据格式	5
参数说明	6
标注	6
数据处理	6
聚类(顺序)/分割	7
主图	8
上注释	8
文字	9
图注	10
图片	10
结果说明	11
主要结果	11
补充结果	11
方法学	12
如何引用	13
常见问题	14

基本概念

- 简易数值热图：热图通过色彩变化来显示数据的数值，可以用于可视化表达矩阵情况。
- 图形构成：

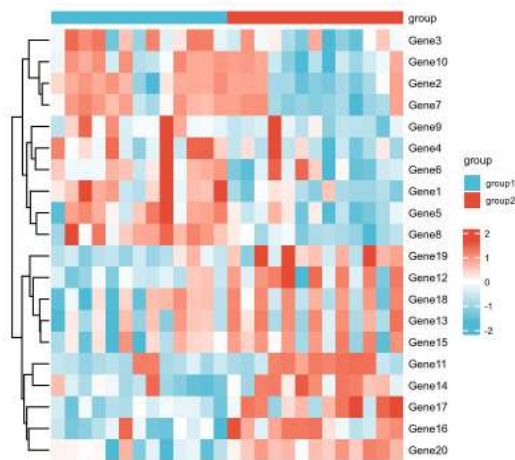


简易数值热图

应用场景

热图经常用于展示多个基因在不同样本中的表达水平，然后通过聚类等方式查看不同组（如疾病组和对照组）特有的 pattern（如 CNV）。另外，热图还可以用于展示微生物的相对丰度、代谢组不同物质的含量等等。

结果解读



简易数值热图

- 横坐标表示样本，纵坐标表示基因，每个方块表示行列所对应的数值
 - 热图部分的数值经过了 zscore 转换后的值的情况
 - zscore 转换是绘制热图中常用的一种对数据进行转换的方法（每个值减去行均值后，再除以标准差），可以减少值差异过大而影响整个热图的可视化效果，并且 zscore 转换保留了值在行的差异情况（如果有一个值在样本间很大，另外一个在样本间很小，可以想象到上一个的热图基本都是红色，另外一个基本都是蓝色，另外的一些值样本间的差异就很小，很难看出差异的效果）

如果不想要进行 zscore 转换，仅使用原始的值，可以在参数中进行调整

- 如果想要调整(列)样本的顺序，可以选择不同的聚类方法或者不对列进行聚类

聚类后行名会根据聚类后的情况进行重新排序，不再是上传数据的行的顺序，如果要保留行的顺序，需要不对行进行聚类。

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1	group1
id	sample1	sample2	sample3	sample4	sample5	sample6	sample7	sample8	sample9	sample10	sample11	sample12	sample13	sample14	sample15	sample16	sample17	sample18	sample19	sample20
Gen#1	1.0857	2.2235	6.8563	3.0051	3.5615	-1.4768	2.1033	1.669	6.5741	0.259	2.5695	2.3185	7.0977	-2.5332	-4.2016	0.259	1.2427	1.0115		
Gen#2	-0.6216	1.0857	2.3033	1.1457	2.0032	0.6216	-6.7765	-9.3188	-1.1699	1.0882	1.2241	1.1451	2.3203	1.325	1.4472	1.3492	-0.6784	-6.045		
Gen#3	-0.2429	5.5259	4.0351	5.0517	4.9229	2.4139	-6.668	4.6317	-1.0219	1.1116	0.6481	4.7647	-0.7783	-1.6115	4.0373	-0.2818	0.9812	-2.9093		
Gen#4	4.2098	1.6785	1.6226	0.1889	4.7902	-0.552	-2.8266	-0.1889	6.3114	-1.9983	5	5.0679	1.9206	-1.2862	1.5692	-1.6584	5.8074	-1.278		
Gen#5	-6.3314	0.6453	2.1495	1.5383	-0.0633	-1.4107	0.972	2.3952	3.7752	-0.5297	1.4701	1.5586	2.2286	0.0633	-1.5309	-0.4899	1.2755	0.1624		
Gen#6	1.6555	-0.1864	-0.1317	-0.0404	2.2153	1.8374	-1.2199	0.8238	6.4755	0.0404	2.1011	2.5644	2.9423	-1.3272	-3.8932	1.5361	4.2863	-4.4002		
Gen#7	-0.7368	1.4554	1.9224	1.3962	0.6254	1.3258	-0.1294	-2.891	-0.6254	1.5282	0.7815	0.6482	1.2701	1.0192	1.613	0.6845	-2.8414	-3.5643		
Gen#8	-1.4519	3.6939	0.1097	2.1123	-0.2289	0.8314	1.3655	1.4233	2.3378	1.0932	1.442	1.5834	0.916	0.3628	0.0772	-1.4307	-0.3454	-0.0772		
Gen#9	-1.4127	2.0638	5.3161	0.6544	4.1518	-0.455	0.7313	0.6092	9.0449	3.2835	0.3413	0.552	-0.1036	-1.235	-0.7487	-0.4607	0.9581	0.3372		
Gen#10	-0.1875	3.0719	2.424	3.1562	-0.9029	3.133	-1.7827	-1.4277	0.0958	3.2824	2.4765	2.4523	2.9365	2.7331	2.8994	2.526	-1.1006	-2.8712		
Gen#11	-0.2351	-0.6441	-2.486	-1.2465	-0.8256	-1.6167	0.7845	8.3511	-1.7221	-0.9228	0.9772	-1.8489	0.6625	-1.5645	-0.2005	0.0871	8.1069	8.4481		
Gen#12	-0.3691	-3.8057	-2.7756	0.6089	-3.4537	0.6895	-0.8223	-1.1505	-0.0516	-1.6713	2.7222	2.9907	-1.0317	4.2462	-0.2412	4.5576	5.2584	7.214		
Gen#13	-8.3031	-0.1096	-0.4033	-0.1713	-7.5656	1.5439	-9.5131	1.2661	-5.6858	2.9047	1.2727	0.9686	-3.9109	3.7148	0.1096	3.9393	4.2799	3.3582		
Gen#14	1.3715	-0.5592	-1.4096	0.2864	-0.1607	-1.6034	-1.1046	2.8815	-1.6774	-2.1606	-2.7054	-2.79	-2.3845	0.1607	-1.7205	2.4851	2.5571	0.4831		
Gen#15	-0.2134	-0.503	-6.4603	-0.1206	-7.7161	0.2749	-7.7146	0.1206	-3.8329	1.5482	0.3335	0.1457	-1.1982	2.216	0.4878	2.4843	0.9303	2.5199		
Gen#16	-0.7107	-2.0028	-1.697	-0.9882	-0.1341	2.6978	-0.0083	-2.0686	-2.1884	0.2763	-0.4148	0.1918	-2.594	5.0375	1.6124	-0.2729	1.7223	2.4516		
Gen#17	-0.4036	-0.0148	-0.1376	-1.4493	-0.9408	-1.0166	0.3084	0.5586	0.1578	-0.0971	0.1064	-0.4840	1.9541	0.0971	2.4683	0.3739	-1.4769	0.8779		

- 数据至少 2 列以上，每列至少 3 行，最多支持 3000 列和 3000 行数据

- 变量名/列名不能重复

参数说明

(说明：标注了颜色的为常用参数。)

标注



- 分子标注：可以填入数据中对应行名(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；默认是正体

数据处理



- 转换：可以对数据进行 log2 或者 log10 等转换
- 归一化：可以选择对行或列进行归一化

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(euclidean)

列分割

不分割

行分割

不分割

左右颠倒

上下颠倒

- 列聚类：可以选择对列进行聚类分析，可以选择不同距离计算方法
- 行聚类：对行进行聚类分析，方法同列聚类
- 列分割：可以对列进行分割展示
- 行分割：同列分割
- 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
- 色块描边：色块是否展示描边
- 外框：是否展示

上注释



- 注释颜色：可以对图上方的注释内容进行颜色调节

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

如果行数过多，建议改小或者是改成标注少数几个分子的方式

- 标题：填入热图的标题，换行可以在需要换行的位置插入\n
- 标题大小：标题的字体大小
- 列名大小：调节列名的大小或者不显示
- 行名大小：通列名大小

图注

图注

是否展示

☒

图注标题

图注标题内容

图注位置

默认

文字大小

6pt

- 是否展示：图注内容是否展示
- 图注标题：可以填入图注标题，换行可以在需要换行的位置插入\n
- 图注位置：默认是右，也可以选下
- 文字大小：图注文字的大小

图片

图片

宽度 (cm)

10

高度 (cm)

9

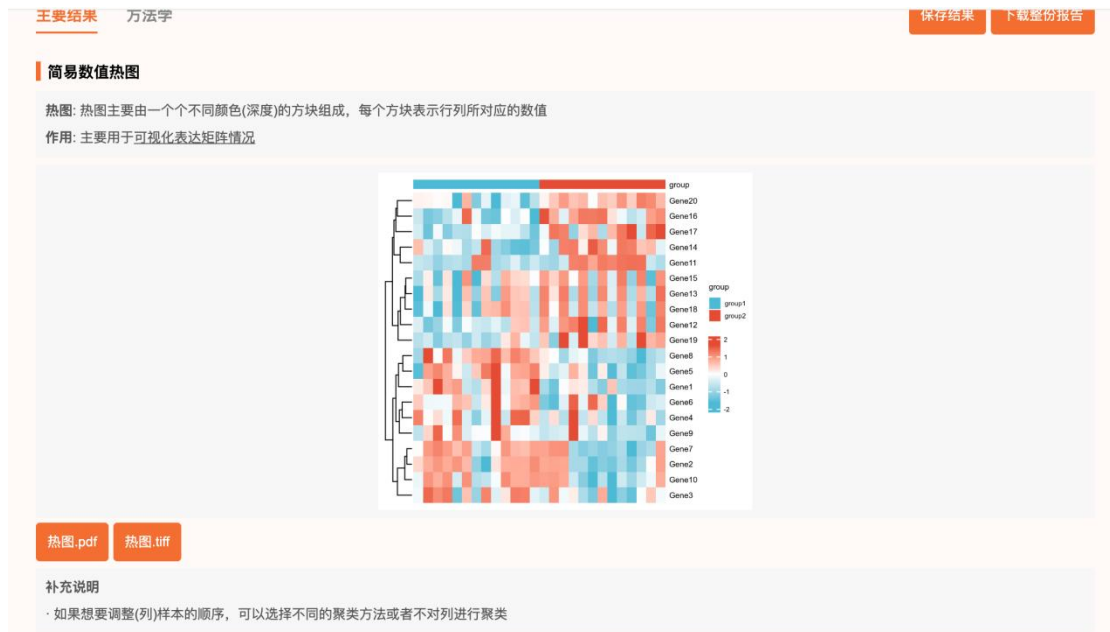
字体

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：可以选择图中文本内容字体

结果说明

主要结果



补充结果

无

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: ComplexHeatmap[2.13.1]包

处理过程:

- 利用 ComplexHeatmap 包进行热图的可视化

补充说明:

- 数据转换: 无
- 归一化处理: 对行归一化
- 行聚类: 欧式距离(euclidean)
- 列聚类: 不聚类

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. 为什么要对数据进行转换?

答：实际应用中，异常值的出现可能会影响可视化结果，对数转换后的数据，看起来清晰很多，也会有更好的可视化效果。而且对数转换后，数据还保留着之前的变化趋势

