Description for Supplementary Materials

Folder “completetableRA” contains the 319 epistatic interactions discovered based on WTCCC data in 319 files “completetable001\_DRR.txt” – “completetable319\_DRR.txt”. The columns are

*epist*: index of epistatic interactions.

*dip*: diplotype, where each number 0, 1 or 2 represents the copy number of second allele at the corresponding SNP in the SNP annotation file.

*control1*: diplotype frequency in controls from 1958 Birth Cohort (58C).

*control2*: diplotype frequency in controls from National Blood Service (NBS).

*case*: diplotype frequency in cases.

*pval.control12*: Fisher’s exact test P-values for this diplotype between two control groups.

*pval.case\_control*: Fisher’s exact test P-values for this diplotype between cases and controls.

*drr*: disease relative risk for the corresponding diplotype.

Folder “SNPinfoRAchr” contains the SNP annotations of 319 epistatic interactions corresponding to the files in folder “completetableRA”. The columns are: SNP id, SNP name, location coordinate, first allele, second allele, and chromosome.

File “bestEpistasesFromWTCCC.txt” lists all significant interactions with P-values less than 4.03e-11 (conservative Bonferroni adjustment at 0.05 level for ~1.24e9 possible diplotypes of the 319 interactions) that have diplotype frequencies larger than 0.05 for either controls or cases (in order to get stable results). The rows were sorted by decreasing genetic disease relative risk.

File “bestEpistasesFromWTCCC\_SNPs.txt” lists the SNP annotations for the epistatic interactions in file “bestEpistasesFromWTCCC.txt”.

Folder “completetableRA\_NARAC” contains files “NARAC\_completetableX.txt”, which correspond to supplementary files “completetableX.txt”, if at least two NARAC SNPs are available for this interaction. The first row gives the NARAC SNPs involved in current interaction.

File “bestEpistasesFromWTCCC\_NARAC.txt” lists the significant NARAC-validated interactions (P-values < 2.94e-5, Bonferroni adjustment at 0.05 level based on 1701 possible diplotypes of these interactions) that also have P-values <1e-4 for WTCCC data, and the diplotypes frequencies > 0.05 in either cases or controls for both NARAC and WTCCC data (after marginalization to the same diplotype as that of NARAC SNPs). The columns are:

*Epist*: index of epistatic interactions.

*dip.N*: diplotype of the NARAC-SNP group.

*control.N*: diplotype frequency in NARAC controls.

*case.N*: diplotype frequency in NARAC cases.

*pval.case\_control.N*: Fisher’s exact test P-values for the diplotype between cases and controls based on NARAC data.

*drr.N*: disease relative risk for the diplotype based on NARAC data.

*control1.W*: frequency of this diplotype in WTCCC controls from 1958 Birth Cohort (58C).

*control2.W*: frequency of this diplotype in WTCCC controls from from National Blood Service (NBS).

*case.W*: frequency of this diplotype in WTCCC cases

*pval.control12.W*: Fisher’s exact test P-values for the this diplotype between two WTCCC control groups.

*pval.case\_control.W*: Fisher’s exact test P-values for the this diplotype between WTCCC cases and controls.

*drr.W*: disease relative risk for this diplotype based on WTCCC data.

File “bestEpistasesFromWTCCC\_NARAC\_SNP.txt” lists the SNP annotations for the epistatic interactions in file “bestEpistasesFromWTCCC\_NARAC.txt”.