

Data Mining Approaches for Identifying Genetic Variation Affecting Complex Phenotypes

Nancy Cox, Ph.D.
University of Chicago

Abstract:

Many of the approaches developed to date for the fine-mapping and positional cloning studies that will be conducted in the effort to identify genetic variation for complex phenotypes are computationally intensive, and cannot easily be applied to large SNP data sets. The problems inherent in analyzing and interpreting data generated in microarray expression studies have many parallels to the problems that arise in relating SNP genotypes to complex phenotypes, and thus, approaches and software tools developed for the analysis of microarray expression studies can be easily subverted for analysis of SNP data. We have developed a number of approaches for the analysis of SNP data using as a foundation approaches and software tools first used for the analysis of microarray expression data. Results are congruent with more computationally intensive analyses, and suggest that these approaches may be useful in initial stages of data mining.