The instruction of heatmap graphics rendering software

Abbreviation: oppHeatmap
Version 1.0

Designer: Zhang Yang

MENU

1. Introduction

Omics Pilot Platform of Heatmap (oppHeatmap) is a kind of heatmap graphics rendering software based on the architecture of MATLAB AppDesigner, aiming to diagram the heatmaps of omics data with a graphical user interface (GUI). oppHeatmap is available for plotting ordinary heatmaps, bilateral hierarchical clustering heatmaps, treemaps, microplates, sample correlation coefficient diagrams (including matrix graphs, upper triangle graphs and lower triangle graphs), gene correlation coefficient diagrams (one table and two tables) and loop heatmaps. oppHeatmap could support the modification of borders, fonts, and colors for consummating the final heatmaps. oppHeatmap could straightforwardly read the data resources from the files of Microsoft Excel to generate particular heatmaps in the environment of MATLAB, and these graphics could be stored into vector diagrams through the command SaveAs.

2. GUI of OPP-Heatmap

Load the program package: Open the MATLAB. In the panel of "APP", click "install APP". And then, select "oppHeatmap.mlappinstall" to load the package of oppHeatmap into MATLAB.

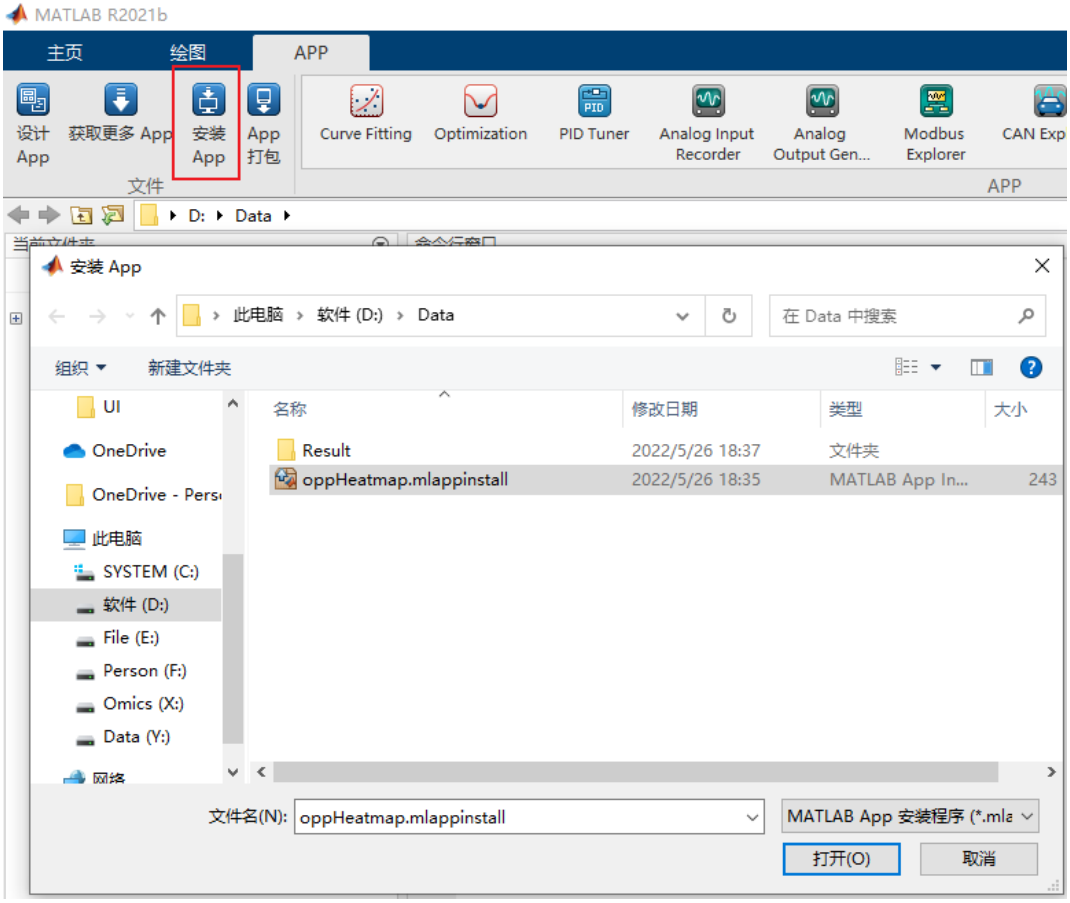

Figure 1. The installation procedure of oppHeatmap.

In the operation interface of MATLAB, execute the command “oppHeatmap”.

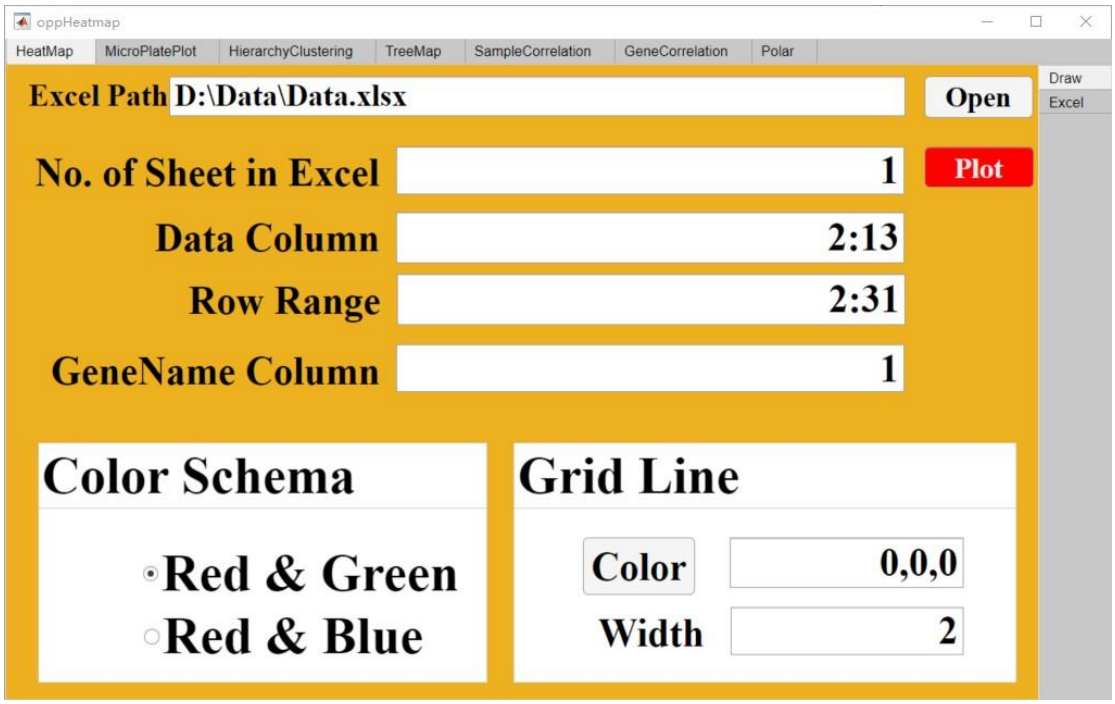

Figure 2. The main interface of oppHeatmap.

Enter the main interface of oppHeatmap. The main interface of oppHeatmap is consist of seven subinterfaces containing Heatmap, MicroPlatePlot, HierarchyClustering, TreeMap, Sample Correlation, Gene Correlation and Polar.

3. Detailed data of cases (Data.xlsx)

Table 1. The details of the data demonstrated in oppHeatmap

Names	Description	Amounts of Samples
Heatmap	Partial outcomes of proteome analysis in previous study	4 samples (including A, B, C and D) with each 3 biological replicates
Proteome	Full outcomes of proteome analysis from *Radices Trichosanthis*	3 samples (including A, B and C) with each 3 biological replicates
TreeMap	Data for displaying function of TreeMap	
T1	Data from previous transcriptome analyzing case	5 samples with each 1 biological replicate
T2	Data from previous proteome analyzing case	5 samples with each 1 biological replicate

4. Common Heatmap

4.1 Draw Common Heatmap

Select the interface of “Heatmap”. In the panel of Heatmap, click “Open” to select the data of Omics sources in the files of Microsoft Excel. The general format is as follows: the column of ID numbers (character string) and columns of each sample expression quantity including the data of biological repetition. For instance, there were four samples containing A, B, C and D together with three biological repeats for each sample.

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
MAAI_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

Figure 3. Import the original data of Omics.

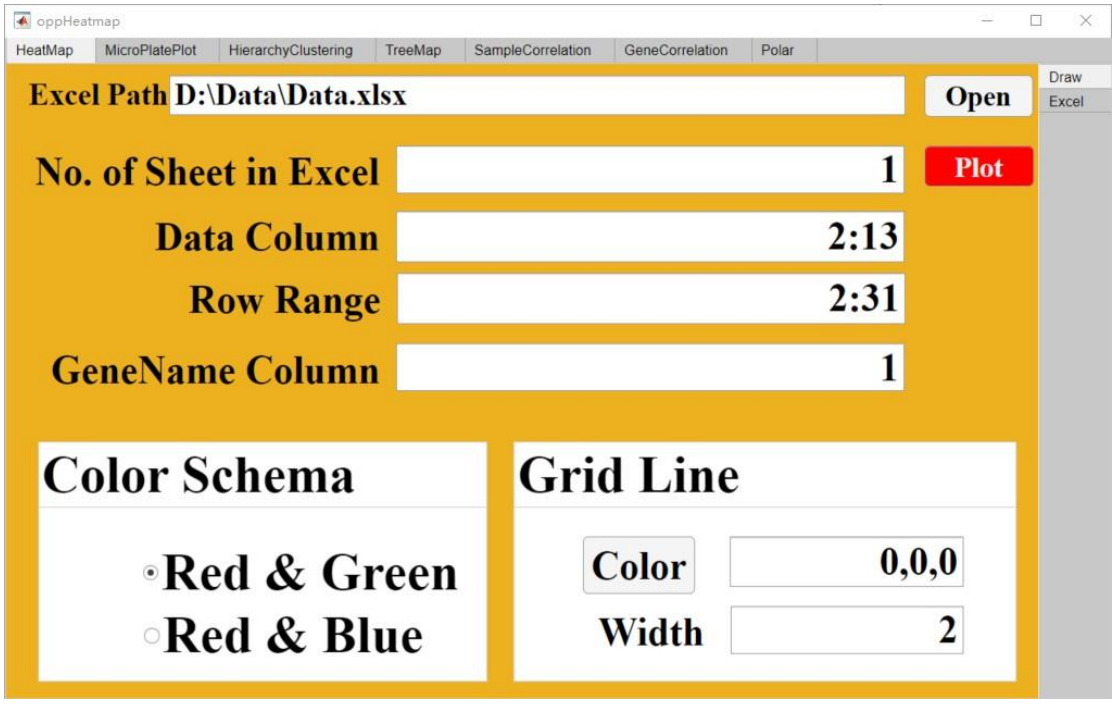

Figure 4. The parameter GUI of operating common heatmap.

In the panel of “No. of Sheet in Excel”, choose the candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular columns as input data. The type-in format is the numbers of first and last columns connected with a doppelpunkt colon “:”. In the panel of “Row Range”, choose the specific lines to plot the heatmaps. In the panel of “yLabel Column”, decide the definite column number of source text for the vertical axis in the heatmap. In the panel of “Grid Line”, decide to add grid line in the heatmap. Click the button “Color” to activate the palette, and then the Red Green Blue (RGB) values of selected color will be displayed in the box of right side. In the inputbox of the right side of “Width”, type in numbers to adjust the width of grid lines. In the panel of “Color Schama”, “Red & Green” presents the heatmap displayed in the color scheme of red and green. The color red reveals upregulation, and the color green reveals downregulation. The more significantly the alterations exist, the deeper the colors will be demonstrated in the heatmap. Click the button “Plot” to diagram the final heatmap.

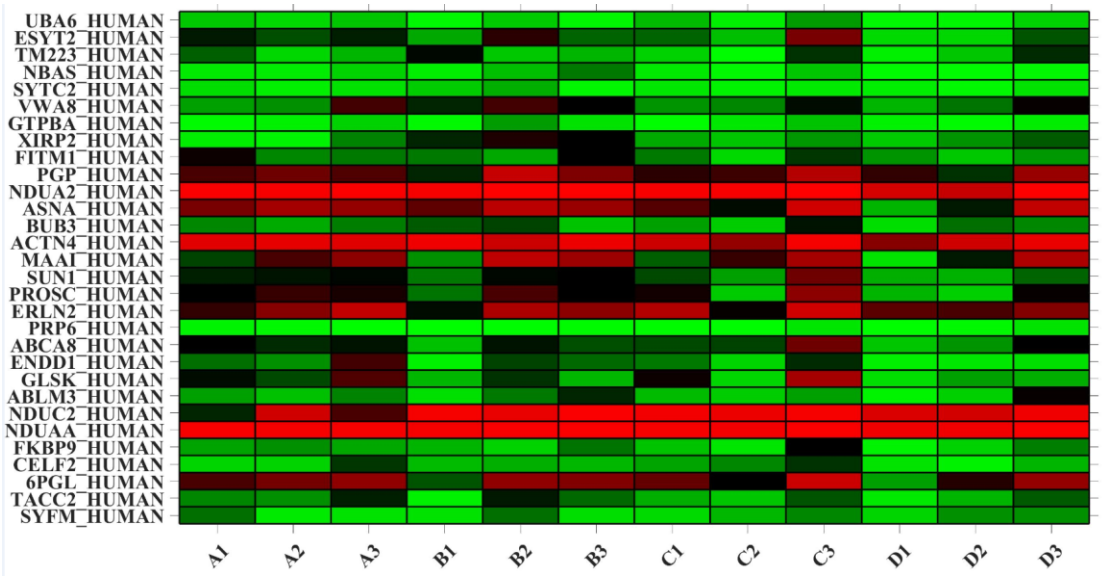

Figure 5. Common heatmap in the color schema of *Red & Green*.

In the panel of "Color Schama", if selecting "Red & Blue", it will present the heatmap displayed in the color scheme of red and blue. The color red reveals upregulation, and the color blue reveals downregulation. The more significantly the alterations exist, the deeper the colors will be demonstrated in the heatmap. In the MALTAB, type in "colormap redgreencmap" or "colormap redbluecmap" could also help the users to convert the color schema of ***Red & Green*** or ***Red & Blue.*** These two schemas of color representation were consistent throughout the software.

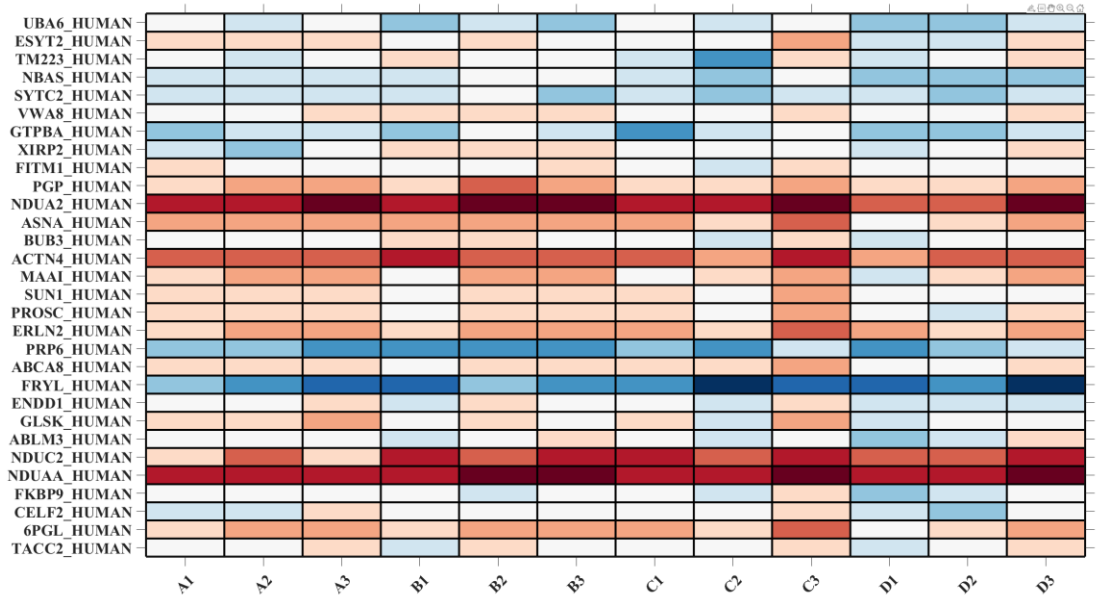

Figure 6. Common heatmap in the color schema of *Red & Blue*.

4.2 Plot to Excel

In the oppHeatmap, the data of input tables could be imported into Microsoft Excel as the format of heatmaps, which replaces the conditional formatting function of diagraming heatmaps from data in Microsoft Excel. Select the function "Row Standardize", the data of input tables could be respectively row standardized by z-score, which means that the row data would be standardized in the extent (-1,1) to optimize visual effects of heatmaps.

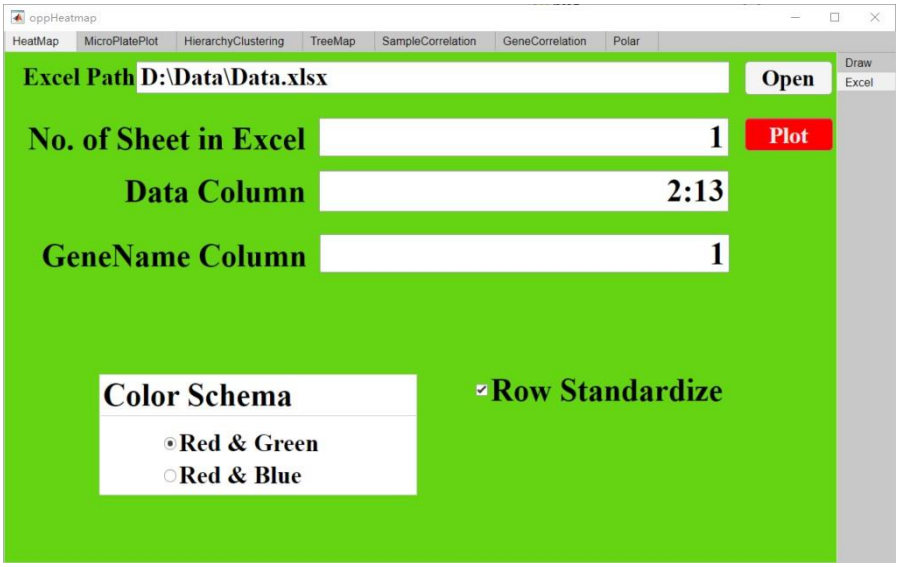

Figure 7. The interface of drawing heatmaps for exporting heatmaps in the format files of Microsoft excel.

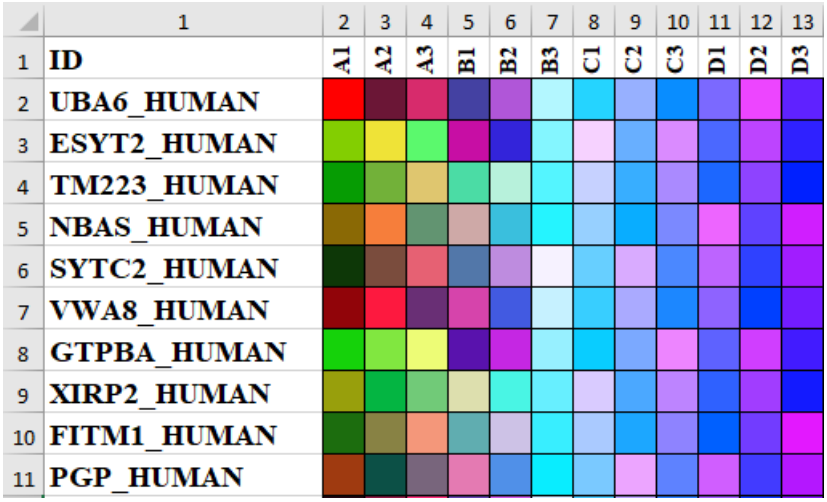

Figure 8. The exported visualization of heatmap in the format files of Microsoft excel.

5. MicroPlatePlot

Select the interface of “MicroPlatePlot”, click “Open” to select the data of Omics sources in the files of Microsoft Excel. In the panel of “No. of Sheet in Excel”, choose the candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular columns as input data. The type-in format is the numbers of first and last columns connected with a doppelpunkt colon “:”. In the panel of “Name Column”, choose the source text for the vertical axis in the microplate. Click the button “Plot” to diagram the final MicroPlatePlot.

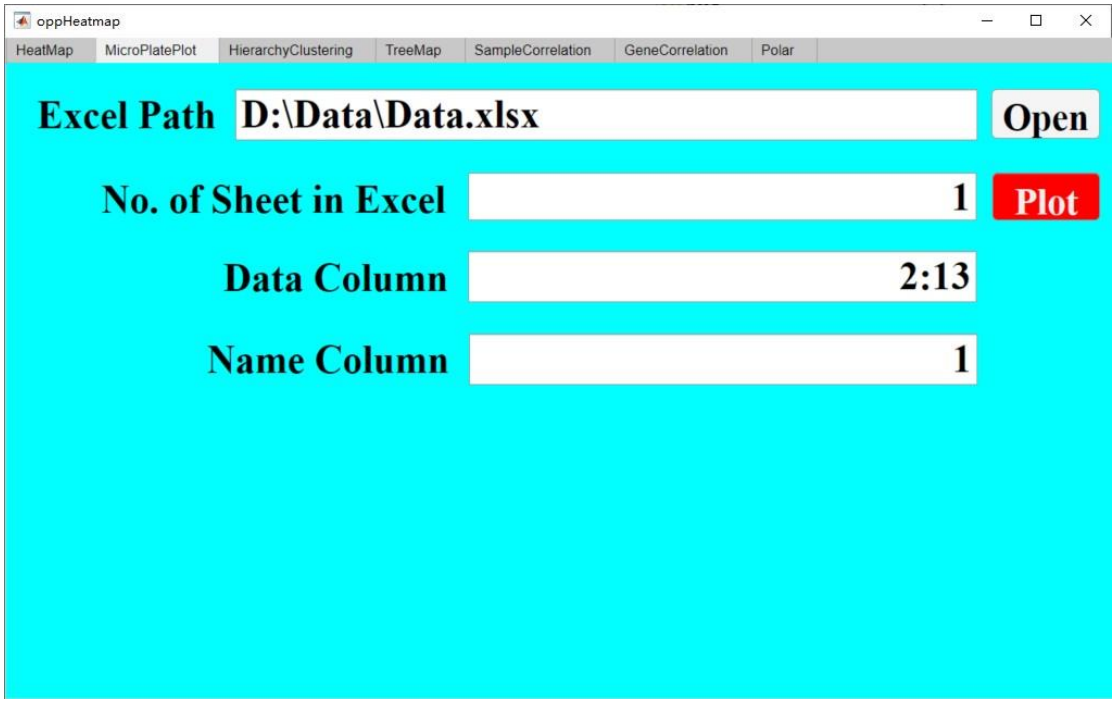

Figure 9. The parameter interface of MicroPlatePlot.

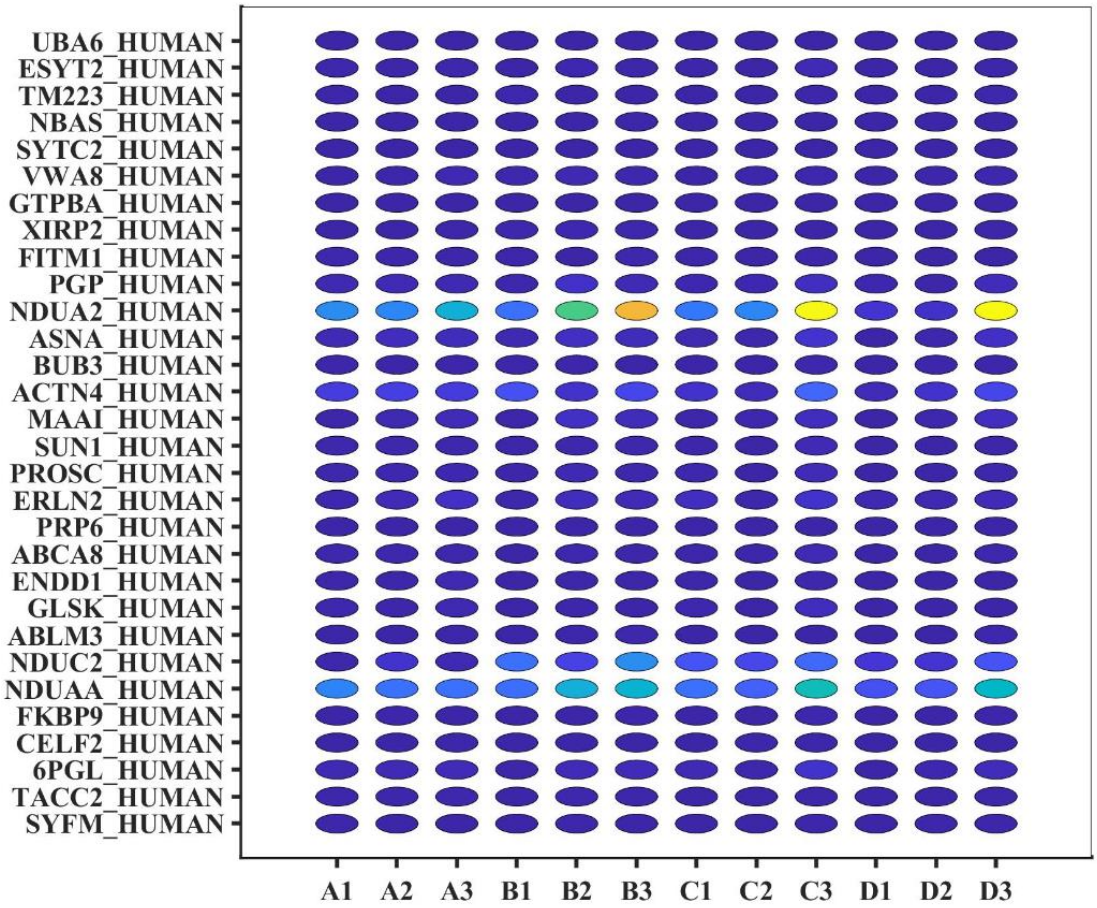

Figure 10. The outcomes displayed by MicroPlatePlot. The depth of color was correlated to the expression quantity in the original data.

6. Hierarchy Clustering

Select the interface of "Bilateral Hierarchy Clustering", click "Open" to select the data of Omics sources in the files of Microsoft Excel. In the panel of "No. of Sheet in Excel", choose the candidate sheet according to the need of user. In the panel of "Data Column", choose the particular columns as input data. The type-in format is the numbers of first and last columns connected with a doppelpunkt colon ":". In the panel of "GeneName Column", choose the source text for the vertical axis in the heatmap. In the interface of "Calculation", the panels of "Distance for Row" and "Distance for Column" are the select boxes to choose the models for calculating the formula of distance in clustering calculation, usually through the protocol of Euclidean distance or Pearson correlation coefficient. In the panel of "Linkage", select the appropriate formula for merging multiple distances. In the panel of "Color Schama", "Red & Green" and "Red & Blue" respectively present the heatmap displayed in the color scheme of red and green or red and blue. The color red reveals upregulation, and the color green or blue reveals downregulation. The more significantly the alterations exist, the deeper

the colors will be demonstrated in the heatmap. In the panel of “Row Standardize”, choose whether standardize the data by row before displaying the heatmap. In the panel of “Export Result”, decide whether to export the results of clustering rearrangement into new file as HC.txt in current path. Click the button “Plot” to diagram the final heatmap.

Figure 11. The parameter interface of Bilateral Hierarchy Clustering.

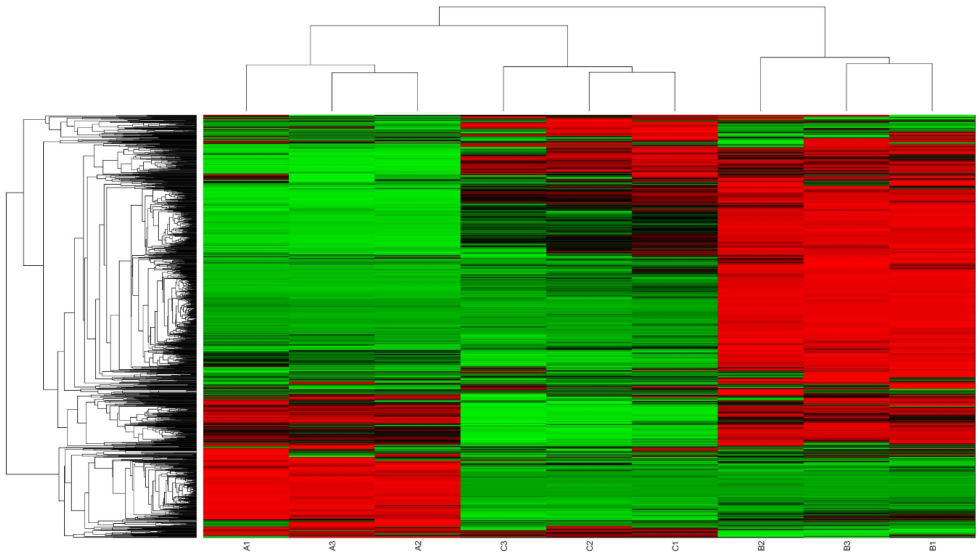

Figure 12. Bilateral Hierarchy Clustering heatmap in the color schema of *Red & Green*.

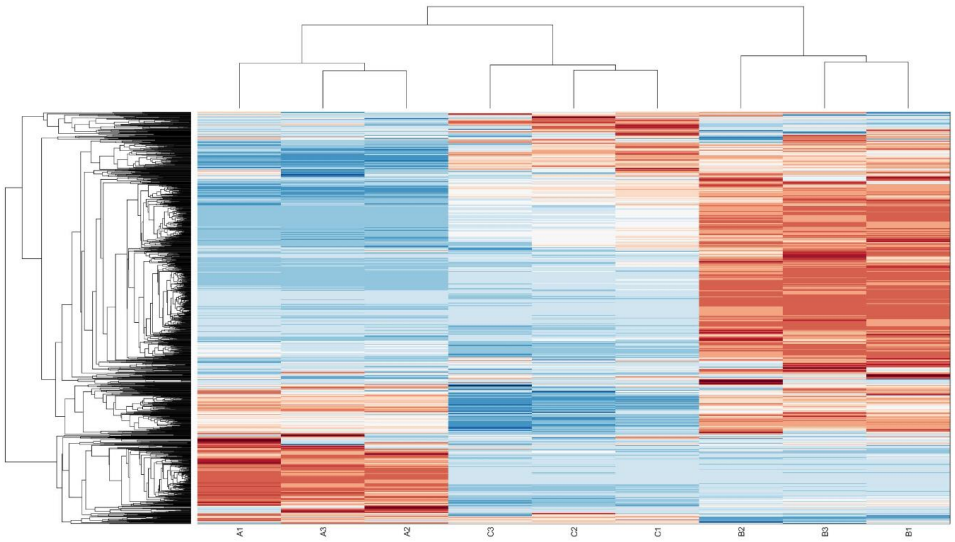

Figure 13. Standardized Bilateral Hierarchy Clustering heatmap in the color schema of *Red & Blue*.

7. TreeMap

Select the interface of “TreeMap”, click “Open” to select the data of Omics sources in the files of Microsoft Excel. In the panel of “Single TreeMap”, choose to diagram single treemaps while in the panel of “Nested TreeMap”, choose to draw multiple levels of treemaps. In the panel of “Excel Path”, select the source text of omics analysis in the file of Microsoft Excel. In the panel of “No. of Sheet in Excel”, choose the candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular column number with a single number. In the panel of “Name Column”, choose the source text displayed in the treemap. Likewise, the operation in the panel of “Nested TreeMap” is very similar except for the selection of data columns. In the panel of “Data Column”, choose the particular columns as input data. The type-in format is the numbers of first and last columns connected with a doppelpunkt colon “:”. When pick the “Add Color”, the treemaps will be colored, otherwise, the treemaps will be monochrome when not pick the “Add Color”. In the panel of “Font Size”, adjust the font size in the grid lines. In the panel of “Font Color”, alter the font color in the grid lines. Click the button “Plot” to diagram the final TreeMap.

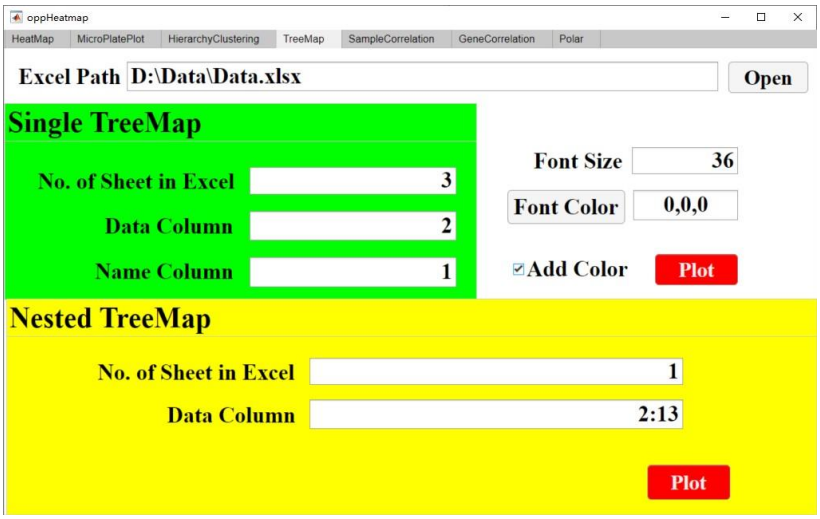

Figure 14. The parameter interface of TreeMaps.

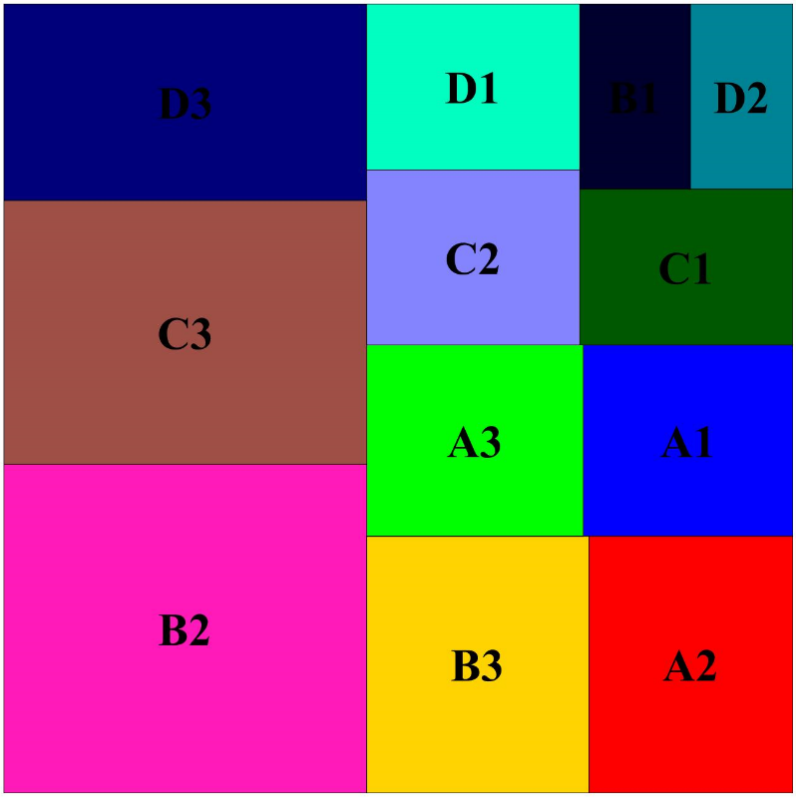

Figure 15. Single TreeMap in the color schema of colored.

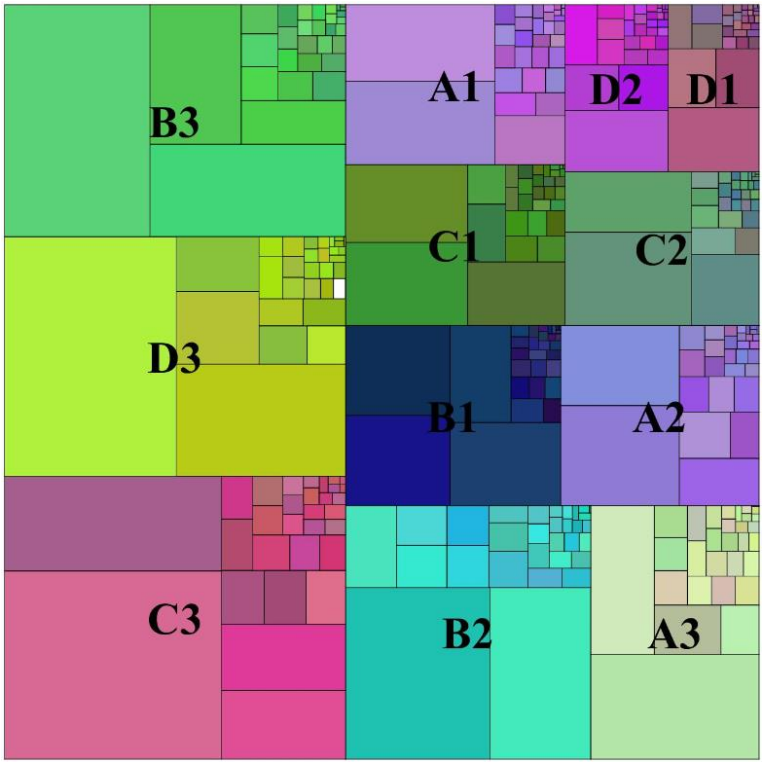

Figure 16. Nested TreeMap in the color schema of colored.

8. Sample Correlation

8.1 Total graphs of correlation coefficient

Used to generate the heatmaps through calculating the correlation coefficient between each column in the single table.

In the “Full” interface of “SampleCorrelation”, click “Open” to select the data of Omics sources in the files of Microsoft Excel. In the panel of “No. of Sheet in Excel”, choose the candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular columns as input data. In the panel of “Replicates Column”, type in the numbers of biological repeats for each sample. The type-in format is the numbers of biological repeats for each sample connected with commas “,”. The correlation coefficients of expression quantity in each two samples will be calculated. The select box of “Distance” was designed to select the method of calculating correlation coefficient.

In the panel of “Grid Line”, decide to visual effects of grid line in the heatmap. Click the button “Color” to activate the palette, and then the RGB values of selected color will be displayed in the box of right side. In the inputbox at the right side of “Width”, type in numbers to adjust the width of grid lines. Click “Sort Column”, decide whether to rearrange the order of approximate samples together when displaying the heatmaps. In the panel of “Add SampleLine”, decide whether to exhibit the barcode of

samples. In the panel of “Export Result”, decide whether to outout the results of calculated correlation coefficient into the text.

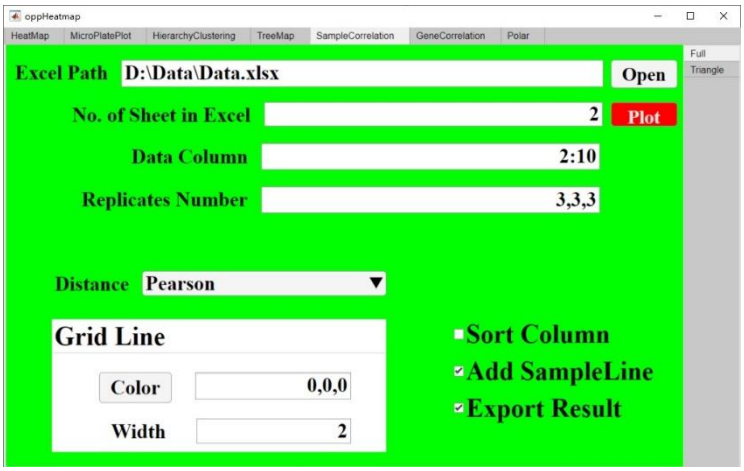

Figure 17. The parameter interface of sample correlation coefficient total graphs.

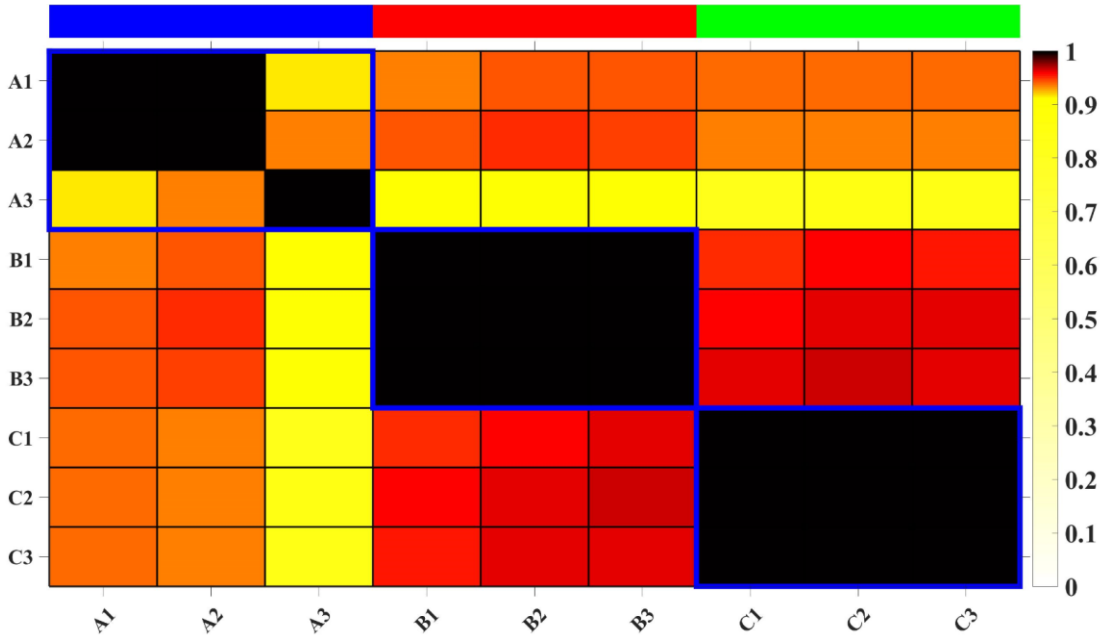

Figure 18. The visual results of sample correlation coefficient total graph.

The blue rectangles represented the correlation coefficients between repeated samples. The users could modify the color contrast of heatmap by right clicking the ColorBar in the right side.

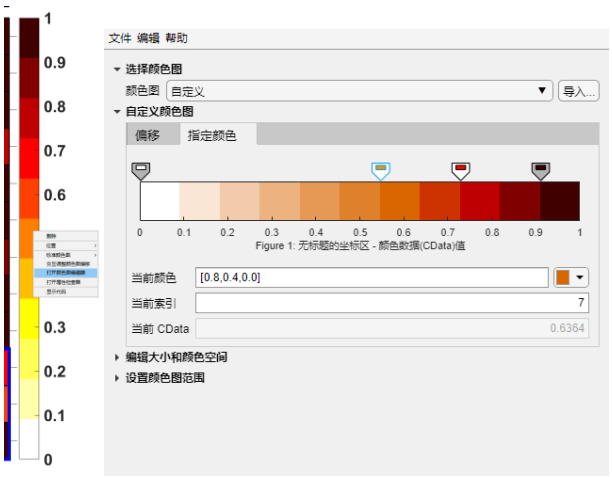

Figure 19. Color Contrast Selector.

8.2 Semi-graphs of correlation coefficient

In the “Triangles” interface of “SampleCorrelation”, click “Open” to select the data of Omics sources in the files of Microsoft Excel. In the panel of “No. of Sheet in Excel”, choose the candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular columns as input data. In the panel of “Replicates Column”, type in the numbers of biological repeats for each sample. The type-in format is the numbers of biological repeats for each sample connected with commas “,”. The correlation coefficients of expression quantity in each two samples will be calculated. The select box of “Distance” was designed to select the method of calculating correlation coefficient.

In the panel of “Grid Line”, decide to visual effects of grid line in the heatmap. Click the button “Color” to activate the palette, and then the RGB values of selected color will be displayed in the box of right side. In the inputbox at the right side of “Width”, type in numbers to adjust the width of grid lines. In the panel of “Orientation”, select the “Upper Triangle” to display the final heatmap in the shape of upper triangle or the “Lower Triangle” to display the final heatmap in the shape of lower triangle. Click “Sort Column”, decide whether to rearrange the order of approximate samples together when displaying the heatmaps. In the panel of “Add SampleLine”, decide whether to exhibit the barcode of samples. In the panel of “Export Result”, decide whether to output the results of calculated correlation coefficient into the text.

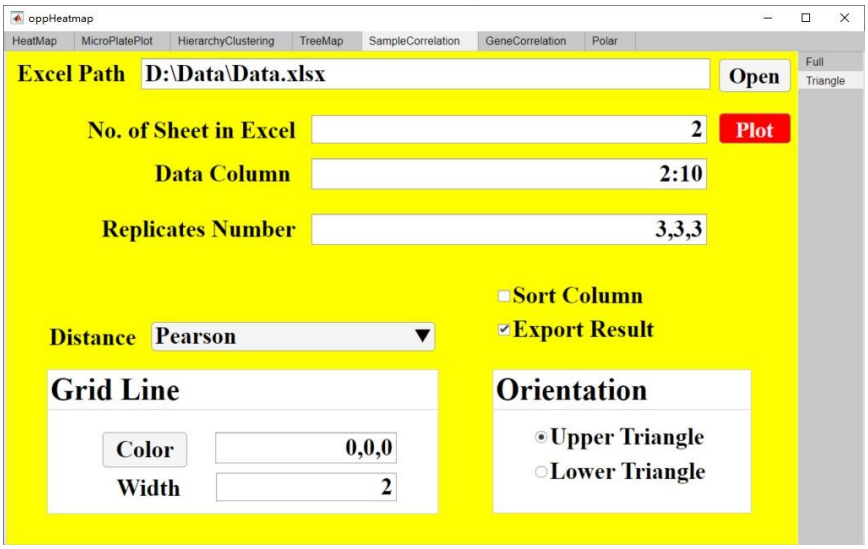

Figure 20. The parameter interface of sample correlation coefficient semi-graphs.

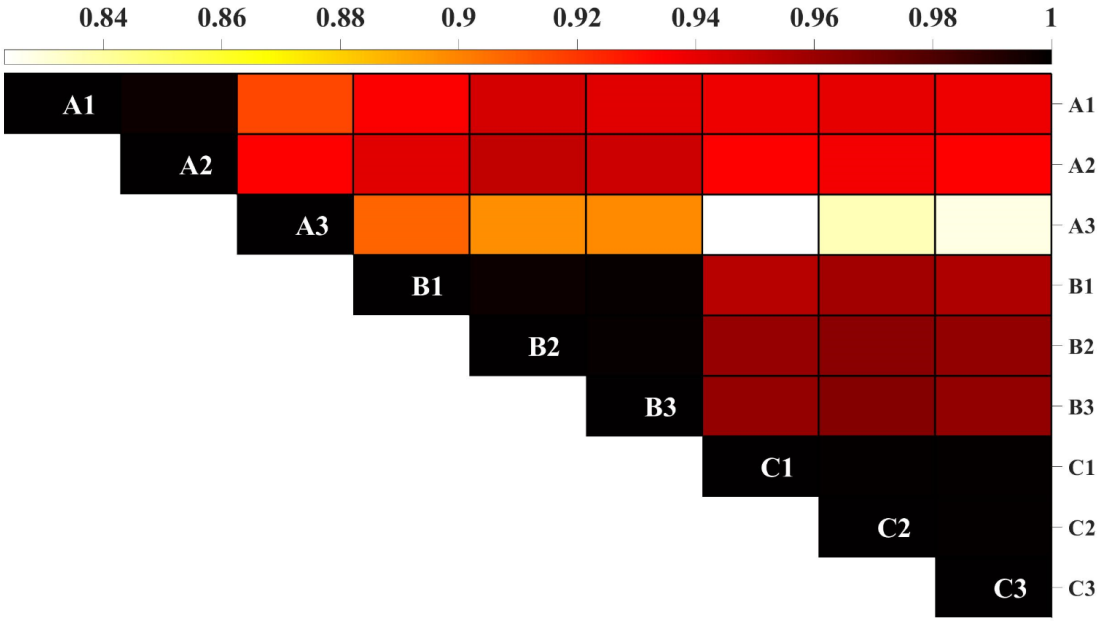

Figure 21. The visual results of sample correlation coefficient semi-graph as upper triangle.

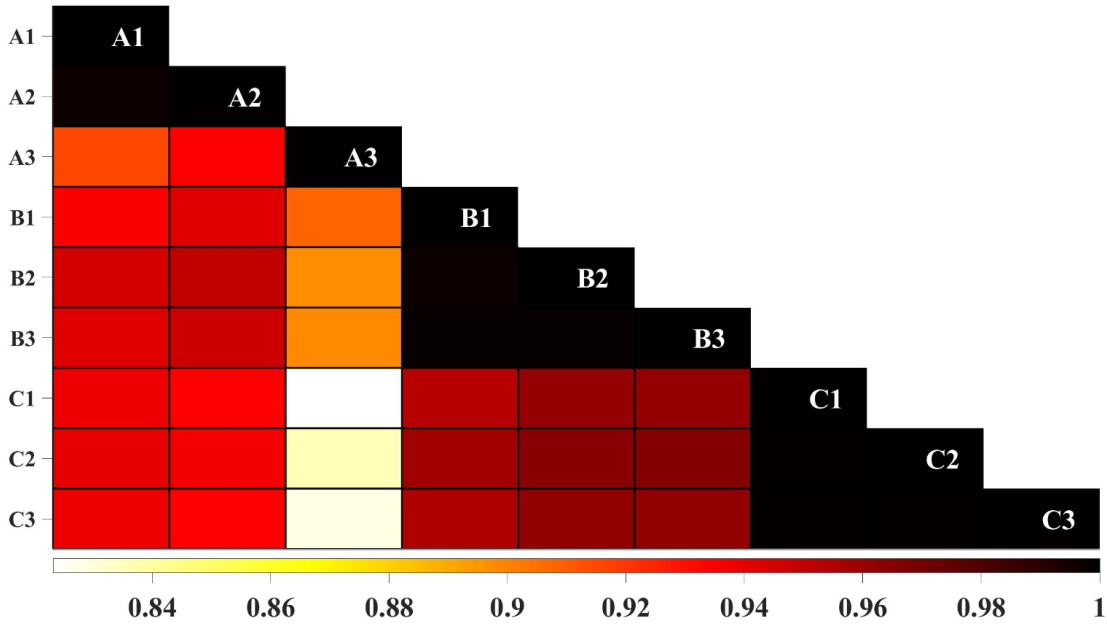

Figure 22. The visual results of sample correlation coefficient semi-graph as lower triangle.

9. Gene Correlation

9.1 Correlation between rows in one single table

Used to generate the heatmaps through calculating the correlation coefficient between each row in the single table.

In the "1 Table" interface of "GeneCorrelation", click "Open" to select the data of Omics sources in the files of Microsoft Excel. In the panel of "No. of Sheet in Excel", choose the candidate sheet according to the need of user. In the panel of "Data Column", choose the particular columns as input data. In the panel of "GeneName Column", choose a particular column as the source of gene names. The correlation coefficients of expression quantity in each two gene (two rows) will be calculated. The select box of "Distance" was designed to select the method of calculating correlation coefficient. Click "Sort Column", decide whether to rearrange the order of approximate gene expression together when displaying the heatmaps. Click the select box of "Show Label", decide whether to show the labels of genes near the heatmaps. In the panel of "Export Result", decide whether to output the results of calculated correlation coefficient (matrix paired with the gene IDs) into the text.

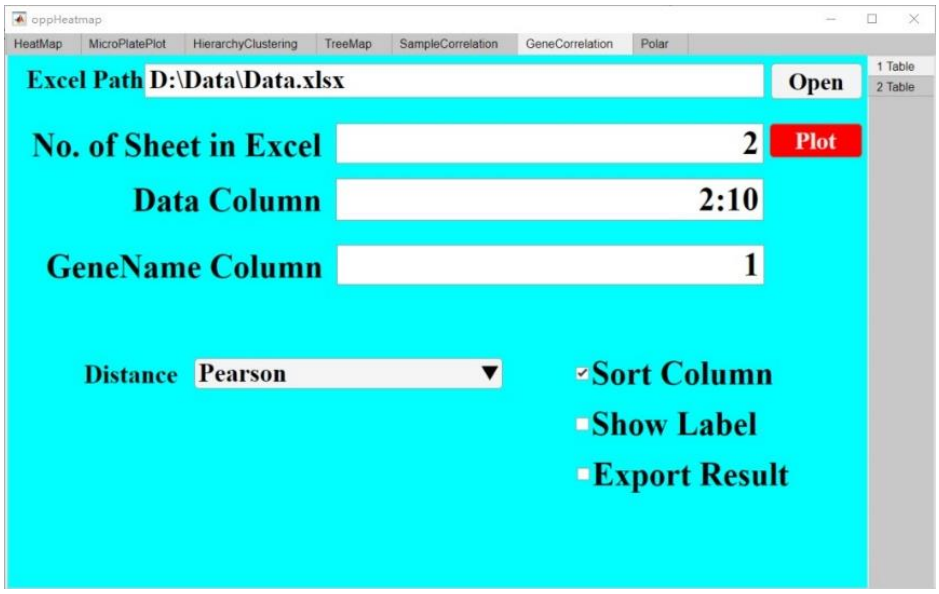

Figure 23. The parameter interface of gene correlation coefficient heatmap in one single table.

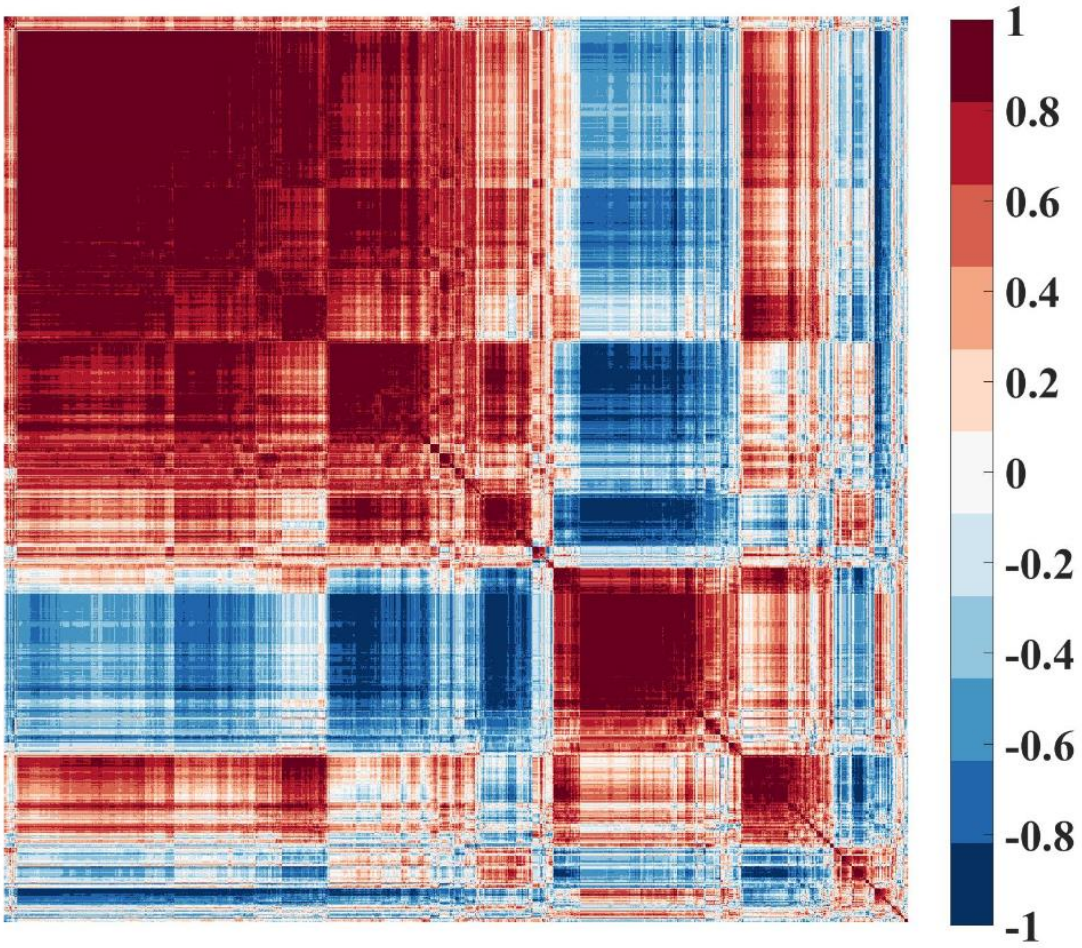

Figure 24. The visual results of gene correlation coefficient heatmap in one single table.

The color values of selected rows and columns correspond to the output results of correlation coefficients between specific rows and columns. The text results of calculated correlation coefficient can be exported by clicking the button “Export Result”.

9.2 Correlation between selected each row across two tables

Used to generate the heatmaps through calculating the correlation coefficient between selected each row across two tables.

In the “2 Table” interface of “GeneCorrelation”, click “Open” to select the data of Omics sources in the files of Microsoft Excel. In the panel of “No. of Sheet1”, choose the candidate sheet according to the need of user. In the panel of “No. of Sheet2”, choose the other candidate sheet according to the need of user. In the panel of “Data Column”, choose the particular columns in each sheet as input data. In the panel of “GeneName Column”, choose a particular column in each sheet as the text source of gene names. Click “Sort Column”, decide whether to rearrange the order of approximate gene expression together when displaying the heatmaps. Click the select box of “Show Label”, decide whether to show the labels of genes near the heatmaps. In the panel of “Export Result”, decide whether to output the results of calculated correlation coefficient (matrix paired with the gene IDs) into the text.

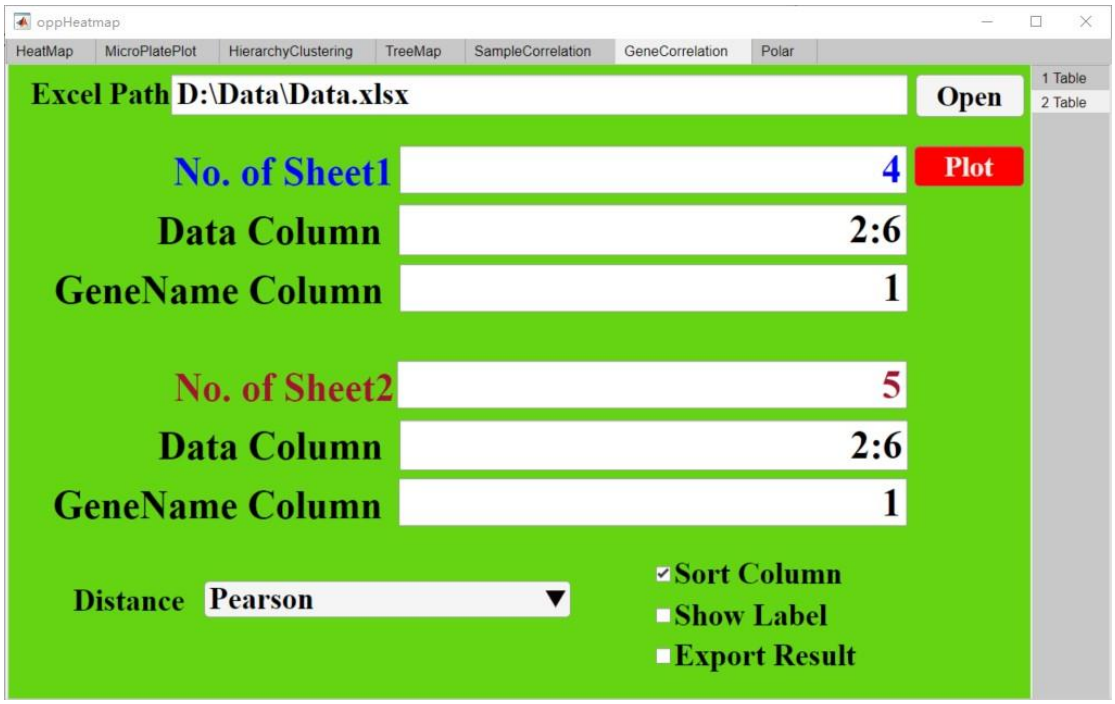

Figure 25. The parameter interface of gene correlation coefficient heatmap across two tables.

The color values of selected rows and columns correspond to the output results of correlation coefficients between specific rows and columns. The text results of calculated correlation coefficient can be exported by clicking the button “Export Result”.

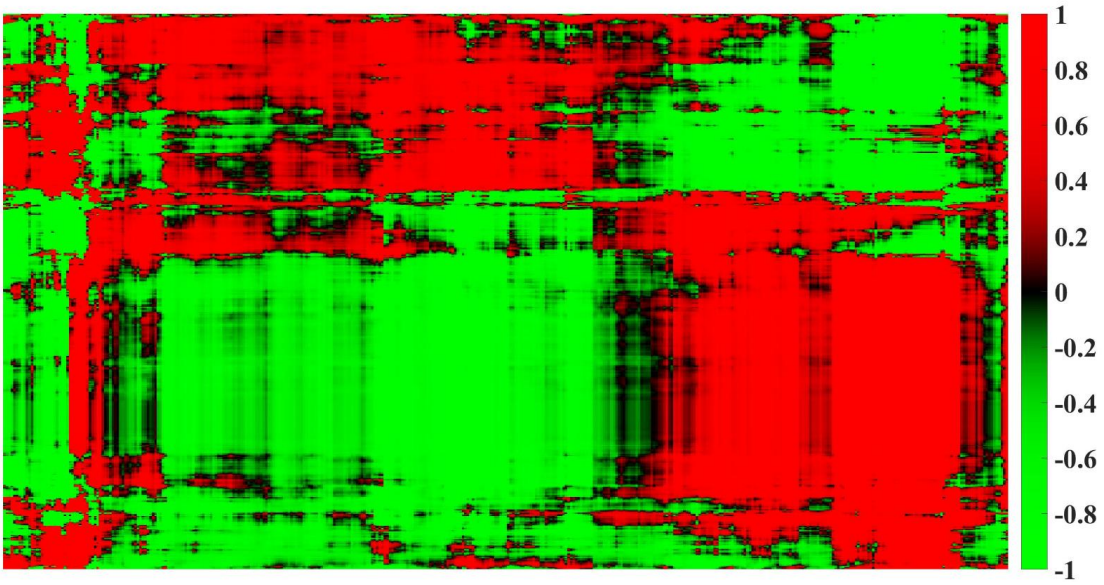

Figure 26. The visual results of gene correlation coefficient heatmap across two tables.

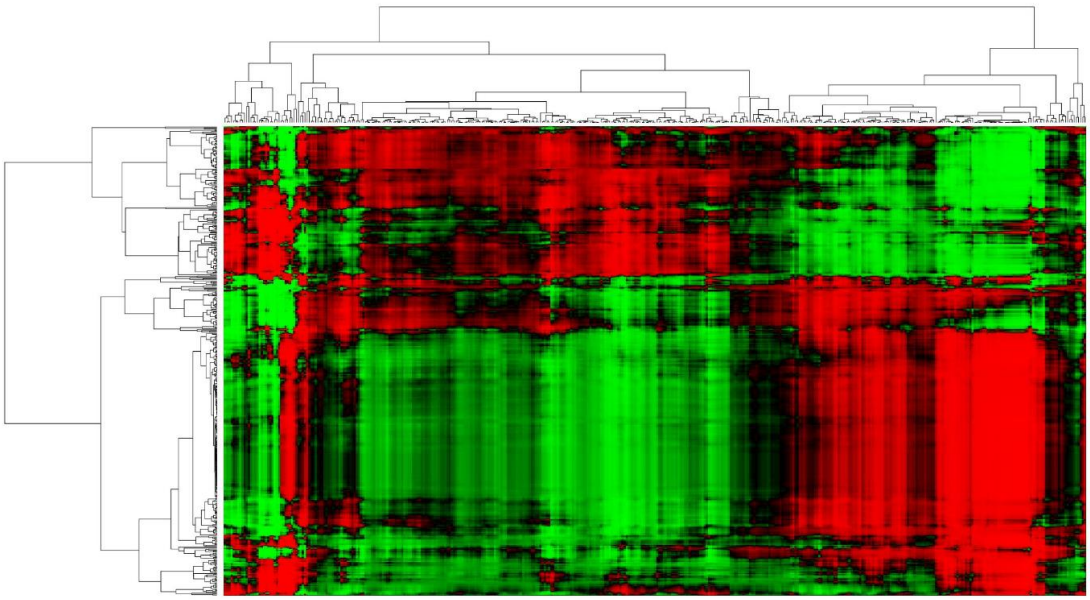

Figure 27. The hierarchy clustering visual results of gene correlation coefficient heatmap across two tables.

10. Polar

Used to generate the heatmaps in polar coordinates.

In the interface of "Polar", click "Open" to select the data of Omics sources in the files of Microsoft Excel. In the panel of "No. of Sheet in Excel", choose the candidate sheet according to the need of user. In the panel of "Data Column", choose the particular columns as input data. In the panel of "GeneName Column", choose a particular column as the source of gene names. In the panel of "Replicates Column", type in the numbers of biological repeats for each sample. In the panel of "Structure", decide the shape structure of the heatmaps. In the panel of "Genes", modify the font size, font bold and inner space of the Gene ID texts. In the panel of "Groups", revise the font size, font bold and bar width between groups. Click the button "Grid Color", set the grid color in the heatmaps. In the textbox of "Offset", type in numbers to adjust the distance between labels and figures.

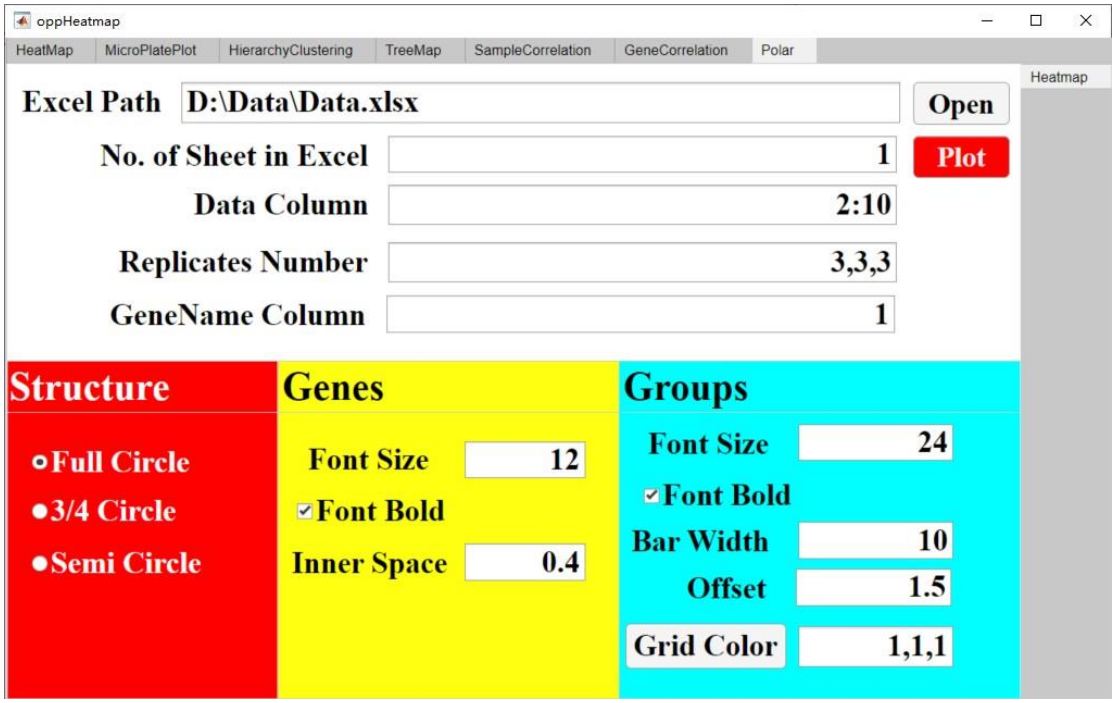

Figure 28. The parameter interface of polar heatmaps.

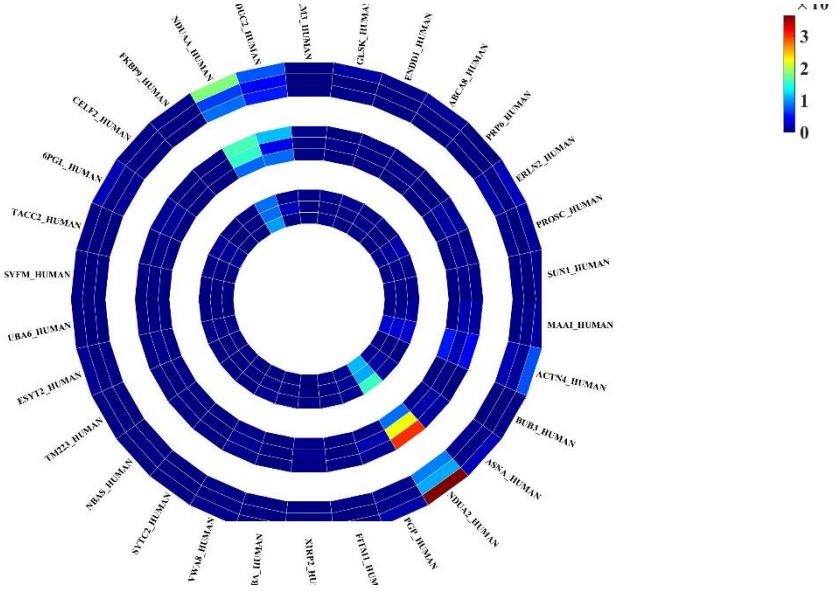

Figure 29. The visual results of polar heatmap in the shape structure of full circle.

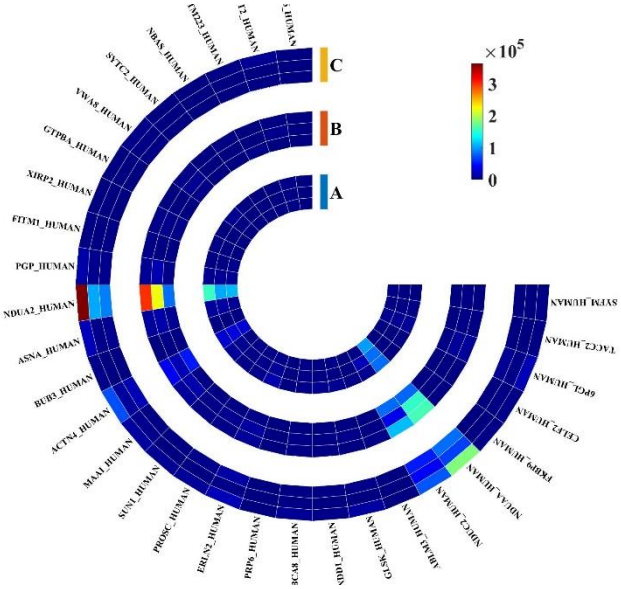

Figure 30. The visual results of polar heatmap in the shape structure of 3/4 circle.

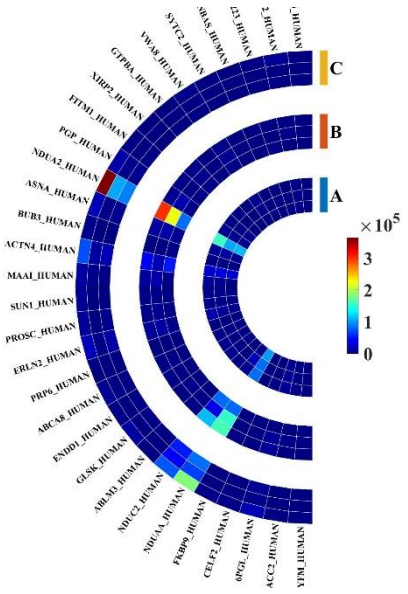

Figure 31. The visual results of polar heatmap in the shape structure of semi circle.