Detecting Rare Haplotype-Environment Interaction under Uncertainty of Gene-Environment Independence Assumption

Rare variants and gene-environment interactions (GXE) have been suggested in the literature as potential causes for “missing heritability” in complex diseases. We consider the challenging problem of detecting GXE where G is a rare haplotype variant and E is a non-genetic factor. A common assumption made in such analyses is independence of G and E in controls or in target population. As this assumption does not hold in many situations, developing methods that do not make this assumption and yet retains good powers to detect GXE is of important practical interest. To this end, we consider the recently proposed method based on logistic Bayesian LASSO (LBL) for detecting GXE using case-control sample, which assume G-E independence – we refer to it as LBL-GXE-I. It has inflated type I errors when G-E independence assumption is violated. We propose a way of relaxing this assumption by modeling the haplotype frequencies as functions of E through a multinomial logistic regression model – we refer to it as LBL-GXE-D. It overcomes the aforementioned limitation of LBL-GXE-I by being able to control type I error rates in all situations. However, LBL-GXE-D has reduced power than LBL-GXE-I when G-E independence holds. So, to optimize power without sacrificing type I error in any scenario, we propose a unified approach by employing reversible jump Markov chain Monte Carlo method. It allows moves between G-E dependence and independence models of different dimensionalities and thus incorporates uncertainty in G-E independence assumption – we refer to this method as LBL-GXE. Our extensive simulation studies show that LBL-GXE has power similar to that of LBL-GXE-I in case of G-E independence and at the same time has controlled type I errors in all situations. Finally, we analyze a lung cancer dataset and found significant interactions between specific (rather than pooled) rare haplotypes and smoking in presence of G-E dependence for the first time in lung cancer literature.