Labeled Volcano Plot DSP DA Script Vignette

Introduction

The Labeled Volcano Plot DSP DA script creates publication-ready labeled volcano plots based on the researchers' input and statistical study results. The script also creates a table of tagged genes.

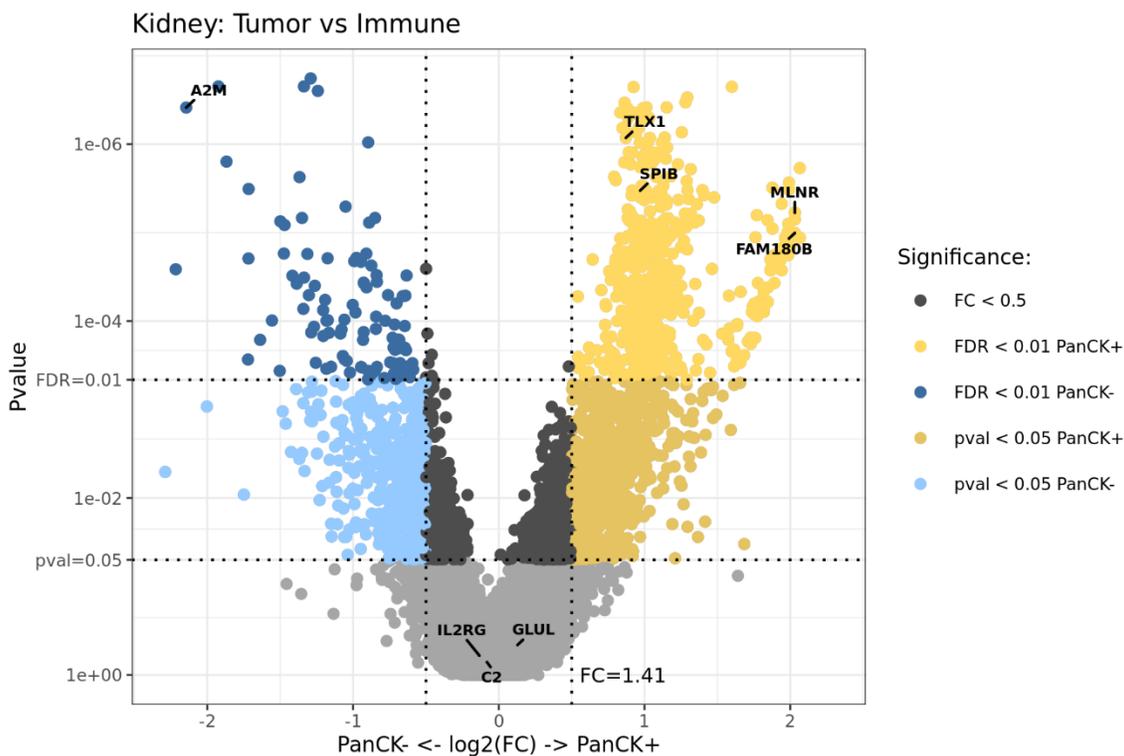## Use of the "LabeledVolcanoPlot" DSP DA script

Intended use

The **Labeled Volcano Plot DSP DA script** supports the **GeoMx nCounter** (*protein or RNA*) or *GeoMx NGS (CTA)* readout applications. The script creates publication-ready labeled volcano plots based on user input and statistical test results. The script also generates a table with the tagged genes.

Contents

For more visual example graphs see page 6

Importing into the DSP DA

The **LabeledVolcanoPlot.R** script requires an extra file input from the DSP DA. After running a statistical test and creating a volcano plot in DSP DA, **export** the results as a **.xlsx** file. Statistical test results are under the dataset the test ran on.
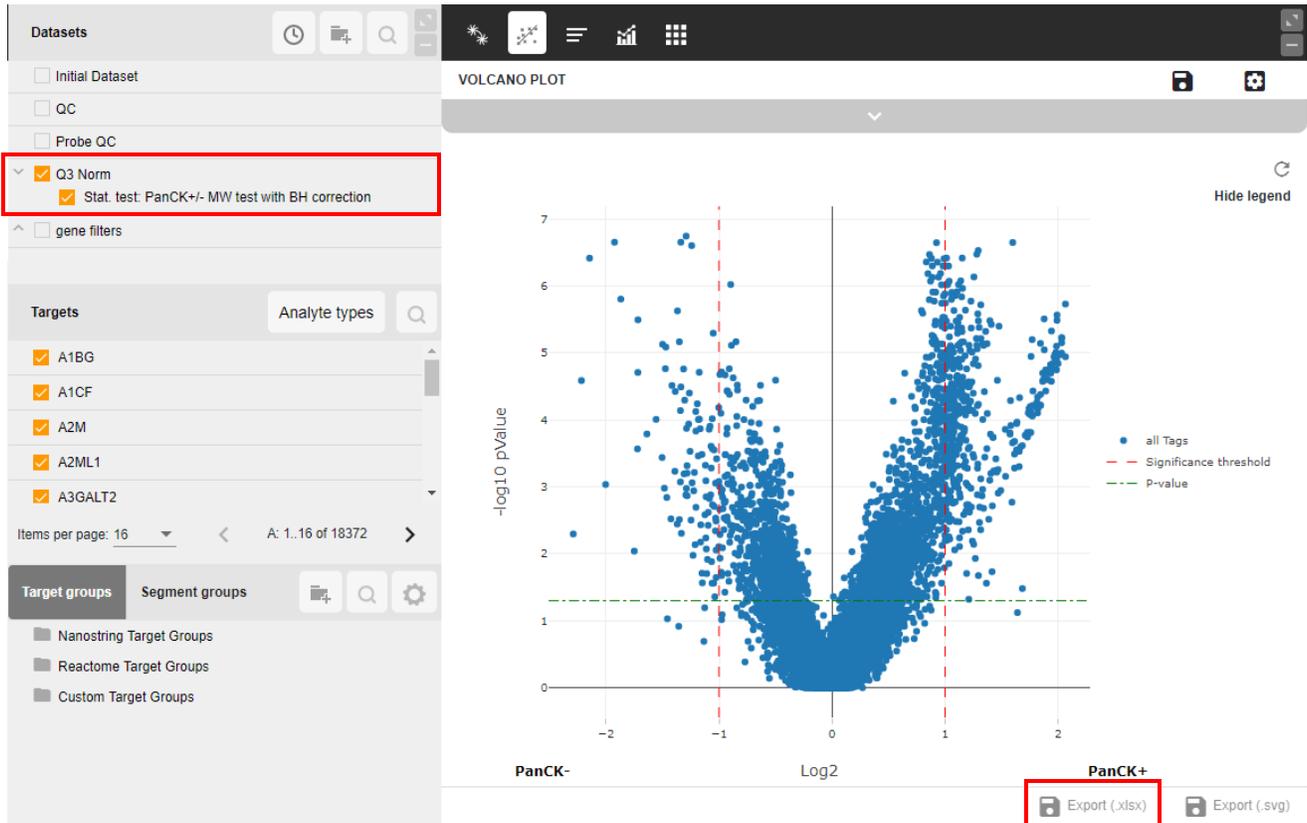

Change the file to a *tab-delimited file .txt* before running the script. The script will **NOT** run if file is in .xlsx format. **Open** the *Exported VOLCANO PLOT.xlsx file* in Excel and click **Save As**, *to change the format* to **Text (Tab delimited) (*.txt).**
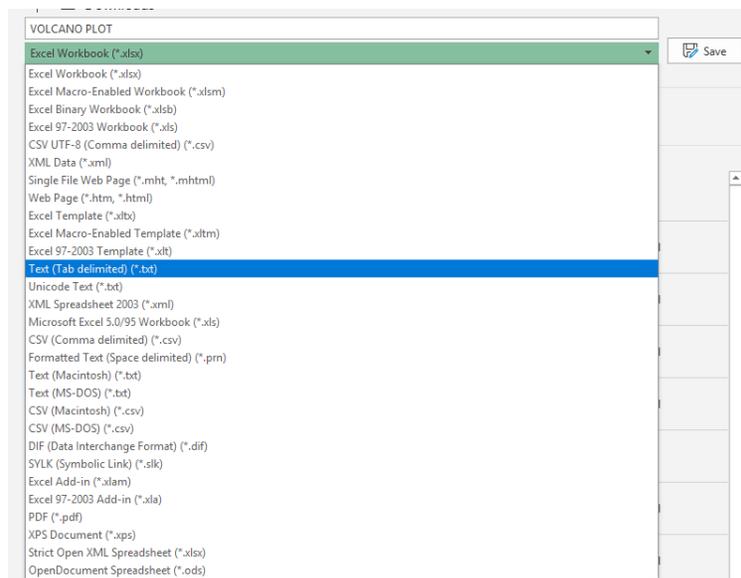
The **LabeledVolcanoPlot.R** loads into the custom scripts section of the DSP DA after you have a dataset processed and ready for analysis. To do so, **open** the custom script section by clicking on the unmarked statistical test dataset.
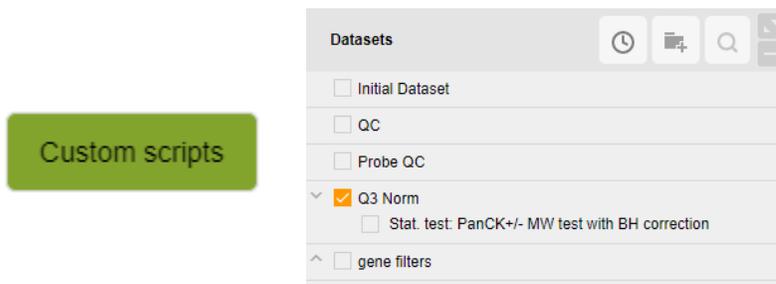
Select the "**Manage**" tab to open the area to load and edit scripts:
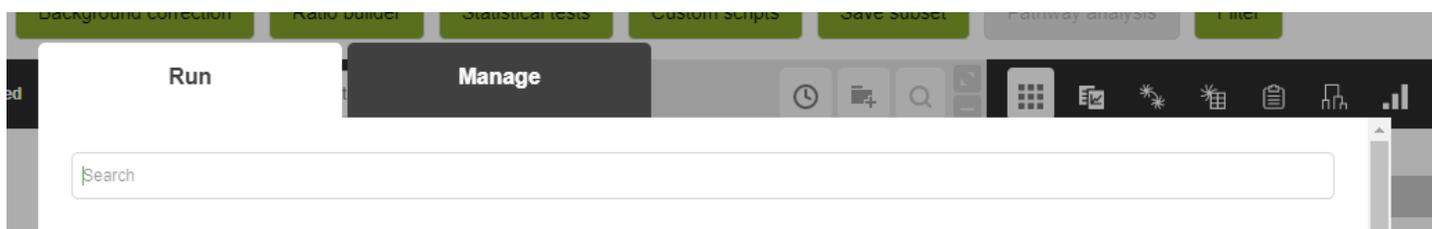

In the **Management** tab to add a new script and adjust parameters, fill out and then scroll to the bottom of the page. Use the **"+"** button to add the **LabeledVolcanoPlot.R** file to the script:

Use the **"+"** button to add the **VOLCANO PLOT.txt file**. Ensure the **LabeledVolcanoPlot.R** file is selected in the **dropdown menu**, indicated by a **star.**

Once the desired parameters of the script are selected, the user can start making modifications, editing the top lines in the script by selecting the **"Save"** button. No need to check the *Create* new dataset button.
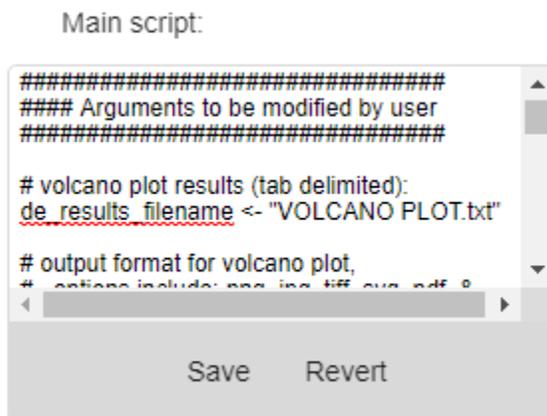
Main script:

Setting and Modifying the User Parameters:

There are twenty-one settings the user can modify at the top of the plug-in script. These settings include:

Files

1. *de_results_filename:* (String) Name of the tab-delimited file you have uploaded to the DSP DA.
2. *output_format:* (String) Desired output format for the volcano plot figure.
    - Options: *PNG, JPG, TIFF, SVG, PDF, and BMP*

Labeling

The user must add labels from the DSP DA volcano plot.

3. *plot_title*: (String) Title for figure
4. *negative_label*: (String) Matching negative (left) x-axis label to the volcano plot in the DSP DA
5. *positive_label*: (String) Matching positive (right) x-axis label to the volcano plot in the DSP DA
6. *show_legend:* (Boolean) A color legend will appear
7. *n_genes:* (Numeric) Number of top genes by pvalue/fdr to label on figure. *gene_list* overrides this variable if set.
8. *gene_list*: (String) Labeled list of specified genes are on figure. Default labeling method over *n_genes*.

Thresholds

A line will appear on the figure to show the thresholds. Remove the line by setting the thresholds to **NULL**.

9. *pval_thresh*: (Numeric) p-value threshold on the y-axis
10. *fdr_thresh:* (Numeric) *f*alse discovery rate threshold on the y-axis
11. *fc_thresh:* (Numeric) log2 fold change cutoff on the x-axis.
12. *label_fc*: (Boolean) Should genes below the FC threshold be labeled if they are also above the significance threshold

Fonts

13. *font_size:* (Numeric) Font size on the figure
14. *label_size:* (Numeric) Size of the font for the gene labels

15. *font_family:* (String) Font family for all text on figure

   o Options: *serif, sans, mono*

Plot Size

16. *plot_width:* (Numeric) Width of the saved figure in inches

17. *plot_height:* (Numeric) Height of saved figure in inches

Coloring

Colors that R can recognize should be either named colors (e.g., "orange2") or hexadecimal colors ("#ABABAB"). See below for a chart of all named R colors.

18. *default_color:* (String) Color of points not in target group or above significance threshold

19. *fc_color:* (String) Color of points below *fc_thresh* but above the significance threshold(s); change to same as default to not call out these targets

20. *target_groups:* (String) Color-specific gene target groups to be colored in plot. Labeled target groups are labeled in the *VOLCANO PLOT.xlsx* file. All genes in the given *target_group* are colored no matter where they are in the figure. If no group (*NULL*) will appear, locate the colored targets above the *pval/fdr* threshold.

21. *color_options:* (String) List of colors to use in the figure. Must have at least the number of *target_groups.*

Named R Colors Chart:
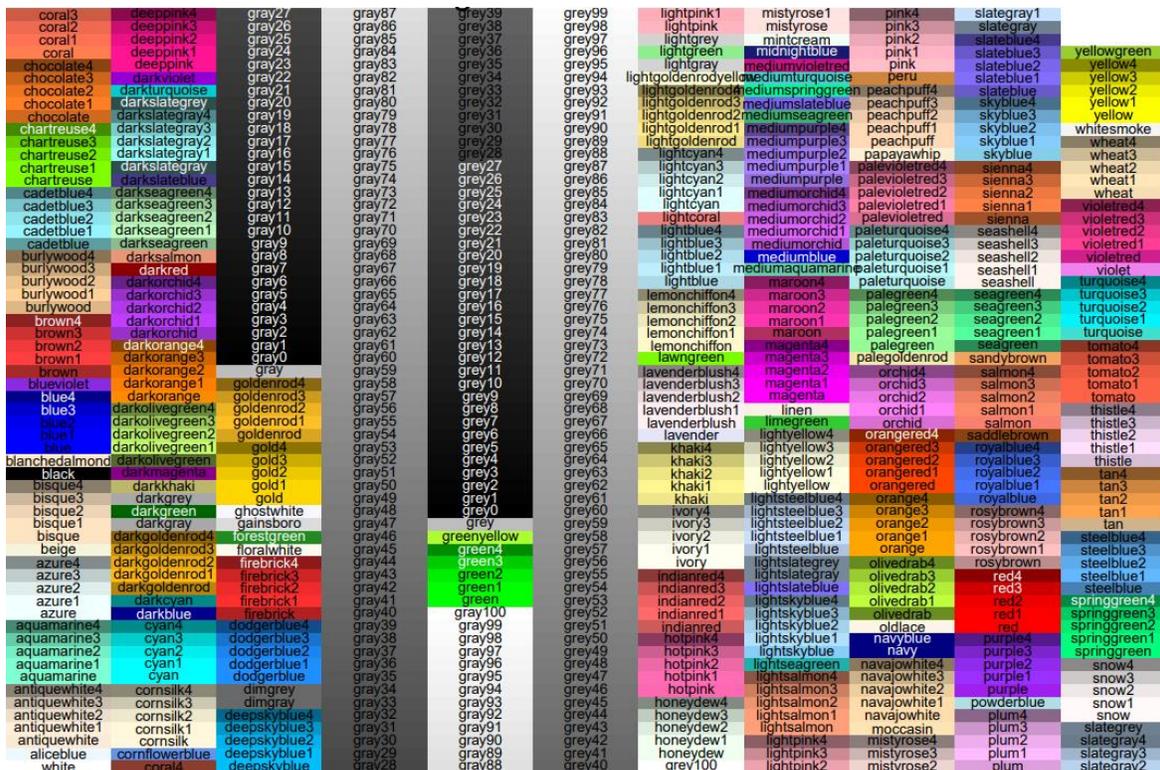
Example Parameter Set-up

The **LabeledVolcanoPlot.r** script outputs a typical volcano plot figure with log2 fold change on the x-axis and the pvalue or FDR on the y-axis for each target. A table of labeled genes in the figure is also output.
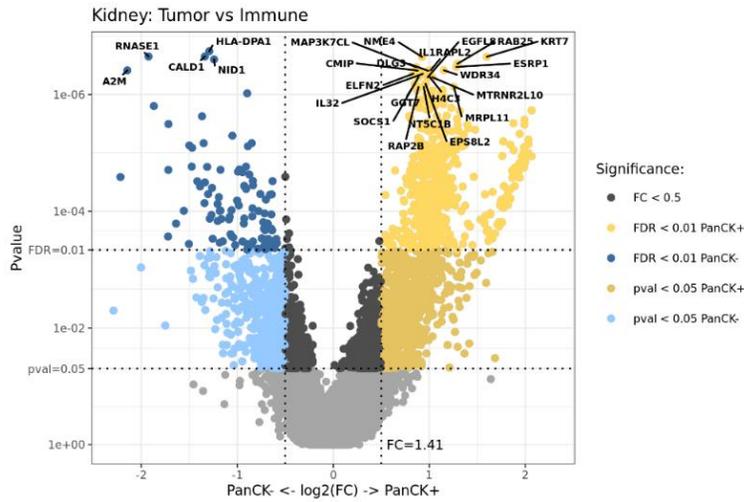
Example figures with different input arguments.

Example 1:

n_genes = 25	**pval_thresh = 0.05**	**label_fc = FALSE**
fdr_thresh = 0.01	**fc_thresh = 0.5**	**target_groups = NULL**

Example 2:

gene_list = c("IL2RG", "GLUL", "SPIB", "C2", "A2M","MLNR", "TLX1", "FAM180B")

target_groups = c("Hemostasis", "DNA Repair")

pval_thresh = NULL

fc_thresh = 0.5

fdr_thresh = 0.01