Improved Modeling and Analysis of Gene Expression

Annamarie Bair, Matthew McDermott, Professor Peter Szolovits

MIT EECS

Angle Undergraduate Research and Innovation Scholar

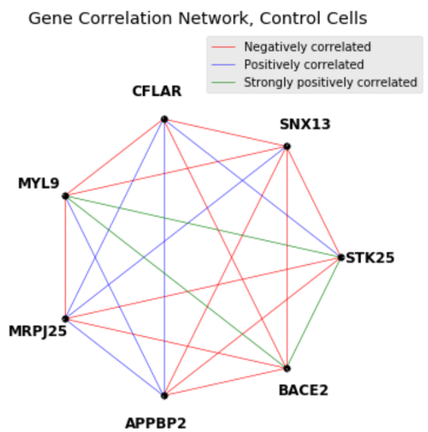
Goals

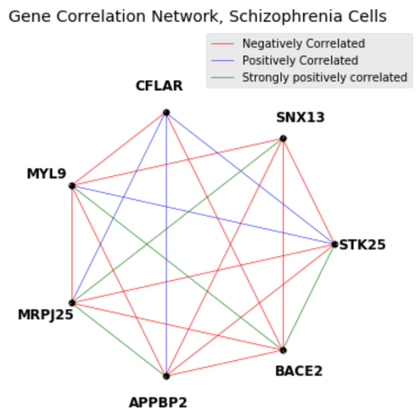
Differential gene extraction: develop methods to extract biologically significant genes that might not be detected by current methods

- *Zeroth order:* develop improved hypothesis testing
- *First order:* find which linear regulatory relationships between genes differ significantly in different conditions

Methods

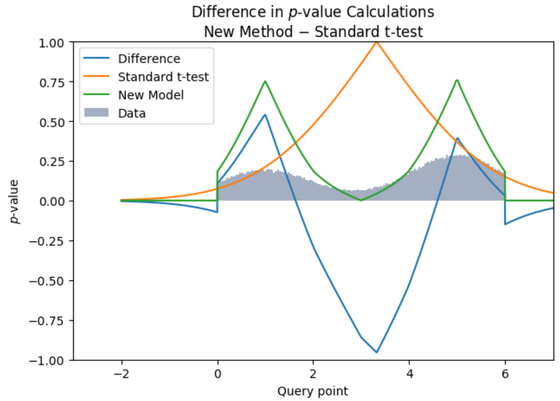
1. Implement mixture model including normal and truncated distributions.
2. Implement model selection algorithm and improved measure of significance
3. Perform correlational analysis to determine differentially expressed genes in two disease states: Control and Schizophrenia





Results

- Significant (p = 0.05) genes discovered through correlational analysis:
  - ACLY, FOXJ3, MPZL1 [3], STXBP1
- New model detects significant values in synthetic dataset which are not detected by standard t-test



Difference in calculated p-value on bimodal data.

Conclusion and Future Work

- This model uses hypothesis testing and correlational analysis to better detect significant gene expression levels.
- Next steps: test zeroth order model on real data; apply more refined network algorithms to model relationships between genes

References

[1] Gyemin Lee, Clayton Scott. EM algorithms for multivariate Gaussian mixture models with truncated and censored data. Computational Statistics \& Data Analysis, Volume 56, Issue 9, 2012, Pages 2816-2829, ISSN 0167-9473, https://doi.org/10.1016/j.csda.2012.03.003.

[2] James W. Jawitz. Moments of truncated continuous univariate distributions}, Advances in Water Resources, Volume 27, Issue 3, 2004, Pages 269-281, ISSN 0309-1708, https://doi.org/10.1016/j.advwatres.2003.12.002.

[3] He G, Liu X, Qin W, Chen Q, Wang X, Yang Y, Zhou J, Xu Y, Gu N, Feng G, Sang H, Wang P, He L. MPZL1/PZR, a novel candidate predisposing schizophrenia in Han Chinese. Mol Psychiatry. 2006 Aug;11(8):748-51. Epub 2006 May 9. PubMed PMID: 16702974.

SuperUROP
ADVANCED UNDERGRADUATE RESEARCH OPPORTUNITIES PROGRAM

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

CENTER FOR GENOMIC MEDICINE

Thanks To: Matthew McDermott, Peter Szolovits, Roy Perlis, Steve Haggarty, Wen-Ning Zhao, Jennifer Wang, Ting Fu, Steve Sheridan, Isaac Kohane, Susanne Churchill, Kamber Hart, & Overall MGH and CEGS Teams; NIH (Grant P50-MH106933)

Massachusetts Institute of Technology
School of Engineering