Accurate Estimation of Heritability in Case-Control GWAS

David Golan

Linear mixed effects models have recently gained popularity as the method of choice for estimating heritability from GWAS data, i.e. quantifying how much of the variability of a phenotype can be explained by the genotyped SNPs.
However, most of the interesting diseases and disorders studied are rare (typically affecting <1% of the population), and so the proportion of cases in a study is usually considerably higher than the proportion of cases in the population. This over-representation of cases invalidates several key assumptions of linear mixed models, e.g. the normality and independence of the random effects. Ignoring these problems results in shrunken estimates of heritability.
We propose an alternative approach for estimating heritability. We derive the relationship between the genetic similarity and the phenotypic similarity of any two individuals as a function of the heritability, while explicitly conditioning on the fact that both individuals were selected for the study. Our method then entails regressing the pairwise phenotypic similarities on the pairwise genetic similarities and using the slope to obtain an estimate of the heritability.
We show, using simulations, that our method yields unbiased estimates which are considerably more accurate than the current state-of-the-art methodology.
Applying our method to several well-studied GWAS yields heritability estimates which are considerably higher than previously published results.