**Determining a reference plasma metabolite interactome for stable heart disease utilizing molecular structure for informative model selection**

Patrick J. Trainor1,2,\*, Samantha M. Carlisle3, Shesh N. Rai4,5,†, Andrew P. DeFilippis1,2,†

1Department of Medicine, Division of Cardiovascular Medicine, University of Louisville,

2Diabetes and Obesity Center, University of Louisville

3Department of Pharmacology and Toxicology, University of Louisville

4Biostatistics Shared Facility, James Graham Brown Cancer Center

5Department of Bioinformatics and Biostatistics, University of Louisville

\*Correspondence:

Email: patrick.trainor@louisville.edu

Address: 580 S. Preston St., Louisville, KY 40202

**Background**

As disease states are either precipitated by or result in metabolic dysregulation, metabolite concentrations can be utilized for determining physiological processes that are differentially impacted across disease states. To detect states of dysregulation, a reference model of probabilistic metabolite interactions in a non-pathological or referential state is desired. We sought to infer a probabilistic interactome for stable Heart Disease using blood plasma to serve as a reference for detecting acute states such as Myocardial Infarction.

**Materials and Methods**

Plasma metabolites were detected and quantified by UPLC-MS/MS and GC/MS from *N*=32 human subjects undergoing follow-up evaluation post-myocardial infarction or post-procedure for the treatment of stable Heart Disease that required cardiac catheterization. We employed the Adaptive Bayesian Graphical Lasso (AdBGL) for determining a plasma metabolite interactome. The AdBGL assumes a multivariate Gaussian model with double exponential prior distributions for off-diagonal concentration matrix entries and exponential prior distributions for the diagonal entries. We developed a method for defining shrinkage hyperparameters that incorporates *a priori* knowledge of molecular structure extracted from databases including PubChem and ChemSpider. This framework increases (decreases) the amount of shrinkage applied to entries corresponding to compounds that are dissimilar (similar) with respect to molecular structure. We developed a software implementation of a sampler for Markov Chain Monte Carlo simulation of the posterior distribution.

**Results and Conclusions**

We constructed a reference plasma metabolite interactome for a stable Heart Disease phenotype. A graphical topology was learned via the posterior median of concentration matrix entries and a thresholding procedure.