

Visualizing Network Data to Reflect State Change

Jocelyn Petitto

1 Introduction

Networks are widely used for visualizing biological systems. These systems are developed from the bottom-up, using experimentally acquired data to construct systems, and from the top-down, using experimental omics data to assemble hypothetical systems that are subsequently confirmed experimentally [1]. The relationship between the mitochondrial and nuclear genome plays out through interactions at multiple levels that include: metabolite concentration, protein availability and localization, and, potentially, noncoding mitochondrial RNA interaction with the nuclear genome [2]. The need for combining research efforts is palpable among those suffering from mitochondrial diseases, a class of dysfunction that is as diverse as cancer and with equally diverse genetic causes, including many that are polygenic.

We believe visualizing the complex interactions between these genomes provides a means of organizing existing research. By changing to a non-linear organization of research materials, we hope to establish which connections require further experimentation and/or validation of existing results. The need for this is based on the limitations in availability or quality of current data. To do so, we propose using methods of network visualization currently utilized to track collaboration such as OnGraX [3]. Visualizing causality is an area of weak-

ness in biological data visualization [4]. The current model of mitochondrial metabolomics, MitoCore, uses [5] uses flux balance analysis to model change, but this model is not visual. Flux analysis as a method requires less in terms of computational power, but is only used to model steady-states and rarely includes regulatory effects [6]. Ultimately, our goal is to create a searchable network integrating our new approach, which would include the effects of perturbations of equilibrium conditions, with existing data. This will provide a framework for exploring the effects of mutations in either genome on the tissue specific state of a cell as the network re-establishes equilibrium.

2 One-sentence description

Drawing from existing tools, we aim to prototype a network that displays both multiple omics data and the supporting research.

3 Project Type

Visualization (as a research tool)

4 Audience

The tool proposed will impact scientists researching mitochondrial function and disease. Access to research related to specific metabolic

pathways or gene expression presented visually will allow for a more robust review of available literature that may be overlooked by traditional search methods. Additionally, the proposed tool will help groups assess what areas are in need of investigation so that resources may be distributed accordingly.

As the non-coding products of DNA transcription are increasingly acknowledged as having significance in cellular structure and function, the lines between various omics will likely grow more fuzzy and the need to combine previously existing networks will increase. This makes visualization of existing research a potentially useful tool for other areas of research where complex dynamics may need to be reassessed in terms of multiple networks.

5 Approach

5.1 Details

Mitochondrial function requires the coordination of two genomes. Their dysfunction is the result of any number of mutation combinations with varying degrees of severity. These degrees of severity are based on the changes to the network that must be made to try to achieve a new steady-state. Visualizing these changes will make the associations between genetic mutation and metabolic effect more apparent. Overlaying the available research on the pathways between and within existing networks will lead to a better understand the function (and dysfunction). Additionally, this approach will help researchers account for potential confounders when designing experiments or identify areas in need of further research.

5.2 Evidence for Success

This approach will work because humans are inclined to see change, but only in situations with which they are not familiar. Unexpected changes in a known realm must be pointed

out or they will go overlooked. The tendency toward the familiar may also dissuade researchers from exploring new territory, when unbeknownst to them, there may be connection to previous work that makes responsible research in a new area more plausible.

6 Best-case Impact Statement

In the best-case, we would demonstrate how the state of a system changes based on an alteration of the attributes of a node in the network. In doing so, we would provide a prototype of visualizing change within a network. Whether or not these results lead to perceptual understanding by the intended user would require additional research.

In addition to visualizing change, the user will be able to retrieve the annotation that indicates that the perceived change is associated with the initial alteration in node attributes.

7 Major Milestones

- Curate a small dataset to use a portion of the target network as to build a prototype.
- Make the network visualization reactive to change in node attributes such as a mutation that renders its protein product non-viable.
- Implement a method of visualizing that change in a lasting way (possibly similar to a topological map?).
- Include interactivity so the user may view the research supporting the visualized relationship.

8 Obstacles

8.1 Major obstacles

- D3 is not the ideal tool to create the final visualization, but it is a viable option for a small prototype given its flexibility and ability to display within a browser. The later increases the ability to demonstrate functionality to foster future collaborations.
- Effectively reflecting causal change in a visualization.

8.2 Minor obstacles

- Defining a subset of relationships (to be represented by nodes and edges) for use developing the prototype.
- Learning to utilize the existing flux balance data using MitoCore within MatLab.
- Hard coding dummy data in a structure usable for this purpose.
- Performing the appropriate data manipulations in R, a relatively new environment for me, to come up with datasets usable to construct visualizations in D3.

9 Resources Needed

- Flux data
- Genetic information correlating mitochondrial DNA and gene products
- Code to make a network with edges reflective of reaction rates
- More investigation into how to visually represent change in a system.

10 5 Related Publications

Go presented a review of the omic levels involved in oxidative physiology and disease in mitochondrial given existing data and the complexity of an adequate, network-based model could be constructed[1].

Uppa introduces xMWAS, a tool that incorporates multiple omics datasets including visualization, clustering of functionally related biomolecules, and analysis to characterize nodes that undergo state change under different conditions[7].

Smith built MitoCore, a constrain based model for analyzing flux within mitochondria based on metabolites and proteonomics, using the COBRA package in MatLab [5].. Such a model is less computationally expensive than a kinetics based approach, though the lack of kinetics does come with limitations[6].

Murray constructed a taxonomy that outlines tasks performed by biologists who work with pathway data and the associated challenges of each data type. These findings were based on interview with scientists in the field at various specializations and levels of expertise [4]. need for research in displaying casual relationships

Lex constructed Entourage, a network visualization tool that aims to retain attributes lost in large datasets through the use of subsetting when displaying biological information using complex networks [8].

11 Define Success

At this point in prototyping, success would be demonstrating how perturbing a sample network changes the steady state.

References

- [1] Y. M. Go, J. Fernandes, X. Hu, K. Uppal, and D. P. Jones, "Mitochondrial network

- responses in oxidative physiology and disease,” feb 2018.
- [2] R. Vendramin, J.-C. Marine, and E. Leucci, “Non-coding RNAs: the dark side of nuclear-mitochondrial communication,” *The EMBO Journal*, vol. 36, pp. 1123–1133, mar 2017.
 - [3] B. Zimmer and A. Kerren, “OnGraX: A Web-Based System for the Collaborative Visual Analysis of Graphs,” *Journal of Graph Algorithms and Applications*, vol. 21, no. 1, pp. 5–27, 2017.
 - [4] P. Murray, F. McGee, and A. G. Forbes, “A taxonomy of visualization tasks for the analysis of biological pathway data,” *BMC Bioinformatics*, vol. 18, p. 21, feb 2017.
 - [5] A. C. Smith, F. Eyassu, J. P. Mazat, and A. J. Robinson, “MitoCore: A curated constraint-based model for simulating human central metabolism,” *BMC Systems Biology*, vol. 11, p. 114, dec 2017.
 - [6] J. D. Orth, I. Thiele, and B. O. Palsson, “What is flux balance analysis?,” mar 2010.
 - [7] K. Uppal, C. Ma, Y. M. Go, and D. P. Jones, “XMWAS: A data-driven integration and differential network analysis tool,” *Bioinformatics*, vol. 34, no. 4, pp. 701–702, 2018.
 - [8] A. Lex, C. Partl, D. Kalkofen, M. Streit, S. Gratzl, A. M. Wassermann, D. Schmalstieg, and H. Pfister, “Entourage: Visualizing relationships between biological pathways using contextual subsets,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2536–2545, 2013.