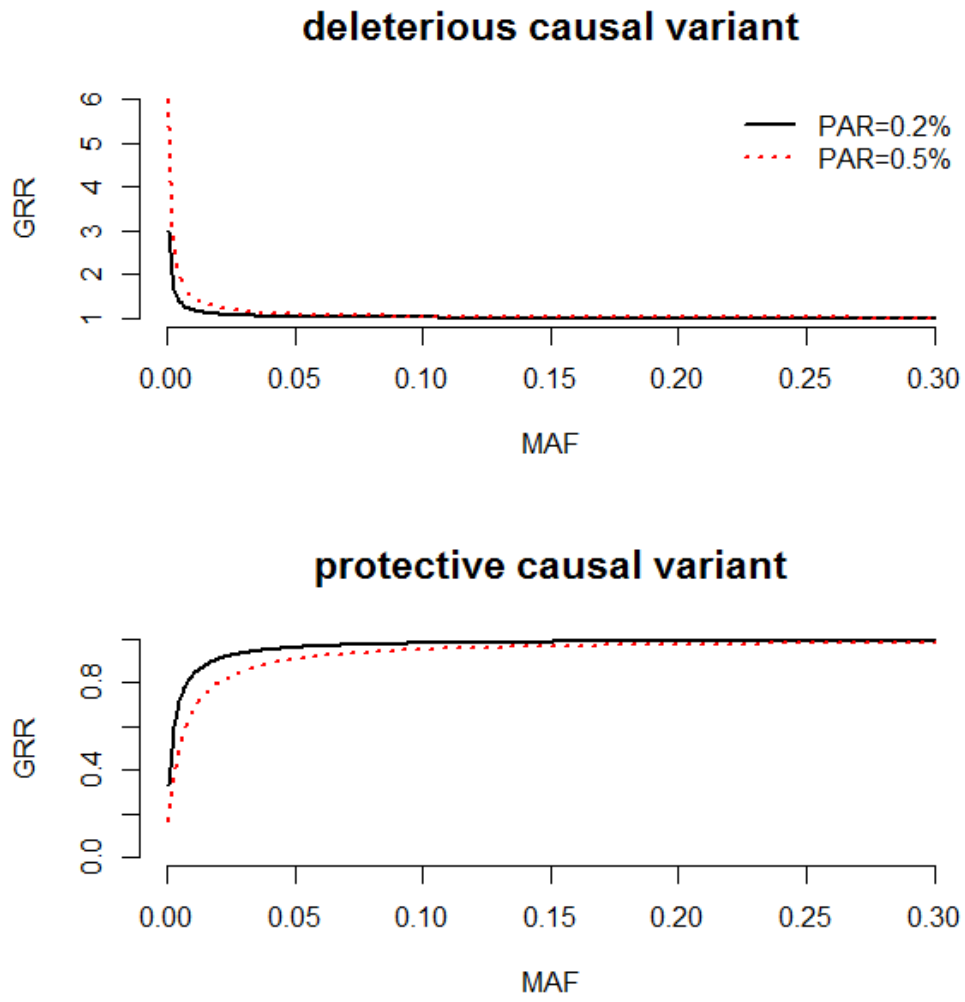
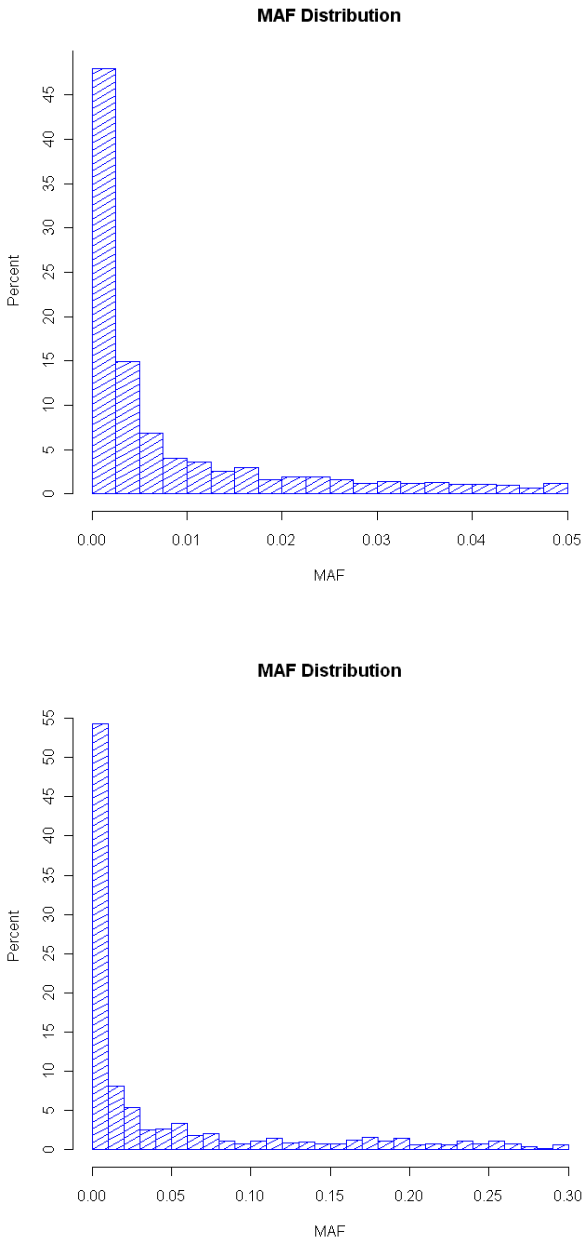
**Supporting Information**

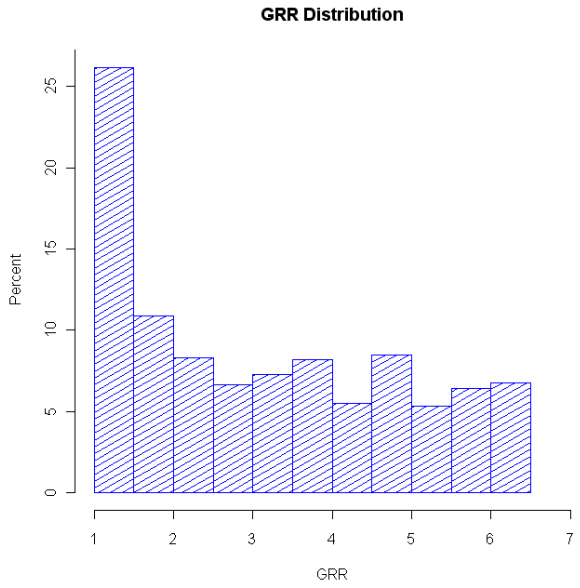
# Haplotype Kernel Association Test as a Powerful Method to Identify Chromosomal Regions Harboring Uncommon Causal Variants

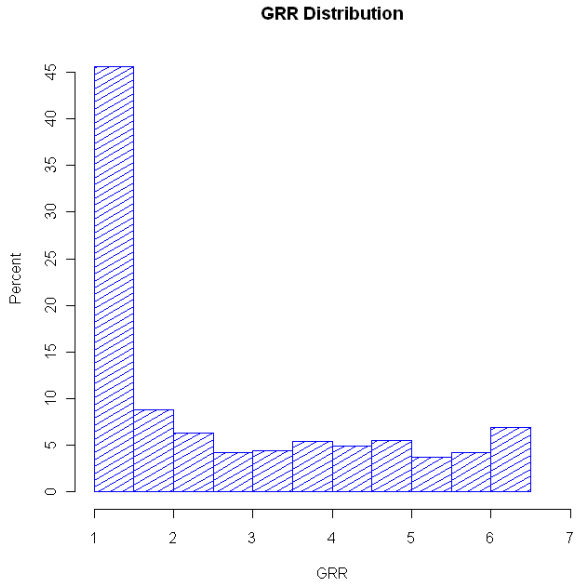



Supplementary Figure S1. The relationship between minor allele frequency (MAF) and genotype relative risk (GRR) in simulations for dichotomous traits, given the population attributable risk (PAR) = 0.5% or 1%.

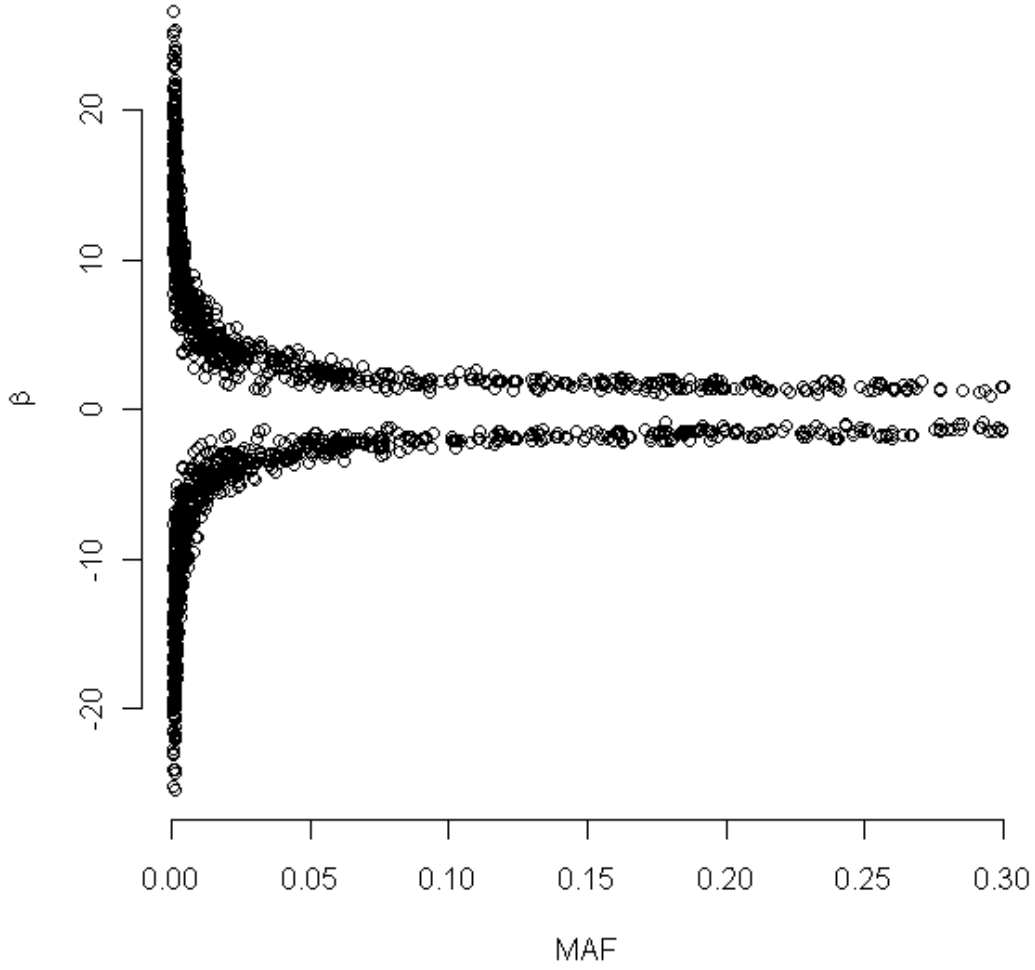



Supplementary Figure S2. The distribution of the minor allele frequencies (MAFs) of the causal variants in our 200 simulated data sets. The top row is for the scenario that causal variants have MAFs $\in [0.1\%,\ 5\%]$, and the bottom row is for causal variants with MAFs $\in [0.1\%,\ 30\%]$.

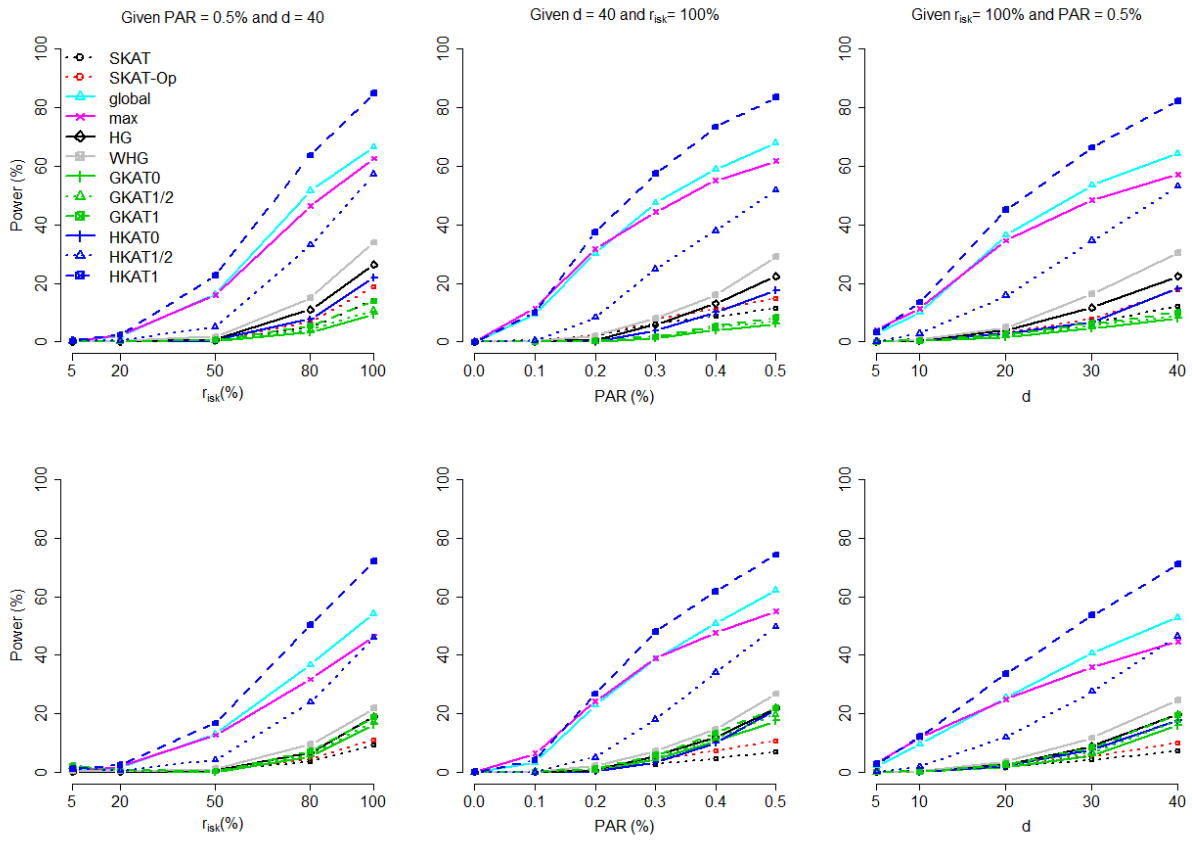



Supplementary Figure S3. The distribution of the genotype relative risks (GRRs) of the deleterious causal variants for dichotomous traits, given PAR = 0.5% for each causal variant. The top row is for the scenario that causal variants have MAFs $\in [0.1\%,\ 5\%]$, and the bottom row is for causal variants with MAFs $\in [0.1\%,\ 30\%]$.

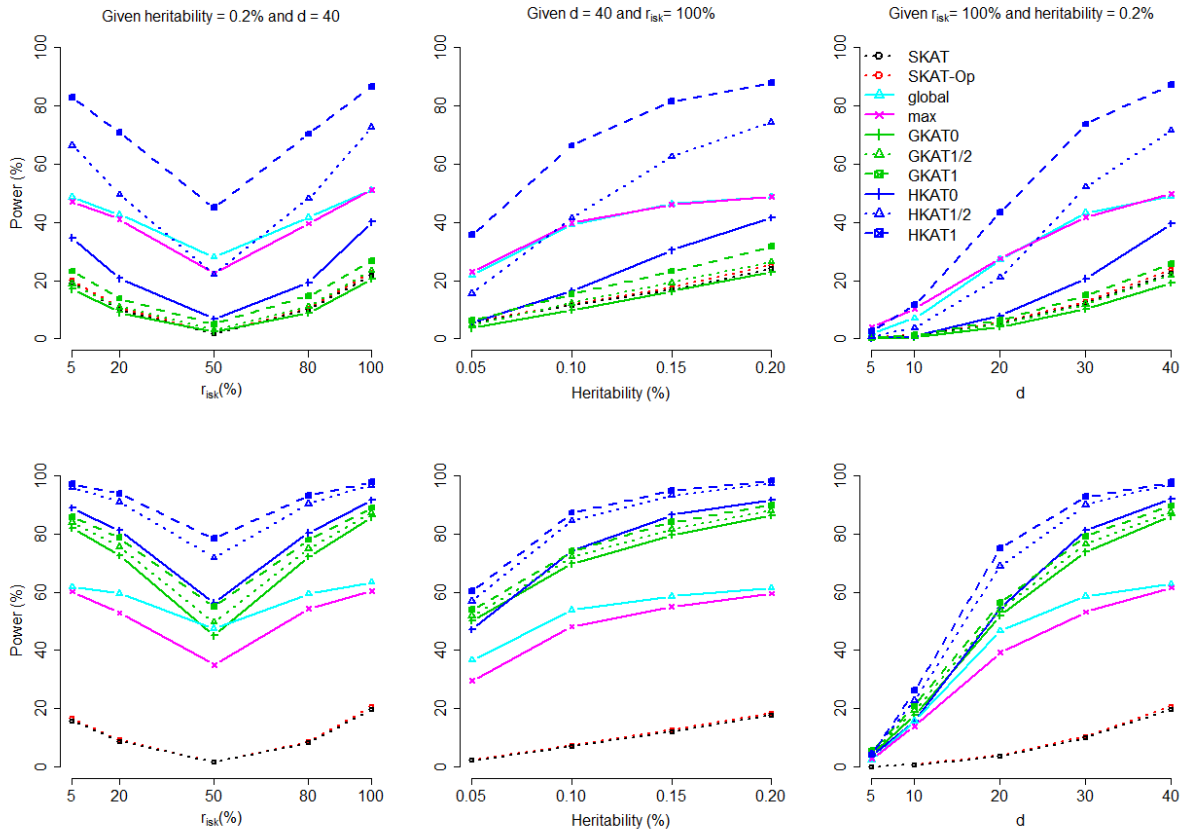



Supplementary Figure S4. The relationship between minor allele frequencies (MAFs) and $\beta_j$'s [in Eq. (9)] when simulating the continuous traits, given the marginal heritability = 0.2%, *d* = 40, and $r_{isk}$ = 50%. The average of $V_e$ is 307.04.




**Supplementary Figure S5 - Dichotomous trait - Comparison of power by $r_{isk}$ (the percent of deleterious variants among the $d$ causal variants), PAR, and $d$ (the number of causal variants)**

The figure shows the power comparison by $r_{isk}$ (the left column, given PAR = 0.5% and $d$ = 40), PAR (the middle column, given $d$ = 40 and $r_{isk}$ = 100%), and $d$ (the right column, given $r_{isk}$ = 100% and PAR = 0.5%), respectively. The nominal significance level was set at $10^{-4}$. The top row is the result given 'uncommon' causal variants with MAFs $\in [0.1\%,\ 5\%]$; the bottom row is the result given 'uncommon + common' causal variants with MAFs $\in [0.1\%,\ 30\%]$.




**Supplementary Figure S6 - Continuous trait - Comparison of power by $r_{isk}$ (the percent of variants among the *d* causal variants that increase the trait value), the marginal heritability, and *d* (the number of causal variants)**

The figure shows the power comparison by $r_{isk}$ (the left column, given the marginal heritability = 0.2% and *d* = 40), the marginal heritability (the middle column, given *d* = 40 and $r_{isk}$ = 100%), and *d* (the right column, given $r_{isk}$ = 100% and the marginal heritability = 0.2%), respectively. The nominal significance level was set at $10^{-4}$. The top row is the result given 'uncommon' causal variants with MAFs $\in [0.1\%,\ 5\%]$; the bottom row is the result given 'uncommon + common' causal variants with MAFs $\in [0.1\%,\ 30\%]$.