

## A Bayesian Partition Model for Detecting eQTL Modules

Wei Zhang<sup>1</sup>, Jun Zhu<sup>2</sup>, Eric E. Schadt<sup>2</sup> and Jun S. Liu<sup>1§</sup>

<sup>1</sup>Department of Statistics, Harvard University, Cambridge, MA

<sup>2</sup>Rosetta Inpharmatics, LLC, a wholly owned subsidiary of Merck & Co., Inc., 401 Terry Avenue North, Seattle, WA 98109

Studies of the relationship between genomic DNA variation and gene expression variation, often referred to as “expression quantitative trait loci (eQTL) mapping”, has been conducted in many species and resulted in many significant findings. Because of the large number of genes and genetic markers in such analyses, it is extremely challenging to discover how a small number of eQTLs interact with each other to affect mRNA expression levels for a set of (most likely co-regulated) genes. We present a Bayesian method to facilitate the task, in which co-expressed genes mapped to a common set of markers are treated as a module characterized by latent indicator variables. A Markov chain Monte Carlo algorithm is designed to search simultaneously for the module genes and their linked markers. We show by simulations that this method is much more powerful for detecting true eQTLs and their target genes than traditional QTL mapping methods. We applied the procedure to a data set consisting of gene expression and genotypes for 112 segregants of *S. cerevisiae* and identified modules containing genes mapped to previously reported eQTL hot spots, dissected these large eQTL hot spots into refined modules with biological implications, and discovered a few epistasis modules.

If time permits, I will also discuss a few ideas regarding Bayesian modeling and discovery of interactions among a large number of variables in a classification or regression framework.