

## **Final Project Competition Spring 2019**

### **MED283 / BINF 268**

#### **Identification of hierarchical community structure in biological networks**

##### **Objective**

Build a hierarchical model of the synapse

##### **Details**

Given a provided set of known synapse genes, assemble a network of experimentally derived molecular interactions relevant to these and other genes that data suggest are synapse-related. Use community detection in molecular network data to construct a data-driven ontology model of the synapse. This model should be able to recapitulate and expand current biological knowledge and inform the interpretation of genetic variants.

##### **Evaluation Metrics**

- 1) How well does the model capture novel synapse proteins? (some of which have been recently uncovered by our collaborators using AP/MS/MS)
- 2) How well does the model capture the known structure of the synapse? (evaluated by alignment to the Gene Ontology)
- 3) How well does the model organize psychiatric disease genes? (your data-driven ontology will be used to provide gene sets for functional enrichment)

##### **Inputs**

Known synapse genes: synapse\_seed\_proteins.txt

Molecular interaction data: Any experimental networks of your choice, either one network dataset or any combination, integrated using your procedure of choice. In the lectures and problem sets we have gained hands-on experience with several large molecular networks that may be of relevance, including protein-protein, protein-DNA, DNA-DNA, and genetic interaction datasets. You are free to use any or all of these, or to access experimental datasets not covered in class. Whatever you do, you **MUST BE ABLE** to verifiably trace the nodes and edges of your network to primary experimental sources of interaction data. Directly encoding interactions extracted from literature or expert knowledge, and not traceable to experiments, is a disqualification.

##### **Outputs**

A hierarchy of network modules formatted as an OBO file. See <http://www.obofoundry.org/>

##### **Timeline**

Week 4: Form teams

Week 6: Network org. & modularity lectures

Week 7: (Midterm) Work for project has begun

Weeks 8 - 11: Work on project

Week 9: Submit draft hierarchies; evaluation on first gold-standard

Week 10: Oral presentations

Week 11 (Finals): Written presentation due Friday June 14 @ midnight

##### **Materials Due**

Oral & written formats follow Nature Methods research paper. For the written portion, the main

text (excluding abstract, Methods, references and figure legends) is approximately 3,000 words. The abstract is 100-150 words. Analyses have no more than 6 display items (figures and/or tables). An introduction (without heading) is followed by sections headed Results, Discussion and online Methods. The Results and Methods should be divided by topical subheadings; the Discussion does not contain subheadings. Up to 50 references. Oral presentations are approximately 12 minutes in length. Design your presentation and slides to roughly follow the flow of the paper format.