

# **Title: Powerful Multilocus Association Testing of Complex Traits**

Michael Epstein, PhD  
Emory School of Medicine

## **Abstract:**

Association mapping of complex traits typically employs tagSNP genotype data to identify a trait locus within a region of interest. However, considerable debate exists regarding the most powerful strategy for utilizing such tagSNP data for inference. A popular approach tests each tagSNP within the region individually, but such tests could lose power due to incomplete linkage disequilibrium between the genotyped tagSNP and the trait locus. Alternatively, one can jointly test all tagSNPs simultaneously within the region (using genotypes or haplotypes), but such multivariate tests have large degrees of freedom that can also compromise power. Here, we consider a semiparametric model for complex-trait mapping that uses genetic information from multiple tagSNPs simultaneously in analysis but produces a test statistic with reduced degrees of freedom compared to existing multivariate approaches. We fit this model using a dimension-reducing technique called least-squares kernel machines, which we show is identical to analysis using a specific linear mixed model (which we can fit using standard software packages like SAS and R). Using simulated SNP data based on real data from the International HapMap Project, we demonstrate our approach often has superior performance for association mapping of complex traits compared to the popular approach of single-tagSNP testing. Our approach is also flexible, as it allows easy modeling of covariates and, if interest exists, high-dimensional interactions among tagSNPs and environmental predictors.