

Tutorial v1.0

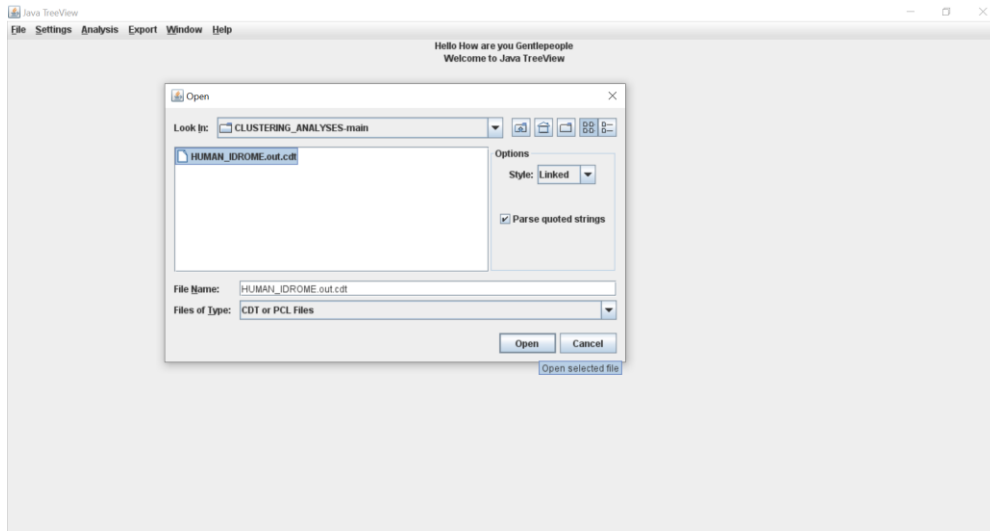
The tutorial explains the steps needed to:

1. analyze a clustering map of the human IDR-ome based on evolutionary signatures using Java TreeView
2. conduct gene and feature enrichment analyses based on the signatures using a python script

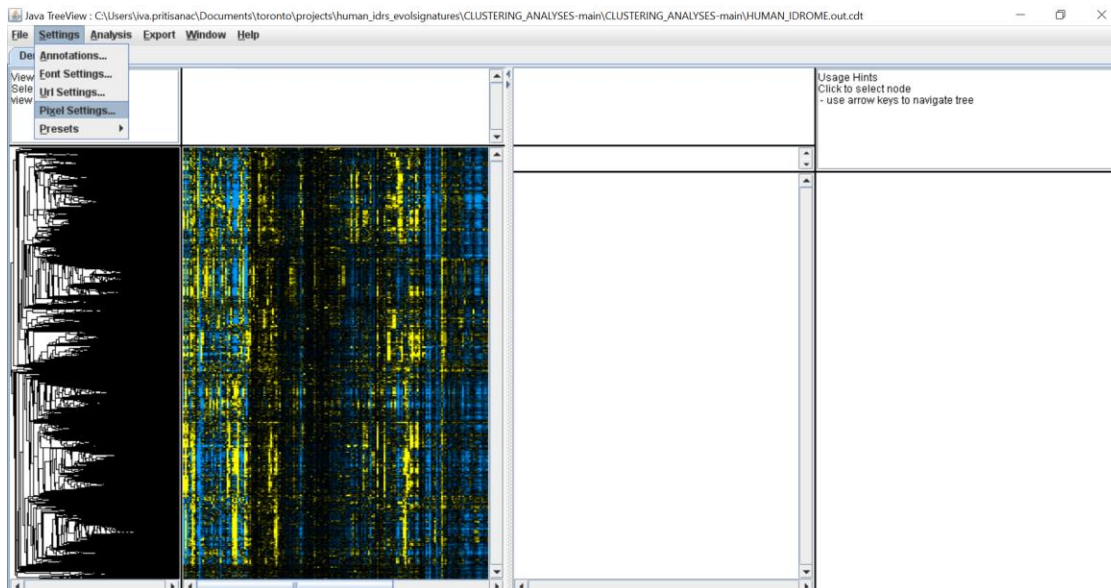
Setting up Java TreeView

Install Java TreeView (follow the instructions on this webpage: <http://jtreeview.sourceforge.net/>).

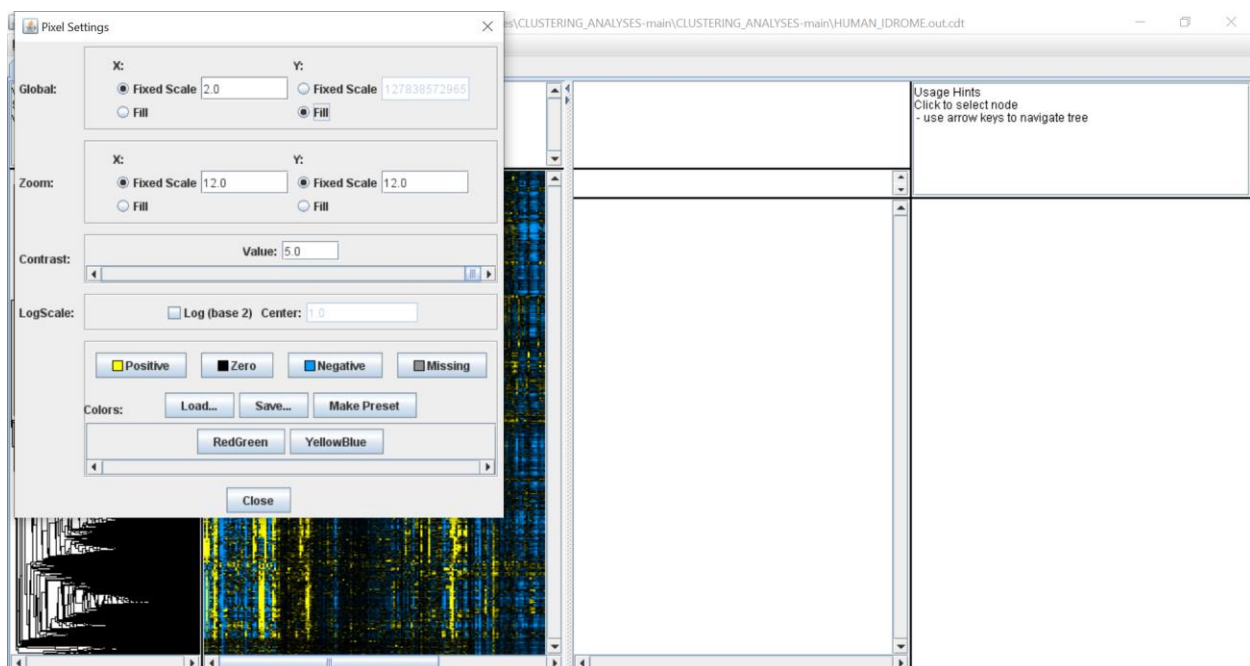
Open Java TreeView. Select “File” Open and navigate to directory with .cdt and .gtr files



Once the clustering data are loaded, go to “Settings/Pixel Settings”



Under “Global Y” select “Fill”. Next, choose “Contrast Value”. It is good to set the contrast to a higher value (e.g., 10) when doing a global analysis, as the higher threshold helps highlight the most prominent features.

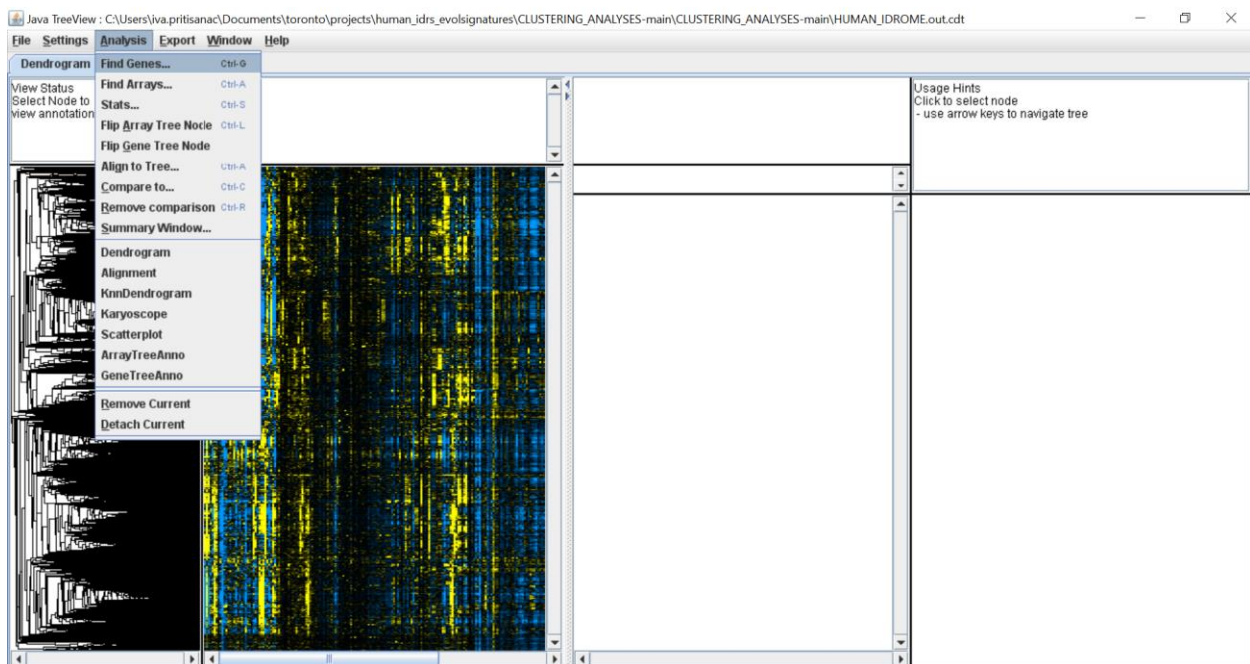


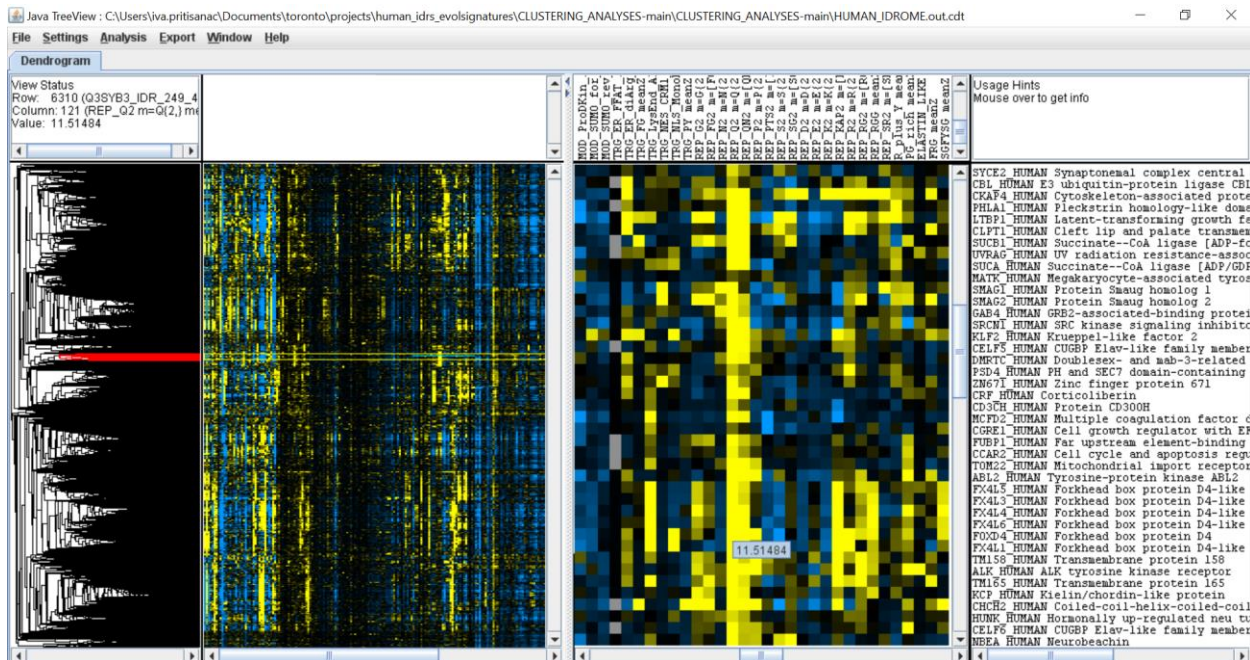
Under *Colors*, adjust colorbar by selecting Positive and Negative color values (e.g., yellow for positive values of Z-score and blue for negative).

Finding IDRs of interest

Go to “*Analysis/Find Genes*” and type in the name of the gene of interest (alternatively search for Uniprot ID, or ENSEMBL Gene ID). Double click the IDR of interest in the pop-out window (a gene can have multiple IDRs; names of entries are given in a format IDR_#start_#end, where #start is the starting amino acid of an IDR and #end is the final amino acid of the IDR). Selection takes you to the place in the clustering where the selected IDR resides. Use mouse to expand from the gene of interest into the clustering neighborhood. Use the middle “zoom in” window to explore the features that define the cluster around the IDR of interest.

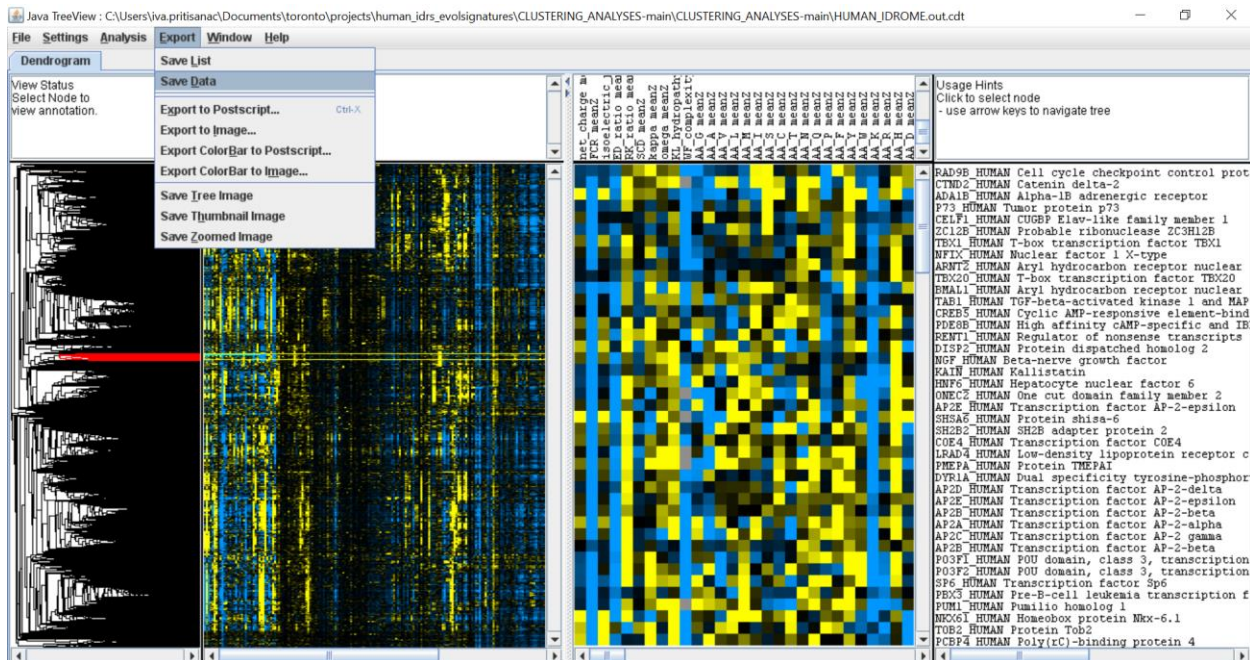
You can also left click on the tree in the region of an IDR/IDRs of interest and leave the cursor at the tree portion of the screen, while moving up or down using keyboard. The area of selection will start expanding to include the IDRs clustering in the neighborhood. You can take a note of the neighbors with similar signatures, export them as a list of data (see below) and look for their relationship to other IDRs of interest within the map.





Analyzing and exporting clusters

To export the data containing only IDRs or clusters of IDRs of interest go to “*Export/Save Data*”.



This will save the selected region in a .cdt format (the same format as the complete dataset). Data saved in this format can be used for a subsequent GO and feature enrichment analysis (see below). Save all data of interest in one directory/folder to make the subsequent analysis easier. Give clusters meaningful names to remember your selection.

Statistical analyses of GO enrichment and associated features

Install python3.0 through Anaconda application (For Windows: <https://docs.anaconda.com/anaconda/install/windows/>, check the same website for Linux or Mac OS specific instructions). Make sure your PYTHON PATH is set to where your es_analyses.py script is (e.g., for Windows see <https://geek-university.com/python/add-python-to-the-windows-path/>) or give a full path to your python installation when running the script. If running on Windows, open command line tool by typing *cmd* in the start menu. To run the script type:

```
> python3 es_analyses.py HUMAN_IDROME.cdt PATH_TO_MY_DATA_DIR/
```

where "PATH_TO_MY_DATA_DIR" contains all .cdt files of interest for GO and features enrichment analyses. The script will output two text files ("CLUSTER_NAME_GO_ANALYSIS.txt" and "CLUSTER_NAME_FEAT_ANALYSIS.txt").

If you have any questions e-mail at: iva.pritisanac@gmail.com

Have fun 😊