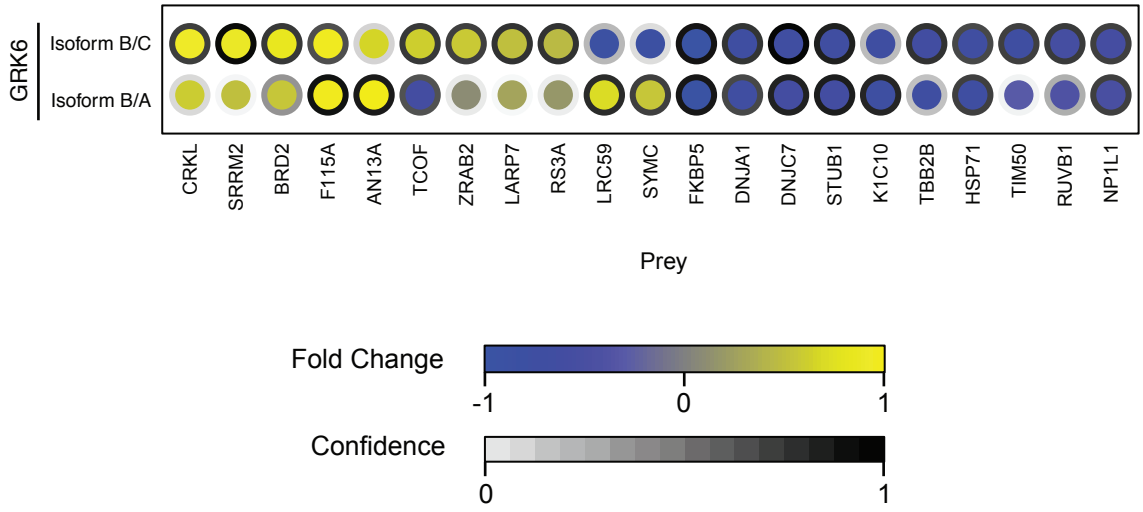
Fold Change Dot Plot User's Guide
v1.0 (September 3rd, 2014)

This manual provides usage information for the fold change tool at http://prohitstools.mshri.on.ca. This tool takes quantitative protein-protein interaction data and generates a 2D dot plot summarizing the fold change information, similar to the example below. This tool was developed to plot both the fold change and the confidence information generated by the fold change calculation in Affinity Purification coupled to SWATH Data Independent Acquisition as published in [1]. The tool is be compatible with other scoring based on fold change calculations (including for MS1 intensity), providing that a confidence value on the measurements is provided with these tools.

For more details about the tool, see JDR Knight, G Liu, JP Zhang, A Pasculescu, H Choi and AC Gingras (Submitted) "A web-tool for visualizing quantitative protein-protein interaction data," *Proteomics*.



Inputting data

Two data files must be supplied as input in tab-delimited format. One file should contain the fold change information and the other the confidence values (on a scale of zero to one; one being the highest confidence level). Both files are matrices with baits listed by column and preys by row as shown below. Baits and preys must be ordered identically in each file.

1. Fold change file

	Bait 1/Bait 2	Bait 1/Bait 3
DNJC7	-0.4946	-0.566
SRRM2	0.4615	0.8213
F115A	0.8891	0.8697
FKBP5	-0.8339	-0.8558
AN13A	0.8979	0.6086
STUB1	-0.512	-0.5557
K1C10	-0.6599	-0.5934
LRC59	0.6647	-0.7433
LARP7	0.2782	0.4674
RS3A	0.1869	0.425
TBB2B	-0.5837	-0.525
DNJA1	-0.5503	-0.6373
BRD2	0.5183	0.7838
TCOF	-0.4926	0.5668
RUVB1	-0.3222	-0.4527
CRKL	0.5638	0.858
HSP71	-0.5868	-0.5599
NP1L1	-0.4129	-0.4826
SYMC	0.5209	-0.6914
TIM50	-0.2395	-0.6004
ZRAB2	0.1177	0.5358

2. Confidence in the fold change file

	Bait 1/Bait 2	Bait 1/Bait 3
DNJC7	0.86	0.98
SRRM2	0.054	0.97
F115A	0.93	0.69
FKBP5	0.9	0.92
AN13A	0.89	0.174
STUB1	0.88	0.87
K1C10	0.85	0.258
LRC59	0.83	0.282
LARP7	0.048	0.8
RS3A	0.084	0.8
TBB2B	0.24	0.8
DNJA1	0.72	0.79
BRD2	0.426	0.78
TCOF	0.68	0.78
RUVB1	0.318	0.77
CRKL	0.156	0.76
HSP71	0.75	0.72
NP1L1	0.75	0.75
SYMC	0.75	0.138
TIM50	0.06	0.75
ZRAB2	0.096	0.75

Clustering Options

1. Agglomerative Hierarchical (Hierarchical): This is the only clustering option currently available and is executed using R. There are several options for calculating the distance metric and for the clustering criterion. Euclidean is the default distance metric and Ward's method the default clustering type. From our own experience we have found Euclidean to be a good metric to use although other options are available and may produce more desirable results on different data sets. The default clustering type is Ward's, which acts to minimize variance within clusters, although many of the types available will produce comparable results.

Available distance metrics: binary, Canberra, Euclidean, Manhattan, maximum and Minkowksi

Available clustering criteria: average, centroid, complete, McQuitty, median, single and Ward's

2. No-clustering: The user can generate dot plots without clustering if desired. For these cases, the ordering as provided in the input files will be used to order the output dot plot.

Output

After the data has been processed, the user will be prompted to download a .zip file that contains the results in a folder. There are several files in this folder.

1. foldchange_dotplot.pdf: This is the dot plot and the file can be opened and edited in Adobe Illustrator or a similar program. In some cases you may get the following warning on opening the file in Illustrator: "The font AdobePiStd is missing. Affected text will be displayed using a substitute font." If this occurs, the dot plot will not display correctly.

To fix this issue on a Mac, copy the file AdobePiStd.otf from /Library/Application/Support/Adobe/PDFL/*Current Version*/Fonts/ and transfer it to the folder

/Library/Fonts/. The *Current Version* folder refers to your version of Adobe. On Windows, the font file is located in C:\Program Files\Common Files\Adobe\PDFL*Current Version*\Fonts\ and needs to be placed in C:\Program Files\Adobe\Adobe Illustrator CS5\Support Files\Required\Fonts\. If the AdobePiStd.otf file is missing, it can be downloaded from a number of sites on the web for free in either Mac or Windows format.

2. foldchange_heatmap_x.pdf: A heatmap will also be produced as an alternative to the dot plot. The data presented will be identical to what is shown in the dot plot but in a standard heatmap format without the confidence values. For very large datasets, the dot plots are often unsuitable to use in publications because too much visual information will be compressed into a very small space. We generate heatmaps for these situations. There will be two heatmaps generated: heatmap_border.pdf and heatmap_no_border.pdf. The only difference between the files is that the first will have black borders drawn around the cells. The choice between which image to use is purely based on aesthetic preferences.

3. process.log: Log file that contains information on the input parameters that were selected for the user's future reference (e.g. to assist with writing the Methods section for a manuscript).

References

[1] Lambert, J. P., Ivosev, G., Couzens, A. L., Larsen, B.*, et al.*, Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature methods* 2013, *10*, 1239-1245.

Troubleshooting

Problems generally result from errors in the input file format, and we encourage users to compare their input files against the examples provided on the web page.

Any other issues should be sent to jknight@lunenfeld.ca.