
软件说明书

热图绘制系统

简称：oppHeatmap

Version 1.0

作者：张扬

目录

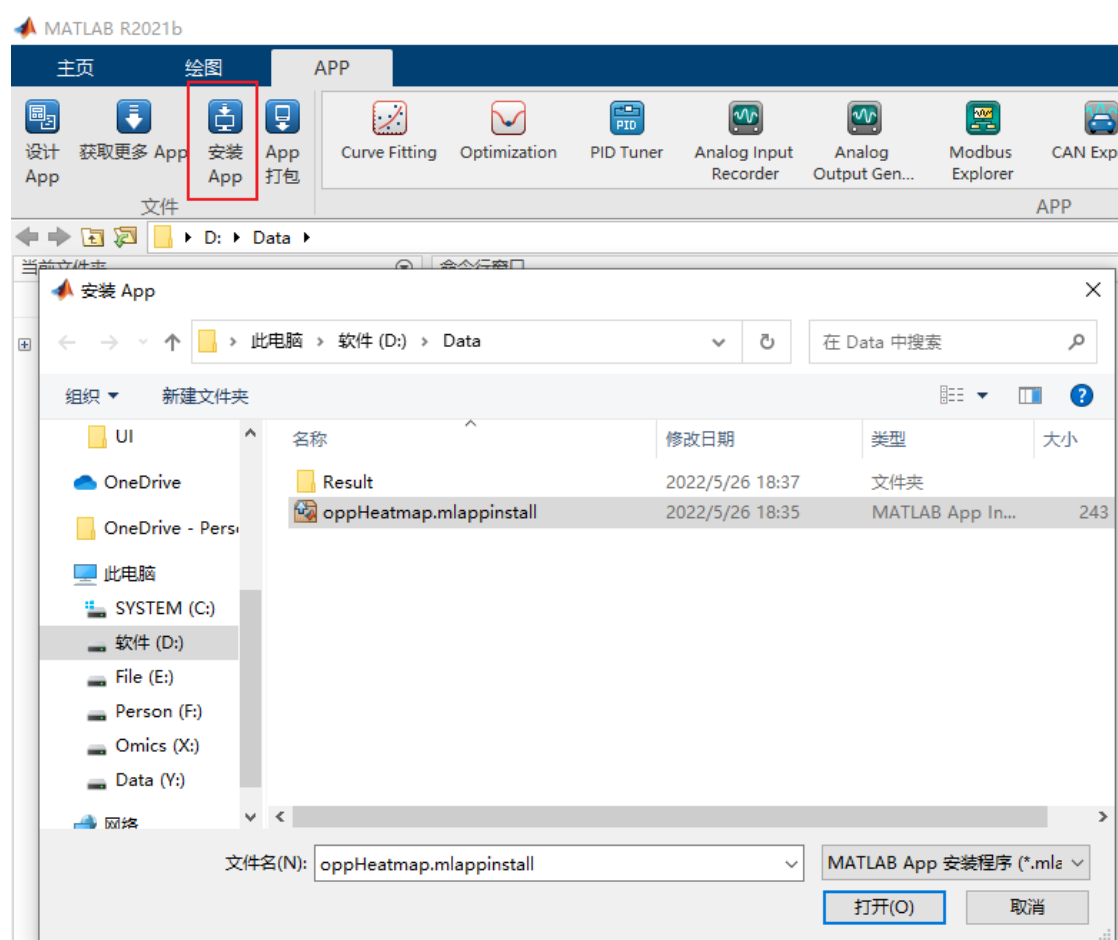
1. 简介	2
2. 软件操作主界面	2
3. 案例数据 (Data.xlsx)	3
4. 普通热图 (Heatmap)	3
5. 微板热图 (MicroPlatePlot)	7
6. 双边分层聚类 (HierarchyClustering)	8
7. 树图 (TreeMap)	10
8. 样品相关系数图 (SampleCorrelation)	11
9. 基因相关系数图 (GeneCorrelation)	14
10. 环形热图 (Polar)	17

1. 简介

热图绘制系统，简称 oppHeatmap，是基于 MATLAB 的 AppDesigner 开发的，带有图形界面的针对组学来源数据的热图绘制软件。oppHeatmap 支持普通热图、双边分层聚类热图、树图（Treemap）、微板图（Microplate）和样品相关系数图（方阵图、上三角图和下三角图）、基因相关系数图（单表和双表）和环状热图的绘制。oppHeatmap 支持热图边框、字体和颜色的修改。oppHeatmap 读取 Excel 来源的组学数据，并在 MATLAB 环境中生成热图，用户可以通过 SaveAs 命令保持成矢量图。

2. 软件操作主界面

先加载程序包：打开 MATLAB，在 APP 界面下点击“安装 APP”，选择“oppHeatmap.mlappinstall”，可以把 oppHeatmap 程序包加载进入 MATLAB。



在 MATLAB 中执行命令“oppHeatmap”。

进入 oppHeatmap 主界面。oppHeatmap 包含 Heatmap、MicroPlatePlot、HierarchyClustering、TreeMap、Sample Correlation、Gene Correlation 和 Polar 共 7 个主界面。

3. 案例数据 (Data.xlsx)

表：oppHeatmap 操作所用到的案例数据。

表名	说明	样品数目
Heatmap	蛋白质组部分结果案例	有 4 个样品（A、B、C 和 D），各三次生物学重复。
Proteome	天花粉蛋白质组全部结果	有 3 个样品（A、B 和 C），各三次生物学重复。
TreeMap	TreeMap 图案例数据	
T1	转录组数据案例	有 5 个样品，各 1 次生物学重复。
T2	蛋白组数据案例	有 5 个样品，各 1 次生物学重复。

4. 普通热图 (Heatmap)

4.1 Draw Common Heatmap

在 Heatmap 面板下，点击“Open”，选择组学来源的表格数据，一般存在 Excel 中。一般格式如下：具有 ID 列（是字符串）、每列都是每个样本的表达量，包含重复样品。如下表：有 4 个样品（A、B、C 和 D），各三次生物学重复。

ID	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3
UBA6_HUMAN	823.7	593.33	863.77	155.2	780.71	183.53	966.26	294.44	1356.9	171.57	204.13	722.93
ESYT2_HUMAN	3244	2408.4	3167.3	1216.6	5012.2	2105	2048.3	957.98	8168.4	596.77	630.09	2308.9
TM223_HUMAN	2117.7	640.63	1071	3611.8	816.26	1043.4	707.25	89.022	2810.4	302.56	845.15	2925.4
NBAS_HUMAN	371.27	366.72	696.96	315.98	940.35	1812.4	390.05	152.67	850.45	119.5	112.22	160.77
SYTC2_HUMAN	560.42	305.94	503.43	775.84	1121.3	180.2	433.7	240.25	386.35	311.56	150.92	479.71
VWA8_HUMAN	1286.5	1522.8	5861.8	3054.3	5830.8	3932.7	1461.1	1616.7	3528.7	1045.1	1864.2	4007.4
GTPBA_HUMAN	119.56	284.73	711.08	143.93	1353.7	493.78	69.883	412.48	868.78	259.11	131.15	338.72
XIRP2_HUMAN	337.78	250.59	1643.1	3011.6	4706.4	3874.7	1164.6	849.2	1376.1	783.39	1426.3	2185.7
FITM1_HUMAN	4288.4	1601.6	1802.4	1788.1	1216.8	3821.6	1814	623.98	2792.4	1468.2	808.23	1375.6
PGP_HUMAN	5972.4	7800	6176.9	3054.7	17778	8502.3	4955.5	5562.1	14266	5263.1	2823.7	10624
NDUA2_HUMAN	111160	105730	156890	84013	220170	300200	90270	105150	363490	22654	17287	366000
ASNA_HUMAN	8412.7	11904	10140	6885.1	14979	10775	6377.3	3545.5	20434	1039.7	3229.9	16335
BUB3_HUMAN	1591.5	1189.1	1717.4	2205.6	2527.2	880.83	1310.6	753.57	3460.5	546.51	1934.9	1617.7
ACTN4_HUMAN	30645	33573	29262	52697	18986	40494	19679	10342	73003	9314	19989	40326
MAAL_HUMAN	2578.2	6090.2	9451.3	1528.3	15811	11139	2155.2	5319.8	11691	465.02	3263.2	13129
SUN1_HUMAN	3179.2	3447.9	3606.7	1788.8	3685.1	3807.6	2473.8	1315.7	7839.7	1113.7	1081	2048.2
PROSC_HUMAN	3751.5	5484	4601.8	1874	5920.7	3865.1	4303.7	790.15	9424.2	1035.2	735.64	4126.3
ERLN2_HUMAN	5145	9210.4	16706	3558.2	12608	9522.3	13412	4134	19961	6709.2	6045.2	9047.2
PRP6_HUMAN	252.89	150.79	45.776	55.294	83.288	98.244	124.76	88.799	451.81	40.655	162.56	459.82
ABCA8_HUMAN	3956.7	2927.4	3441.9	902.22	3427.1	2391.7	2500.1	2527.6	7834.8	865.62	1463.4	3779.4
FRYL_HUMAN	120.06	40.967	20.883	35.202	161.21	55.629	49.192	4.684	29.449	34.372	77.989	
ENDD1_HUMAN	1961.6	1494.7	5840	282.51	2518.2	2000.6	1769.4	669.06	2992.2	303.96	429.49	525.84
GLSK_HUMAN	3482.1	2485.9	6168.8	1017.4	2848	1018.6	4167.5	638.71	11923	545.21	1302.4	1127.3
ABLM3_HUMAN	1310	925.58	1667.8	516.81	1787.7	3104.1	975.45	754.26	1298.8	233.27	732.13	4087.8
NDUC2_HUMAN	3085.1	20996	6076.1	83540	37061	113270	55218	42730	75815	25025	21734	53547
NDUAA_HUMAN	101610	84579	84193	81880	154200	160990	84328	68095	182940	51920	56142	169130
FKBP9_HUMAN	1265.6	1502.3	1210	1051.2	689.98	1851.4	891.59	448.69	3770	240.75	710.36	1747.3
CELF2_HUMAN	660.06	698.41	2794.7	964	1149.9	1013.3	1255.6	1595.3	2855.5	466.54	260.33	1035.6
6PGL_HUMAN	6009.9	8063.7	9731.8	2337.2	9907.5	8998.7	7233.3	3943.4	18919	1289.7	4903.8	10233
TACC2_HUMAN	1629.4	1505	3131.3	283.78	3253	2003.3	1130	858.48	2327.5	393.98	1045.4	2190.8
SYFM_HUMAN	1906	394.18	494.81	477.47	1948.2	559.67	596.74	1033.6	1601.4	665	1483	1500
LYPA2_HUMAN	6526	3572.6	3015.8	686.33	4048.1	209.55	2473	738.28	8567.6	920.36	652.77	811.78
IPO7_HUMAN	1799.7	2102.1	3478.4	1034.4	1740.6	2373.6	1783.3	997.31	3483.5	1625.7	1851.3	2101.3
ARI2_HUMAN	158.82	297.74	416.41	96.123	2053.4	186.67	283.48	404.63	1238.1	267.8	485.2	218.25

图. 导入的组学原始数据。

oppHeatmap

HeatMap

MicroPlatePlot

HierarchyClustering

TreeMap

SampleCorrelation

GeneCorrelation

Polar

Excel Path

D:\Data\Data.xlsx

Open

Draw Excel

No. of Sheet in Excel

1

Plot

Data Column

2:13

Row Range

2:31

GeneName Column

1

Color Schema

Red & Green

Red & Blue

Grid Line

Color

0,0,0

Width

2

图. 普通热图的参数面板。

“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“Row Range”是截取哪些行进行热图绘制；“yLabel Column”是热图中纵坐标的文字来源。

Grid Line 是在热图中添加网格线。“Color”按钮可以激活调色板，选择的颜色 RGB 值会出现在右侧方框中，是网格线的颜色；“Width”是网格线的粗细。缺省以红绿方式显示热图，即红色表示上调，绿色表示下调。整个软件颜色的表述方式一致。

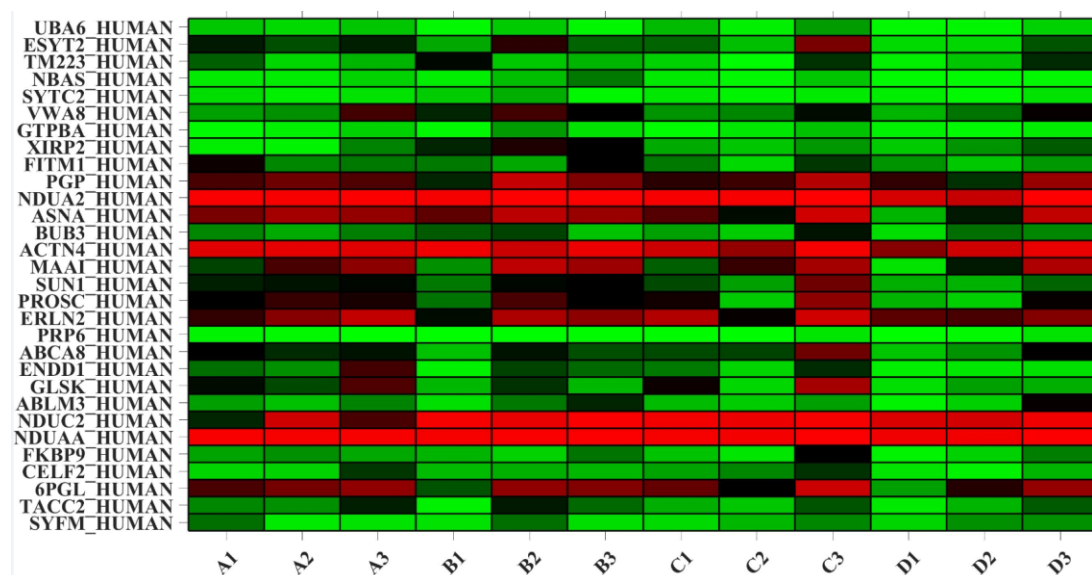


图. 利用红绿方式显示普通热图。

若在“Color Schema”中选择“Red&Blue”，便以红蓝方式显示热图。即红色表示上调，蓝色表示下调。在 MATLAB，键入“colormap redgrencmap”或“colormap redbluecmap”也可以实现红绿和红蓝转换。

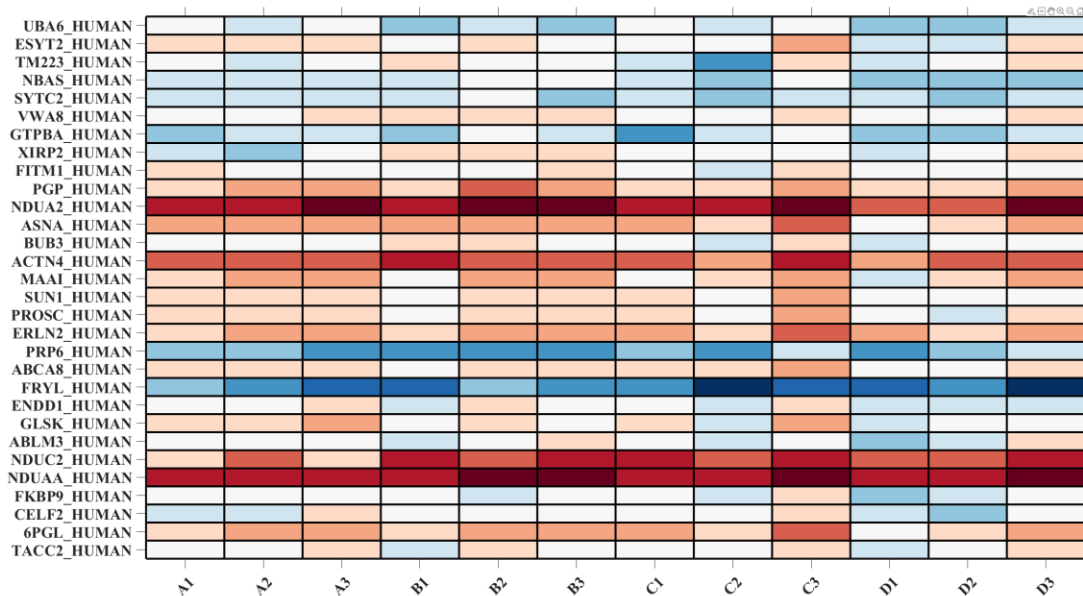


图. 利用红蓝方式显示普通热图。

4.2 Plot to Excel

oppHeatmap 可以实现将表格数据以热图的方式输出到 Excel 中，以代替 Excel 中根据数据产生热图的“条件格式”功能。“Row Standardize”表示输入表格进行 zscore 的行归一化，即行数据控制在 $(-1,1)$ 之间，以优化热图的可视化效果。

oppHeatmap

HeatMap MicroPlatePlot HierarchyClustering TreeMap SampleCorrelation GeneCorrelation Polar

Excel Path

No. of Sheet in Excel

Data Column

GeneName Column

Color Schema

☒ Red & Green

☐ Red & Blue

☒ Row Standardize

图. 在 Excel 中产生热图的面板。

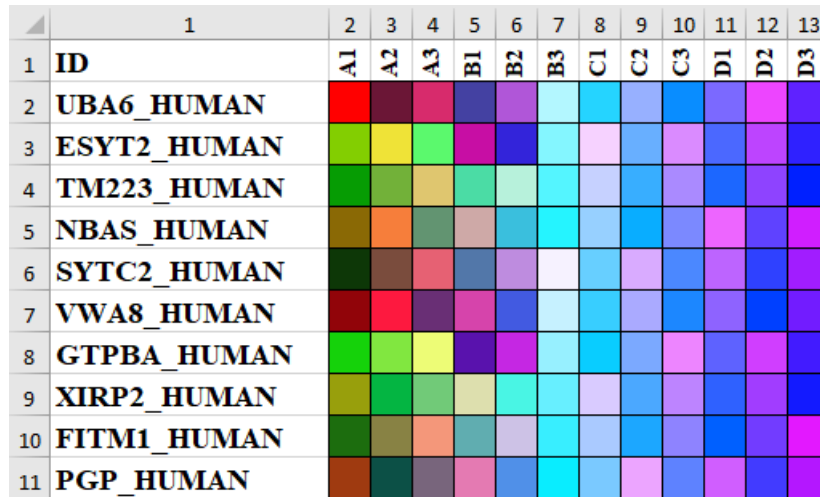


图. 在 Excel 中产生热图的效果。

5. 微板热图 (MicroPlatePlot)

在 MicroPlatePlot 面板中，“Excel Path”填写组学的来源数据；“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“Name Column”是微板图纵坐标中显示的文字来源列。

oppHeatmap

HeatMap MicroPlatePlot HierarchyClustering TreeMap SampleCorrelation GeneCorrelation Polar

Excel Path

D:\Data\Data.xlsx

Open

No. of Sheet in Excel

1

Plot

Data Column

2:13

Name Column

1

图. 微板图的参数界面。

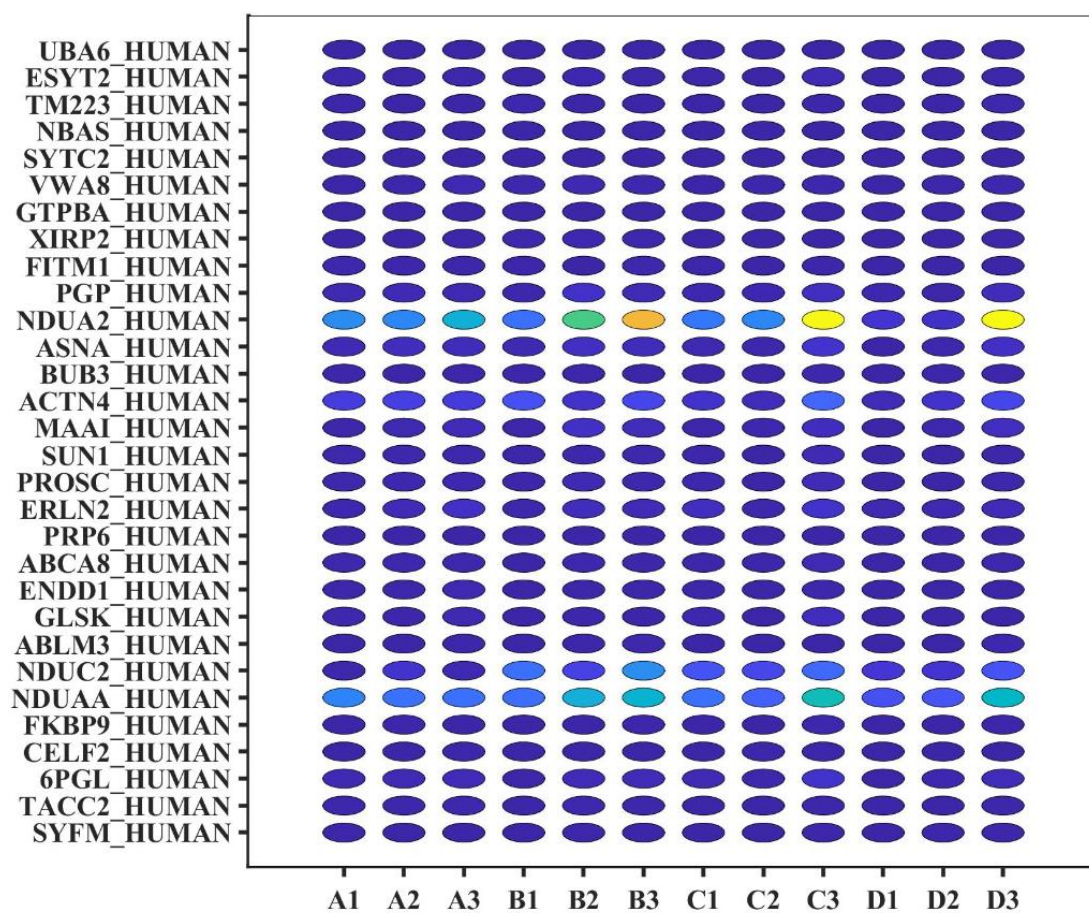


图. 微板图的显示结果。颜色深浅与原始数据的表达量相关。

6. 双边分层聚类 (HierarchyClustering)

“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“GeneName Column”是热图中纵坐标的文字来源；“Replicates Number”是样品重复性的分组情况。

Calculation 面板中“Distance for Row”和“Distance for Column”是聚类计算中距离的公式，常用欧氏距离或 Pearson 相关系数；“Linkage”表示多个距离合并时所采用的公式。

Color Schema 面板中“Red & Green”和“Red & Blue”分别是红绿和红蓝方式显示聚类热图。红色表示上调；绿色和蓝色表示下调。

“Row Standardize”是原始数据经过行归一化后再显示热图。

“Export Result”是将聚类重排的结果输出到当前路径下的 HC.txt 文件中。

oppHeatmap

HeatMap

MicroPlatePlot

HierarchyClustering

TreeMap

SampleCorrelation

GeneCorrelation

Polar

Excel Path

D:\Data\Data.xlsx

Open

No. of Sheet in Excel

2

Plot

Data Column

2:10

GeneName Column

1

Calculation

Distance for Row

Euclidean

Distance for Column

Euclidean

Linkage

Average

Color Schema

☒ Red & Green
 ☐ Red & Blue

☒ Row Standardize
 ☒ Export Result

图. 分层聚类热图参数面板。

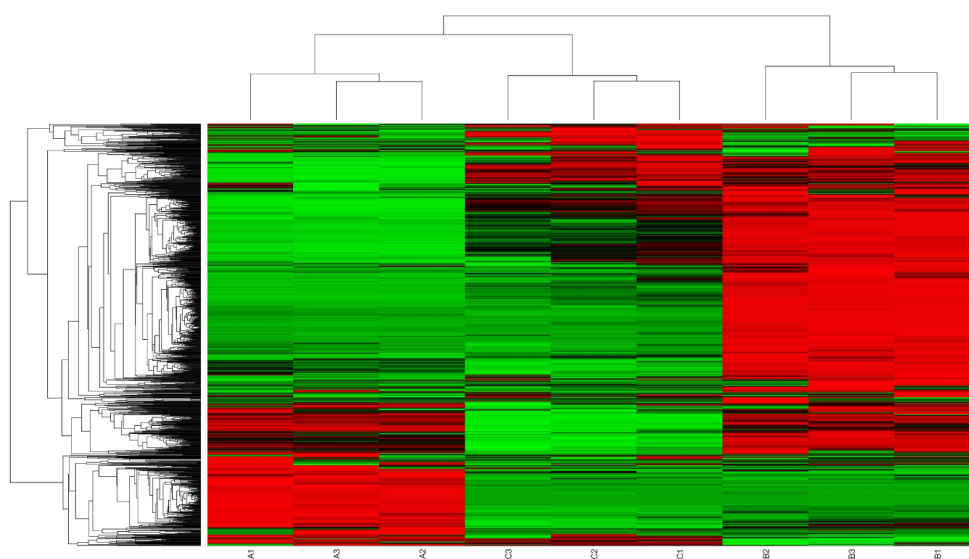


图. 红绿显色的双边聚类热图。

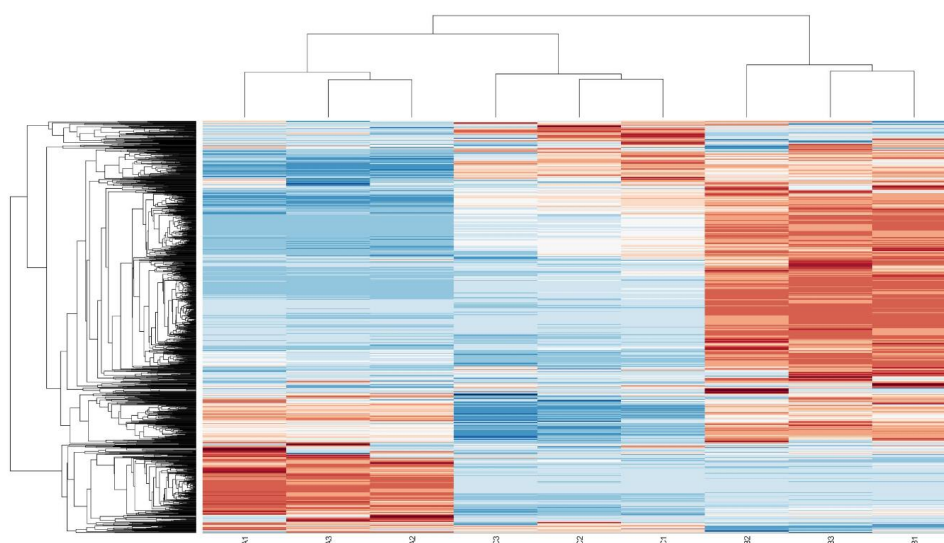


图. 行归一化后的红蓝显色的双边聚类热图。

7. 树图 (TreeMap)

在 TreeMap 面板中, Single TreeMap 是做一级的树图, Nested TreeMap 是多级树图。“Excel Path”填写组学的来源数据;“No. of Sheet in Excel”是填写获取 Excel 中哪张表;“Data Column”是获取具体哪些列;“Name Column”是树图中每个格子中显示的文字来源列。“Add Color”不选时,树图是黑白的;选择时,树图是彩色的。“Font Size”是网格中文字的大小;“Font Color”是网格中文字的颜色。

oppHeatmap

HeatMap MicroPlatePlot HierarchyClustering **TreeMap** SampleCorrelation GeneCorrelation Polar

Excel Path

Single TreeMap

No. of Sheet in Excel

Data Column

Name Column

Font Size

Font Color

☒ Add Color

Nested TreeMap

No. of Sheet in Excel

Data Column

图. 树图的参数面板。

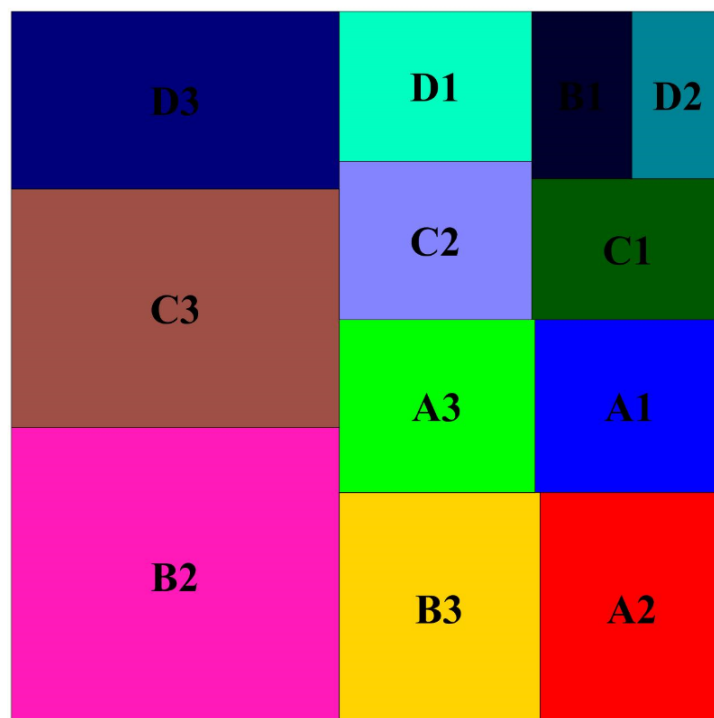


图. 彩色方式显示的一级树图。

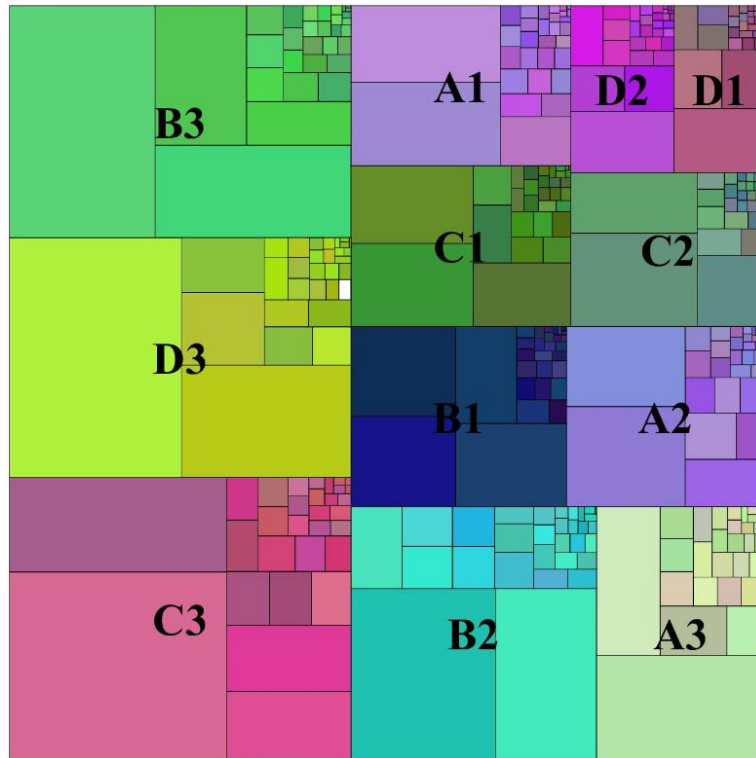


图. 彩色方式显示的多级树图。

8. 样品相关系数图 (SampleCorrelation)

8.1 相关系数全图

用以计算单表格内，每列之间的相关系数，形成热图。

在 SampleCorrelation 面板的 Full 面板中，“Excel Path”填写组学的来源数据；“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“Replicates Column”是样品的重复性分类。每两个样品的表达量都会做相关系数，采用 Distance 中选择的参数进行计算。

Grid Line 面板中，“Color”表示网格线的颜色；“Width”表示网格线的粗细。“Sort Column”表示热图展示的时候，接近的样品是否要重新排序，显示在一起；“Add SampleLine”表示热图的上方是否要显示样品 barcode，重复的样品会出现相同的颜色。“Export Result”是将计算得到的相关系数结果输出到文本中。

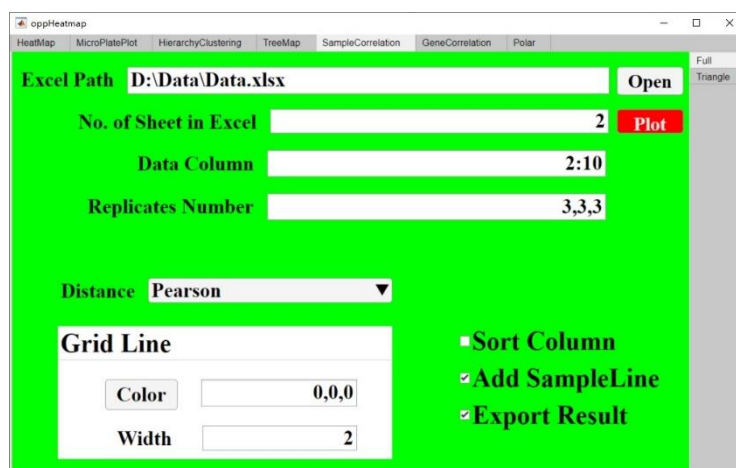


图. 样品相关系数全图的参数界面。

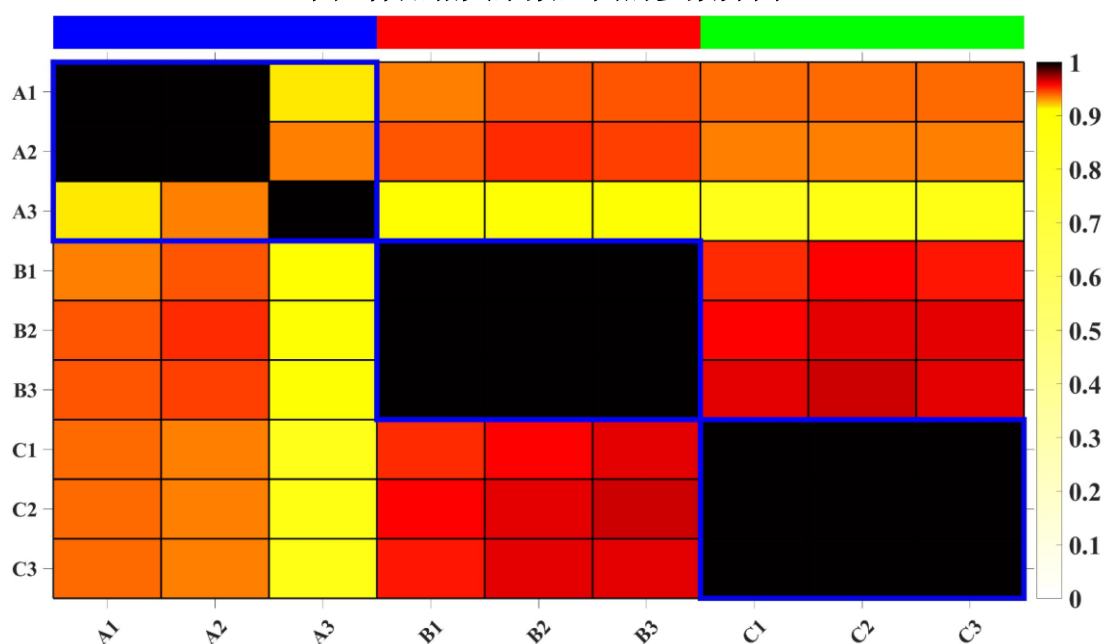


图. 样品相关系数图的显示结果。蓝色的方框表示重复的样品之间的相关系数。用户可以右键点击右侧的 ColorBar，进行热图的颜色对比度修改。

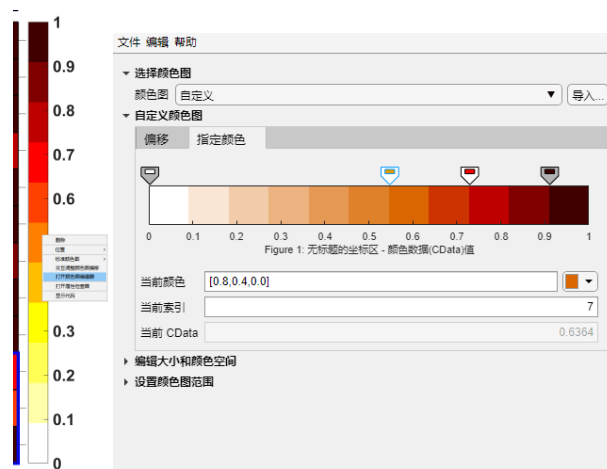


图. 对比度选择器。

8.2 相关系数半图

在 SampleCorrelation 面板的 Triangle 面板中，“Excel Path”填写组学的来源数据；“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“Replicates Column”是样品的重复性分类。每两个样品的表达量都会做相关系数，采用 Distance 中选择的参数进行计算。

Grid Line 面板中，“Color”表示网格线的颜色；“Width”表示网格线的粗细。Orientation 面板中，“Upper Triangle”表示上三角热图；“Lower Triangle”表示下三角热图。“Sort Column”表示热图展示的时候，接近的样品是否要重新排序，显示在一起；“Add SampleLine”表示热图的上方是否要显示样品 barcode，重复的样品会出现相同的颜色。“Export Result”是将计算得到的相关系数结果输出到文本中。

oppHeatmap

HeatMap MicroPlatePlot HierarchyClustering TreeMap SampleCorrelation GeneCorrelation Polar

Excel Path

No. of Sheet in Excel

Data Column

Replicates Number

Distance

☐ Sort Column

☒ Export Result

Grid Line

Color

Width

Orientation

☒ Upper Triangle

☐ Lower Triangle

Full Triangle

图. 样品相关系数半图的参数界面。

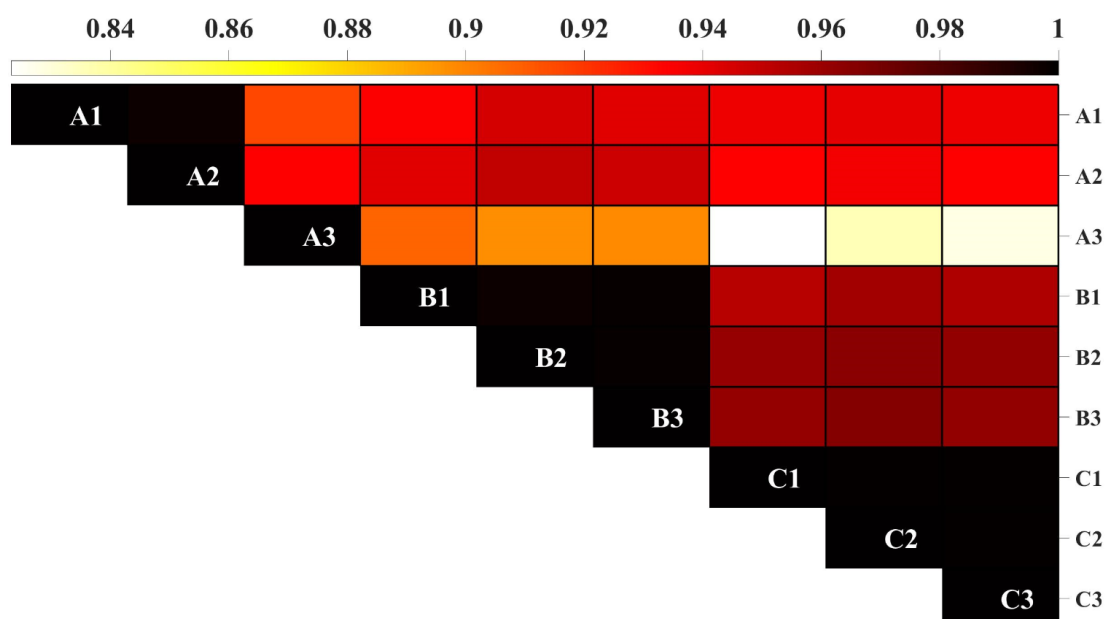


图. 样品相关系数半图，上三角热图。

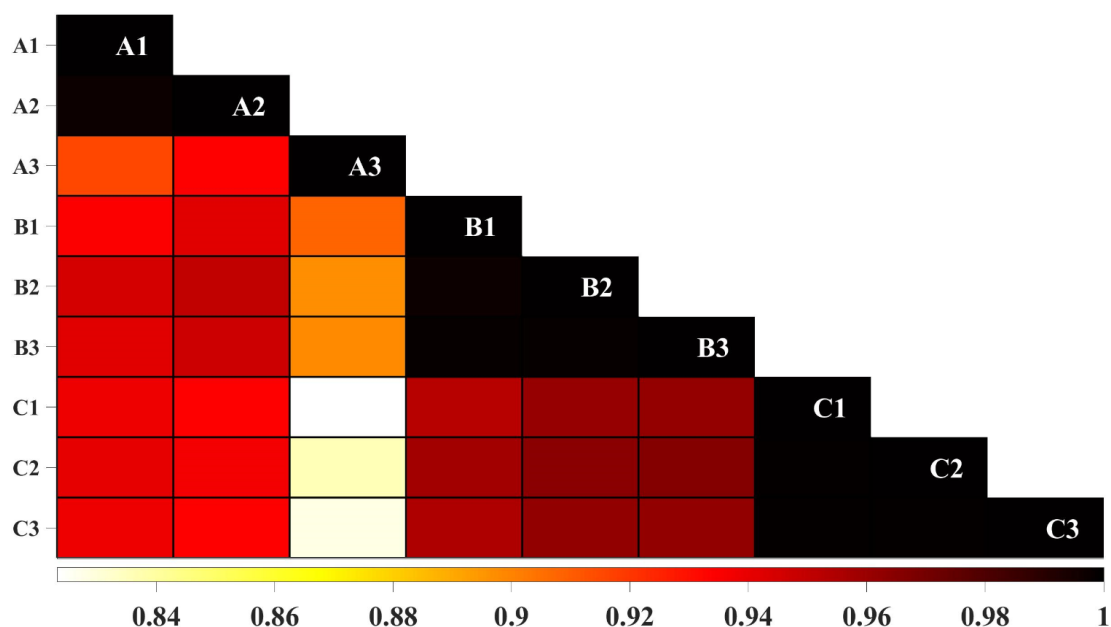


图. 样品相关系数半图，下三角热图。

9. 基因相关系数图 (GeneCorrelation)

9.1 单表格自相关 (1 Table)

用以计算单表格内，每行之间的相关系数，形成热图。

在 GeneCorrelation 面板的 1 Table 面板中，“Excel Path”填写组学的来源数据；“No. of Sheet in Excel”是填写获取 Excel 中哪张表；“Data Column”是获取具体哪些列；“GeneName Column”是哪一列是行名称列。输入表格中每两个基因（每两行）之间的表达量都会做相关系数，采用 Distance 中选择的参数进行计算。

“Sort Column”表示热图展示的时候，接近的样品是否要重新排序，显示在一起；“Show Label”表示是否在热图旁显示基因和样品名称。“Export Result”是将计算得到的相关系数结果（矩阵和基因 ID 配对）输出到文本中。

The screenshot shows the 'oppHeatmap' application window. The 'HeatMap' tab is selected. The interface includes the following fields and controls:

- Excel Path:** D:\Data\Data.xlsx (with an 'Open' button)
- No. of Sheet in Excel:** 2 (with a 'Plot' button)
- Data Column:** 2:10
- GeneName Column:** 1
- Distance:** Pearson (dropdown menu)
- Sort Column:** ☒ (checked)
- Show Label:** ☐ (unchecked)
- Export Result:** ☐ (unchecked)

On the right side, there is a sidebar with '1 Table' and '2 Table' options, where '2 Table' is currently selected.

图. 基因相关系数图的参数界面，单表格。

输出结果的行列对应的颜色值是特定行与列之间的相关系数。文本结果可以通过点选“Export Result”输出。

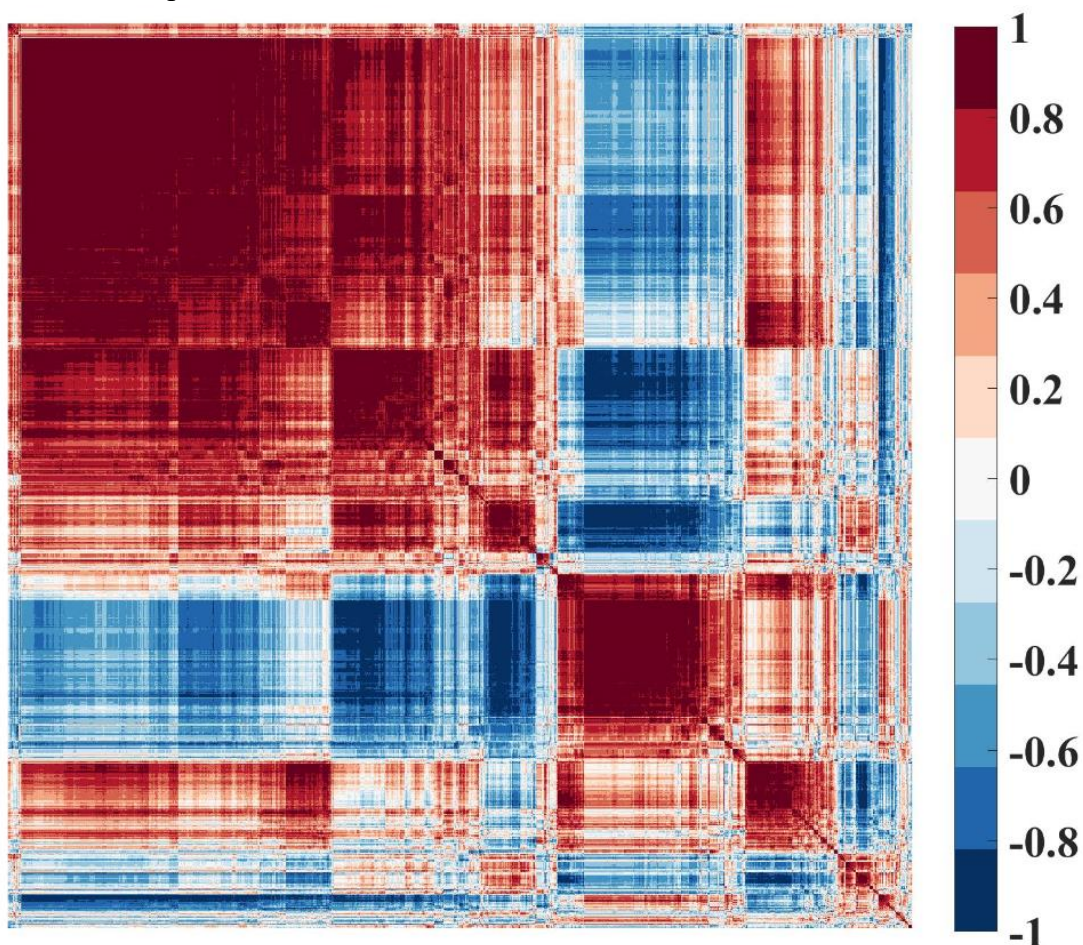


图. 单表格基因相关系数图的输出案例。

9.2 单表格自相关 (2 Table)

用以计算 1 张表格内的某行与另一张表格中的某行之间的相关系数，形成热图。

在 GeneCorrelation 面板的 2 Table 面板中，“Excel Path”填写组学的来源数据；“No. of Sheet1”是填写第 1 个 Excel 的 Sheet 名；“No. of Sheet2”是填写第 2 个 Excel 的 Sheet 名；“Data Column”是获取具体哪些列；“GeneName Column”是哪一列是行名称列。输入表格中每两个基因（每两行）之间的表达量都会做相关系数，采用 Distance 中选择的参数进行计算。

“Sort Column”表示热图展示的时候，接近的样品是否要重新排序，显示在一起；“Show Label”表示是否在热图旁显示基因和样品名称。“Export Result”是将计算得到的相关系数结果（矩阵和基因 ID 配对）输出到文本中。

图. 基因相关系数图的参数界面，双表格。

输出结果的行列对应的颜色值是特定行与列之间的相关系数。文本结果可以通过点选“Export Result”输出。

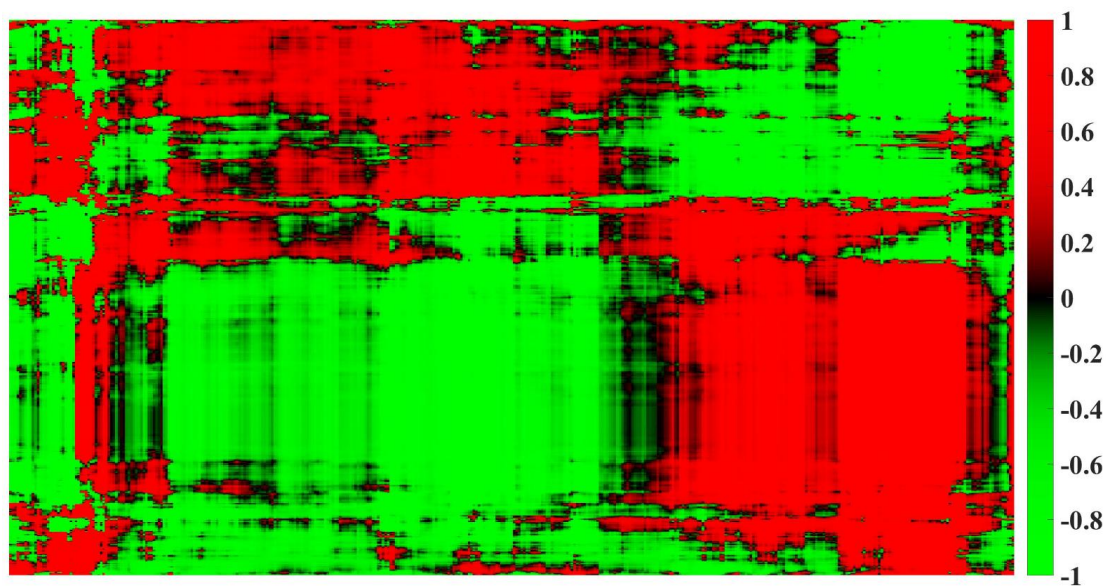


图. 双表格基因相关系数图的输出案例。

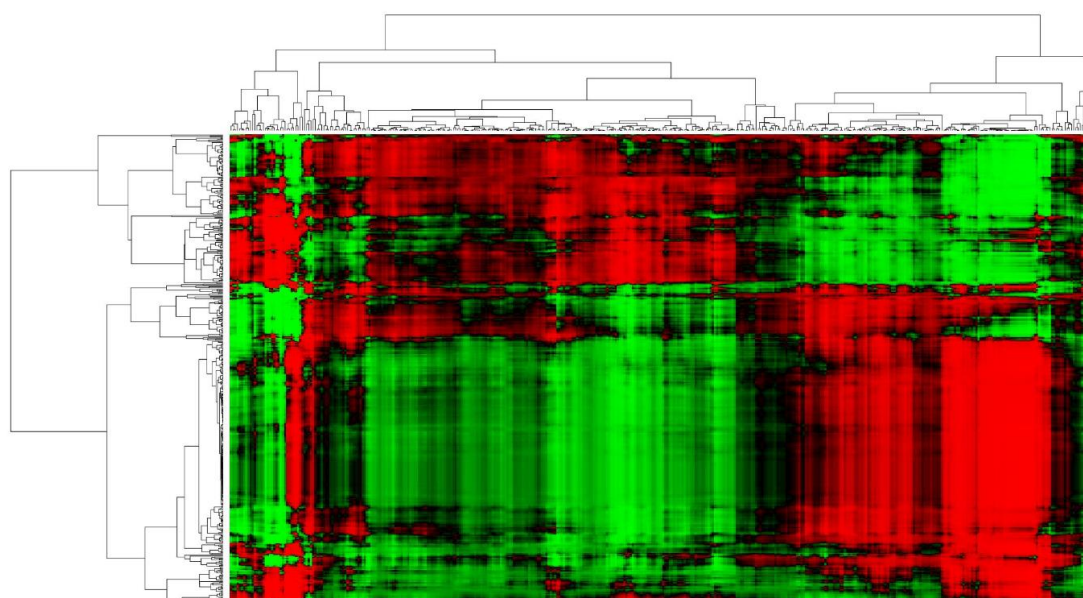


图. 双表格基因相关系数的分层聚类图的输出案例。

10. 环形热图 (Polar)

用极坐标的方式显示热图。

在 Polar 的面板中，“Excel Path”填写组学的来源数据；“No. of Sheet in Excel”是填写 Excel 的 Sheet 名；“Data Column”是获取具体哪些列；“GeneName Column”是哪一列是行名称列；“Replicates Number”表示每组样品的重复数。“Structure”面板表示形成环状热图的结构；“Genes”面板表示对基因文字的修改；“Groups”面板表示组内样品的颜色、字体和宽度等修改。

“Grid Color”表示热图中网格线的颜色设置；“Offset”表示文字标签与图之间的间隔。

oppHeatmap

HeatMap

MicroPlatePlot

HierarchyClustering

TreeMap

SampleCorrelation

GeneCorrelation

Polar

Heatmap

Excel Path

D:\Data\Data.xlsx

Open

No. of Sheet in Excel

1

Plot

Data Column

2:10

Replicates Number

3,3,3

GeneName Column

1

Structure

Full Circle

3/4 Circle

Semi Circle

Genes

Font Size

12

Font Bold

☒

Inner Space

0.4

Groups

Font Size

24

Font Bold

☒

Bar Width

10

Offset

1.5

Grid Color

1,1,1

图. 环形热图参数界面。

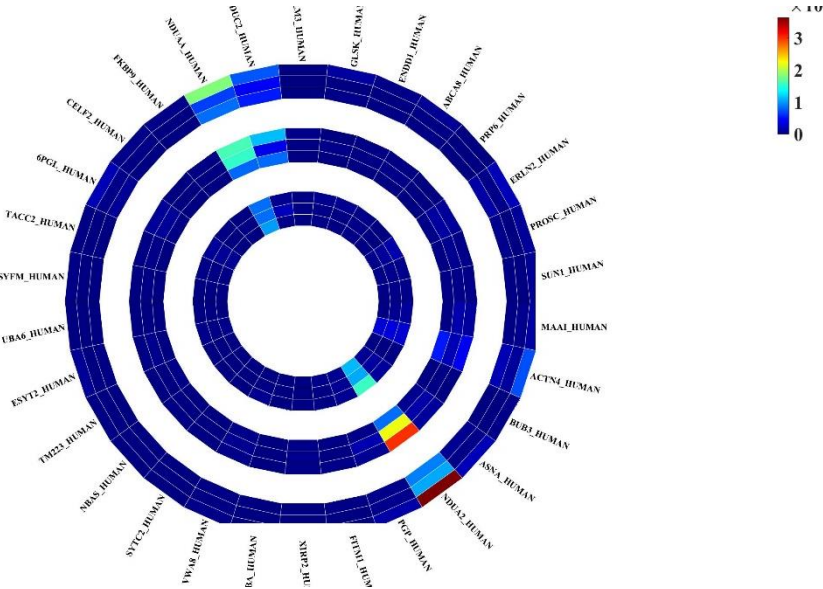


图. 全环形热图案例。

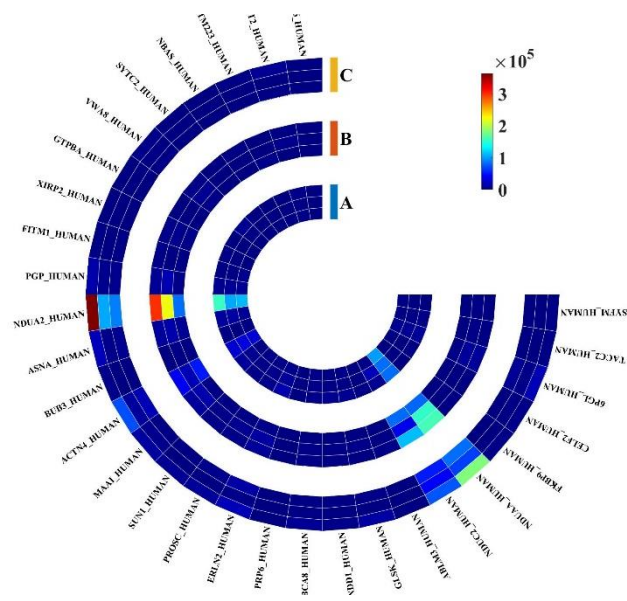


图. 3/4 形热图案例。

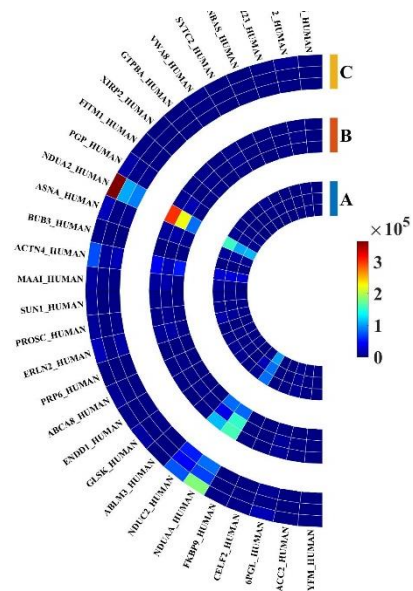


图. 半热图案例。