**Use of the** **“LabeledCombinedVolcanoPlot” DSP DA script**

**Intended use**

The Labeled Combined Volcano Plot DSP DA script was designed to create a combined volcano plot from protein data analysis and RNA data analysis for data GeoMx Spatial Proteogenomics data from a single study. It may also be used to overlay volcano plots from serial protein and rna scans in separate studies. It creates publication ready labeled volcano plots based on user inputs and statistical test results. A table of labeled genes is also created.

**Contents**

*Page 2*: Loading into the DSP DA

*Page 4*: Setting User Parameters

*Page 6*: Example Parameter Set-up

A graph with blue and yellow dots

Description automatically generated

*For more example graphs see page 6*

**Loading into the DSP DA**

The LabeledCombinedVolcanoPlot script requires an extra file input from DSP DA. After running a statistical test and creating a volcano plot in DSP DA for protein and RNA, the results file should be Exported as an .xlsx file. Statistical test results are under the dataset the test was run on. You have to export 2 files, one from the RNA data analysis and another from your protein data analysis.

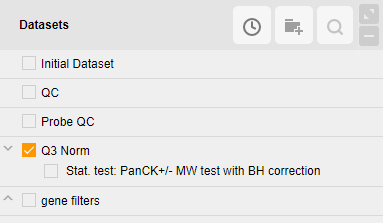
A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

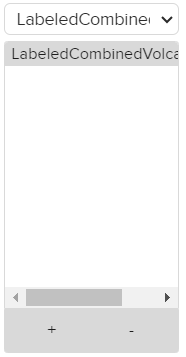
No need to change the format of the downloaded files. The script works for .xlsx files.

The LabeledCombinedVolcanoPlot.R file may be loaded into the custom scripts section of the DSP DA after you have a dataset processed and ready for analysis. To do so open the custom script section by clicking on the button shown below with the statistical test dataset unmarked:

Select the “Manage” tab to open the area to load and edit scripts:



In the Management tab to add a new script and adjust parameters, fill out and then scroll to the bottom of the page. Use the “+” button to add the LabeledCombinedVolcanoPlot.R file to the script:

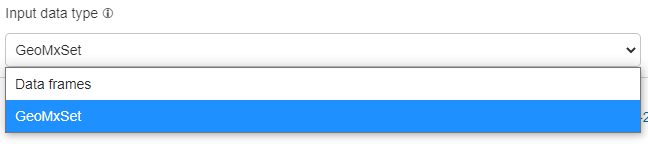


Use the “+” button to add the VOLCANO PLOT.xlsx files. Ensure the LabeledCombinedVolcanoPlot.R file is selected in the dropdown menu, indicated by star.

A screenshot of a cell phone

Description automatically generated

This script will work with both DataFrame and GeoMxSet inputs.



Once added the parameters of the script can be adjusted by editing the top lines in the script and hitting the “Save” button. You do **not** need to check the Create new dataset button.

A screenshot of a computer program

Description automatically generated

**Setting User Parameters:**

There are 21 settings that can be adjusted by the user at the top of the plug-in script. These include:

**Files**

1. *output\_format*: (String) Desired output format for the volcano plot figure.
   * Options: png, jpg, tiff, svg, pdf, bmp

**Labeling**

Labels from DSP DA volcano plot are not transferred to results file so must be user added

1. *plot\_title*: (String) Title for figure
2. *negative\_label*: (String) Matching negative (left) x-axis label to the volcano plot in DSP DA
3. *positive\_label*: (String) Matching positive (right) x-axis label to the volcano plot in DSP DA
4. *show\_legend:* (Boolean) Should a color legend be shown
5. *n\_genes:* (Numeric) Number of top targets by pvalue/fdr to label on figure. gene\_list overrides this variable if set.
6. *n\_genes\_per\_dataset:* (Boolean) Set to TRUE to label top n\_genes targets for each dataset, gene\_list overrides this variable if set.
7. *gene\_list*: (String) List of specified targets that will be labeled no matter what on figure. Default labeling method over n\_genes.

**Thesholds**

If thresholds are set a threshold line will appear on figure, set thresholds to NULL if no line is desired.

1. *pval\_thresh*: (Numeric) p-value threshold on y-axis
2. *fdr\_thresh:* (Numeric) *f*alse discovery rate threshold on y-axis
3. *fc\_thresh:* (Numeric)log2 fold change cutoff on x-axis.
4. *label\_fc*: (Boolean)Should genes below the FC threshold be labeled if they are also above the significance threshold

**Fonts**

1. *font\_size:* (Numeric) Font size on figure
2. *label\_size:* (Numeric) Size of font for the gene labels
3. *font\_family:* (String) Font family for all text on figure
   * Options: serif, sans, mono

**Plot Size**

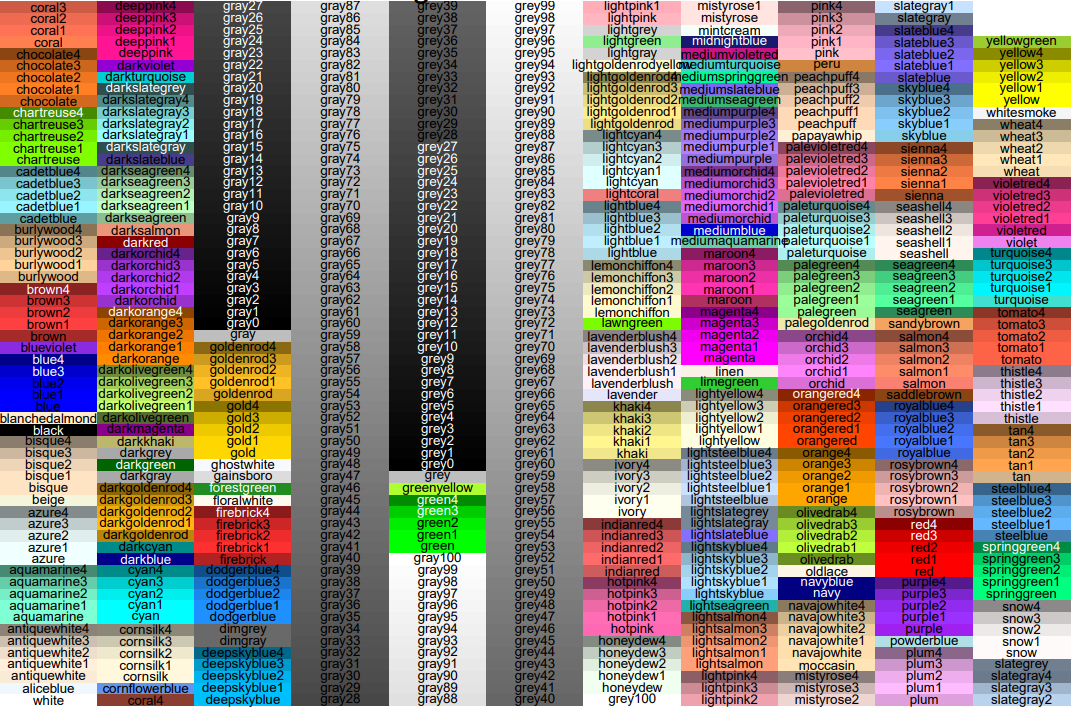
1. *plot\_width:* (Numeric) Width of saved figure in inches
2. *plot\_height:* (Numeric) Height of saved figure in inches

**Coloring**

Colors that can be recognized by R should be either named colors (e.g. “orange2”) or hexadecimal colors (“#ABABAB”). See below for a cheat sheet of all named R colors.

1. *default\_color:* (String) Color of points not in target group or above significance threshold
2. *fc\_color:* (String) Color of points below fc\_thresh but above significance threshold(s); change to same as default to not call out these targets
3. *target\_groups:* (String) Specific gene target groups to be colored in plot. Target groups are labeled in the VOLCANO PLOT.xlsx file. All genes in given target\_group are colored no matter where they are in the figure. If no group is given (NULL), targets are colored if they are above pval/fdr threshold.
4. *color\_options:* (String) List of colors to use in figure. Must have at least the number of target\_groups.
5. *color\_by\_dataset:* (Boolean) Should genes be colored by dataset name.

**Named R Colors:**



**Example Parameter Set-up**

The LabeledCombinedVolcanoPlot script outputs a typical volcano plot figure with log2 fold change on the x-axis and pvalue or FDR on the y-axis for each target. A table of labeled genes in the figure is also output.

*Example figures with different input arguments.*

**Example 1:**

**n\_genes = 10**

**fdr\_thresh = 0.01**

**pval\_thresh = 0.05**

**fc\_thresh = 0.5**

**label\_fc = FALSE**

**color\_by\_dataset = TRUE**

A graph with blue and yellow dots

Description automatically generated