**Clustergrammer Visualization of Transcriptomic NURSA Data**

****

Legend: Above is an example Clustergrammer visualization of transcriptomic data from a dataset available through NURSA. The dataset name is: Time course transcriptomic analysus of all-trans retionoic acid (ATRA)-treated NB4 cells. A subset of the data, ATRA treated TGM2 knockdown NB4 cells, is shown in the clustergram. Enrichment analysis has been performed on the genes (rows) and the top ten enriched mouse phenotypes after gene knockdown are shown – we can see that many of the top enriched phenotypes are immune-related and are associated with down-regulated genes (blue matrix cells) after ATRA treatment.

**Data Visualization using Clustergrammer**

The NURSA Transcriptomine search engine allows users to access over 30 million transcriptomic datapoints, but enabling researchers to easily engage with data and obtain insights remains difficult. Web-based visualization can be a powerful tool to enable user-friendly interaction with and analysis of complex data. We developed the web-based visualization tool, Clustergrammer, that enables users to interactively visualize high-dimensional biological datasets such as: genome-wide expression and proteomics data. Clustergrammer allows users to interactively reorder, filter, perform enrichment analysis, and export data to other tools. Clustergrammer is specifically built to allow users to intuitively explore and identify functional clusters in complex biological data, such as the transcriptomic data provided by NURSA.

We plan on using Clustergrammer to enable users to visualize user-specified subsets of the NURSA transcriptomic data as well as obtain down-sampled global views of the entire transcriptomic data. Specifically, for NURSA transcriptomic Dataset pages we will add interactive heatmap visualizations of differentially expressed genes, which are currently shown as fold-change values in interactive gene list tables (see example [dataset](https://www.nursa.org/nursa/datasets/dataset.jsf?doi=10.1621/z04egxZq1P)). The Clustergrammer visualizations will enable users to easily identify clusters of differentially expressed genes and experiments and perform enrichment analysis on these genes using the Ma’ayan lab tool Enrichr.

To demonstrate

[choose subsets of experiments, bring in genome-wide information to help understand]

**Integration with Gene Signature Database**

Transcriptomic data can be used to assign gene-expression signatures to specific biological states, such as disease-states and drug-perturbation states, and these signatures can be used to discover novel relationships between biological states. For instance, we can use gene-signature comparison to draw connections between genes and diseases. We have combined and curated a diverse set of gene signatures from the several resources, including the LINCS L1000 project and the Gene Expression Omnibus, to build several large databases of gene-signatures. This resource of gene-signatures can be used in combination with the NURSA transcriptomic data to identify novel associations between: genes, tissues, and nuclear receptors. We have built several tools, including L1000CDS2 and CREEDs, that allow users to easily query these data.