Mihir Samdarshi

January 25, 2019

BIOL 588 – Advanced Systems Biology Research

**Refactor of code to Model-View-Controller paradigm improves** **GRNsight: a web application and service for visualizing small- to medium-scale gene regulatory networks**

Annotated Bibliography

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *Icwsm*, *8*(2009), 361-362.

Gephi is a commonly used software for easy visualization and manipulation of networks. Many of the features of GRNsight are similar in nature to that of Gephi’s. Gephi can be considered a “competitor” to GRNsight, however it can be referred to as primarily being targeted to large graphs, as is mentioned in the paper. One of the main file formats for importing data into GRNsight is GraphML added because of the commonality with Gephi.

Dahlquist, K. D., Fitzpatrick, B. G., Camacho, E. T., Entzminger, S. D., & Wanner, N. C. (2015). Parameter estimation for gene regulatory networks from microarray data: cold shock response in Saccharomyces cerevisiae. *Bulletin of mathematical biology*, 77(8), 1457-1492.

GRNmap is a necessity to preface the creation of GRNsight; GRNsight was spawned to serve as a tool to visualize the data that was being produced in the difficult to interpret tabular format. The algorithms powering GRNmap allows scientists to estimate parameters of gene expression from sets of experimental data. The visualization provided by GRNsight allows scientists, especially those working with data derived from GRNmap to easily view and share their data with other scientists.

Franz, M., Lopes, C. T., Huck, G., Dong, Y., Sumer, O., & Bader, G. D. (2015). Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*, *32*(2), 309-311.

Cytoscape.js is an important JavaScript library that is being viewed as a candidate for potential integration into GRNsight. It allows the rendering of interactive graphs and provides the ability to carry out graph operations server-side. Although GRNsight is based on d3.js, the visualization capabilities that Cytoscape.js provides are worth mentioning as a benchmark for comparison to that of GRNsight’s capabilities.

Futschik, M. E., & Crompton, T. (2004). OLIN: optimized normalization, visualization and quality testing of two-channel microarray data. *Bioinformatics*, *21*(8), 1724-1726.

This article provides information regarding the method of normalization and visualization of microarray data used by this particular research group. The normalization function that they used is a complicated algorithm, as opposed to the normalization that is used in GRNsight, which is relative to the node with the smallest weight. Additionally, the visualization functions that OLIN provides are features that are now commonly available and used in many commercial microarray analysis software.

Kohl, M., Wiese, S., & Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. *In Data Mining in Proteomics* (pp. 291-303). Humana Press.

Cytoscape is a software tailor-made for the visualization of biological pathways represented as networks of interactions. This article provides a few of the motivations behind the creation of Cytoscape, and also contains a list of properties key for a successful software in the field of biological network data representation. Additionally, this article provides additional use cases for the software, which can be mentioned in my paper.

Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., ... & Davis, R. W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proceedings of the National Academy of Sciences*, *94*(24), 13057-13062.

This article provides insight into the first major creation and application of DNA microarrays for the representation of activation of *Saccharomyces cerevisiae* open reading frames. DNA microarrays are used in the Dahlquist lab’s wet lab to obtain data about the activation or repression of genes under cold shock conditions. This particular study used heat shock to demonstrate differences between microarrays.

Leff, A., & Rayfield, J. T. (2001). Web-application development using the model/view/controller design pattern. In *Enterprise Distributed Object Computing Conference, 2001. EDOC'01. Proceedings. Fifth IEEE International* (pp. 118-127). IEEE.

Application of the Model-View-Controller paradigm to GRNsight is going to be my big focus during the course of this semester. This article provides a roadmap to implementing this archetype; it provides solutions to potential problems that may arise, such as state synchronization, and, despite its age, provides examples of successful implementations. It also espouses the many benefits that this architecture provides, making it a worthwhile source to cite.

Teller, S. (2013). Data visualization with d3.js. Birmingham: Packt Publishing. Retrieved from https://the-eye.eu/public/Books/qt.vidyagam.es/library/Web Development/Data Visualization With D3.js/Data Visualization With D3.js - Swizec Teller.pdf

D3.js provides the backbone for all the visualization that GRNsight does. It’s role as a lightweight, easy to use, open source visualization tool also makes it one of the most popular visualization libraries used in modern software development. This book explores the many avenues that a developer can take with d3.js and also provides a guide on familiarizing oneself with all of d3.js’s tools, especially those used on the GRNsight project. This book is a useful tool to cite, especially when discussing GRNsight’s new MVC architecture’s interactions with d3.js.

Pavlopoulos, G. A., Malliarakis, D., Papanikolaou, N., Theodosiou, T., Enright, A. J., & Iliopoulos, I. (2015). Visualizing genome and systems biology: technologies, tools, implementation techniques and trends, past, present and future. *Gigascience*, *4*(1), 38.

Visualization of genomic data in Systems Biology is a large field, with many software tools available, each filling or competing to fill a particular niche. GRNsight’s niche, so far seriously unopposed, is the visualization of small- to medium-scale gene regulatory networks in a lightweight, easy to use package. This review goes over the other software tools and technologies that are available and widely used in the field of Systems Biology. Some of the tools and techniques mentioned here could be mentioned in my introduction or conclusion, as areas that we could improve on, or that we are already better at.

Saraiya, P., North, C., & Duca, K. (2005). Visualizing Biological Pathways: Requirements Analysis, Systems Evaluation and Research Agenda. *Information Visualization*, 4(3), 191–205. https://doi.org/10.1057/palgrave.ivs.9500102

The main task of GRNsight is visualization of gene regulatory networks, and this article, written by experts in Human-Computer Interaction and Bioinformatics provides rules and best practices for the visualization of biological pathways. The five best practices of biological pathway visualization that this article provides can be easily cited as GRNsight accomplishes each one of them.

Schade, B., Jansen, G., Whiteway, M., Entian, K. D., & Thomas, D. Y. (2004). Cold adaptation in budding yeast. *Molecular biology of the cell*, 15(12), 5492-5502.

This article is the classically cited article that provide the basis of the research of the Dahlquist lab. It provides the fundamental transcriptional data of genes most regulated when exposed to a cold shock treatment comparing it to a control treatment as well as other environmental stressors. The genes identified as being most regulated are subject to further, smaller scale analysis in experiments by the Dahlquist lab.

Theocharidis, A., Van Dongen, S., Enright, A. J., & Freeman, T. C. (2009). Network visualization and analysis of gene expression data using BioLayout Express 3D. *Nature protocols*, 4(10), 1535.

BioLayout Express 3D is a commonly used software tool to visualize gene expression data. However, while BioLayout Express 3D is meant to visualize large-scale networks in both three-dimensional and two-dimensional formats, GRNsight works with small-scale networks in simply a two-dimensional format. This paper can be cited, as any of the features and functionalities that BioLayout Express 3D can be compared to those of GRNsight. Additionally, as a multi-year project from a full-time lab, it has functionalities that surpass those of GRNsight, allowing us to derive future directions from it.