STOCHASTIC SEARCH GENE SUGGESTION: A BAYESIAN HIERARCHICAL MODEL FOR GENE MAPPING

M. D. Swartz[1], M. Kimmel[2], P. Mueller[1], C. I. Amos[1]

[1]*University of Texas M. D. Anderson Cancer Center, Houston, TX, USA*
[2]*Rice University, Houston, TX, USA*

Email: *mdswartz@mdanderson.org*

Mapping the genes for a complex disease, such as diabetes or rheumatoid arthritis (RA), involves finding multiple genetic loci that may contribute to the onset of the disease. Pairwise testing of the loci leads to the problem of multiple testing. Looking at haplotypes, or linear sets of loci, avoids multiple tests but results in a contingency table with sparse counts, especially when using marker loci with multiple alleles. We propose a hierarchical Bayesian model for case–parent triad data that uses a conditional logistic regression likelihood to model the probability of transmission to a diseased child. We define hierarchical prior distributions on the allele main effects to model the genetic dependencies present in the human leukocyte antigen (HLA) region of chromosome 6. First, we add a hierarchical level for model selection that accounts for both locus and allele selection. This allows us to cast the problem of identifying genetic loci relevant to the disease into a problem of Bayesian variable selection. Second, we attempt to include linkage disequilibrium as a covariance structure in the prior for model coefficients. We evaluate the performance of the procedure with some simulated examples and then apply our procedure to identifying genetic markers in the HLA region that influence risk for RA. Our software is available on the website http://www.epigenetic.org/Linkage/ssgs-public/.