ATELIER « MÉTHODES STATISTIQUES COMPUTATIONNELLES EN GÉNOMIQUE ET EN BIOLOGIE SYSTÉMIQUE »

18–22 AVRIL 2011

WORKSHOP ON "COMPUTATIONAL STATISTICAL METHODS FOR GENOMICS AND SYSTEMS BIOLOGY"

APRIL 18–22, 2011

Conditional independence structures in the analysis of genetic data in pedigrees and populations

Elizabeth Thompson *

eathomp@u.washington.edu

Mendel's first law, in modern terminology, states that all meioses are independent. Additionally, to a good approximation, the meiosis process of transmission of maternal or paternal genome is Markov across the genome. Gene identity by descent (ibd) underlies all genetically mediated similarities among relatives. The ibd graph, defined among observed individuals and across the genome, specifies the segments of genome shared ibd among these individuals. The ibd graph is a function of of the outcome of the meiosis process in meioses ancestral to these observed individuals. Conditional on the ibd graph at a locus, all data for a trait affected by allelic variation only at this locus on individuals in disjoint components of the graph are independent. Conditional on the ibd graph across the genome, the distributions among relatives for traits affected by allelic variation at distinct loci are independent, in the absence of allelic associations among the loci.

If pedigree relationships among individuals are known, relatively sparse genetic marker data serve to provide realizations of the latent ibd graph. For each realization of the ibd graph, analyses of quantitative trait data may be carried out conditionally on the ibd graph, and the pedigree relationships and genetic marker data are no longer relevant. Efficient methods to determine when ibd graphs are the same, both among realizations and across the genome, are then key to efficient analysis. More remote relationships among members of different pedigrees are usually not known. The ibd resulting from these remote relationships can be estimated using denser genetic marker data and a population-genetic based ibd model. Algorithms to merge the ibd graphs inferred within and among pedigrees provide a combined ibd graph. This combined graph is a sufficient statictic for subsequent trait-data analysis. Combining information among pedigrees via the ibd graph increases both the power and the resolution of mapping of genes contributing to complex quantitative traits.

*Department of Statistics, University of Washington, Seattle, WA 98195-4322, USA.