Reconstructability Analysis Detects Genetic Variation Associated with Gene Expression

Beth Wilmot, Martin Zwick, Shannon McWeeney

Gene expression is a quantitative trait which varies among individuals and can be affected by polymorphisms in DNA sequence. Integrative analysis of gene expression and genetic variation patterns provides a functional context for discovery of complex trait susceptibility genes. Reconstructability analysis (RA) is an information-theoretic methodology for multivariate categorical data that overlaps log-linear modeling and Bayesian networks. With the simplest RA analysis, namely the calculation of mutual information, we detected associations between SNP alleles and binned gene expression values. We used RA to replicate and extend SNP--gene expression associations in a public data set (Myers et al 2007, Nature Genetics 39:1494). For the purposes of this pilot study, only the data from SNPs and expression from the published positive cis-acting SNP – gene expression associations was included. True epistatic interactions were identified by low linkage disequilibrium (LD) between the SNP pairs and a significant information gain.

RA identified all but two cis-acting SNPs (false negative rate = 2.7%) and detected additional cis-acting SNPs. In addition, RA identified cis-, trans-acting SNP pairs significantly associated with gene expression among these genes, indicating RA may be well suited for detection of epistatic interactions.

Wilmot, B., Zwick, M., and McWeeney, S. (2008). "Reconstructability Analysis Detects Genetic Variation Associated with Gene Expression." Presented at the 12th QTL-MAS Workshop in Computational Genetics, Uppsala, Sweden, May 15-16.