**Supplementary Methods**

**1. Methods for enrichment test**

Functional variants have direct impact on gene expression and function. Putative set of functional variants containing polymorphisms with direct impact on the gene expression and function that is related to the phenotype are expected to be enriched in association signals. We test enrichment in low p-values of functional variants using two complementary statistics: i) Simes ([Levin, 1996](#_ENREF_49)) test and ii) a VEGAS-like ([Liu et al., 2010](#_ENREF_53)) sum of squares test (SST). Intuitively, by using univariate p-values, SIMES is useful to detect enrichment in groups of variants having a few, but strong, signals. By summing the individual SNPs square statistics, SST is in fact a multivariate statistic. Thus, SST should be more powerful to detect enrichment in a set of variants having many signals of small-to-moderate effect. It is well known that Simes tests can be rather conservative for SNPs in high linkage disequilbrium. However, for testing enrichment in a set of variants (e.g. promoter and methylation SNPs) with a relatively small number of SNPs spread over many regions in the genome, the conservativeness of this test should be minimal.

GWAS of large studies, e.g. PGC([2014](#_ENREF_1)), are known to yield a fair number of statistically significant signals. Besides these significant signals, GWAS also harbor many small and moderately large signals spread over the entire genome([Purcell et al., 2009](#_ENREF_73)). Thus, the statistics from a large GWAS, are already enriched in small p-values. Consequently, to be considered candidate to containing causal variants, a putative set of variants should be enriched for association signals “above” the background enrichment of the GWAS. For the Simes test, this object is simply achieved by substituting the relative ranks of the absolute statistics for the GWAS p-values. Assume are the univariate statistics from the GWAS. Let be the rank of taken in decreasing order. Then, the relative rank “p-values” are and their relevant subset is used in deriving the background adjusted Simes test for the putative set of variants.

The adjustment for background enrichment of SST is much more complex that the one for the Simes test. Let SST statistic for the entire genome be . It follows that , i.e. SST statistics behave like the sum of non-central variables having non-centrality parameter . Based on the scan statistics, the non-centrality pararmeter is estimated as . Consequently, under the null hypothesis of no enrichment above the GWAS background, the distribution of the square of a univariate scan statistic should be assumed to be , not a central . Assume that the statistics associated with SNPs in the putative set of variants are the first statistics in the scan, i.e. . Then the SS statistic associated with this set of variants is . If the number of SNPs is large enough for the Central Limit Theorem to provide a reasonable distributional assumption, under the null hypothesis of no enrichment above GWAS background, is approximately a normal variable with mean and variance , i.e. . To test for enrichment above background in the putative set of variants, we compute the normally distributed statistic and test vs. . The only unknown is which can be estimated with reasonable accuracy from LD reference genotype data, e.g. European subjects from the 1000 Genomes ([Joober, 2011](#_ENREF_39)) available in Mach ([Li et al., 2010](#_ENREF_51)) and Impute ([Williams et al., 2012](#_ENREF_94)) databases. However, most tested set of variants do not have the very large number of SNPs needed to ensure that the Central Limit Theorem provides a good approximation for the distribution of the enrichment statistic,. Consequently, we use 50,000 simulations to assess the statistical significance of under a more restrictive "competitive" , i.e. we test for an enrichment beyond the enrichment of the remaining SNPs in the GWAS (or the average GWAS enrichment). For these simulations we assume that i) the LD structure between selected SNPs is the one estimated from the European subjects of 1000 Genomes project and ii) the distribution of the square statistics is a non-central chi-square ( ), i.e. it is not central chi-square as used in a more liberal self contained . In more detail, similar to VEGAS ([Liu et al., 2010](#_ENREF_53)) , the statistics associated with SNPs on a chromosome are assumed to have a multivariate normal distribution with the variance matrix equal to the correlation between SNPs' genotypes.

**2. Additional Quality Control and Imputation of SMRI genotypes.**

Rather than relying on the QC procedures employed in each of the original SMRI GWAS and miRNA expression studies ([Kim and Webster, 2010](#_ENREF_44), [Kim et al., 2010](#_ENREF_43)), we chose to apply additional QC steps that eliminated all SNPs and subjects with less than 100% genotyping efficiency for the SNPs we were studying. SNPs were not considered if they were not represented by each of the three allelic states, and a minor allele frequency < =0.02 and if they were located in probes used in the SMRI miRNA expression study.

Additionally, in order to increase the GWAS resolution of the SMRI GWAS panel and to make the SMRI platform comparable to the platform utilized by the PGC group, we imputed the original SMRI genotypes using IMPUTE([Howie et al., 2009](#_ENREF_35)) and Shape-IT ([O'Connell et al., 2014](#_ENREF_64)). The original SMRI GWAS panel was the Affymetrix Genome-wide Human SNP Array 5.0 which contained 900,568 SNPs whereas the GWAS panel used by the PGC consortium was the Affymetrix 6.0 array that contains > 1.8 million SNPs.

Briefly, imputation was performed using a heterogeneous reference panel (1000 genomes, phase 1) for genomic regions of 5MB with overlap interval of 500Kb between adjacent segments. Imputed SNPs having aninfo metric> 0.7 and a confidence interval > 0.95 were retained for eQTL evaluation by Matrix EQTL.