The Impact of Population Stratification on Commonly Used Statistical Procedures in Populayion-based QTL Association Studies

Zhaohai Li

Department of Statistics, George Washington University

Abstract

Population-based association study using unrelated individuals is a powerful strategy for detecting association between markers and quantitative trait loci (QTLs). However, association test using unrelated individuals may suffer from confounding due to population structure. In this paper, we examine the impact of confounding due to population substructure on commonly used statistical procedures. Two study designs for genetic association study are considered: 1) retrospective sampling of cases and controls according to two cutoff points of the quantitative trait values (high or low) with allele frequency based test statistic; 2) random sample of individuals with regression analysis. For the first design, we consider the impact of confounder on association analysis between markers and QTLs. It is found that the false positive rate (or type I error) could be inflated substantially in the presence of both population stratification and trait heterogeneity; under other situations (e.g., in the presence of population stratification without trait heterogeneity), the inflated false positive rate is relatively minor. For the second design, we consider candidate marker association analysis. It is found that the inflated false positive rate could also be considerable in the presence of both population stratification and trait heterogeneity. Simulation studies confirm the theoretical results.