Abstract:

Genome-Wide Association Studies (GWAS) have been successful in discovering thousands of statistically significant, reproducible, genotype-phenotype associations in humans. However, the discovered variants (genotypes) explain only a small fraction of the phenotypic variance in the population for most human traits. In contrast, the heritability, defined as the proportion of phenotypic variance explained by all genetic factors, was estimated to be much larger for those same traits using indirect population-based estimators. This gap is referred to as “missing heritability.”

Mathematically, heritability is defined by considering a function $F$ mapping a set of (Boolean) variables, $(x_1, \ldots, x_n)$ representing genotypes, and additional environmental or “noise” variables $\epsilon$, to a single (real or discrete) variable $z$, representing phenotype. We use the variance decomposition of $F$, separating the linear term, corresponding to additive (narrow-sense) heritability, and higher-order terms, representing genetic-interactions (epistasis), to explore several explanations for the “missing heritability” mystery. We show that genetic interactions can significantly bias upwards current population-based heritability estimators, creating a false impression of “missing heritability.” We offer a solution to this problem by providing a novel consistent estimator based on unrelated individuals. We also propose novel estimators for the different variance components (beyond additive) of heritability from GWAS data.

Finally, we use the Wright–Fisher process from population genetic theory to study the relative contributions of rare and common variants to heritability.