The American Journal of Human Genetics, Volume *87*

Supplemental Data

To Identify Associations with Rare Variants,

Just WHalT: *W*eighted *H*aplotype *a*nd *I*mputation-Based *T*ests

Yun Li, Andrea E. Byrnes, and Mingyao Li

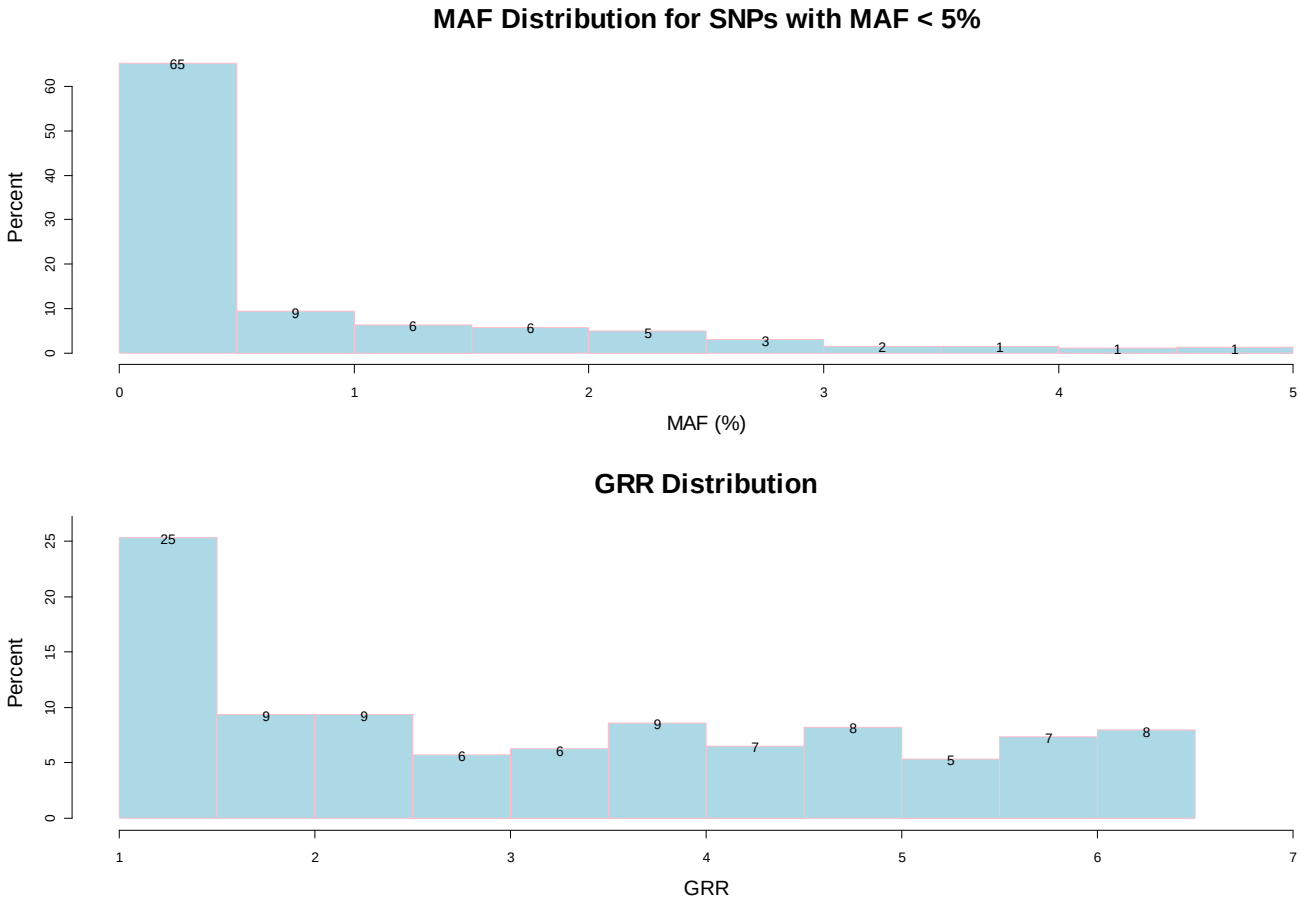

Figure S1. MAF and GRR Distribution for Simulated Rare Variants (MAF < 5%)

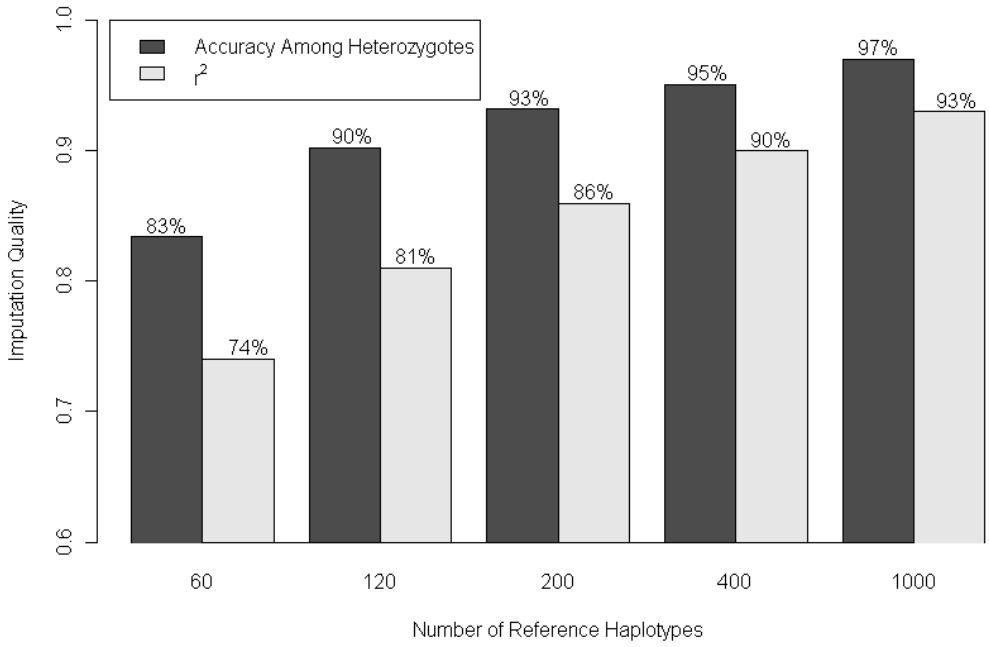

Figure S2. Imputation Quality for Rare Variants (FUSION Data)
We constructed a set of 1,000 reference haplotypes encompassing both common (MAF >= 5%) and rare (MAF < 5%) SNPs using 500 FUSION individuals. We masked and imputed genotypes at rare (MAF < 5%) SNPs for an independent set of 500 FUSION individuals using a subset of 60, 120, 200, 400, or the full set of 1,000 haplotypes. We calculated two measures of imputation quality by comparing imputed genotypes with their mask experimental counterparts: allelic accuracy among the heterozygous genotypes, and r^2. r^2 is defined as the squared correlation between imputed genotype dosages (ranging from 0 to 2) and the experimental dosages (taking values 0, 1, or 2).