The American Journal of Human Genetics, Volume *101*

Supplemental Data

Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits

Huwenbo Shi, Nicholas Mancuso, Sarah Spendlove, and Bogdan Pasaniuc

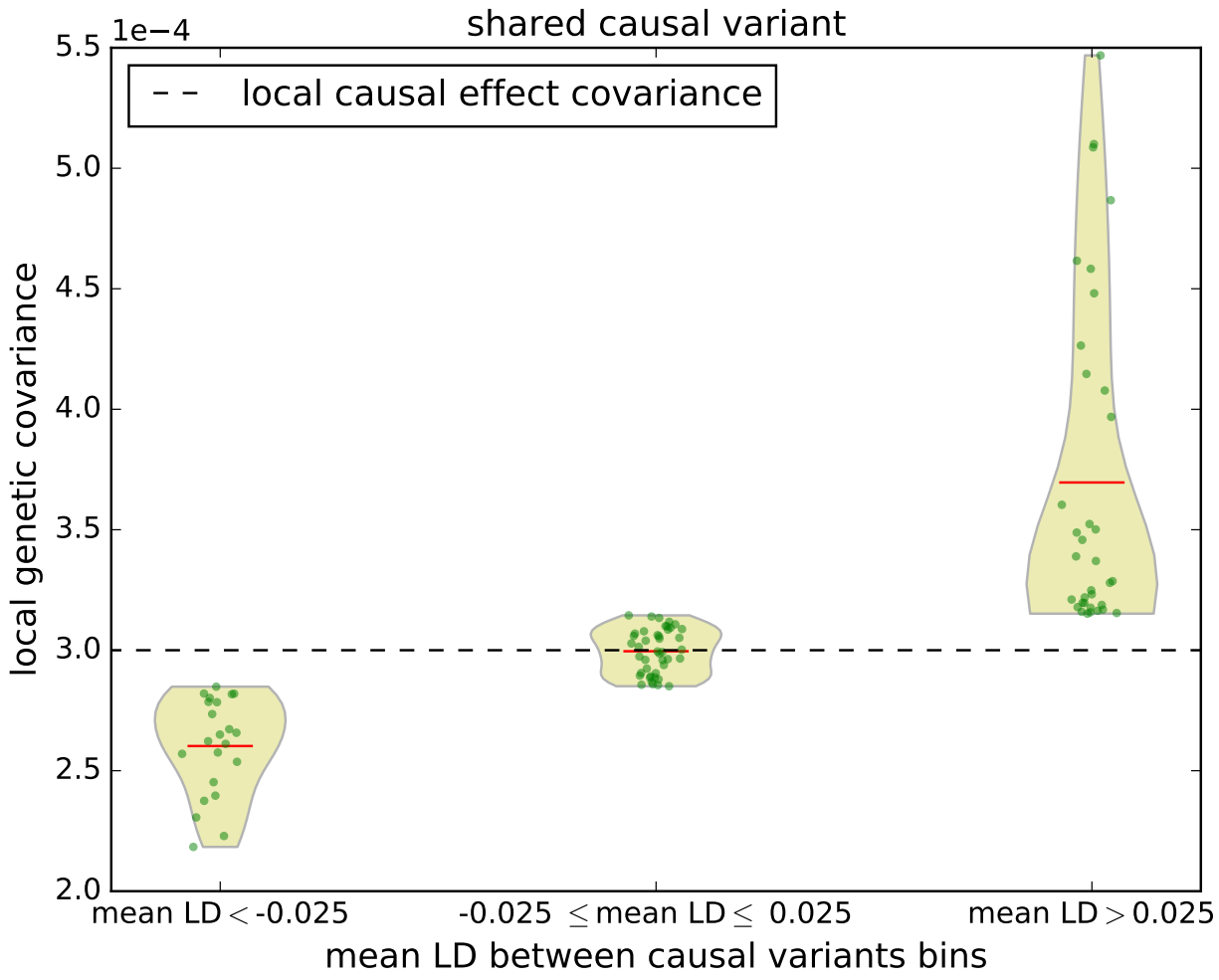

Figure S1: **Distribution of simulated local genetic covariance and causal effect covariance across 100 LD-independent regions on chromosome 1 binned by average LD between causal variants.** The red lines represent the average local genetic covariance in each bin. Both traits each have 3 causal variants with effect size set to 0.01, and share all the causal variants. Here, local genetic covariance varies with respect to LD whereas local causal effect covariance is fixed at 0.0003.

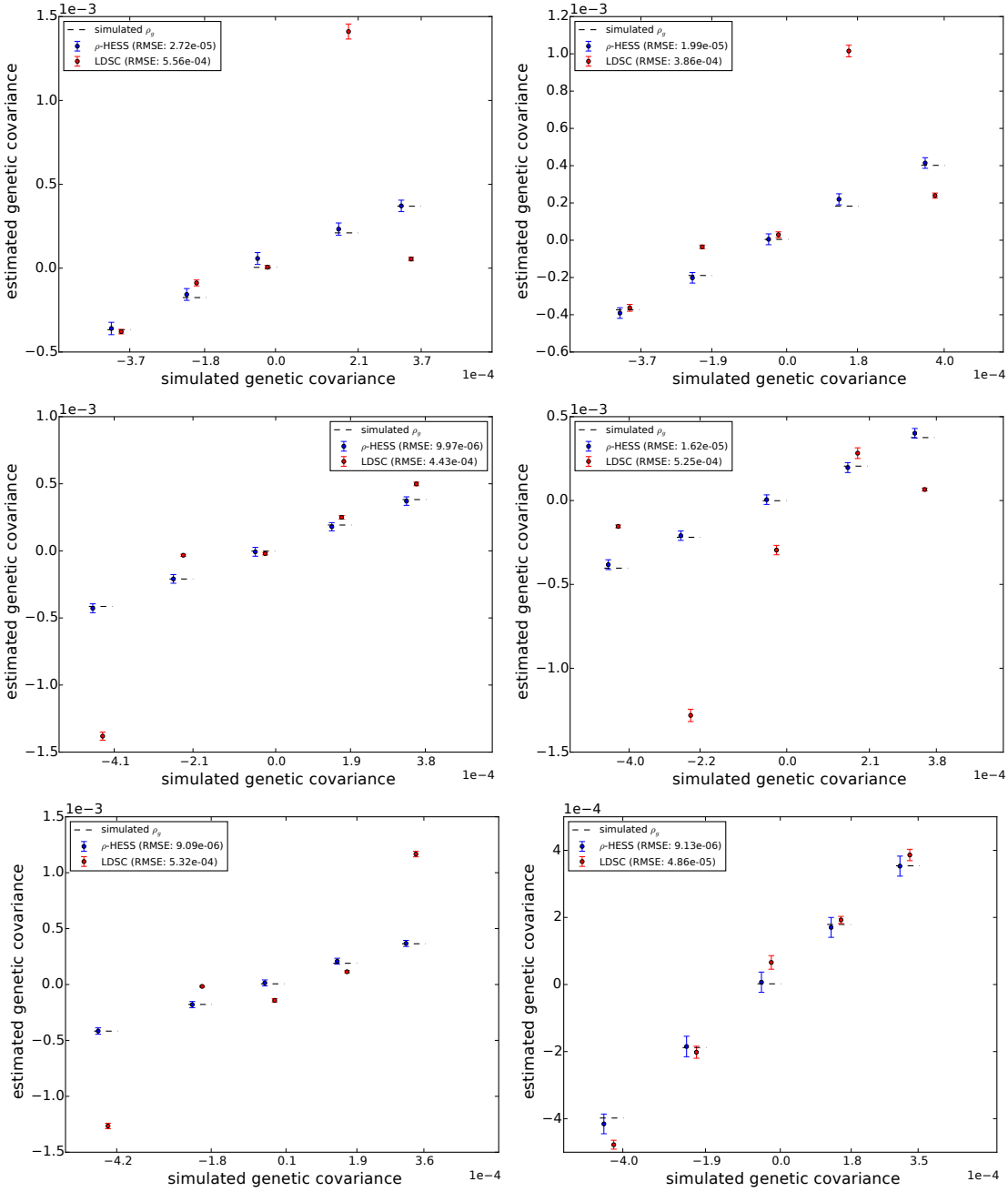

Figure S2: **Performance of ρ-HESS and cross-trait LDSC in estimating local genetic covariance at 6 randomly selected LD-independent loci.** ρ-HESS obtains unbiased estimates of local genetic covariance ρ_g in simulations when in-sample LD matrix is available. Here, the mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

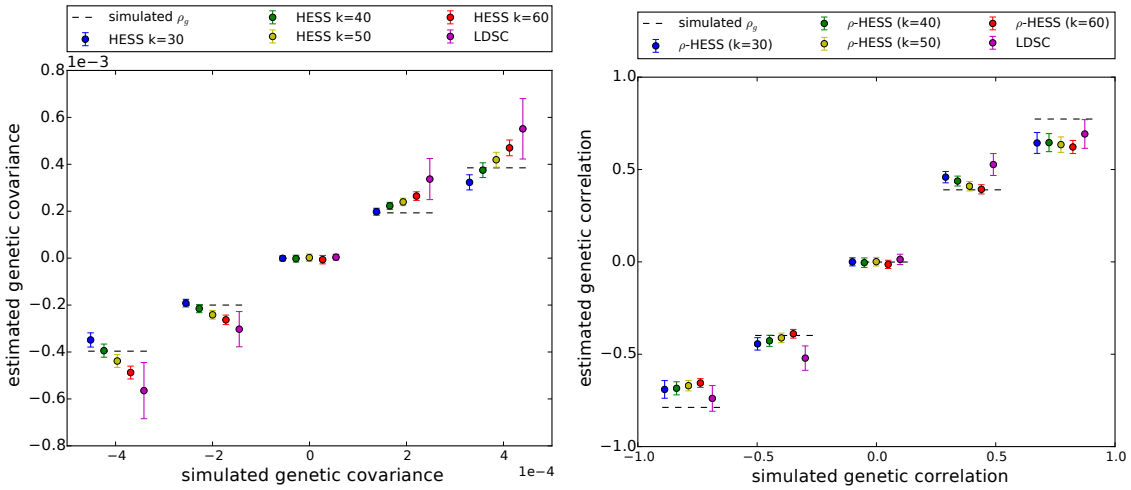

Figure S3: **Performance of ρ-HESS and cross-trait LDSC using external reference LD across 50 LD-independent regions, with each region having 1000 simulations.** Here, all the causal variants are drawn from DHS regions. Each dot represents the mean (over 50 regions) of the average performance (over 1000 simulations per region), with error bars representing 1.96 times the standard error on both sides.

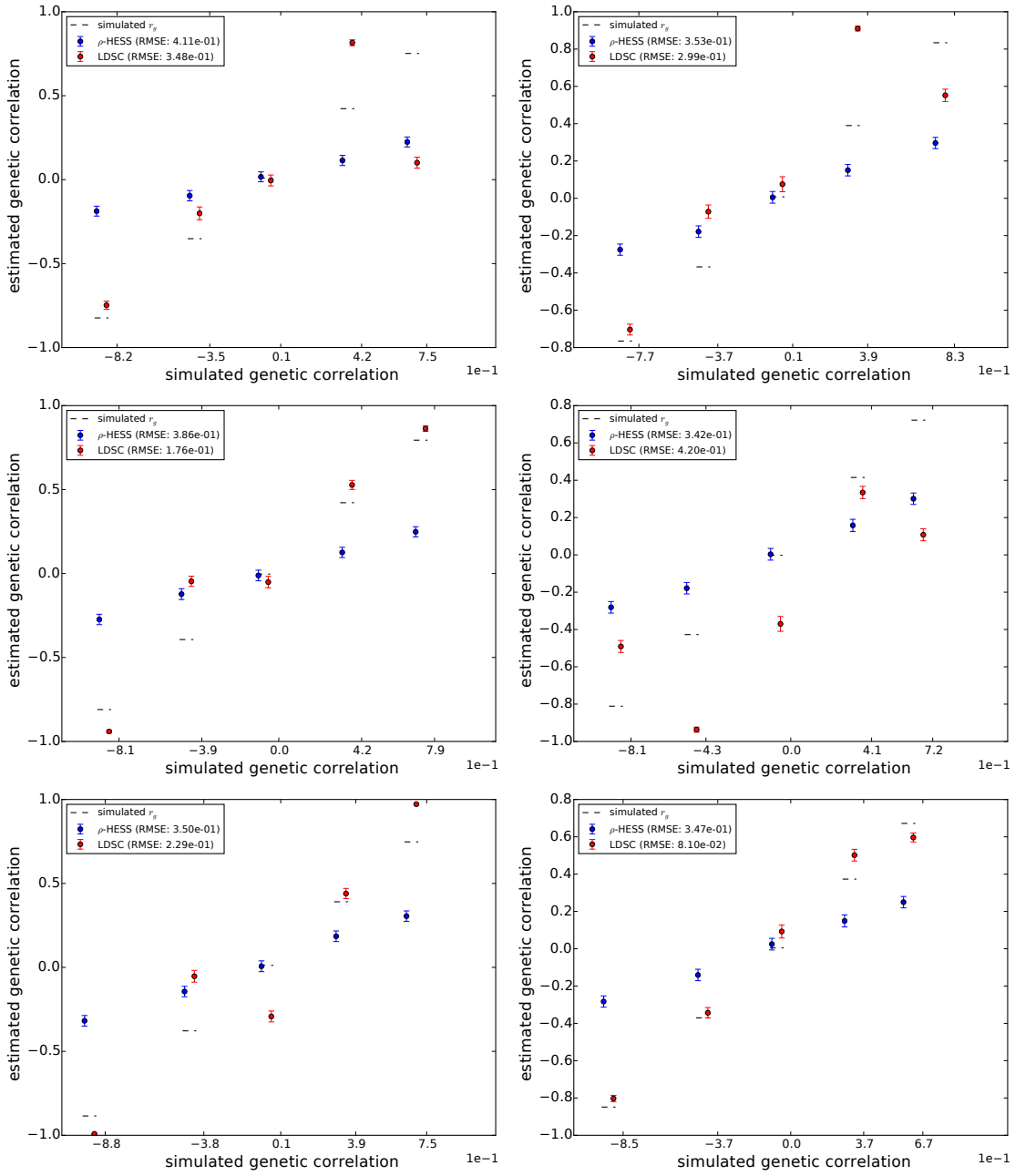

Figure S4: **Performance of ρ-HESS and cross-trait LDSC in estimating local genetic correlation at 6 LD-independent loci.** Mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

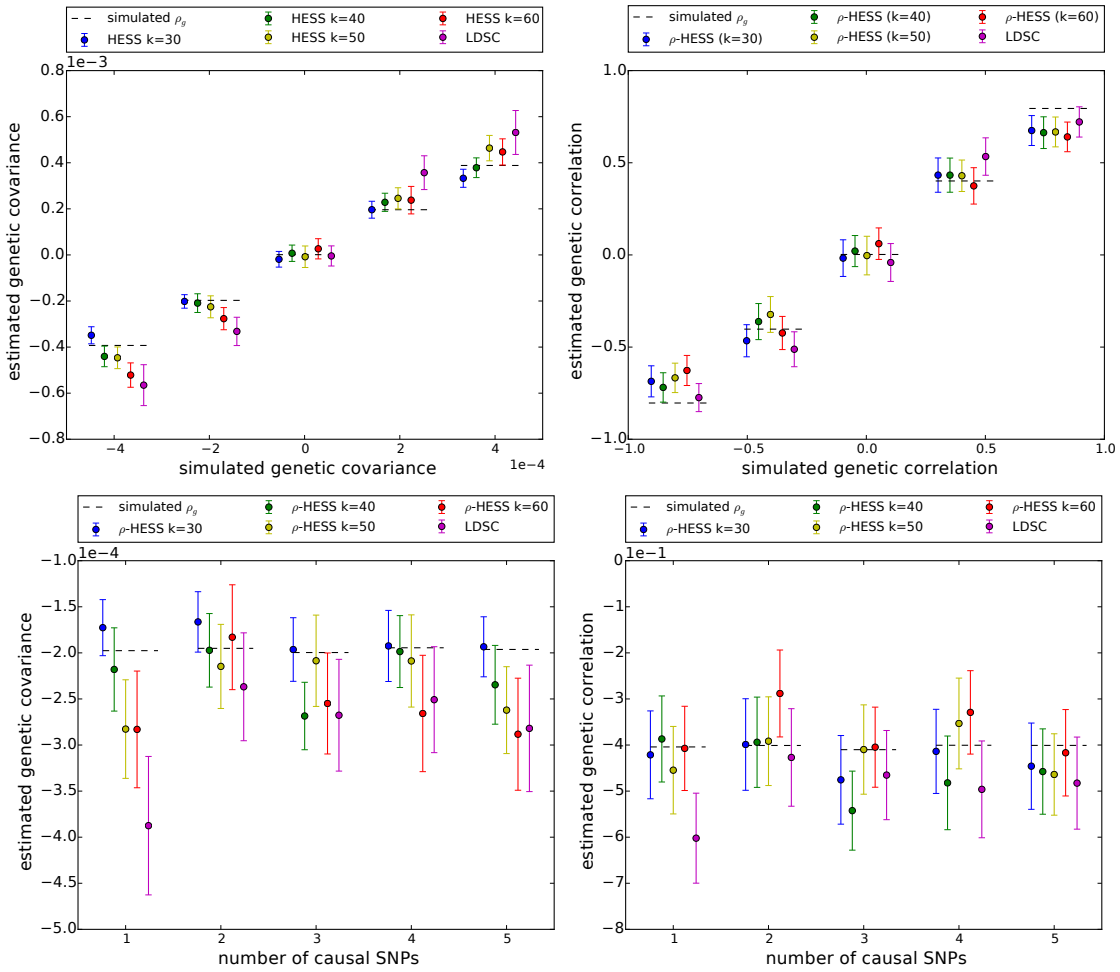

Figure S5: **Performance of ρ-HESS and cross-trait LDSC in estimated local genetic covariance (left panel) and correlation (right panel) across different simulated genetic covariance (correlation) and various degrees of polygenicity.** Here, each dot represents the mean over 100 simulations at 100 regions, with error bars representing 1.96 times the standard error on both sides. Overall, ρ-HESS provides less biased and more consistent estimates of genetic covariance and correlation than cross-trait LDSC.

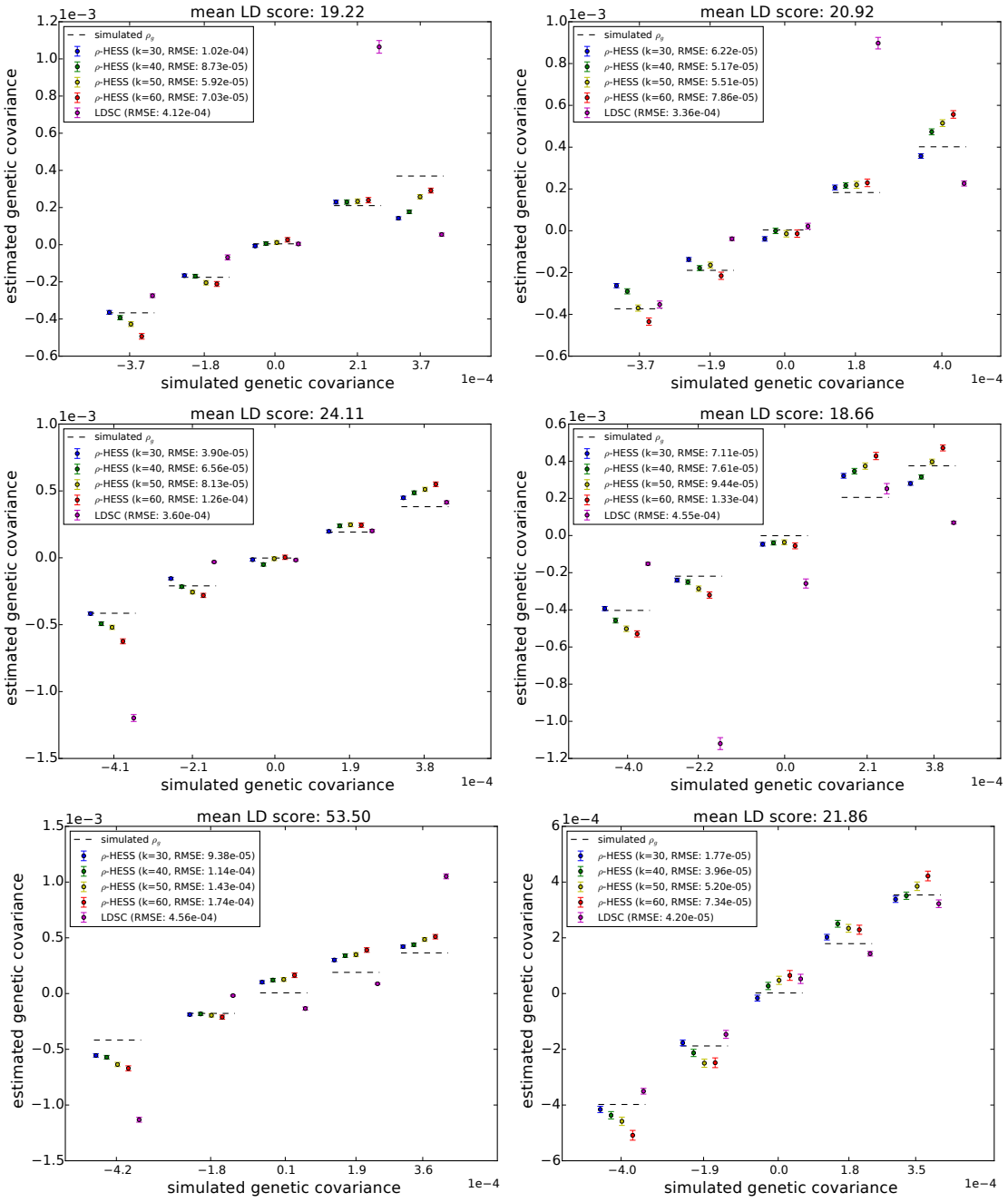

Figure S6: **ρ-HESS obtains more accurate estimates of local genetic covariance (ρ_g) than LDSC at different simulated local genetic covariance. Here, we compare the performance of ρ-HESS at 6 independent loci.** ρ-HESS is approximately unbiased and consistently achieves smaller root mean squared error than LDSC at different number (30 to 60) of eigenvectors in the truncated-SVD regularization. Mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

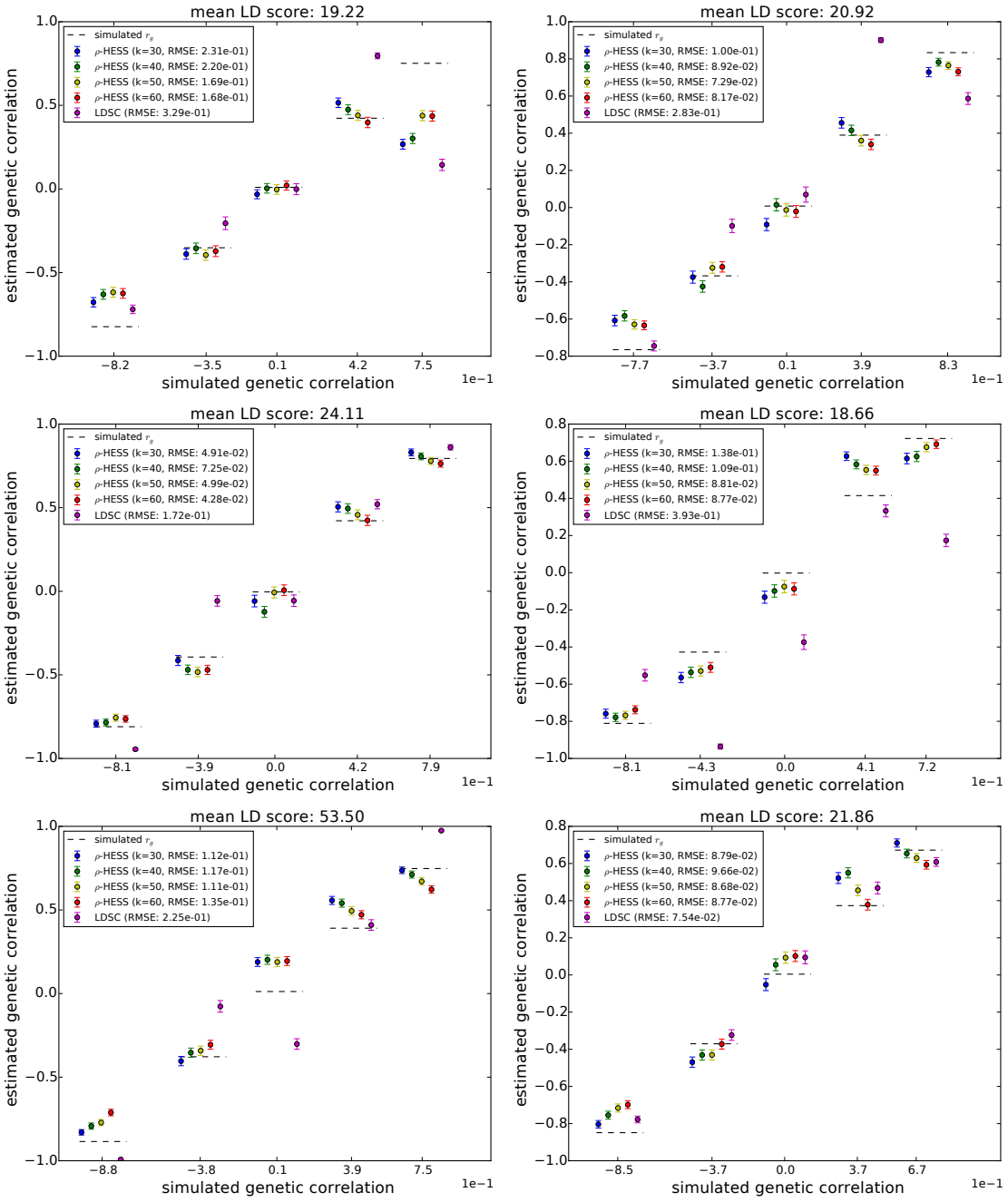

Figure S7: **ρ-HESS obtains more accurate estimates of local genetic correlation (r_g) than LDSC at different simulated local genetic correlation.** Here, we compare the performance of ρ-HESS at 6 independent loci. ρ-HESS is approximately unbiased and consistently achieves smaller root mean squared error than LDSC at different number (30 to 60) of eigenvectors in the truncated-SVD regularization. Mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

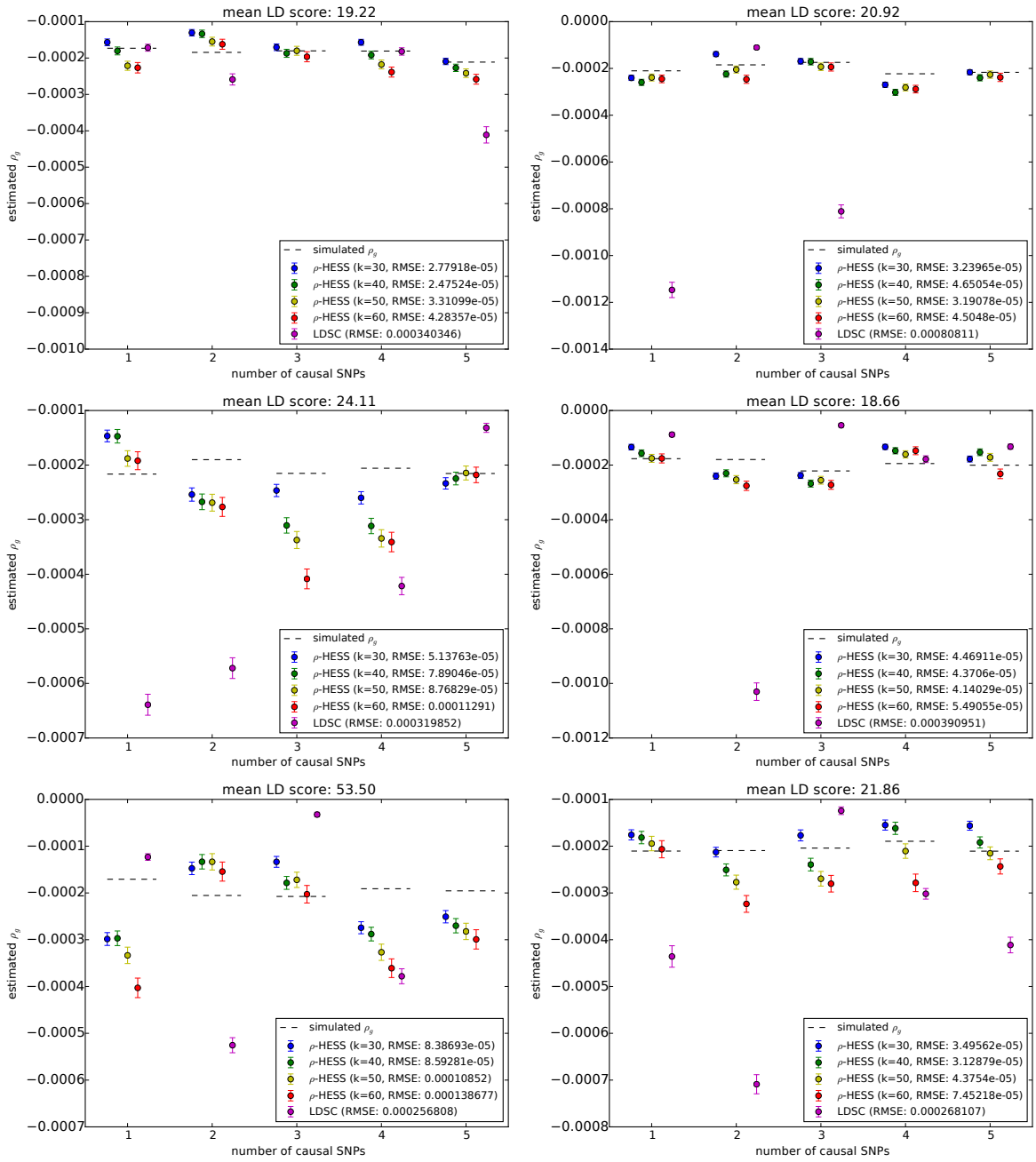

Figure S8: **ρ-HESS obtains more accurate estimates of local genetic covariance (ρ_g) than LDSC in simulations where the number of causal SNPs varies.** Here, we compare the performance of ρ-HESS at 6 independent loci. ρ-HESS is approximately unbiased and consistently achieves smaller root mean squared error than LDSC at different number (30 to 60) of eigenvectors in the truncated-SVD regularization. Mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

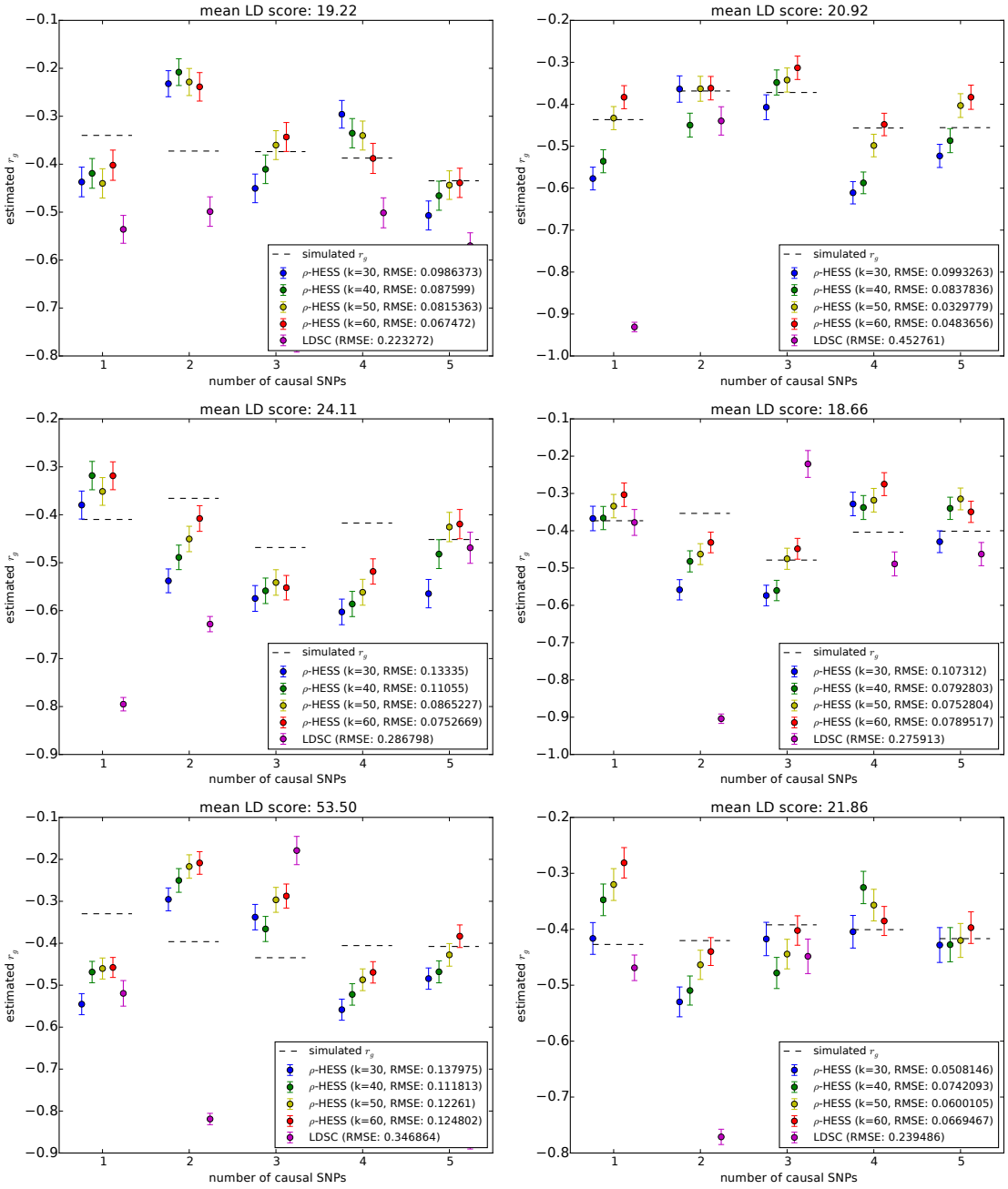

Figure S9: **ρ-HESS obtains more accurate estimates of local genetic correlation (r_g) than LDSC in simulations where the number of causal SNPs varies.** Here, we compare the performance of ρ-HESS and LDSC in recovering simulated local genetic correlation at 6 independent loci. ρ-HESS consistently achieves smaller root mean squared error than LDSC at different number (30 to 60) of eigenvectors in the truncated-SVD regularization. Mean and standard error are obtained over 1000 simulations. Error bars represent 1.96 times the standard error on each side.

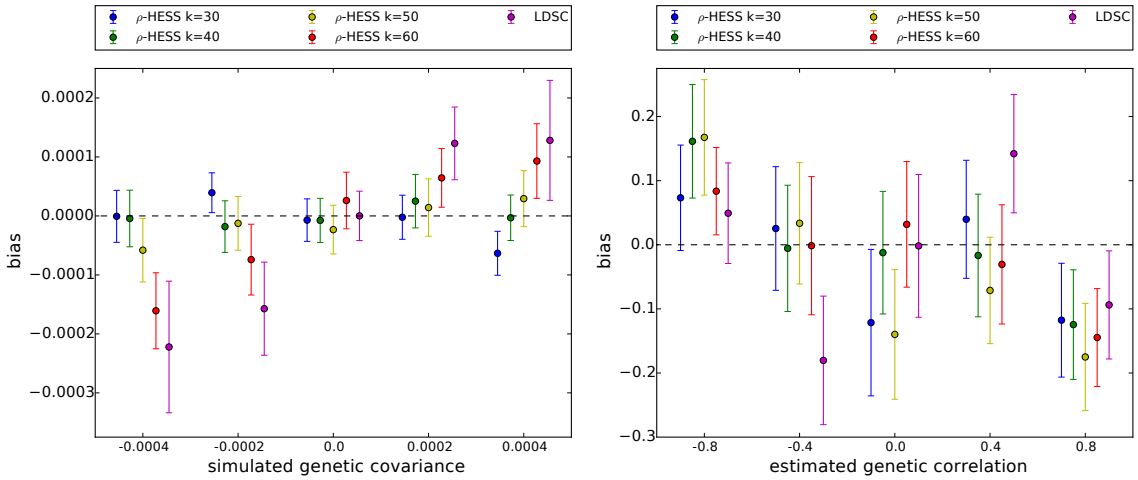

Figure S10: **Distribution of the bias (estimated minus simulated) of the local genetic covariance / correlation estimates obtained by ρ-HESS and cross-trait LDSC.** Mean and standard error are obtained across 100 simulations at 100 LD-independent regionts. Error bars represent 1.96 times the standard error on each side.

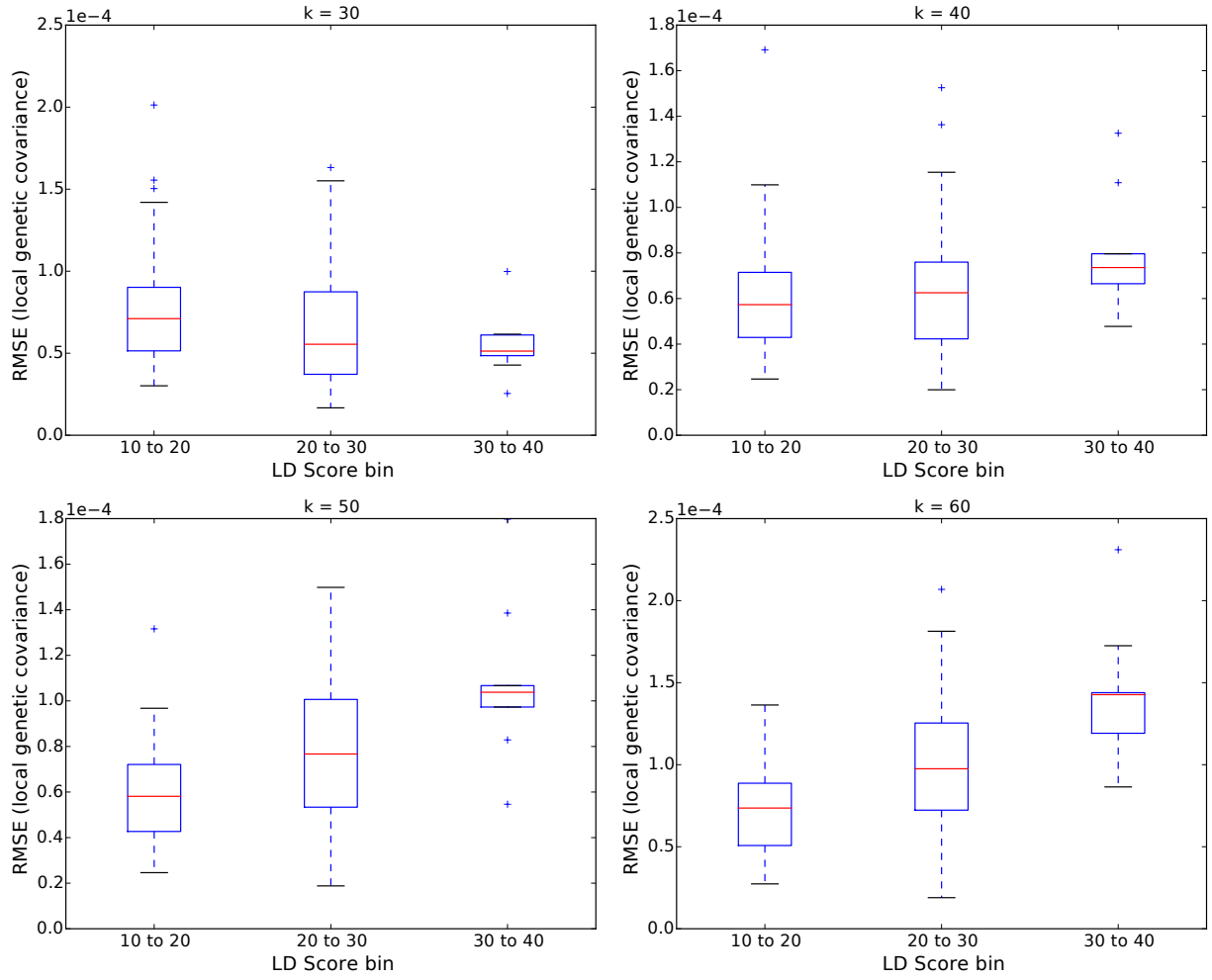

Figure S11: **Accuracy (measured as RMSE) of ρ-HESS estimates of local genetic covariance decreases as the amount of LD (measured as mean LD score) in the locus increases.** Here, the bottom and top of each box represent the 1st and 3rd quartile of the distribution, respectively. The red line represents the median. And the bottom and top of the whiskpers represent 1.5 times the interquartile range below and above the first and third quartile, respectively.

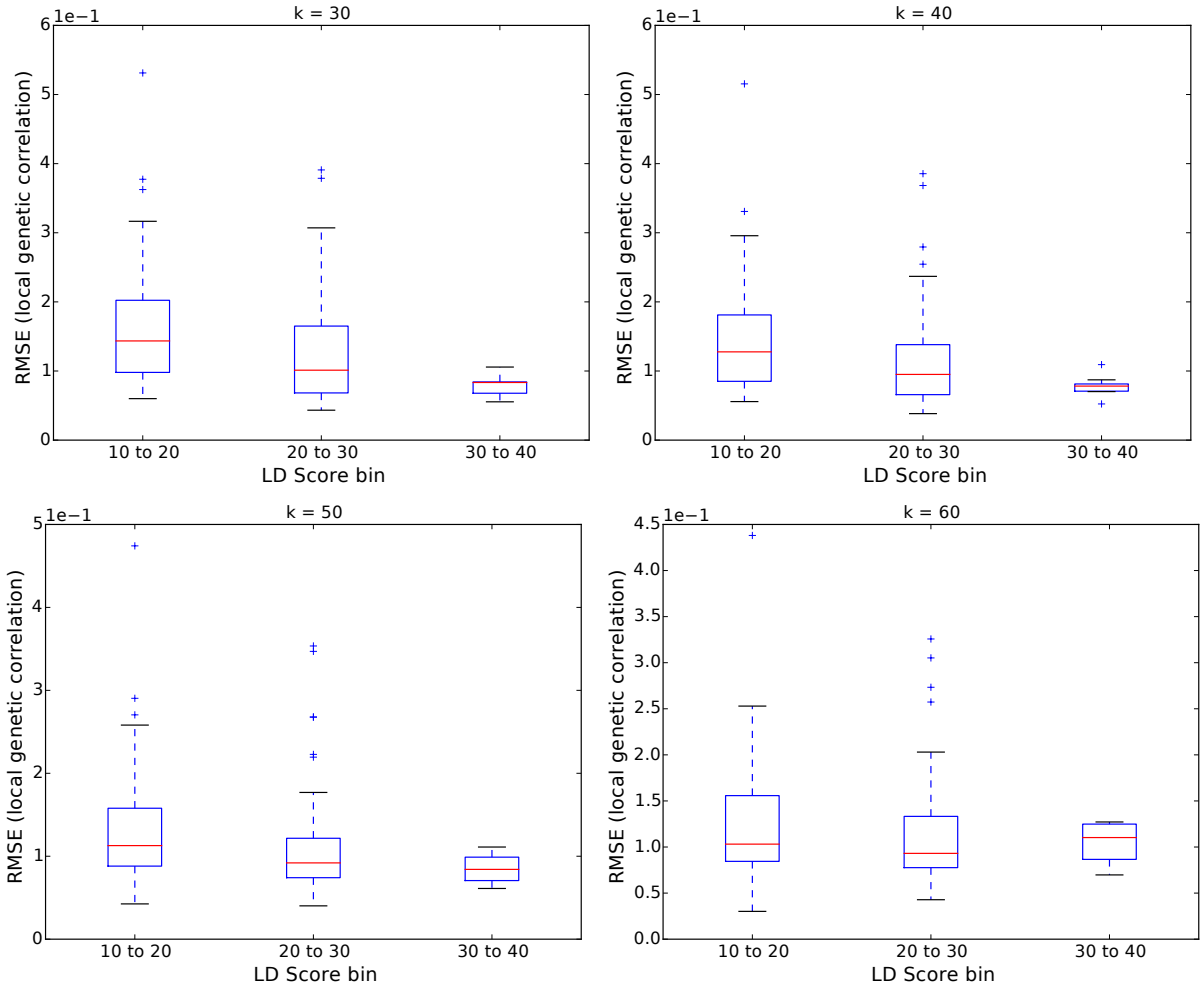

Figure S12: **Accuracy (measured as RMSE) of ρ-HESS estimates of local genetic covariance decreases as the amount of LD (measured as mean LD score) in the locus increases.** Here, the bottom and top of each box represent the 1st and 3rd quartile of the distribution, respectively. The red line represents the median. And the bottom and top of the whiskpers represent 1.5 times the interquartile range below and above the first and third quartile, respectively.

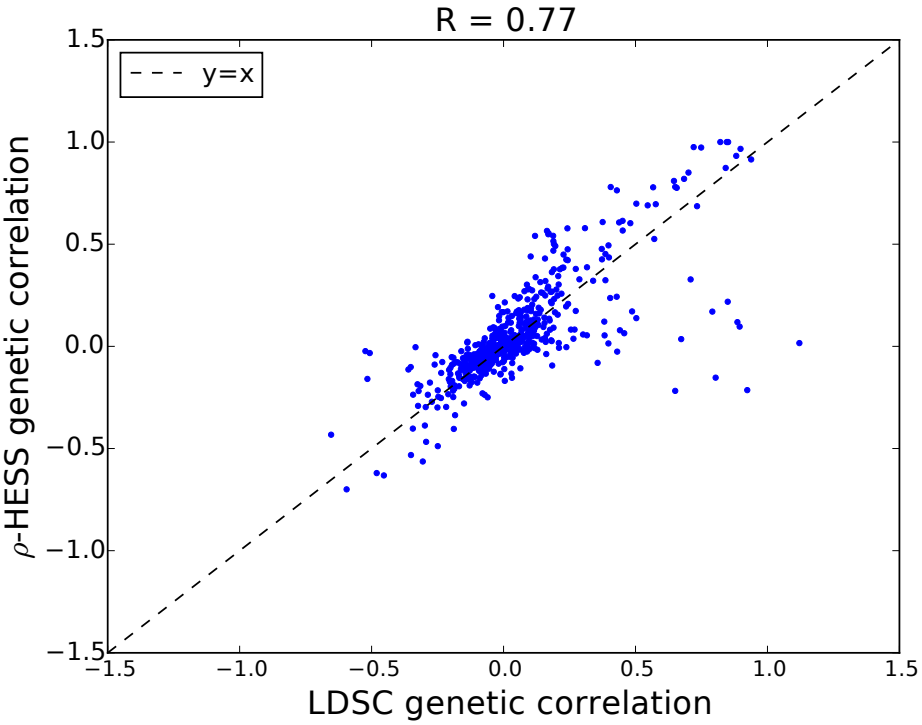

Figure S13: **ρ-HESS and cross-trait LDSC yields similar estimates of genome-wide genetic correlation.** Correlation between estimates of genome-wide genetic correlation obtained by ρ-HESS and cross-trait LDSC is 0.77.

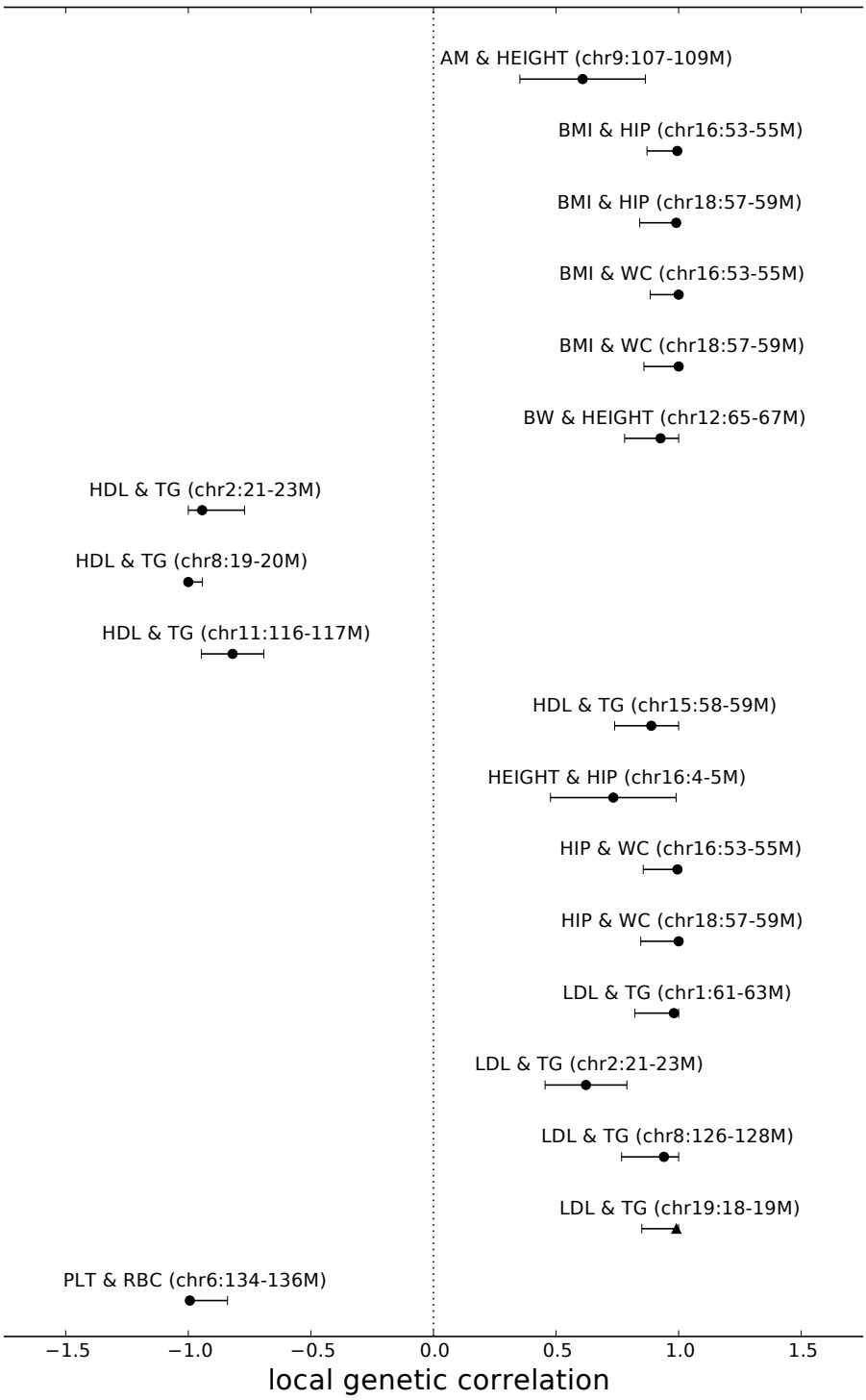

Figure S14: **Local genetic correlation at loci displaying significant local genetic covariance and SNP-heritability for pairs of traits with significant genome-wide genetic correlation.** Here, we excluded all pairs of traits involving TC. We obtain standard error estimates through parametric bootstrapping. Error bars represent 1.96 times the standard error on both sides. Here, triangle represents loci that lack GWAS risk variant for both traits, and circle represents loci that contain GWAS risk variants for both traits.

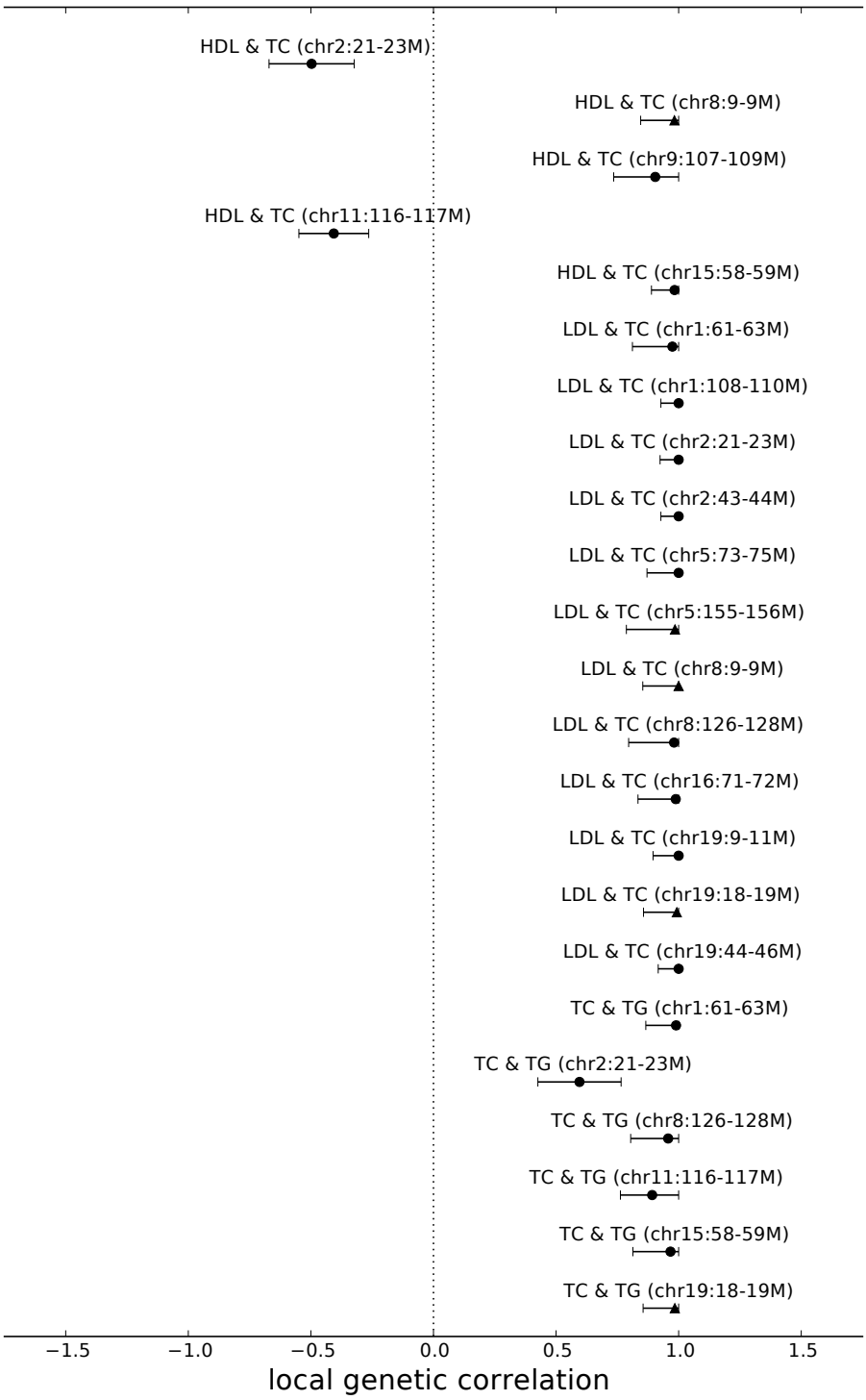

Figure S15: **Local genetic correlation at loci displaying significant local genetic covariance and SNP-heritability for pairs of traits with significant genome-wide genetic correlation.** Here, we focus only on the pairs of traits involving TC. We obtain standard error estimates through parametric bootstrapping. Error bars represent 1.96 times the standard error on both sides. Here, triangle represents loci that lack GWAS risk variant for both traits; diamond represents loci that harbor GWAS risk variants for one of the traits; and circle represents loci that contain GWAS risk variants for both traits.

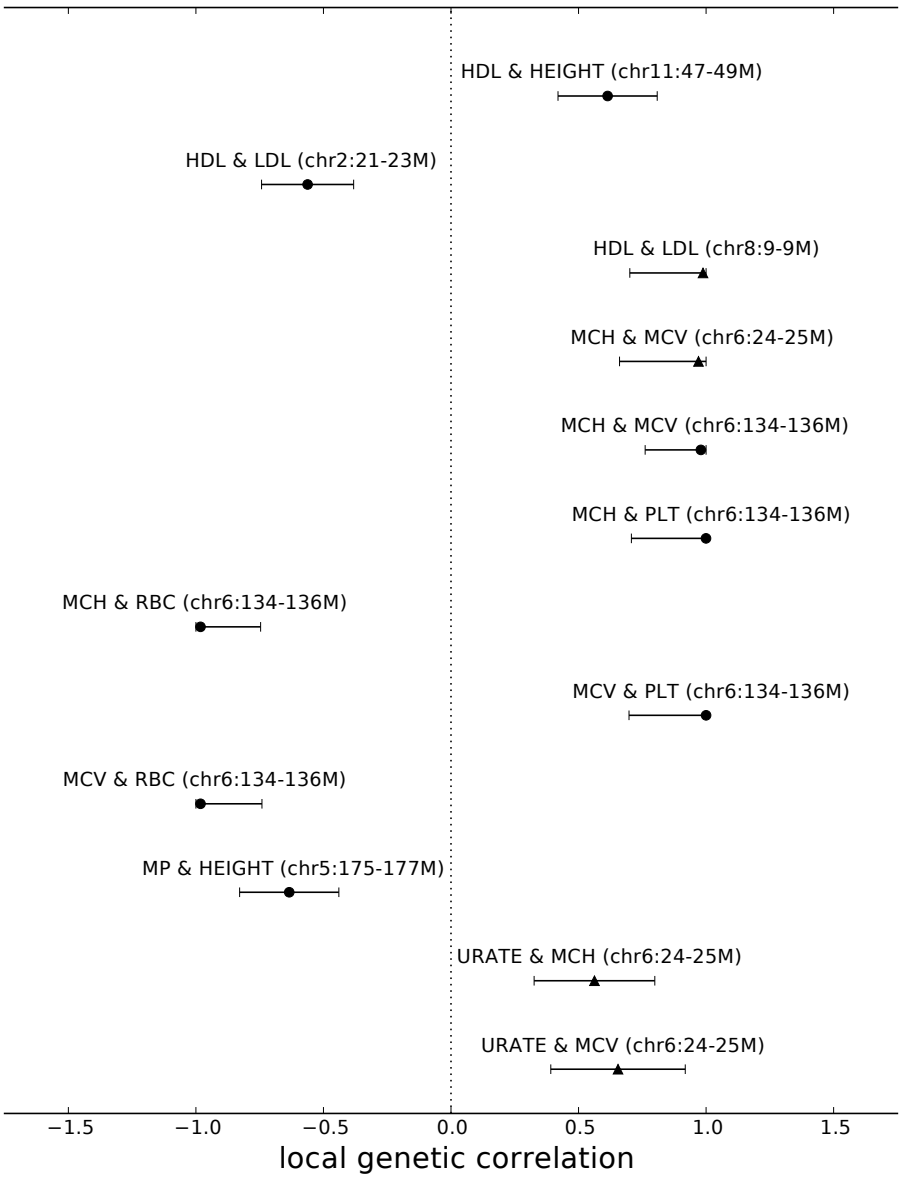

Figure S16: **Local genetic correlation at loci displaying significant local genetic covariance and SNP-heritability for pairs of traits with negligible genome-wide genetic correlation.** We obtain standard error estimates through parametric bootstrapping. Error bars represent 1.96 times the standard error on both sides. Here, triangle represents loci that lack GWAS risk variant for both traits; diamond represents loci that harbor GWAS risk variants for one of the traits; and circle represents loci that contain GWAS risk variants for both traits.

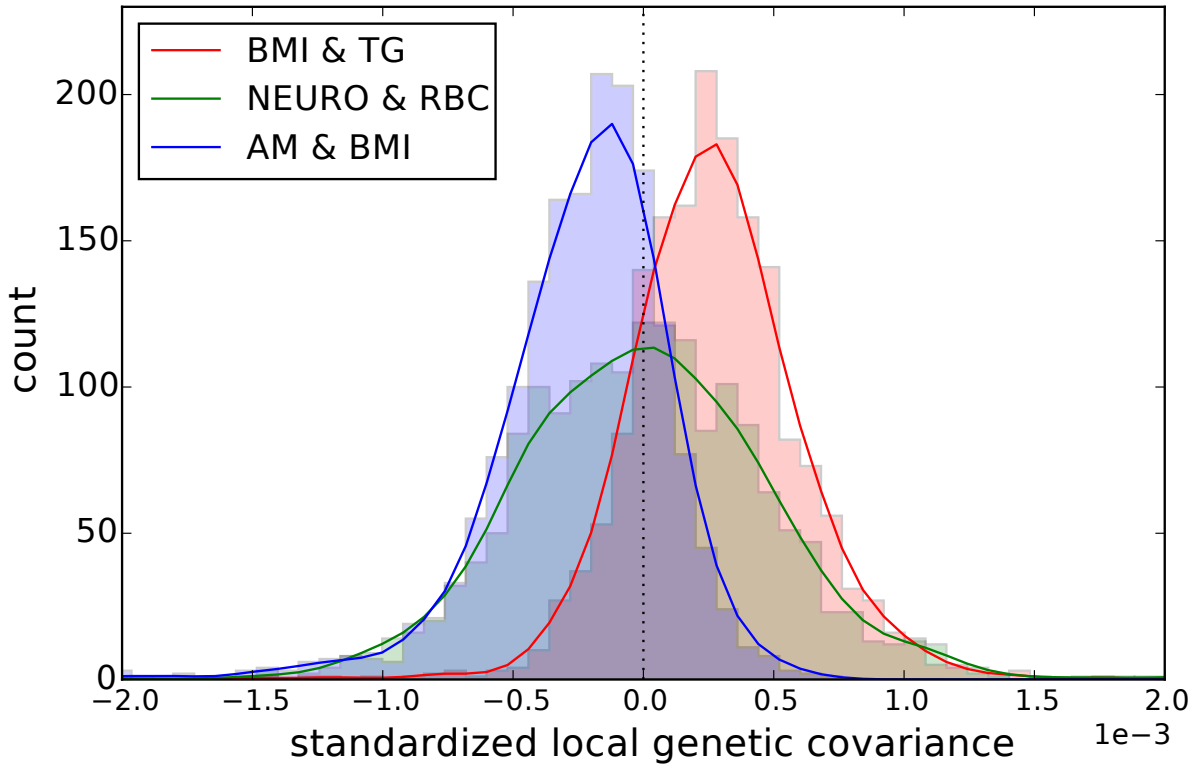

Figure S17: **Distribution of standardized local genetic covariance (local genetic covariance standardized by the square roots of total SNP-heritability of two traits) for the pairs of traits BMI and TG, NEURO and RBC, AM and BMI.** Pairs of traits with positive (negative) genome-wide genetic correlation show a shift in the distribution of standardized local genetic covariance away from 0.

Table S1: Loci that show significant local genetic covariance (two-tailed $p < 0.05/1703/630$) and local SNP heritability (one-tailed $p < 0.05/1703/36$) for both traits.

Trait1	**Trait2**	**Locus**	$\mathbf{h^2_{g,local,trait1}}$	$\mathbf{h^2_{g,local,trait2}}$	$\mathbf{r_{g,local}}$
HDL	TC	chr2:21-23M	0.16 (0.03)	0.63 (0.04)	-0.50 ([-0.67,-0.33])
HDL	TC	chr8:9-9M	0.14 (0.02)	0.16 (0.02)	0.98 ([0.74,1.00])
HDL	TC	chr9:107-109M	0.28 (0.03)	0.19 (0.03)	0.90 ([0.65,1.00])
HDL	TC	chr11:116-117M	0.40 (0.04)	0.27 (0.03)	-0.41 ([-0.55,-0.27])
HDL	TC	chr15:58-59M	1.18 (0.06)	0.31 (0.03)	0.98 ([0.83,1.00])
LDL	TC	chr1:61-63M	0.14 (0.03)	0.28 (0.03)	0.97 ([0.67,1.00])
LDL	TC	chr1:108-110M	0.74 (0.05)	0.52 (0.04)	1.00 ([0.88,1.00])
LDL	TC	chr2:21-23M	0.84 (0.05)	0.63 (0.04)	1.00 ([0.87,1.00])
LDL	TC	chr2:43-44M	0.31 (0.03)	0.31 (0.03)	1.00 ([0.91,1.00])
LDL	TC	chr5:73-75M	0.28 (0.03)	0.24 (0.03)	1.00 ([0.76,1.00])
LDL	TC	chr5:155-156M	0.11 (0.02)	0.13 (0.02)	0.98 ([0.58,1.00])
LDL	TC	chr8:9-9M	0.12 (0.02)	0.16 (0.02)	1.00 ([0.72,1.00])
LDL	TC	chr8:126-128M	0.16 (0.03)	0.19 (0.03)	0.98 ([0.61,1.00])
LDL	TC	chr16:71-72M	0.19 (0.03)	0.19 (0.03)	0.99 ([0.70,1.00])
LDL	TC	chr19:9-11M	0.49 (0.04)	0.33 (0.03)	1.00 ([0.81,1.00])
LDL	TC	chr19:18-19M	0.18 (0.03)	0.26 (0.03)	0.99 ([0.74,1.00])
LDL	TC	chr19:44-46M	0.77 (0.05)	0.43 (0.04)	1.00 ([0.86,1.00])
TC	TG	chr1:61-63M	0.28 (0.03)	0.28 (0.03)	0.99 ([0.77,1.00])
TC	TG	chr2:21-23M	0.63 (0.04)	0.22 (0.03)	0.60 ([0.43,0.76])
TC	TG	chr8:126-128M	0.19 (0.03)	0.32 (0.04)	0.96 ([0.69,1.00])
TC	TG	chr11:116-117M	0.27 (0.03)	1.27 (0.06)	0.89 ([0.73,1.00])
TC	TG	chr15:58-59M	0.31 (0.03)	0.18 (0.03)	0.97 ([0.69,1.00])
TC	TG	chr19:18-19M	0.26 (0.03)	0.21 (0.03)	0.98 ([0.75,1.00])

Here, we focus only on the pairs of traits involving TC. The genome-wide genetic correlation of each pair of traits is also significant (two-tailed $p < 0.05/630$). Numbers in parentheses represent standard errors for local SNP heritability estimates and 95% confidence intervals for local genetic correlation estimates.

Table S2: Bi-directional analysis of local genetic correlation identifies 40 pairs of traits for which one is likely a causal factor of the other.

Trait1	$\hat{\mathbf{r}}_{\mathbf{g,local,trait1}}$	**No. loci**	**Trait2**	$\hat{\mathbf{r}}_{\mathbf{g,local,trait2}}$	**No. loci**	**Direction**	**Ratio**
AM	-0.47 (0.06)	54	BMI	-0.49 (0.07)	51	BMI ↓ AM	0.00e+00
AM	0.26 (0.05)	39	HEIGHT	0.09 (0.02)	429	AM ↑ HEIGHT	0.00e+00
AM	-0.18 (0.05)	60	HIP	-0.26 (0.08)	36	HIP ↓ AM	7.00e-05
AM	-0.23 (0.05)	58	WC	-0.36 (0.09)	28	WC ↓ AM	1.09e-04
BMI	-0.25 (0.06)	60	EY	-0.35 (0.04)	133	EY ↓ BMI	0.00e+00
BMI	-0.47 (0.05)	57	HDL	-0.18 (0.04)	81	BMI ↓ HDL	0.00e+00
BMI	-0.02 (0.04)	39	HEIGHT	-0.16 (0.02)	432	HEIGHT ↓ BMI	0.00e+00
BMI	0.95 (0.02)	32	HIP	0.77 (0.11)	11	BMI ↑ HIP	0.00e+00
BMI	0.47 (0.05)	59	TG	-0.02 (0.06)	60	BMI ↑ TG	0.00e+00
URATE	0.07 (0.08)	28	BMI	0.55 (0.05)	64	BMI ↓ URATE	1.80e-05
BMI	0.69 (0.03)	58	WHR	0.13 (0.13)	22	BMI ↑ WHR	0.00e+00
BW	-0.22 (0.05)	41	URATE	-0.08 (0.09)	28	URATE ↓ BW	0.00e+00
URATE	-0.13 (0.05)	22	CRN	-0.36 (0.08)	36	URATE ↓ CRN	1.00e-06
CRN	0.04 (0.06)	41	WHR	0.07 (0.09)	27	CRN ↓ WHR	2.06e-04
EY	-0.19 (0.05)	138	HB	-0.05 (0.10)	15	EY ↓ HB	1.00e-06
EY	0.22 (0.03)	134	HDL	0.08 (0.03)	85	EY ↑ HDL	3.91e-04
EY	0.16 (0.03)	100	HEIGHT	0.16 (0.02)	420	HEIGHT ↑ EY	0.00e+00
EY	-0.24 (0.04)	133	NEURO	-0.14 (0.11)	11	EY ↓ NEURO	0.00e+00
EY	-0.20 (0.03)	134	TG	-0.05 (0.05)	62	EY ↓ TG	0.00e+00
EY	-0.30 (0.03)	134	WC	-0.25 (0.08)	34	EY ↓ WC	0.00e+00
EY	-0.34 (0.03)	136	WHR	-0.17 (0.06)	27	EY ↓ WHR	0.00e+00
HDL	-0.12 (0.04)	81	HIP	-0.50 (0.05)	36	HIP ↓ HDL	2.00e-06
HDL	-0.51 (0.07)	52	TG	-0.48 (0.10)	29	HDL ↓ TG	1.00e-06
HDL	-0.25 (0.04)	82	WC	-0.64 (0.05)	31	WC ↓ HDL	0.00e+00
HDL	-0.27 (0.06)	79	WHR	-0.59 (0.12)	19	WHR ↓ HDL	6.82e-04
HEIGHT	0.44 (0.01)	432	HIP	0.25 (0.06)	18	HEIGHT ↑ HIP	0.00e+00
HEIGHT	-0.09 (0.02)	420	LDL	-0.01 (0.04)	30	HEIGHT ↓ LDL	0.00e+00
HEIGHT	-0.14 (0.02)	446	PLT	-0.08 (0.13)	17	HEIGHT ↓ PLT	0.00e+00
HEIGHT	-0.12 (0.02)	415	TC	-0.05 (0.05)	40	HEIGHT ↓ TC	0.00e+00
HEIGHT	-0.08 (0.02)	429	TG	-0.05 (0.05)	37	HEIGHT ↓ TG	1.00e-06
HEIGHT	0.30 (0.02)	443	WC	0.17 (0.05)	23	HEIGHT ↑ WC	0.00e+00
HIP	0.26 (0.07)	41	TG	-0.09 (0.06)	63	HIP ↑ TG	4.59e-04
HIP	0.44 (0.08)	37	WHR	-0.14 (0.09)	22	HIP ↑ WHR	7.00e-06
LDL	0.93 (0.03)	11	TC	0.80 (0.08)	26	TC ↑ LDL	0.00e+00
URATE	0.05 (0.05)	29	MP	-0.02 (0.05)	62	URATE ↑ MP	1.33e-04
NEURO	0.08 (0.11)	17	PLT	-0.04 (0.07)	30	NEURO ↑ PLT	4.00e-06
TG	0.09 (0.05)	62	WC	0.56 (0.06)	34	WC ↑ TG	0.00e+00
TG	0.28 (0.05)	59	WHR	0.57 (0.10)	22	WHR ↑ TG	6.67e-04
URATE	0.07 (0.05)	30	WC	0.69 (0.06)	39	WC ↓ URATE	9.70e-05
WC	0.80 (0.02)	29	WHR	0.60 (0.07)	20	WC ↑ WHR	0.00e+00

Here, "Trait 1" and "Trait 2" refer to the trait for which the GWAS hit loci were ascertained in the bi-directional analysis. Traits that are likely a causal factor of the other are marked with stars. Numbers in parentheses represent standard errors of the local genetic correlation estimates.