Summary:

Techniques will be developed for genetic analysis of complex diseases segregating in pedigrees of arbitrary size and structure. Cardiovascular, neurological and behavioral traits are among those having both environmental and genetic components. However, identification of genes contributing to increased risk of related disorders has been limited by both computational and statistical constraints. The development of Markov chain Monte Carlo (MCMC) methods has helped overcome these limitations, and the research now proposed concerns extension of these computational and statistical methods in several areas. Improved methods will be developed for the MCMC analysis of gene identity by descent (ibd) in general pedigrees, given data at a dense genome screen of markers, together with the use of this ibd in analyses of complex traits. The ibd inferred from dense genetic markers permits the combination of information within and among pedigrees, increasing the power and resolution of linkage analyses. The ibd framework also provides for far more computationally efficient trait analyses, enabling models for discrete and quantitative trait phenotypes to be extended to a variety of complex models. Additionally, use of the same MCMC-realized ibd patterns in the analysis of multiple trait models provide measures of trait model robustness, and of confidence and significance of linkage findings. With the availability of sparser genetic marker data on extended pedigrees, together with dense localized data on a subset of individuals, the problems of haployping, genotype imputation, and error detection, assume increasing importance. Again, MCMC-based realization of ibd at sparser locations provides new methods for imputation of genotypes and haplotypes at denser locations, incorporating linkage disequilibrium and models for next-generation sequence data into the analysis of trait data on pedigrees. Methods will be evaluated by analyses on several simulated and real data sets, including pedigrees segregating cardiovascular disease or behavioral disorders. These real data sets include several on which are available genome-wide marker screens, localized multigene haplotypes, and next-generation sequence data. Finally, software implementing these methods will be developed, documented and released.

RELEVANCE (See instructions): Complex traits including cardiovascular, neurological and behavioral phenotypes are related to human health disorders which carry a significant public health burden. Identification of genes contributing to increased disease risk has been limited by trait complexity. Development of methods for genetic analysis offers the potential for finding the genes contributing to risk, and for resolving the interaction of genes and environment.