A novel method for identifying alternative relationships in eQTL mapping

Zhiyong Wang[1], Jinbo Xu[1] and Xinghua Shi[2]

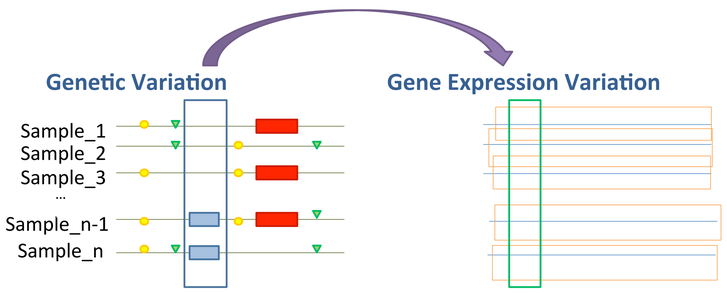
1.Toyota Technological Institute at Chicago 2.Department of Bioinformatics and Genomics, University of North Carolina at Charlotte

Abstract

Biological systems typically contain redundant and backup mechanisms. However, biological redundancy is rarely taken into account in expression quantitative trait locus (eQTL) mapping. We develop a method to capture alternative eQTLs by exploring sparse model space. In addition to mathematical proofs, we illustrate the application of our method to simulation and real datasets.

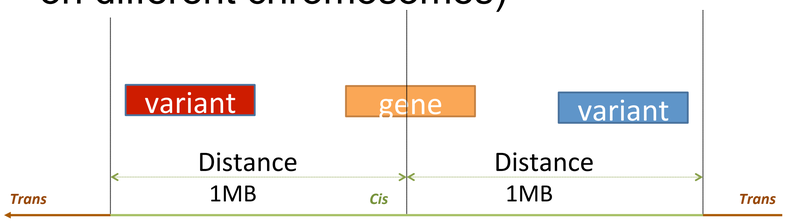
Background

eQTL mapping: identify genetic variants that significantly affect the variation of expression levels of genes.



General Strategy

- Cis associations (±1MB/500kb/200kb window of the target gene)
- Trans associations (beyond the window or on different chromosomes)

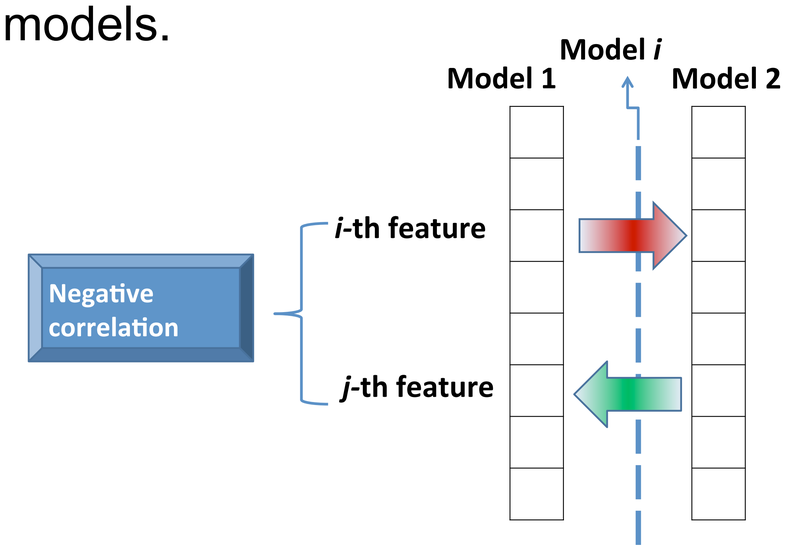


Pairwise Correlation Analysis: statistical test for each (variant, gene) pair. **hallenge: burden of multiple test correction.**

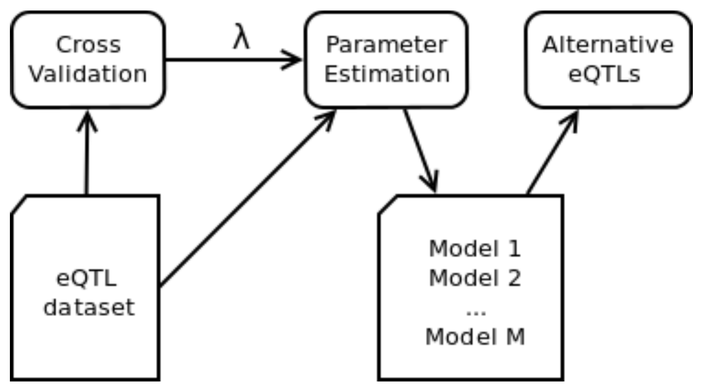
Lasso: minimize the number of non-zero weights in the model. **Challenge: sparse model assumption.**

Methods

Find alternative eQTL pairs in multiple models.



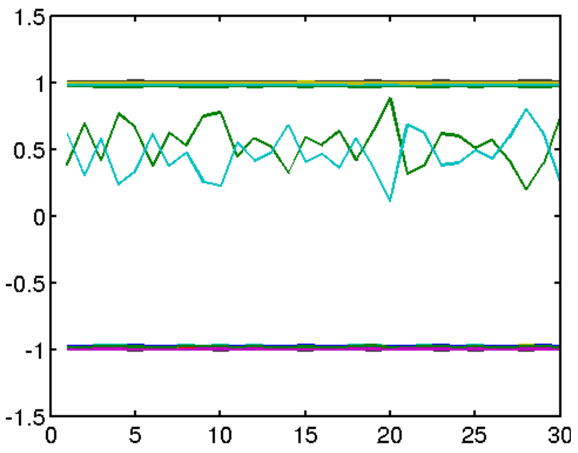
Workflow to find alternative eQTL pairs.



Simulation

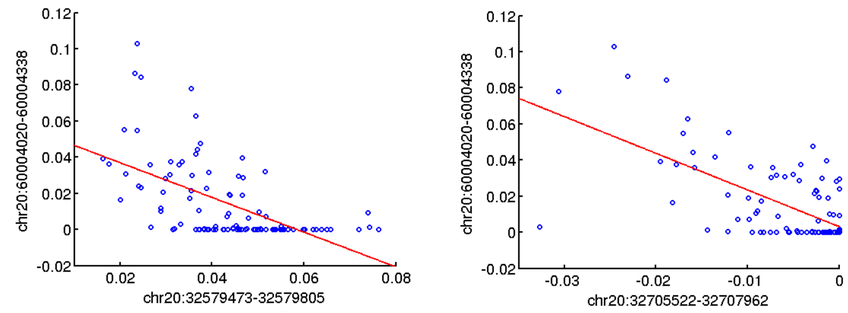
-Suboptimal solutions

-Strong negative correlation of model weights for two alternative eQTLs.

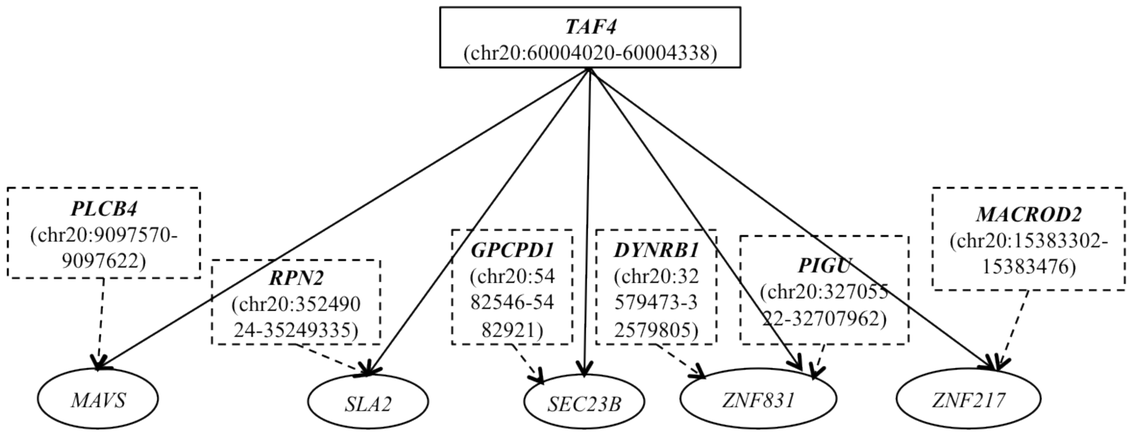


Real dataset

A human eQTL dataset using the copy number variant (CNV) genotypes (chr20) from the 1000 Genomes Project (The 1000 Genomes Project Consortium, Nature, 2012), and the gene expression quantifications measured using RNA sequencing (Montgomery SB et al., Nature, 2010). Alternative CNV eQTLs affect the expression of the same gene but might lie in different pathways.



Alternative eQTLs point to a network view of alternative pathways toward understanding gene regulation.



Summary

-A new method to capture alternative relationship among eQTLs by exploring sparse model space.

-Theoretical proofs with optimal and sub-optimal solutions.

-Results on synthetic and real data sets.

-Future work:

+ Develop fast algorithms for large datasets.
+ Discover more complicated relationships among eQTLs.
+ Extend to other sparse learning models.

Reference:

"Finding Alternative eQTLs by Exploring Sparse Model Space", Wang Z, Xu, J, and Shi X, Journal of Computational Biology, In Press, 2014.

Acknowledgment:

This work was supported by an NSF grant AF-1149811 (CAREER award) and the Alfred P. Sloan Research Fellowship and NIH R01GM089753 grants to JX, the Wells Fargo Foundation Fund for Faculty Excellence from Charlotte Research Institute and University of North Carolina at Charlotte to XS.