Figure S1. Quality Control (QC) in HostSeq. Flowchart describing the multi-step process of sample and variant QC of joint-called HostSeq data. N = 8,474 / 10,059 samples were retained for genetic analysis. PCA was performed on a subset of variants; these PCs are used as covariates in genetic analysis. HWE was performed on the subset of controls with European ancestry [N = 3,876], and variants with P < 1E-50 were removed from all samples.