

**SUPPLEMENT: Gene- and pathway-based association tests for multiple traits
with GWAS summary statistics**

Il-Youp Kwak*, Wei Pan*

*Division of Biostatistics, School of Public Health,
University of Minnesota, Minneapolis, MN 55455

April 29, 2016

Table S1 The numbers of the genome-wide significant genes identified by single SNP–single trait analysis (Single) and the MTaSPUsSet test (MT) for the GIANT data.

		Single					
		Men			Women		
		Significant	Not	Total	Significant	Not	Total
MT	Significant	53	28	81	85	40	125
	Not	9	66	75	12	19	31
	Total	62	94	156	97	59	156

Table S2 The p-values of 17 significant genes identified by MGAS for men and women in the GIANT data. * marks out genome-wide significant genes based on the corresponding thresholds for MGAS and MTaSPUsSet. ^M and ^W mark out the significant genes identified by both MGAS and MTaSPUsSet for men and women respectively.

Gene	Men		Women	
	MGAS	MTaSPUsSet	MGAS	MTaSPUsSet
LCORL ^{M,W}	1.65e-9*	<1e-7*	9.55e-14*	<1e-7*
VTA1 ^W	2.37e-7*	1.5e-6*	1.47e-15*	<1e-7*
BICD2 ^W	3.24e-7*	3.9e-6	5.35e-7*	4e-7*
RASA2 ^{M,W}	4.49e-7*	<1e-7*	2.43e-6*	7e-7*
NDUFS3 ^M	5.56e-6*	4e-7*	2e-4	<1e-7*
PSMD1 ^M	7.35e-6*	1.1e-6*	.01456	.000298
HTR2B	7.35e-6*	9.5e-6	.01456	.106
GNA12 ^W	1.91e-5	<1e-7*	2.89e-12*	<1e-7*
NCOA1 ^W	.00729	.0009	3.89e-9*	<1e-7*
TNS1 ^W	.00063	.0102	3.4e-8*	<1e-7*
CEP112 ^W	.00027	1e-6*	1.2e-7*	<1e-7*
ADAMTSL3	.393	.527	2.13e-7*	5e-6
LPGAT1	.0239	.0071	5.45e-7*	0.0608
FAM65C	.00194	.0011	2.52e-6*	1.97e-5
DNM3 ^W	.00684	.0004	2.91e-6*	<1e-7*
KIF27	.055	.228	3.45e-6*	1.3e-5
RFWD2 ^W	.00752	.116	5.88e-6*	5e-7*

$$\mathbf{P} = \begin{bmatrix} 1.00 & 0.00 & 0.19 & 0.32 & 0.67 & 0.08 \\ 0.00 & 1.00 & 0.02 & 0.00 & 0.03 & -0.04 \\ 0.19 & 0.02 & 1.00 & 0.29 & 0.22 & -0.02 \\ 0.32 & 0.00 & 0.29 & 1.00 & 0.32 & 0.34 \\ 0.67 & 0.03 & 0.22 & 0.32 & 1.00 & 0.04 \\ 0.08 & -0.04 & -0.02 & 0.34 & 0.04 & 1.00 \end{bmatrix} \quad (1)$$

$$\mathbf{P}_m = \begin{bmatrix} 1.00 & -0.02 & 0.12 & 0.27 & 0.56 & 0.07 \\ -0.02 & 1.00 & 0.00 & -0.01 & 0.08 & 0.00 \\ 0.12 & 0.00 & 1.00 & 0.27 & 0.16 & -0.01 \\ 0.27 & -0.01 & 0.27 & 1.00 & 0.25 & 0.34 \\ 0.56 & 0.08 & 0.16 & 0.25 & 1.00 & 0.03 \\ 0.07 & 0.00 & -0.01 & 0.34 & 0.03 & 1.00 \end{bmatrix} \quad (2)$$

$$\mathbf{P}_w = \begin{bmatrix} 1.00 & -0.01 & 0.19 & 0.30 & 0.66 & 0.03 \\ -0.01 & 1.00 & -0.01 & -0.01 & 0.04 & -0.01 \\ 0.19 & -0.01 & 1.00 & 0.29 & 0.21 & -0.02 \\ 0.30 & -0.01 & 0.29 & 1.00 & 0.29 & 0.32 \\ 0.66 & 0.04 & 0.21 & 0.29 & 1.00 & 0.02 \\ 0.03 & -0.01 & -0.02 & 0.32 & 0.02 & 1.00 \end{bmatrix}. \quad (3)$$

Figure S1 Equation (1) shows estimate for \mathbf{P} used in Simulation study. Equation (2) is estimate for \mathbf{P}_m , correlation among traits for GIANT men. Equation (3) is estimate for \mathbf{P}_w , correlation among traits for GIANT men.

Figure S2 QQ plots of the MTaSPUsSet p-values from a Control-Control experiment of the WTCCC data. Different reference panels are used to estimate \mathbf{R} : 1) all 2938 WTCCC control samples for the ideal case, 2) a random subset of 100 WTCCC control samples, and 3) the 1000 Genome Project CEU samples.

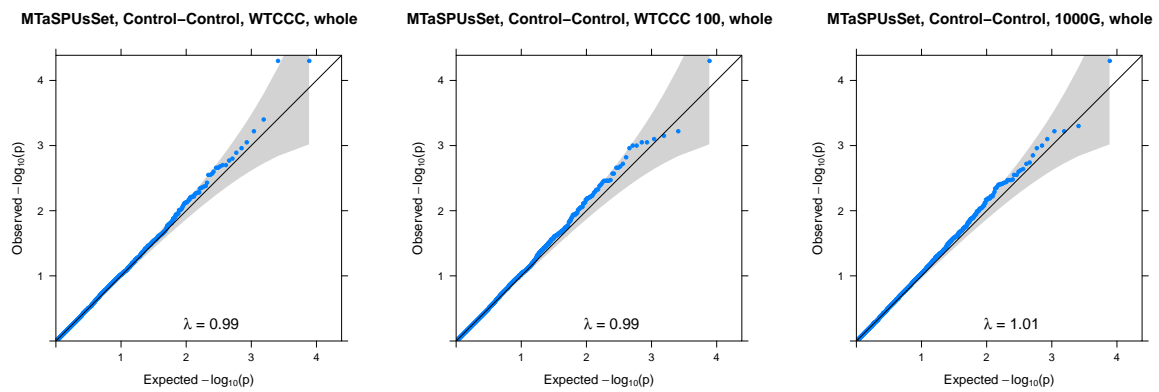


Figure S3 QQ plots of the metaCCA p-values from a Control-Control experiment of the WTCCC data. Six reference panels were used to estimate R : 1) the 1000 Genome Project CEU samples, 2) 100 WTCCC control samples, 3) 500 WTCCC control samples, 4) 1000 WTCCC control samples, 5) 2000 WTCCC control samples, and 6) the whole 2938 WTCCC control samples.

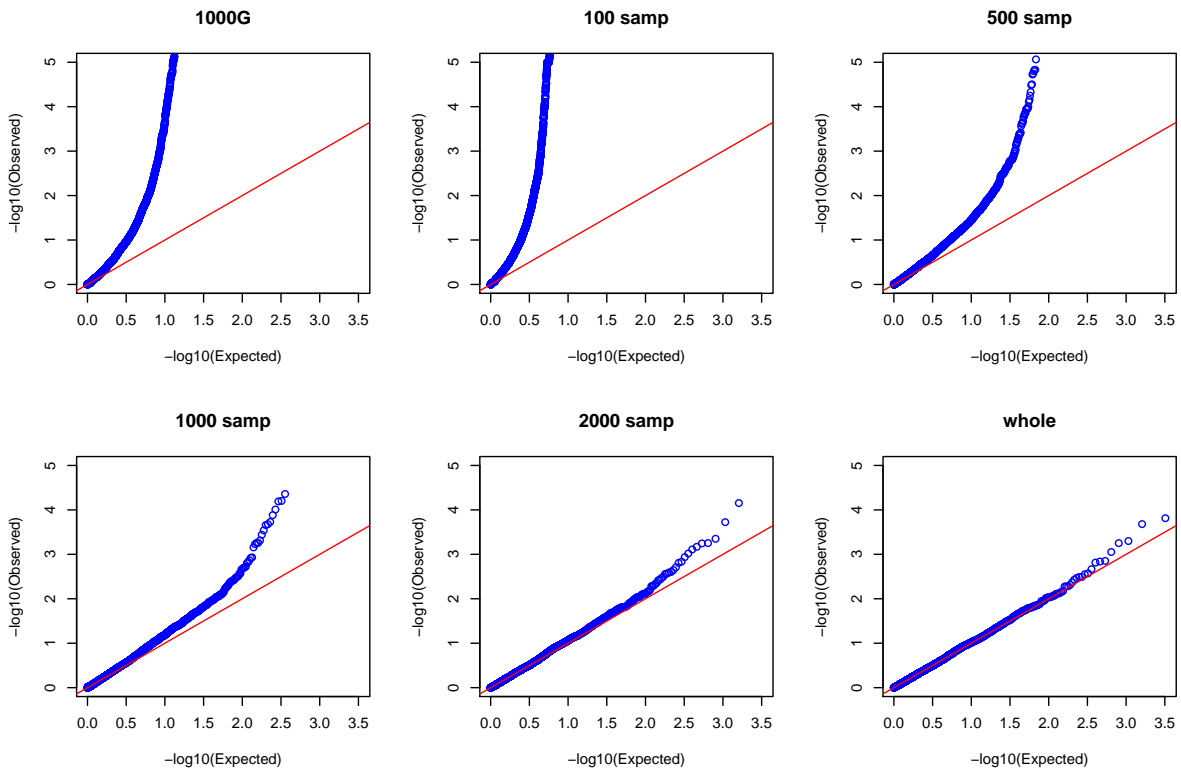


Figure S4 QQ plots of the T_{Sco} p-values from a Control-Control experiment of the WTCCC data. Six reference panels were used to estimate \mathbf{R} : 1) the 1000 Genome Project CEU samples, 2) 100 WTCCC control samples, 3) 500 WTCCC control samples, 4) 1000 WTCCC control samples, and 5) the whole 2938 WTCCC control samples.

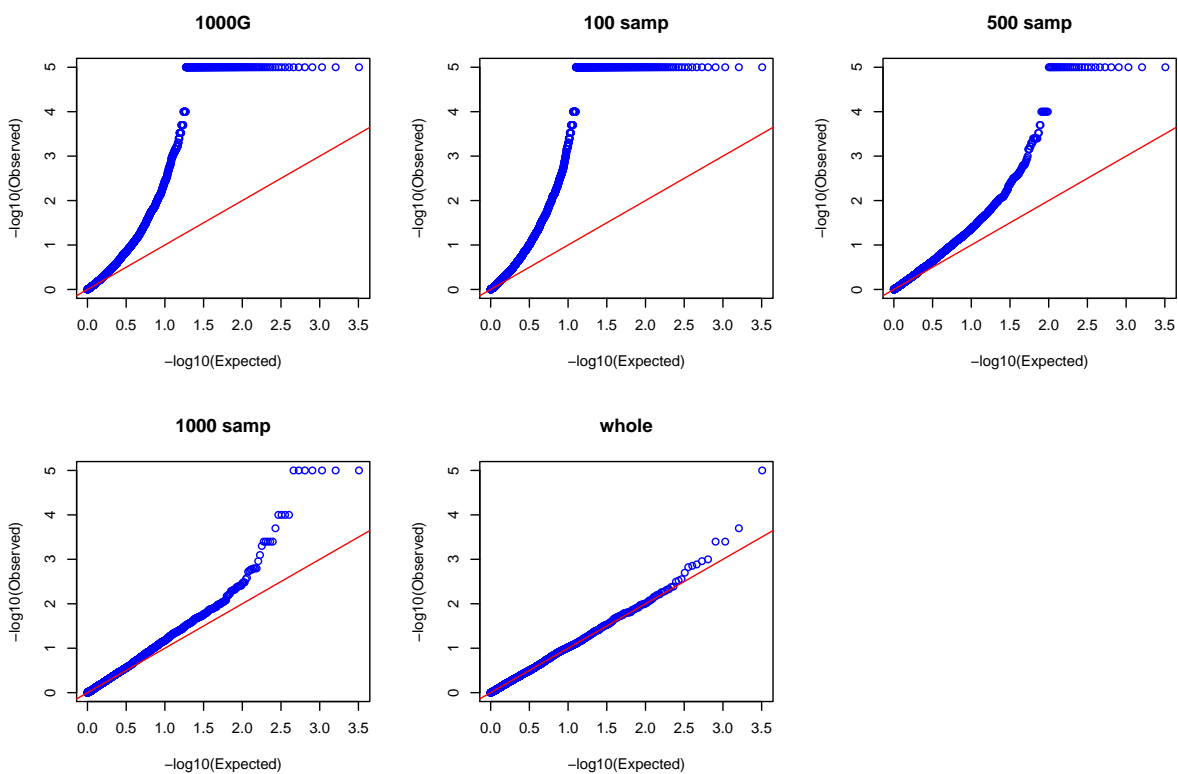


Figure S5 Comparison of the p-values of metaCCA versus T_{SCO} from a Control-Control experiment of the WTCCC data. The whole 2938 WTCCC control samples were used as the reference panel.

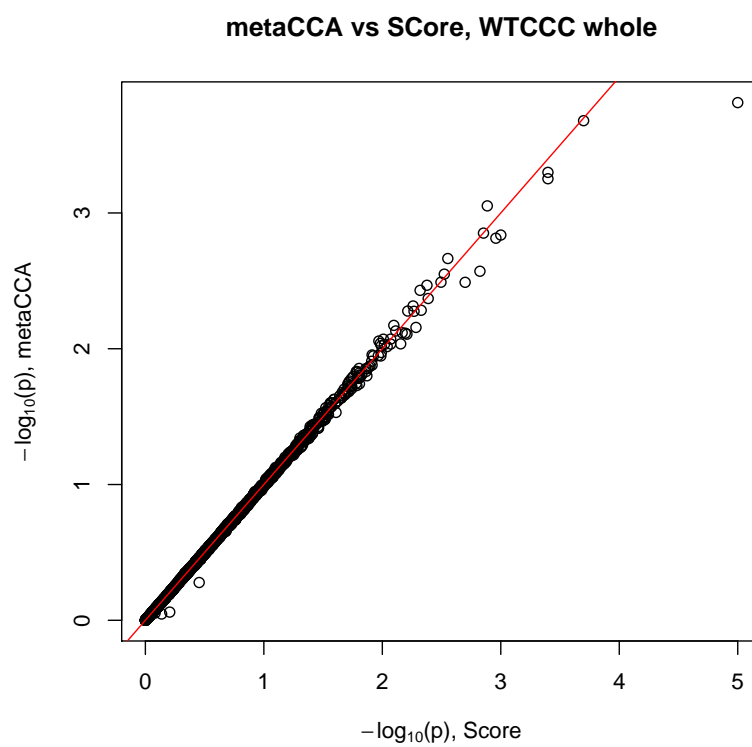


Figure S6 Comparison of the (log-transformed) p-values of MTaSPUsPath using various reference panels and that of GEE-aSPUpaht using individual-level data.

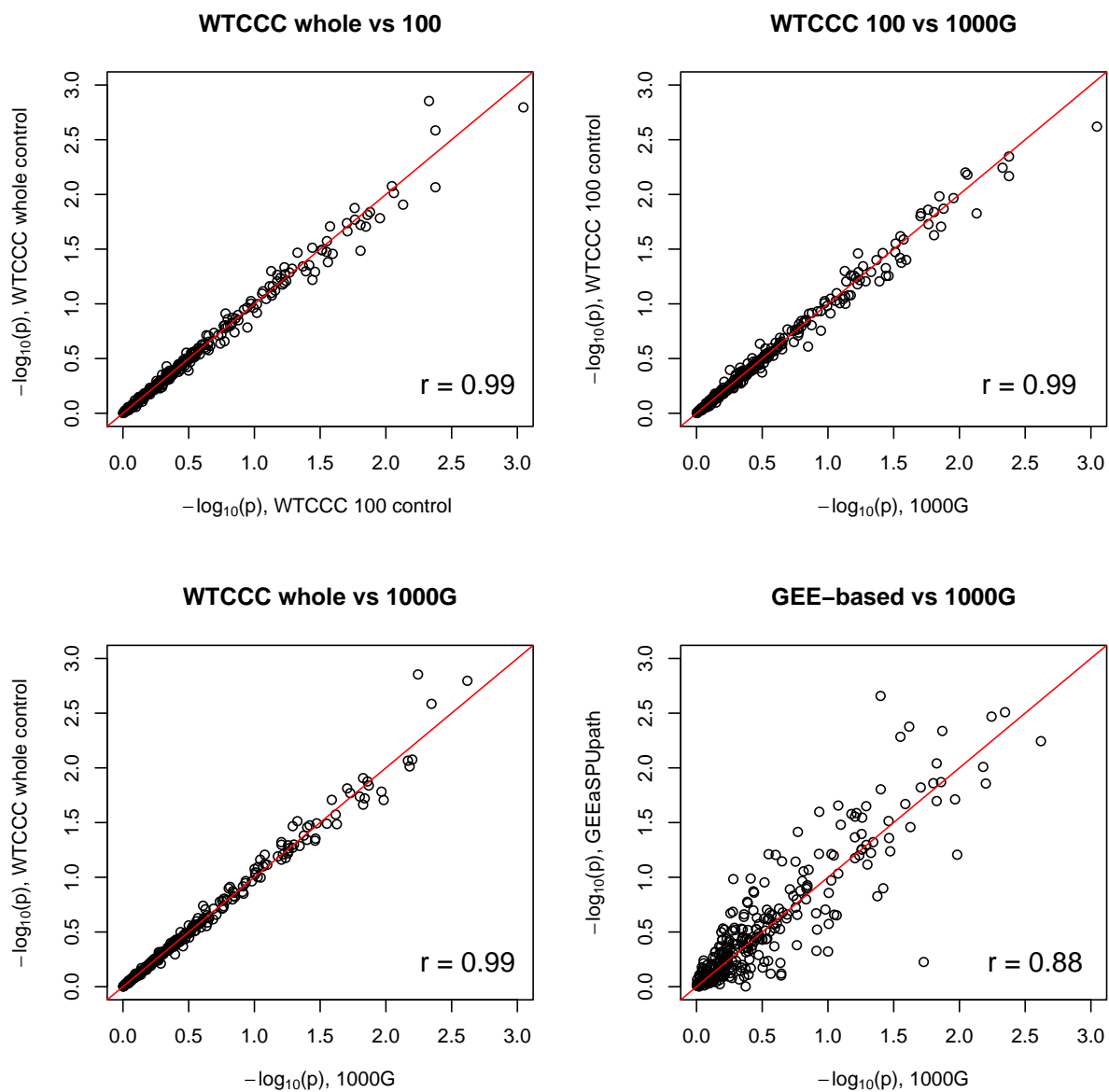


Figure S7 Manhattan Plots for GIANT data using MTaSPUsSet, Men(upper) and Women(lower).

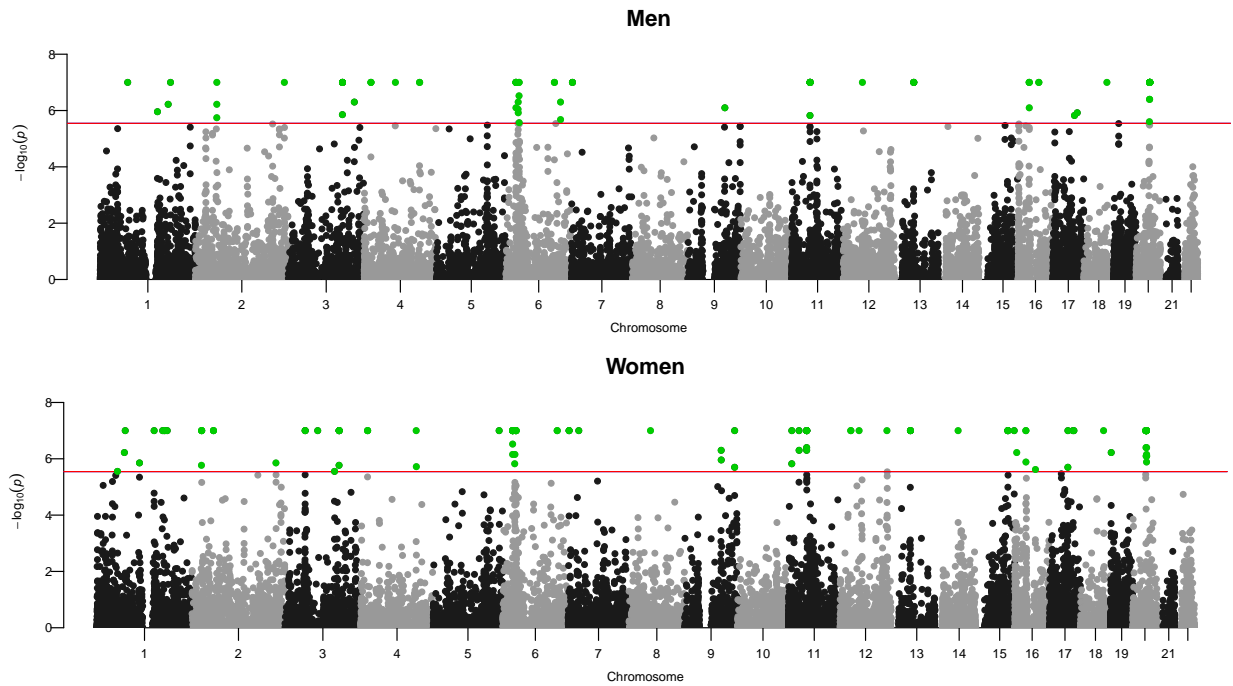


Figure S8 Manhattan Plots for GIANT data using MGAS, Men(upper) and Women(lower).

