The American Journal of Human Genetics, Volume *102*

Supplemental Data

Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood

Guiyan Ni, Gerhard Moser, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Naomi R. Wray, and S. Hong Lee

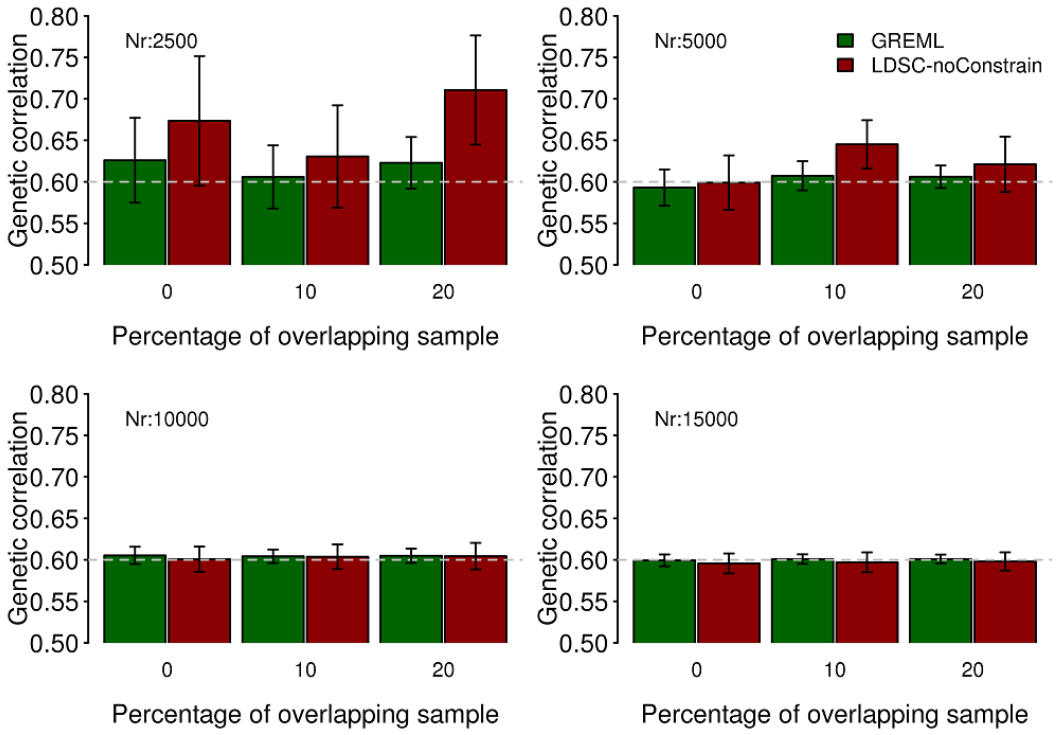

Figure S1. Estimated genetic correlation from GREML and LDSC (without constraining the intercept) in different simulation scenarios based on UK Biobank data.

The x-axis labels of 0, 10 or 20 are the percentage of overlapping individuals between the first and second traits. Nr: The number of available phenotype records in each trait. The horizontal lines stand for the simulated true genetic correlation (0.6). The error bars are 95% CI obtained from 100 replications.

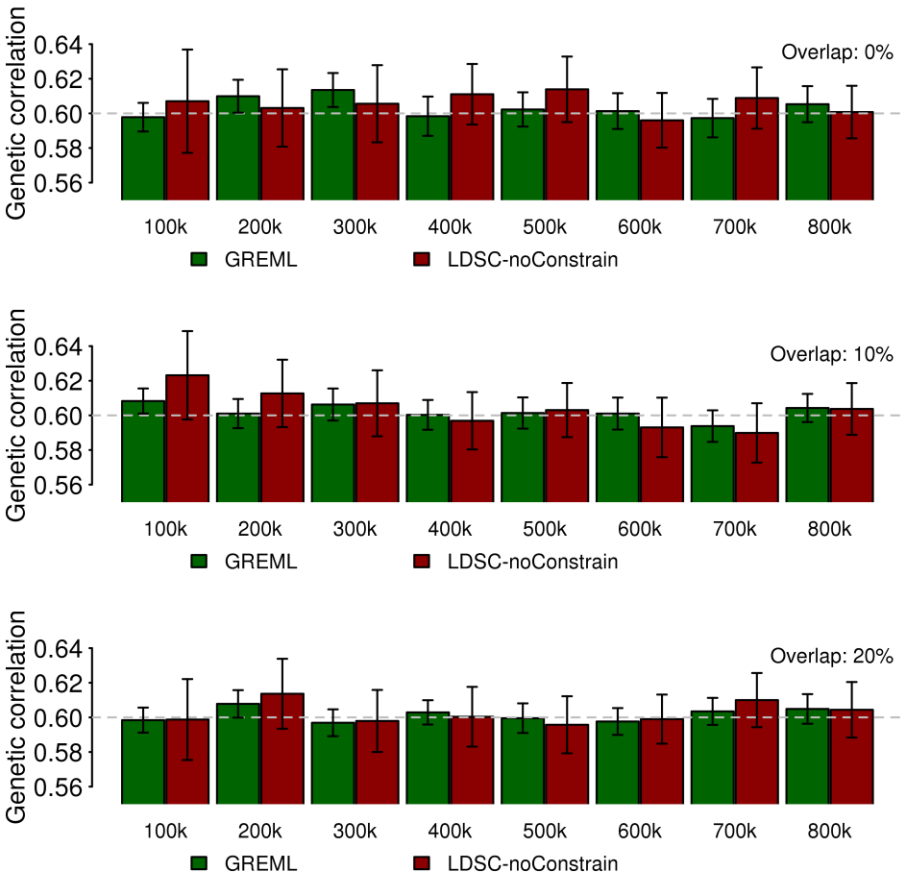

Figure S2. Estimated genetic correlation using simulated phenotypes based on UK Biobank genotypes.

Bars are 95% CI based on 100 replicates. The grey dashed line stands for the true simulated genetic correlation 0.6. This result was based on 858K SNPs (after QC) and 10,000 individuals that were randomly selected from the UK Biobank. SNPs in each bin were randomly and independently drawn from the 858k SNPs. The number of causal SNPs was 10,000 that were randomly selected in each bin. The true simulated value for the genetic correlation was 0.6 and that for the heritability was 0.5 for both traits. Overlap (0%, 10% and 20%) stands for the percentage of overlapping individuals in the first and second traits.

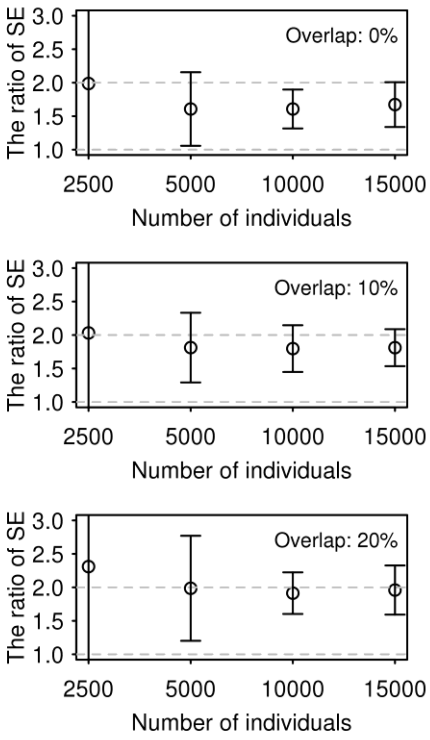

Figure S3. The ratio of SE of LDSC estimate to that of GREML estimate using simulated phenotypes based on UK Biobank genotypes.

Bars are 95% CI based on 100 replicates. The x-axis shows the number of individuals in each trait. Out of 858,991 SNPs, 10,000 SNPs were randomly selected as causal variants. Overlap (0%, 10% and 20%) stands for the percentage of overlapping individuals in the first and second traits. The horizontal dashed lines are the ratios of SE at 1 and 2.

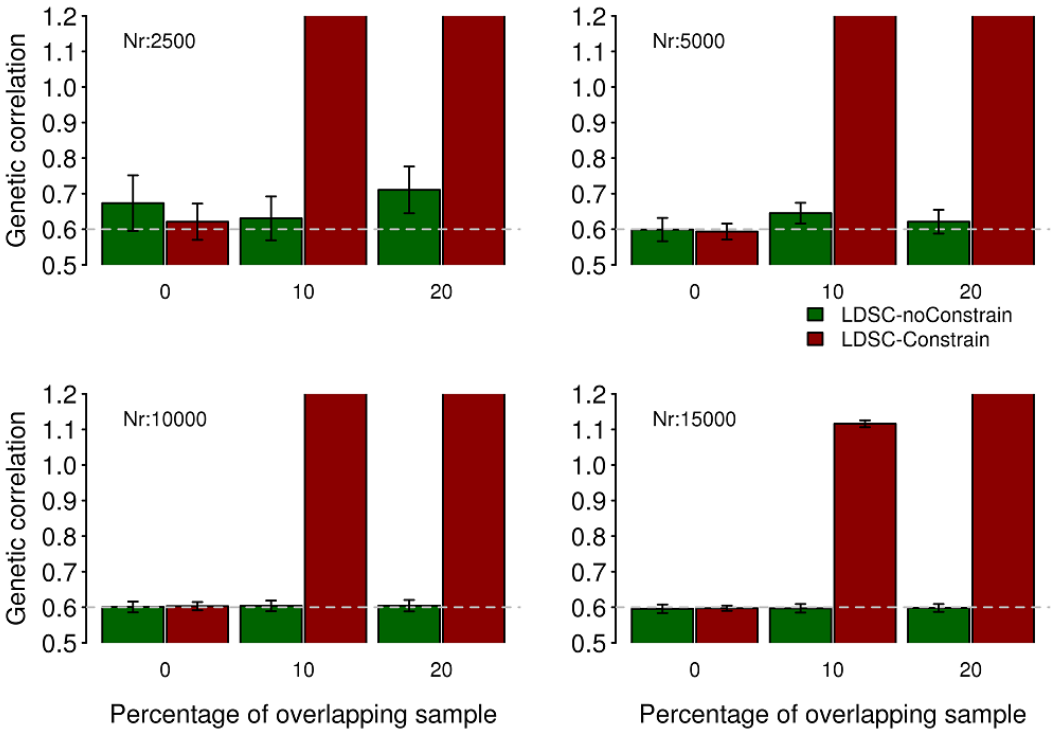

Figure S4. Estimated genetic correlation based on LDSC with or without intercept constrained to zero.

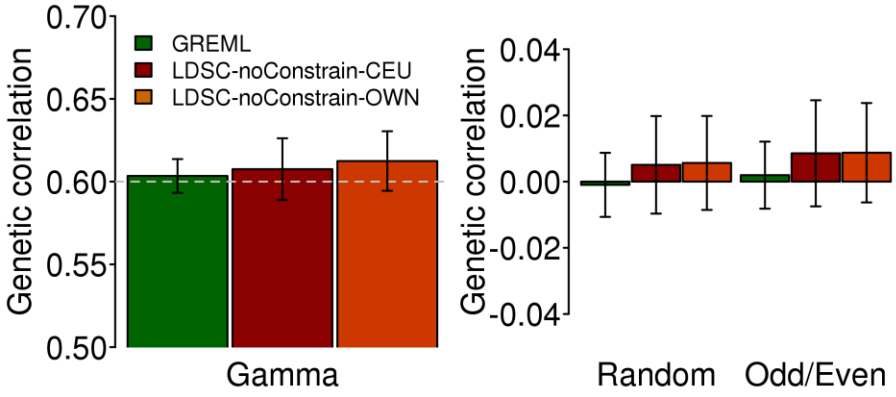

Figure S5. Genetic correlation of GREML and LDSC based on different simulated genetic architectures.

Gamma: the same simulation except that SNP effects were drawn from a multivariate gamma distribution with a shape parameter of one. Random: two sets of non-overlapped SNPs (N=10,000) were randomly selected and each set was assigned as causal SNPs to each trait (SNP effects were generated from a multivariate normal distribution). Odd/Even: two sets of SNPs were randomly selected such that one set was strictly selected from odd and the other set was selected from even number of chromosomes, and each set was assigned as causal SNPs to each trait (SNP effects were generated from a multivariate normal distribution).

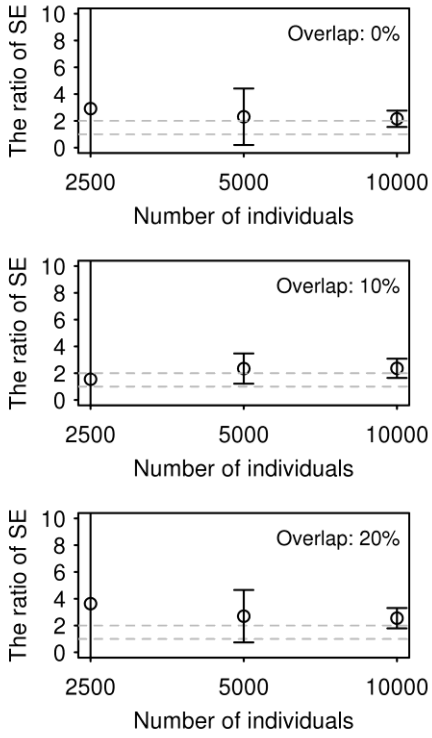

Figure S6. The ratio of SE of LDSC estimate to that of GREML estimate using simulated phenotypes based on WTCCC2 genotypes.

Bars are 95% CI based on 100 replicates. The x-axis shows the number of individuals in each trait. The total sample size of WTCCC2 was 20,659; therefore, the maximum number of individuals for each trait for the analyses was 10,000. Overlap (0%, 10% and 20%) stands for the percentage of overlapping individuals in the first and second traits. The horizontal dashed lines are the ratios of SE at 1 and 2.

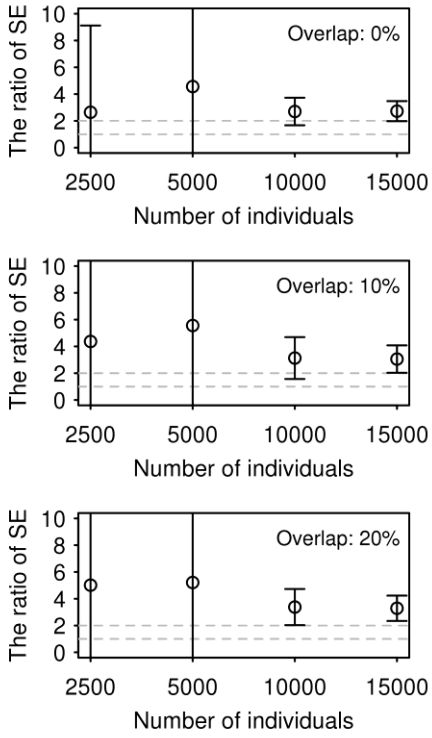

Figure S7. The ratio of SE of LDSC estimate to that of GREML estimate using simulated phenotypes based on GERA genotypes.

Bars are 95% CI based on 100 replicates. The x-axis shows the number of individuals in each trait. Overlap (0%, 10% and 20%) stands for the percentage of overlapping individuals in the first and second traits. The horizontal dashed lines are the ratios of SE at 1 and 2.

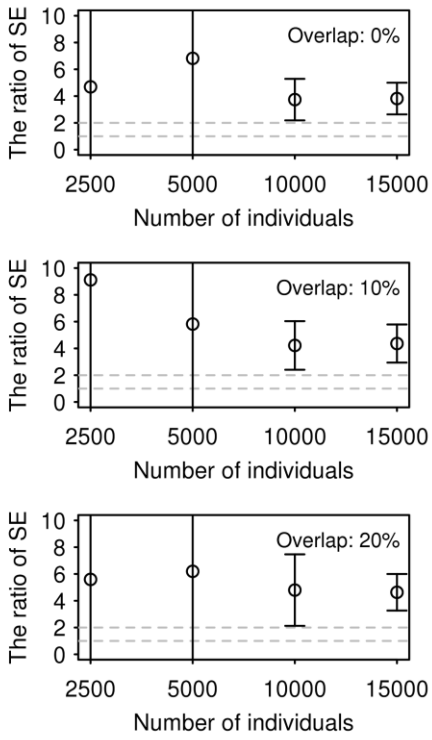

Figure S8. The ratio of SE of LDSC estimate to that of GREML estimate using simulated phenotypes based on raw genotype data of the UK Biobank.

Bars are 95% CI based on 100 replicates. The x-axis shows the number of individuals in each trait. Overlap (0%, 10% and 20%) stands for the percentage of overlapping individuals in the first and second traits. The horizontal dashed lines are the ratios of SE at 1 and 2.

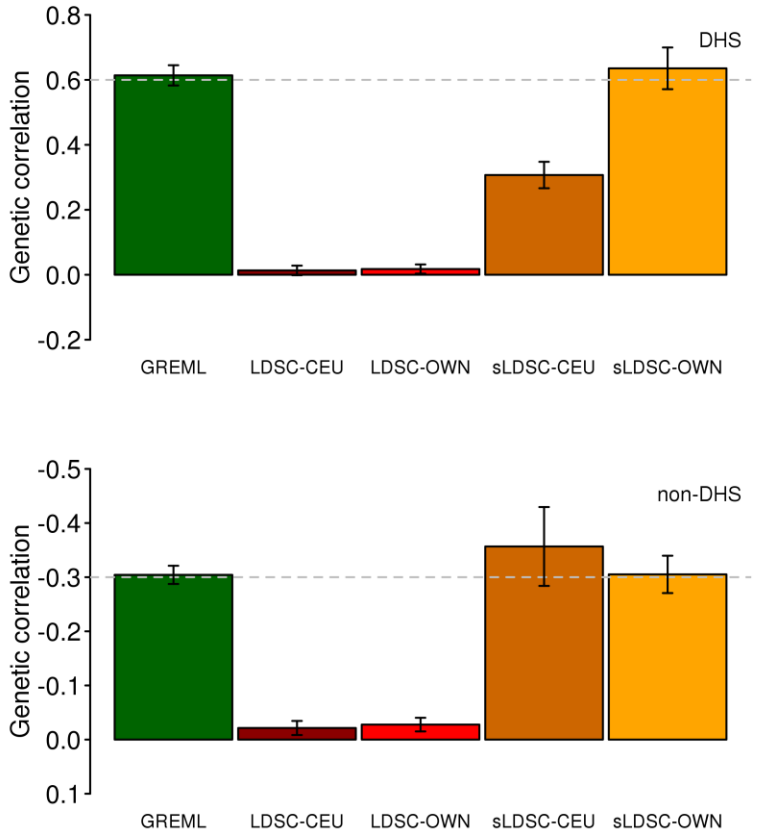

Figure S9 Estimated genetic correlation of simulated data based on a genomic partitioning model.

Simulation was based on 10,000 individuals that were randomly selected from UKBB with 858K SNP. Based on Gusev et al.[1], the 858K SNPs across the genome were stratified as two categories: DHS (194K SNPs with 2268 causal SNPs) and non-DHS (664K SNPs with 7732 causal SNPs). The genetic correlation for the simulated phenotypes between the first and second traits was 0.6 and -0.3 in DHS and non-DHS region, respectively. Bars are 95% CI based on 100 replicates. LDSC-CEU: Using LD-scores estimated from 1KG reference data. LDSC-OWN: Using LD-scores estimated from UKBB. sLDSC-CEU: Using stratified LD-scores estimated from 1KG reference data. sLDSC-OWN: Using stratified LD-scores estimated from UKBB. The presented results were based on 0% overlapping samples between the first and second traits and those based on other scenarios (e.g. 10% and 20%) are presented in Table S1.

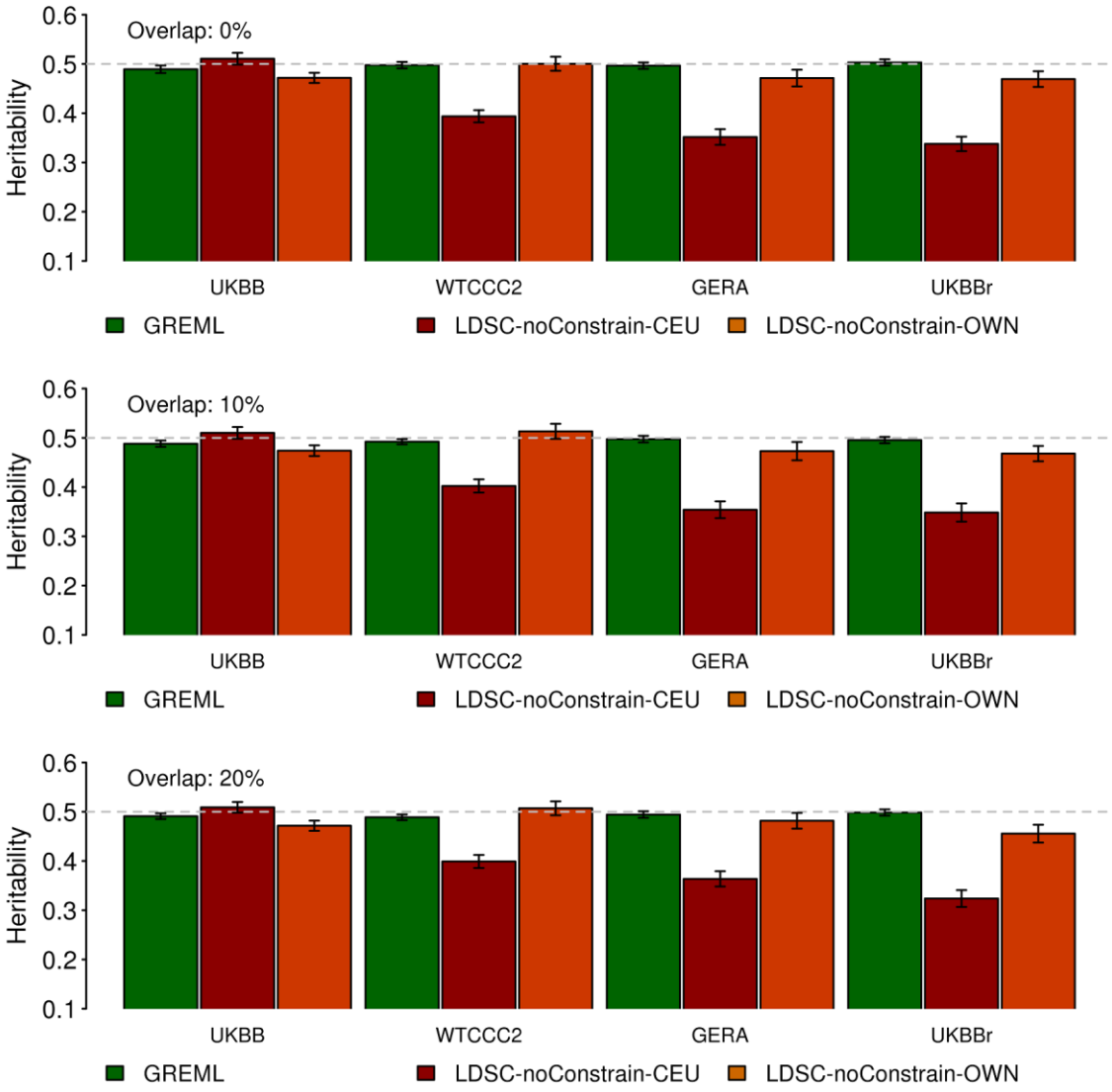

Figure S10. Estimated heritability with GREML and LDSC (without constraining the intercept) based on UKBB, WTCCC2, GERA and UKBBr.

UKBB: Imputed genotype data of UK Biobank sample; WTCCC2: Wellcome Trust Case Control Consortium 2; GERA: Genetic epidemiology research on adult health and aging cohort; UKBBr: Raw genotype data of UK Biobank sample. LDSC-noConstrain-CEU: using LD scores estimated based on the 1KG CEU reference sample. LDSC-noConstrain-OWN: using LD scores estimated based on the target sample (i.e. UKBB, WTCCC2, GERA or UKBBr).

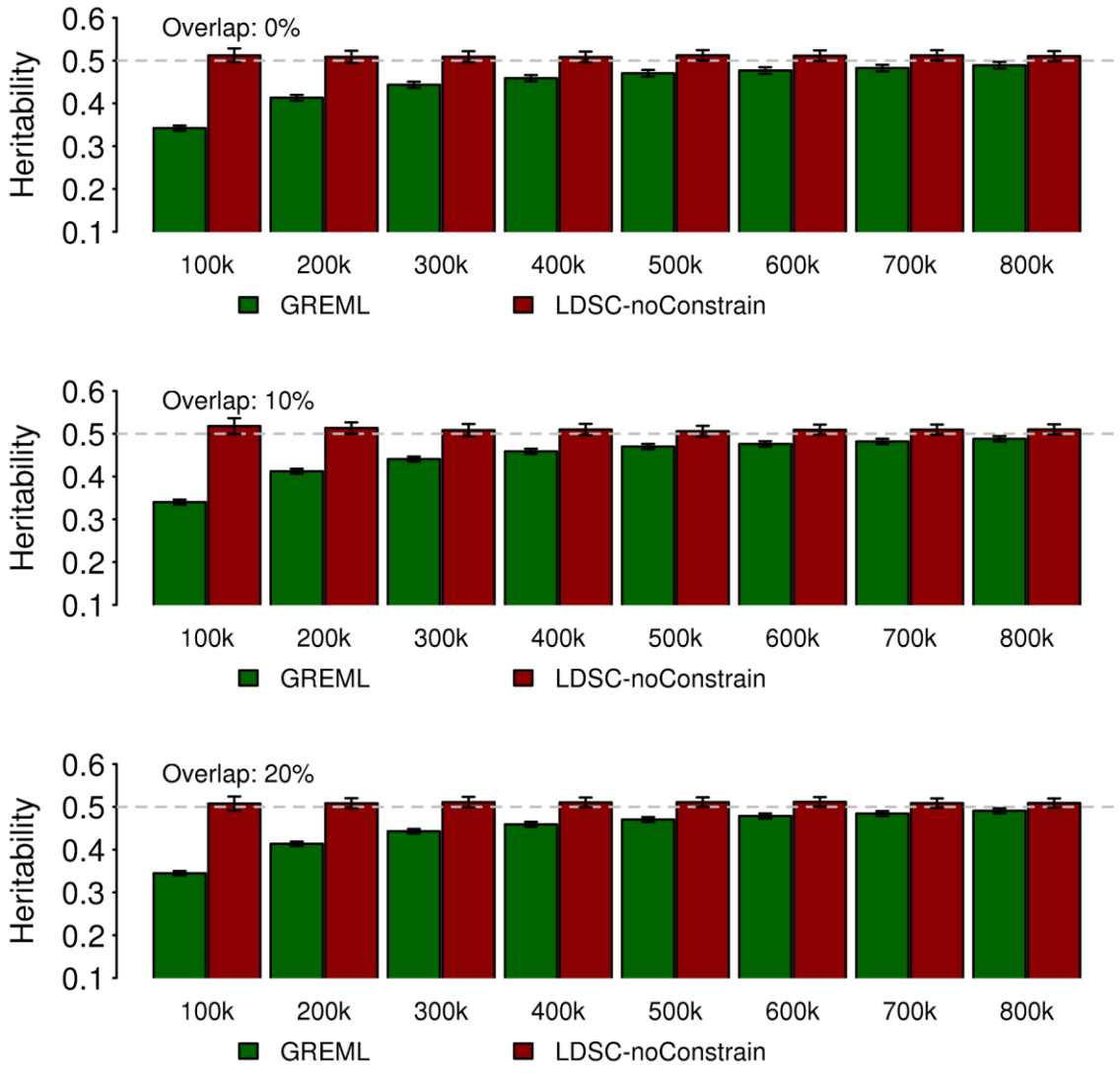

Figure S11. Estimated heritability using simulated phenotypes based on UKBB genotypes.

The number of causal SNPs in each bin reduced proportionally when the total number of SNPs decreased.

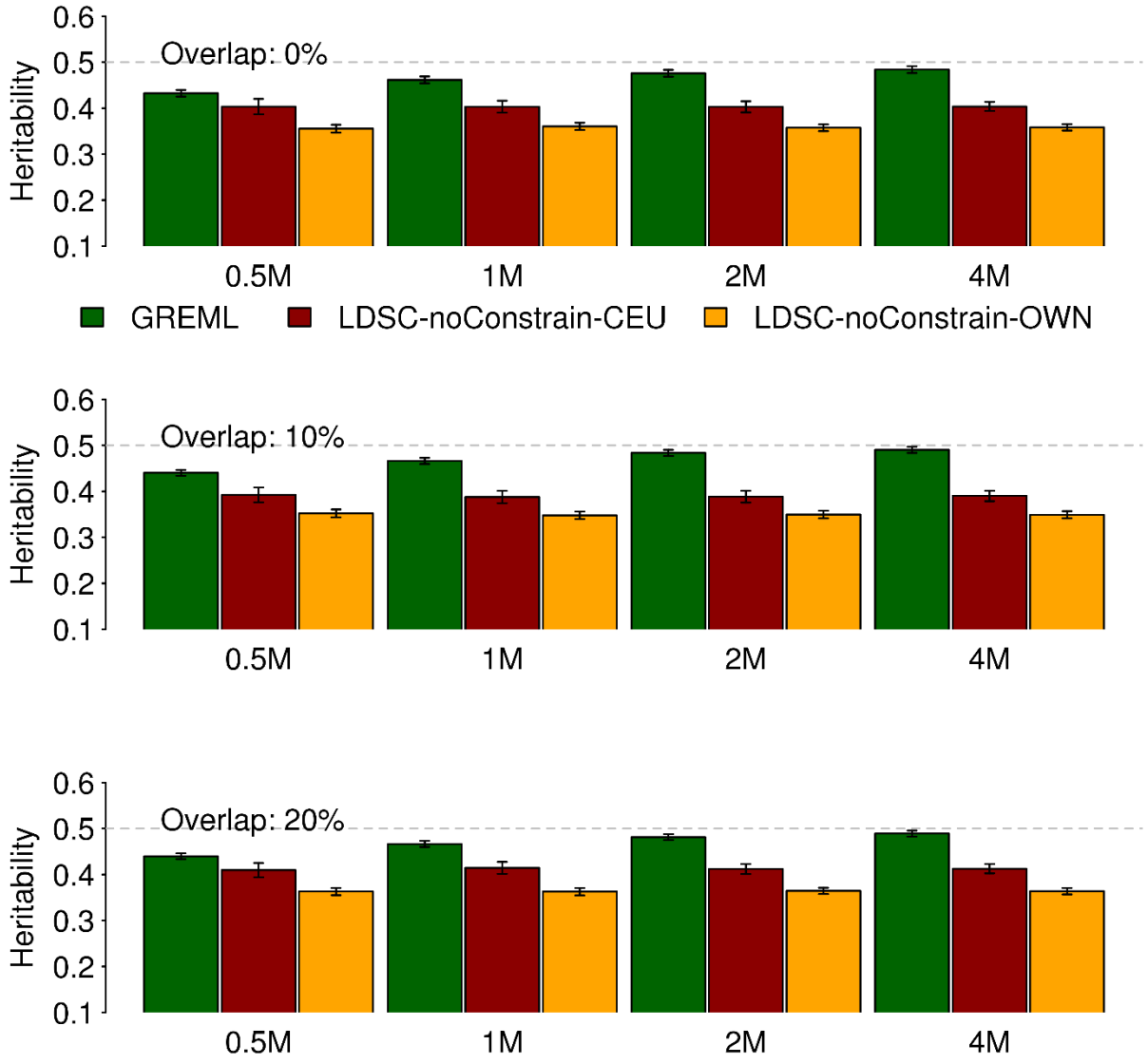

Figure S12. Estimated heritability using simulated phenotypes based on UKBB genotypes.

The number of SNPs in X-axis refers to SNPs used for GREML and LDSC. The simulated phenotypes were based on 6M SNPs (after QC) from which 10,000 SNPs were randomly selected as causal variants. SNPs in each bin were randomly and independently drawn from 6M SNPs. Thus, the number of causal SNPs in each bin reduced proportionally when the total number of SNPs decreased. LDSC-noConstrain-CEU: using LD scores estimated based on the 1KG CEU reference sample. LDSC-noConstrain-OWN: using LD scores estimated based on the target sample (i.e. UKBB).

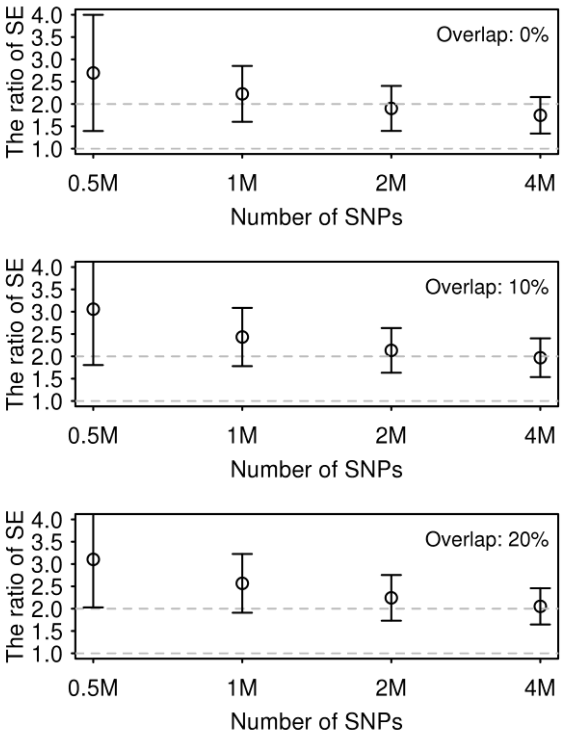

Figure S13. The ratio of SE of LDSC estimate (using LD-scores estimated from 1KG reference sample) to that of GREML based on the results from Figure S10.

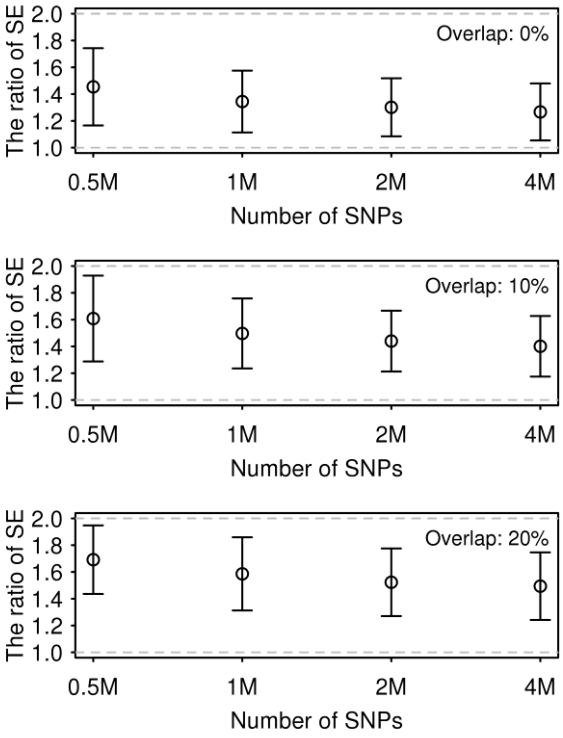

Figure S14. The ratio of SE of LDSC estimate (using LD-scores estimated from UKBB (in-sample)) to that of GREML based on the results from Figure S10.

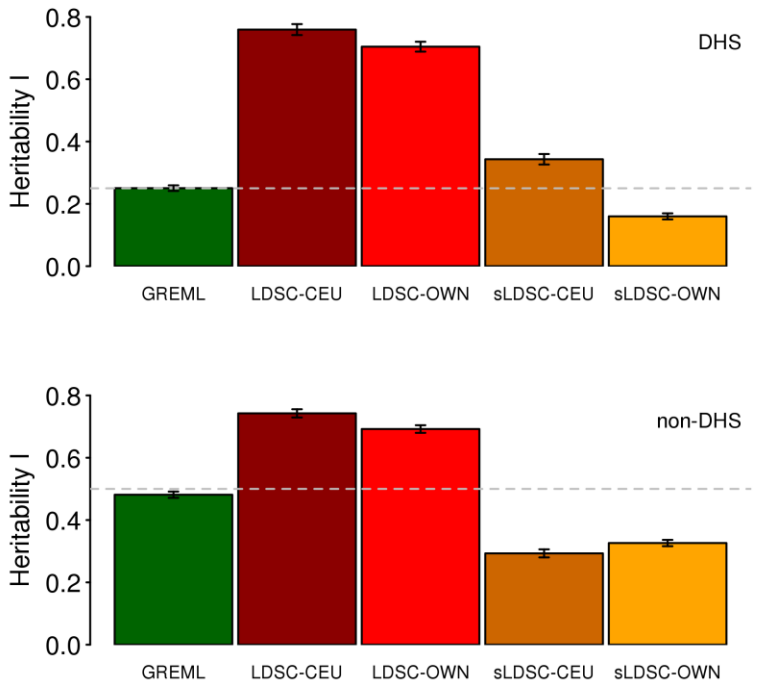

Figure S15. Estimated SNP-heritability of simulated data based on a genomic partitioning model.

LDSC-CEU: Using LD-scores estimated from 1KG reference data. LDSC-OWN: Using LD-scores estimated from UKBB. sLDSC-CEU: Using stratified LD-scores estimated from 1KG reference data. sLDSC-OWN: Using stratified LD-scores estimated from UKBB.

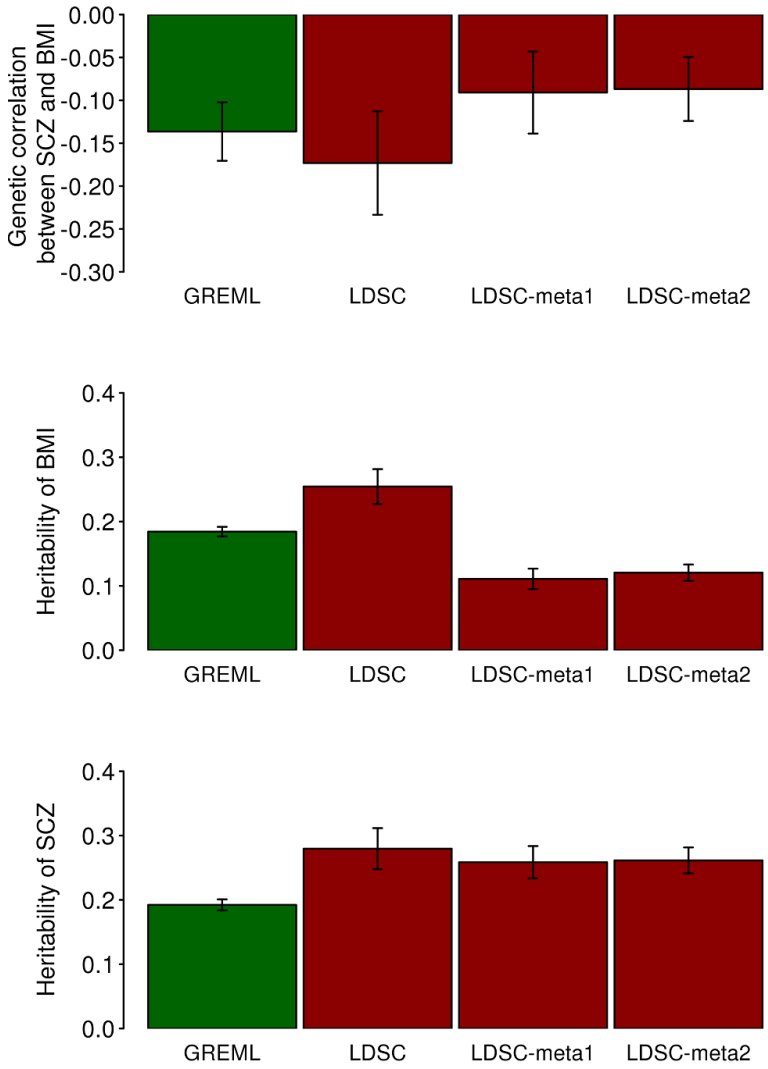

Figure S16. Genetic correlation between BMI and SCZ and heritability estimated with GREML and LDSC.

GREML: Analysis was based on quality controlled genetic data for BMI (from UK Biobank with 111,019 individuals and 518,992 SNPs) and schizophrenia (from PGC with 41,630 individuals and 518,992 SNPs). LDSC: The data sets used in LDSC were the same as in GREML. LDSC-meta1: GWAS summary statistics for BMI were based on meta-analysed GWAS results of UKBB individual-level genetic data (with 111,019 individuals and 518,992 SNPs) and of GIANT (245,051 individuals and 477,163 SNPs). For SCZ, the GWAS summary statistics from the full PGC sample based on 77,096 individuals were used. LDSC-meta2: The data sets used in LDSC-meta2 were the same as in LDSC-meta1 except for the increased number of SNPs (1,011,748) with which its performance was checked. Bars are standard errors.

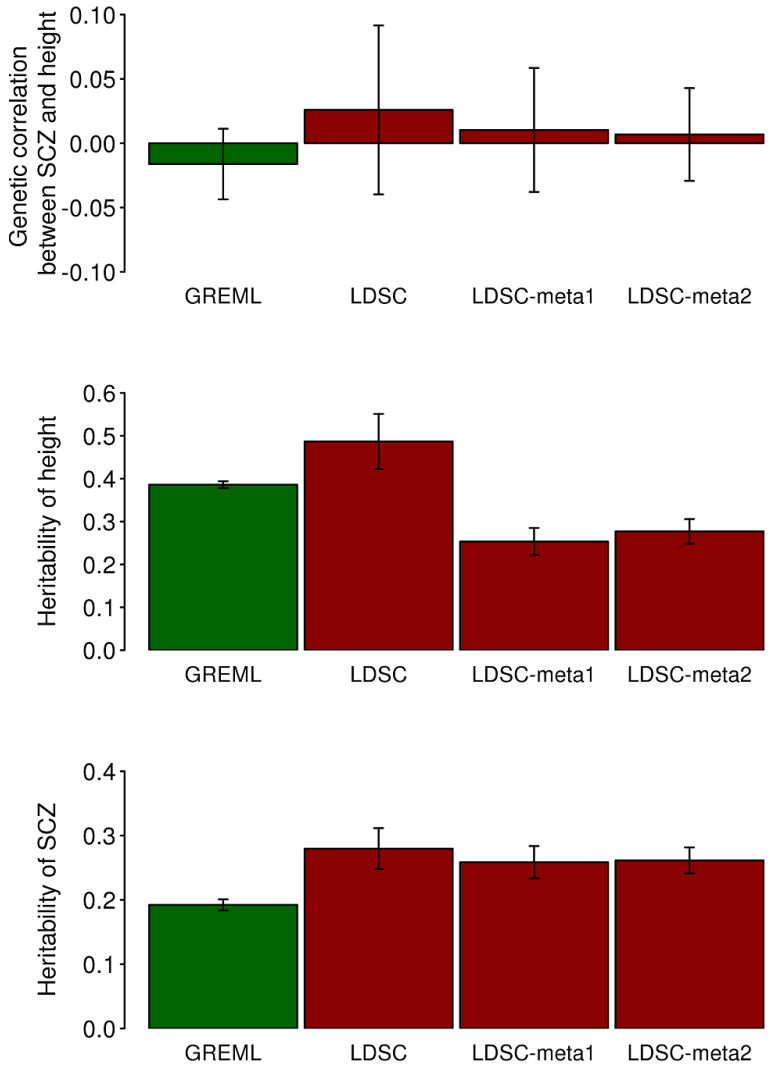

Figure S17. Genetic correlation between height and SCZ and heritability estimated with GREML and LDSC.

GREML: Analysis was based on quality controlled genetic data for height (from UK Biobank with 111,143 individuals and 518,992 SNPs) and schizophrenia (from PGC with 41,630 individuals and 518,992 SNPs). LDSC: The data sets used in LDSC were the same as in GREML. LDSC-meta1: GWAS summary statistics for height were based on meta-analysed GWAS results of UKBB individual-level genetic data (with 111,143 individuals and 518,992 SNPs) and of GIANT (253,280 individuals and 476,824 SNPs). For SCZ, the GWAS summary statistics from the full PGC sample based on 77,096 individuals were used. LDSC-meta2: The data sets used in LDSC-meta2 were the same as in LDSC-meta1 except for the increased number of SNPs (1,010,783) with which its performance was checked. Bars are standard errors.

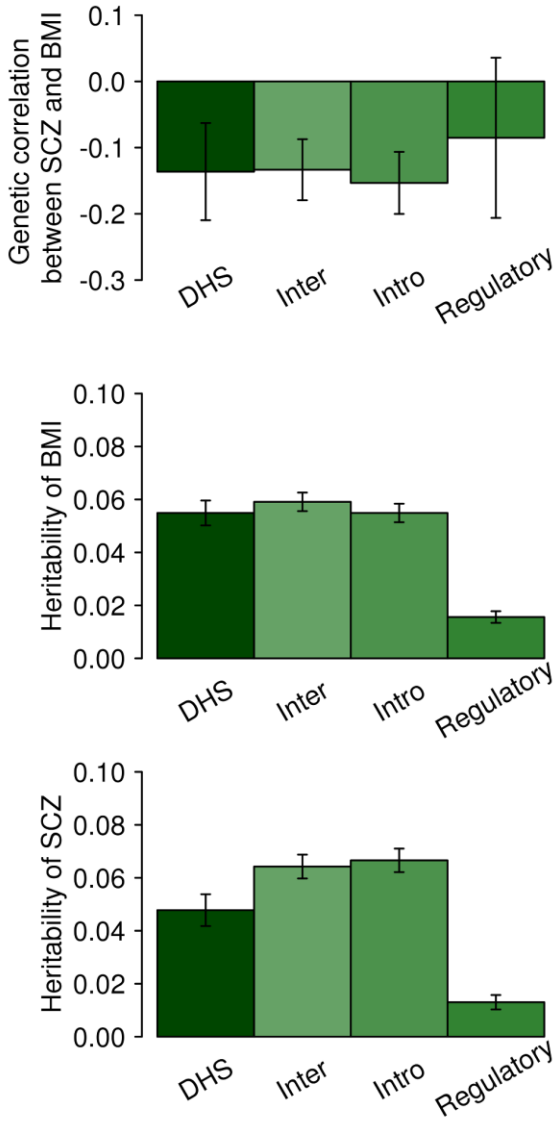

Figure S18. Genetic correlation between SCZ and BMI and heritability based on SNPs in partitioned genomic regions estimated with GREML.

A joint model was applied by fitting four genomic relationship matrices simultaneously, each estimated based on the set of SNPs belonging to each of the functional categories (DHS, intergenic and intronic regions, and regulatory). The bars are standard errors.

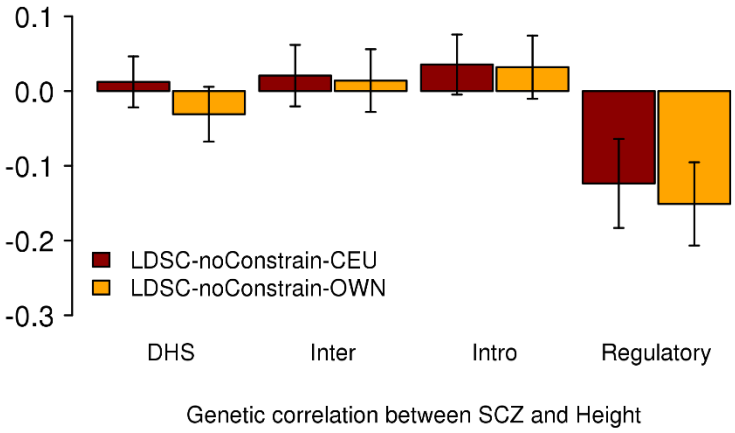

Figure S19. Genetic correlation and standard error between SCZ and height estimated with LDSC.

P-values for the estimate significantly different from 0 were 0.04, 0.38, 0.61 and 0.71 for LDSC-noConstrain-CEU, and 0.007, 0.45, 0.74 and 0.40 for LDSC-noConstrain-OWN for regulatory, intronic, intergenic and DHS regions, respectively. LDSC-noConstrain-CEU: using LD-scores estimated from 1KG reference data. LDSC-noConstrain-OWN: using LD-scores estimated from in-sample.

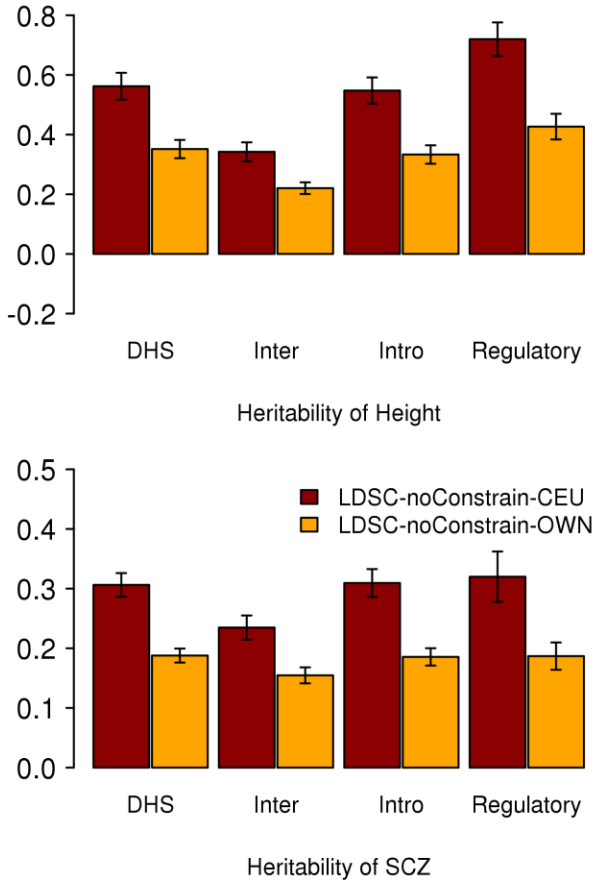

Figure S20. Heritability and standard error of height and SCZ (in liability scale) estimated with LDSC.

LDSC-noConstrain-CEU: using LD-scores estimated from 1KG reference data. LDSC-noConstrain-OWN: using LD-scores estimated from in-sample.

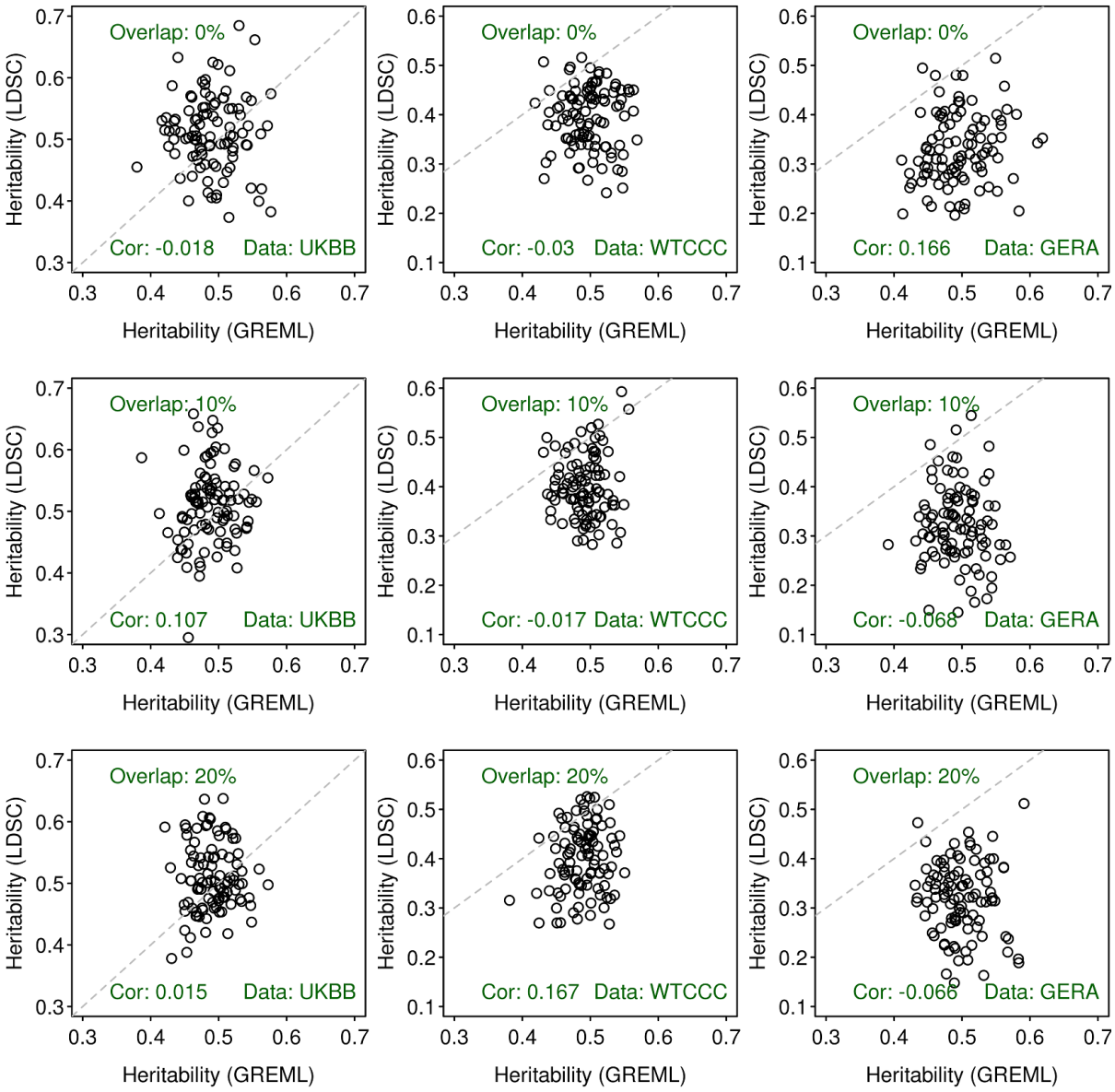

Figure S21. Scatter plot of heritability from GREML and LDSC using UKBB, WTCCC2, or GEAR data.

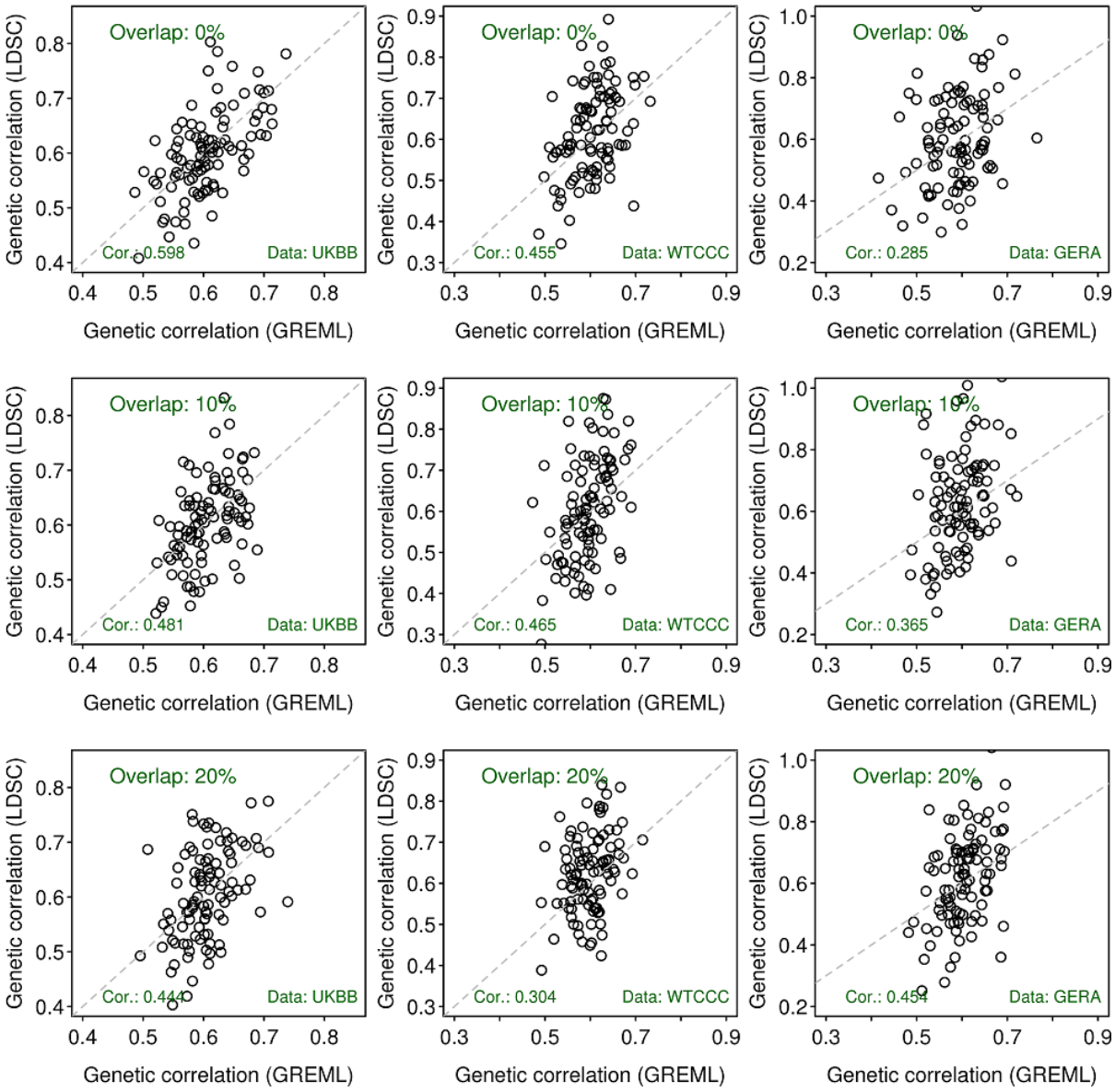

Figure S22. Scatter plot of genetic correlation from GREML and LDSC using UKBB, WTCCC2, or GEAR data.

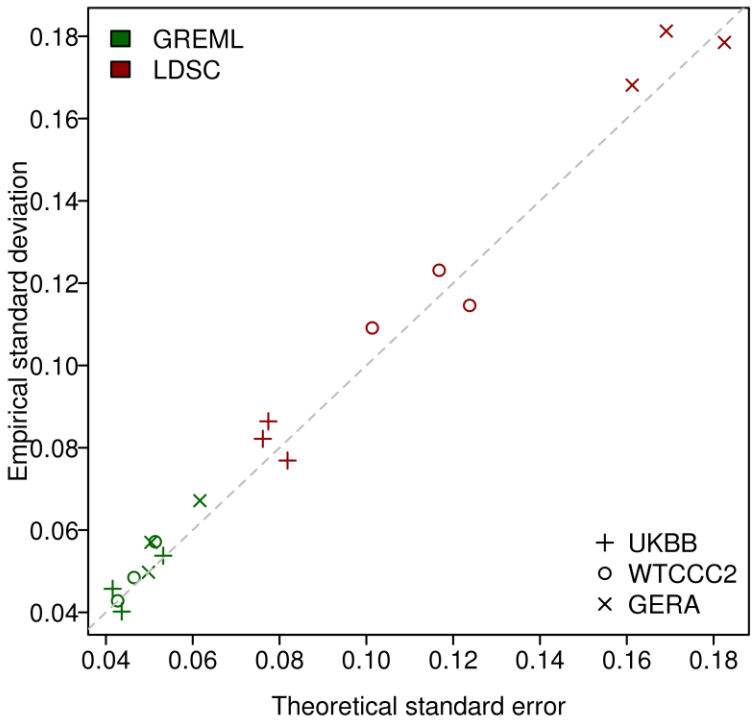

Figure S23. Empirical and theoretical SE for GREML and LDSC.

Table S1a. Covariance and heritability estimated from LDSC and GREML when a simulated genetic covariance between the first and second trait is 0.15 and -0.3 for DHS and non-DHS, respectively[a].

	%Overlapping	Method	LD score	Covariance		Variance I		Variance II	
				Estimate	SE[b]	Estimate	SE	Estimate	SE
True in DHS				0.15		0.25		0.25	
DHS	0	LDSC	CEU	-0.14	0.01	0.76	0.01	0.77	0.01
			OWN	-0.13	0.01	0.70	0.01	0.71	0.01
		SLDSC	CEU	0.09	0.01	0.54	0.31	0.57	0.06
			OWN	0.09	0.00	0.17	0.01	0.17	0.01
		GREML	GREML	0.15	0.00	0.25	0.01	0.24	0.01
	10	LDSC	CEU	-0.14	0.01	0.75	0.01	0.77	0.01
			OWN	-0.13	0.01	0.70	0.01	0.71	0.01
		SLDSC	CEU	0.10	0.01	1.40	0.85	0.52	0.04
			OWN	0.10	0.00	0.17	0.01	0.17	0.01
		GREML	GREML	0.15	0.00	0.25	0.00	0.24	0.00
	20	LDSC	CEU	-0.14	0.01	0.76	0.01	0.77	0.01
			OWN	-0.13	0.01	0.71	0.01	0.71	0.01
		SLDSC	CEU	0.09	0.01	0.64	0.18	0.54	0.05
			OWN	0.09	0.00	0.19	0.01	0.16	0.01
		GREML	GREML	0.15	0.00	0.25	0.01	0.24	0.01
True in nonDHS				-0.30		0.50		0.50	
non-DHS	0	LDSC	CEU	-0.17	0.01	0.74	0.01	0.74	0.01
			OWN	-0.17	0.01	0.69	0.01	0.68	0.01
		SLDSC	CEU	-0.20	0.01	-0.36	0.26	0.02	0.11
			OWN	-0.19	0.00	0.33	0.01	0.31	0.01
		GREML	GREML	-0.29	0.00	0.48	0.01	0.50	0.01
non-DHS	10	LDSC	CEU	-0.18	0.01	0.74	0.01	0.74	0.01

		OWN	-0.17	0.01	0.69	0.01	0.68	0.01
	SLDSC	CEU	-0.22	0.01	-1.19	0.95	0.12	0.09
		OWN	-0.20	0.00	0.31	0.01	0.31	0.01
	GREML	GREML	-0.29	0.00	0.48	0.01	0.49	0.00
	LDSC	CEU	-0.18	0.01	0.75	0.01	0.74	0.01
		OWN	-0.17	0.01	0.69	0.01	0.68	0.01
	SLDSC	CEU	-0.21	0.01	0.03	0.20	0.20	0.08
		OWN	-0.19	0.00	0.33	0.01	0.31	0.01
20	GREML	GREML	-0.29	0.00	0.48	0.01	0.49	0.01

[a] Simulation was based on 10,000 individuals that were randomly selected from UKBB with 858K SNP. Based on Gusev et al.[1] , the SNPs across the genome were stratified into two categories: DHS (194K SNPs with 2268 causal SNPs) and non-DHS (664K SNPs with 7732 causal SNPs). The genetic correlation (genetic covariance) for the simulated phenotypes between the first and second traits was 0.6 (0.15) and -0.6 (-0.3) in DHS and non-DHS region, respectively. The phenotypic variance is one. The genetic variance of two traits and covariance for DHS and non-DHS are presented in the table.

[b] The standard error were estimated based on 100 replicates.

Table S1b. Covariance and heritability estimated from LDSC and GREML when a simulated genetic covariance between the first and second trait is 0.15 and -0.15 for DHS and non-DHS, respectively[a].

	%Overlapping	Method	LD score	Covariance		Variance I		Variance II	
				Estimate	SE[b]	Estimate	SE	Estimate	SE
True in DHS				0.15		0.25		0.25	
DHS	0	LDSC	CEU	0.01	0.00	0.76	0.01	0.76	0.01
			OWN	0.01	0.00	0.71	0.01	0.71	0.01
		SLDSC	CEU	0.12	0.01	0.77	0.18	0.58	0.08
			OWN	0.09	0.00	0.17	0.01	0.17	0.01
		GREML	GREML	0.15	0.00	0.25	0.01	0.24	0.01
	10	LDSC	CEU	0.03	0.00	0.75	0.01	0.77	0.01
			OWN	0.03	0.00	0.70	0.01	0.71	0.01
		SLDSC	CEU	0.13	0.01	1.27	0.45	0.81	0.26
			OWN	0.10	0.00	0.17	0.01	0.17	0.01
		GREML	GREML	0.15	0.00	0.25	0.00	0.24	0.00
	20	LDSC	CEU	0.05	0.00	0.76	0.01	0.77	0.01
			OWN	0.05	0.00	0.71	0.01	0.71	0.01
		SLDSC	CEU	0.12	0.01	1.17	0.76	-0.24	0.80
			OWN	0.09	0.00	0.18	0.01	0.16	0.01
		GREML	GREML	0.15	0.00	0.25	0.00	0.24	0.01
True in nonDHS				-0.15		0.50		0.50	
non-DHS	0	LDSC	CEU	0.00	0.00	0.74	0.01	0.74	0.01
			OWN	0.00	0.00	0.69	0.01	0.68	0.01
		SLDSC	CEU	-0.11	0.01	-0.23	0.30	0.00	0.15
			OWN	-0.09	0.00	0.31	0.01	0.32	0.01

		GREML	GREML	-0.15	0.00	0.48	0.01	0.49	0.01
		LDSC	CEU	0.02	0.00	0.74	0.01	0.74	0.01
			OWN	0.02	0.00	0.69	0.01	0.69	0.01
		SLDSC	CEU	-0.12	0.01	-1.97	1.47	-0.13	0.21
			OWN	-0.10	0.00	0.31	0.01	0.31	0.01
	10	GREML	GREML	-0.15	0.00	0.48	0.00	0.49	0.00
		LDSC	CEU	0.04	0.00	0.75	0.01	0.74	0.01
			OWN	0.04	0.00	0.69	0.01	0.68	0.01
		SLDSC	CEU	-0.11	0.01	-0.61	0.89	-0.02	0.27
			OWN	-0.10	0.00	0.31	0.01	0.33	0.01
	20	GREML	GREML	-0.14	0.00	0.48	0.01	0.49	0.01

[a] Simulation was based on 10,000 individuals that were randomly selected from UKBB with 858K SNP. Based on Gusev et al.[1], the SNPs across the genome were stratified into two categories: DHS (194K SNPs with 2268 causal SNPs) and non-DHS (664K SNPs with 7732 causal SNPs). The genetic correlation (genetic covariance) for the simulated phenotypes between the first and second traits was 0.6 (0.15) in DHS and -0.3 (-0.15) in non-DHS region, respectively. The phenotypic variance is one. The genetic variance of two traits and covariance for DHS and non-DHS are presented in the table.

[b] The standard error were estimated based on 100 replicates.

Table S2 Heritability and genetic correlation based on different data sets

Method	#SNPs	Data	#individuals		h^2 height		h^2 SCZ (liability scale)		Genetic correlation		
			Mean	SD	Estimate	SE	Estimate	SE	Estimate	SE	P
GREML	518992	UKBB+ SCZ	152961.0		0.386	4.10E-03	0.192	4.39E-03	-0.016	1.40E-02	0.247
LDSC	516519	UKBB+ SCZ	151385.1	1433.7	0.487	3.28E-02	0.280	1.63E-02	0.026	3.35E-02	0.439
LDSC-meta1	476824	UKBB+ GIANT+ PGCSCZ	437303.9	8173.6	0.253	1.62E-02	0.259	1.28E-02	0.010	2.46E-02	0.675
LDSC-meta2	1010783	UKBB+ GIANT+ PGCSCZ	428349.1	28211.2	0.277	1.47E-02	0.261	1.03E-02	0.007	1.84E-02	0.712

GREML: Analysis was based on quality controlled genetic data for height (from UK Biobank with 111,143 individuals and 518,992 SNPs) and schizophrenia (from PGC with 41,630 individuals and 518,992 SNPs).

LDSC: The data sets used in LDSC were the same as in GREML.

LDSC-meta1: GWAS summary statistics for height were based on meta-analysed GWAS results of UKBB individual-level genotype data (with 111,143 individuals and 518,992 SNPs) and of GIANT (253,280 individuals and 476,824 SNPs). For SCZ, the GWAS summary statistics from the full PGC sample based on 77,096 individuals was used.

LDSC-meta2: The data sets used in LDSC-meta2 were the same as in LDSC-meta1 except for the increased number of SNPs (1,010,783) with which its performance was checked.

Due to different call rates of each SNP, the numbers of individuals for each SNP used in GWAS were different (see the column of the number of individuals with mean and SD).

Table S3 Heritability estimated from GREML and stratified LDSC based on the real data sets in the genomic partitioning analyses

BMI									
		GREML					**Stratified LDSC[a]**		
	Proportion of SNPs	Estimate	SE (Estimate)	Ratio[b]	SE (ratio)[c]	Enrichment _P-value[d]	Estimate	SE (Estimate)	Ratio
Regulatory	0.065	0.016	0.002	0.085	0.012	5.61E-02	0.012	0.006	0.060
DHS	0.225	0.055	0.005	0.298	0.025	1.74E-03	0.084	0.015	0.411
Intron	0.287	0.055	0.004	0.298	0.018	5.54E-01	0.054	0.008	0.262
Intergenic	0.424	0.059	0.004	0.320	0.017	1.76E-09	0.055	0.009	0.267
Total	1.000	0.185	-	1.000	-	-	0.206	0.012	1.000
Whole genome[e]	1.000	0.184	0.004	-	-	-	0.255	0.014	-
Height									
		GREML					**Stratified LDSC**		
	Proportion of SNPs	Estimate	SE (Estimate)	Ratio	SE (ratio)	Enrichment _P-value	Estimate	SE (Estimate)	Ratio
Regulatory	0.065	0.094	0.004	0.246	0.009	7.60E-92	0.117	0.012	0.261
DHS	0.225	0.167	0.005	0.438	0.014	2.38E-54	0.421	0.031	0.942
Intron	0.287	0.071	0.004	0.187	0.011	7.83E-22	-0.017	0.014	-0.038
Intergenic	0.424	0.049	0.004	0.130	0.009	7.29E-226	-0.074	0.016	-0.165
Total	1.000	0.381	-	1.000	-	-	0.447	0.026	1.000
Whole genome	1.000	0.386	0.004	-	-	-	0.487	0.033	-
SCZ[f]									
		GREML					**Stratified LDSC**		
	Proportion of SNPs	Estimate	SE (Estimate)	Ratio	SE (ratio)	Enrichment _P-value	Estimate	SE (Estimate)	Ratio
Regulatory	0.065	0.013	0.003	0.068	0.014	1.67E+00	0.021	0.008	0.095
DHS	0.225	0.048	0.006	0.249	0.031	3.62E-01	0.068	0.020	0.303
Intron	0.287	0.067	0.004	0.347	0.022	3.72E-03	0.076	0.009	0.340
Intergenic	0.424	0.064	0.005	0.335	0.022	2.41E-05	0.059	0.010	0.262

Total	1.000	0.192	-	1.000	-	-	0.244	-	1.000
Whole genome	1.000	0.192	0.004	-	-	-	0.280	0.016	-

[a]GWAS summary statistics used in the stratified LDSC were based on the same genotype data used in GREML (Table 3). Because we used our own annotation that had discrete categories, for which the stratified LDSC does not provide stratified LD scores, we had to calculate stratified LD scores using the individual-level genotype data.

[b]Ratio is the proportion of each estimate over the sum of all estimates.

[c]SE (ratio) is estimated using the delta method implemented in MTG2. For such discrete annotated categories, the stratified LDSC does not provide the SE of the ratio.

[d]Enrichment p-value was obtained from the Wald test.

[e]Whole-genome analyses were based on all the SNPs across the genome (i.e. Table 3).

[f]The estimates of SCZ are in the liability scale.

Table S4. The computational and memory requirements for bivariate GREML and LDSC for 800k SNPs using a single CPU running at 2.1 GHz

Methods	#individuals/trait	Memory	Time (minute)
GREML	5,000	1333MB	4
LDSC		544MB	1
GREML	10,000	6228MB	32
LDSC		544MB	1
GREML	20,000	18.5GB	198
LDSC		544MB	1

It is noted that MTG2 software used for the bivariate GREML can facilitate parallel computing and increase the computation efficiency approximately by a factor of 10 when using 20 CPUs.

SUPPLEMENTAL METHODS

Genetic data

Schizophrenia (SCZ) data

The SCZ GWAS data were from the second phase of the Psychiatric Genomic Consortium[2] (PGC). Quality control (QC) and imputation of raw autosomal SNPs was performed by PGC using the imputation program IMPUTE2/SHAPEIT[3; 4] with CEU samples from the 1000 Genomes Project dataset[5] as the reference set. Post-imputation quality control was performed for each cohort as described in Mehta et al.[6], and subsequently merged across all cohorts. Based on the merged genotype data, we utilized HapMap3 SNPs with a call rate ≥ 0.9 and individuals with a call rate ≥ 0.9. In addition, one individual in a pair was randomly excluded if their genomic relationship was ≥ 0.05. After QC, 688,145 SNPs and 41,630 individuals (18,987 cases and 22,673 controls) remained.

UK Biobank (UKBB)

Based on the data released in the first version of UKBB[7], which were collected from a community sample, there were initially 152,249 individuals and 72,355,667 imputed SNPs available. Non-ambiguous strand SNPs identified in HapMap3 with imputation INFO ≥ 0.6, minor allele frequency (MAF) ≥ 0.01, call rate ≥ 0.95 and Hardy-Weinberg equilibrium P-value ≥ 10-7 were retained. Individuals with call rate ≥ 0.95 and clustered as Caucasian, which were within six standard deviations from the mean of the EUR reference sample[8] based on the first and second principal components of the genomic relationship matrix were retained. Similar to the SCZ sample, one individual per pair was randomly excluded if their genomic relationship

was > 0.05 within the UKBB sample. In addition, UKBB samples were excluded if their genomic relationship with SCZ samples was > 0.05. After QC, 111,330 individuals, 858,991 SNPs were available. In addition to the imputed UKBB data, the raw genotype data of UKBB (UKBBr) were used for a comparison analysis. Raw genotypes were available for 805,426 SNPs from the UK Biobank and UK BiLEVE Axiom array. After the same QC process as above and matching with the HapMap3 SNPs, 111,330 individuals and 123,921 SNPs were used for the simulation study.

Wellcome Trust Case Control Consortium 2 (WTCCC2)

The WTCCC2 data[9-12] were combined from four disease datasets (ischaemic stroke, multiple sclerosis, primary biliary cirrhosis, and psoriasis) and two controls (1958 Birth cohort and UK Blood Service) which were genotyped with Illumina Human 1M-Duo BeadChip and QCed separately as follows: SNPs were excluded due to either call rate ≤ 0.95, MAF < 0.01, HWE P-value <1E-4 or significantly (P < 1E-5) different call rates between cases and controls. Individuals were excluded due to either individual call rate ≤ 0.97, being duplicated, genomic relationship > 0.185 or being of non-CEU ancestry. In the combined genotype data, we further excluded those SNPs with a call rate ≤ 0.95, HWE P-value ≤ 10E-7, or MAF ≤ 0.01 and one random individual in a pair with high relatedness (> 0.05). After QC, 20,659 individuals each with 432,663 SNPs were available for the simulation study.

Genetic epidemiology research on adult health and aging cohort (GERA)

The details of GERA data and its QC process was described in Lee et al.[13]. Briefly, genetic data were strictly from participants of the Kaiser Permanente GERA cohort with European ancestry[14]. In the QC step, SNPs were excluded due to either call rate ≤ 0.95, HWE p-

value ≤ 10E-4, or MAF ≤ 0.01. Only HapMap3 SNPs were retained for the simulation and analyses. Individuals were excluded due to either call rate ≤ 0.95 or being population outliers (i.e. greater than six SD from the first and second principal components). One individual in a pair with genomic relationship larger than 0.05 was randomly excluded. After QC, 46,345 individuals each with 239,976 SNPs were available for the simulation study.

Simulation

The simulation process was based on the individual-level genotype data of UKBB (111,330 individuals and 858,991 imputed or 123,921 raw genotyped SNPs), WTCCC2 (20,659 individuals and 432,663 SNPs), and GERA (46,345 individuals and 239,976 SNPs), respectively.

Phenotype Simulation

For the bivariate model, 10,000 SNPs were randomly selected as causal variants, and assigned two sets of causal effects following a multivariate normal distribution with mean [0, 0] and (co)variance matrix as $\begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix}$. Thus, the genetic correlation between the first and second traits was 0.6. True breeding values or genetic profile scores were obtained from the products of SNP genotypes and the corresponding SNP effects. The simulated phenotypes were generated as the sum of the true genetic profile scores and the residual effects that were drawn from a multivariate normal distribution with mean [0, 0] and the covariance matrix $\begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$. Therefore, the simulated true heritability was 0.5 for both first and second traits.

For the sample available in each dataset, a random set of 10,000 individuals was made available for the first and second trait, respectively, such that the percentage of overlapping individuals between the first and second traits was 0, 10 or 20%. Genetic correlation between

two simulated traits was estimated using GREML and LDSC. The details of the methods are well documented elsewhere[15-19]. The number of replicates was 100. The simulation process was conducted using MTG2[15; 20].

For sensitivity analyses using UKBB data, a different number of individuals (2500, 5000 or 15000) or a different number of SNPs (100, 200 … 700k) was used. For testing a different number of SNPs, a subset of SNPs was randomly selected from 858,991 SNPs from which 10,000 SNPs were randomly selected as causal variants to simulate phenotypes as above. We also tested a situation where the number of causal SNPs was proportionally reduced in a subset of randomly selected SNPs that used the same causal variants as the original 858,991 SNPs.

*Simulation for genomic partitioning analyses***.**

This simulation was based on the QCed UKBB genotype data with 858,991 SNPs. According to Gusev et al.[1], the genomic regions were divided into two categories: DHS (194,778 SNPs) and non-DHS (664,213 SNPs). Phenotypes were simulated for the categories such that the heritability was 0.25 and 0.5 for the DHS region and non-DHS regions, respectively. The genetic correlation between the first and second traits was 0.6 and -0.6 for DHS and non-DHS regions, respectively. In an alternative scenario, we simulated the genetic correlation between DHS regions being 0.6 and the genetic correlation between non-DHS regions being -0.3. For the genetic correlation, we gave substantially different values between the DHS and non-DHS regions to make the contrast clear in the estimation, which also increased the power to access the performance of the methods. The residual correlation between two traits was 0.8. For the sample, 10,000 individuals were randomly selected for each trait to perform GREML and LDSC with different levels (0, 10, or 20%) of overlapping individuals between the two traits. For the GREML genomic partitioning analysis, genomic relationship matrices[9] (estimated based on the

information of the functional categories) were jointly fitted to estimate SNP-heritability and genetic correlation for each category. For the LDSC partitioning analysis, we used the sLDSC (i.e. --h2 flag in sLDSC software)[21], which could estimate SNP-heritability for each category, following Finucane et al.[21] in the case of non-overlapping categories. We also used sLDSC to estimate genetic correlation between traits for each category (i.e. --rg flag). Unlike --h2 function in sLDSC, there is neither documented instruction nor a publication verifying the function, and the software does not provide SE of estimates. Therefore, we limited our use of the --rg function in sLDSC to simulated data only. Because of the limitation of sLDSC (--rg), we also used LDSC estimates based on a subset of SNPs belonging to each category to obtain a genetic correlation for each category, which also give more insights into the method.

Real data analyses

We used LDSC and GREML to estimate SNP-heritability and genetic correlation for the real data sets. Traits of interest were height, body mass index (BMI), and SCZ. After QC as described above, the number of phenotypic records was 111,143 for height and 111,019 for BMI from the UKBB, and 41,630 for SCZ from PGC. The number of SNPs was 518,992. The SNP number was reduced compared to the simulation study because only the SNPs common to both UKBB and SCZ could be used to build a genomic relationship in GREML. Phenotypic records of height and BMI were adjusted for age at interview, the assessment centre at which participant consented, genotype batch, year of birth and the first 15 principal components (PCs). SCZ were adjusted for sex, cohort and the first 15 PCs. The pre-adjusted phenotypes were used for GREML and LDSC estimations. In addition, publicly available GWAS summary statistics for height, BMI and SCZ were used for LDSC (LDSC-meta). The GWAS results for height and

BMI from GIANT[22; 23] and those for SCZ from the full PGC samples[2] were meta-analysed with SNPs treated as fixed effects.

Genomic partitioning analysis was applied. Based on Gusev et al.[1], we stratified the genome into four categories, i.e. regulatory, DHS, intronic and intergenic region. The regulatory category includes promoters (within 2kbp of a transcription start site), UTR (overlapping a 5' or 3' untranslated region) and coding (overlapping a coding exon). DHS includes chromatin zone that are sensitive to cleavage by the DNase I enzyme, observed in any cell-type. Intronic and intergenic regions include introns and all other intergenic variants except regulatory, DHS and introns. We used the GREML, LDSC and sLDSC to estimate SNP-heritability and genetic correlation for each category in the same manner as in the analyses of the simulated data except that we did not attempt using sLDSC with --rg flag for the real data analysis because of its incompleteness (e.g. SE estimates are not provided).

SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM

Stephan Ripke[1,2], Benjamin M. Neale[1,2,3,4], Aiden Corvin[5], James T. R. Walters[6], Kai-How Farh[1], Peter A. Holmans[6,7], Phil Lee[1,2,4], Brendan Bulik-Sullivan[1,2], David A. Collier[8,9], Hailiang Huang[1,3], Tune H. Pers[3,10,11], Ingrid Agartz[12,13,14], Esben Agerbo[15,16,17], Margot Albus[18], Madeline Alexander[19], Farooq Amin[20,21], Silviu A. Bacanu[22], Martin Begemann[23], Richard A. Belliveau Jr[2], Judit Bene[24,25], Sarah E. Bergen[2,26], Elizabeth Bevilacqua[2], Tim B. Bigdeli[22], Donald W. Black[27], Richard Bruggeman[28], Nancy G. Buccola[29], Randy L. Buckner[30,31,32], William Byerley[33], Wiepke Cahn[34], Guiqing Cai[35,36], Dominique Campion[37], Rita M. Cantor[38], Vaughan J. Carr[39,40], Noa Carrera[6], Stanley V. Catts[39,41], Kimberly D. Chambert[2], Raymond C. K. Chan[42], Ronald Y. L. Chen[43], Eric Y. H. Chen[43,44], Wei Cheng[45], Eric F. C. Cheung[46], Siow Ann Chong[47], C. Robert Cloninger[48], David Cohen[49], Nadine Cohen[50], Paul Cormican[5], Nick Craddock[6,7], James J. Crowley[51], David Curtis[52,53], Michael Davidson[54], Kenneth L. Davis[36], Franziska Degenhardt[55,56], Jurgen Del Favero[57], Ditte Demontis[17,58,59], Dimitris Dikeos[60], Timothy Dinan[61], Srdjan Djurovic[14,62], Gary Donohoe[5,63], Elodie Drapeau[36], Jubao Duan[64,65], Frank Dudbridge[66], Naser Durmishi[67], Peter Eichhammer[68], Johan Eriksson[69,70,71], Valentina Escott-Price[6], Laurent Essioux[72], Ayman H. Fanous[73,74,75,76], Martilias S. Farrell[51], Josef Frank[77], Lude Franke[78], Robert Freedman[79], Nelson B. Freimer[80], Marion Friedl[81], Joseph I. Friedman[36], Menachem Fromer[1,2,4,82], Giulio Genovese[2], Lyudmila Georgieva[6], Ina Giegling[81,83], Paola Giusti-Rodrı´guez[51], Stephanie Godard[84], Jacqueline I. Goldstein[1,3], Vera Golimbet[85], Srihari Gopal[86], Jacob Gratten[87], Lieuwe de Haan[88], Christian Hammer[23], Marian L. Hamshere[6], Mark Hansen[89], Thomas Hansen[17,90], Vahram Haroutunian[36,91,92], Annette M. Hartmann[81], Frans A.

Henskens[39,93,94], Stefan Herms[55,56,95], Joel N. Hirschhorn[3,11,96], Per Hoffmann[55,56,95], Andrea Hofman[55,56], Mads V. Hollegaard[97], David M. Hougaard[97], Masashi Ikeda[98], Inge Joa[99], Antonio Juliá [100], René S. Kahn[34], Luba Kalaydjieva[101,102], Sena Karachanak-Yankova[103], Juha Karjalainen[78], David Kavanagh[6], Matthew C. Keller[104], James L. Kennedy[105,106,107], Andrey Khrunin[108], Yunjung Kim[51], Janis Klovins[109], James A. Knowles[110], Bettina Konte[81], Vaidutis Kucinskas[111], Zita Ausrele Kucinskiene[111], Hana Kuzelova-Ptackova[112], Anna K. Kähler[26], Claudine Laurent[19,113], Jimmy Lee Chee Keong[47,114], S. Hong Lee[87], Sophie E. Legge[6], Bernard Lerer[115], Miaoxin Li[43,44,116], Tao Li[117], Kung-Yee Liang[118], Jeffrey Lieberman[119], Svetlana Limborska[108], Carmel M. Loughland[39,120], Jan Lubinski[121], Jouko Lönnqvist[122], Milan Macek Jr[112], Patrik K. E. Magnusson[26], Brion S. Maher[123], Wolfgang Maier[124], Jacques Mallet[125], Sara Marsal[100], Manuel Mattheisen[17,58,59,126], Morten Mattingsdal[l14,127], Robert W. McCarley[128,129], Colm McDonald[130], Andrew M. McIntosh[131,132], Sandra Meier[77], Carin J. Meijer[88], Bela Melegh[24,25], Ingrid Melle[14,133], Raquelle I. Mesholam-Gately[128,134], Andres Metspalu[135], Patricia T. Michie[39,136], Lili Milani[135], Vihra Milanova[137], Younes Mokrab[8], Derek W. Morris[5,63], Ole Mors[17,58,138], Kieran C. Murphy[139], Robin M. Murray[140], Inez Myin-Germeys[141], Bertram Müller-Myhsok[142,143,144], Mari Nelis[135], Igor Nenadic[145], Deborah A. Nertney[146], Gerald Nestadt[147], Kristin K. Nicodemus[148], Liene Nikitina-Zake[109], Laura Nisenbaum[149], Annelie Nordin[150], Eadbhard O'Callaghan[151], Colm O'Dushlaine[2], F. Anthony O'Neill[152], Sang-Yun Oh[153], Ann Olincy[79], Line Olsen[17,90], Jim Van Os[141,154], Psychosis Endophenotypes International Consortium[155], Christos Pantelis[39,156], George N. Papadimitriou[60], Sergi Papiol[23], Elena Parkhomenko[36], Michele T. Pato[110], Tiina Paunio[157,158], Milica Pejovic-Milovancevic[159], Diana O. Perkins[160], Olli Pietiläinen1[58,161], Jonathan Pimm[53], Andrew J. Pocklington[6], John Powell[140], Alkes Price[3,162], Ann E. Pulver[147], Shaun M. Purcell[82], Digby Quested[163], Henrik B.

Rasmussen[17,90], Abraham Reichenberg[36], Mark A. Reimers[164], Alexander L. Richards[6], Joshua L. Roffman[30,32], Panos Roussos[82,165], Douglas M. Ruderfer[6,82], Veikko Salomaa[71], Alan R. Sanders[64,65], Ulrich Schall[39,120], Christian R. Schubert[166], Thomas G. Schulze[77,167], Sibylle G. Schwab[168], Edward M. Scolnick[2], Rodney J. Scott[39,169,170], Larry J. Seidman[128,134], Jianxin Shi[171], Engilbert Sigurdsson[172], Teimuraz Silagadze[173], Jeremy M. Silverman[36,174], Kang Sim[47], Petr Slominsky[108], Jordan W. Smoller[2,4], Hon-Cheong So[43], Chris C. A. Spencer[175], Eli A. Stahl[3,82], Hreinn Stefansson[176], Stacy Steinberg[176], Elisabeth Stogmann[177], Richard E. Straub[178], Eric Strengman[179,34], Jana Strohmaier[77], T. Scott Stroup[119], Mythily Subramaniam[47], Jaana Suvisaari[122], Dragan M. Svrakic[48], Jin P. Szatkiewicz[51], Erik Söderman[12], Srinivas Thirumalai[180], Draga Toncheva[103], Sarah Tosato[181], Juha Veijola[182,183], John Waddington[184], Dermot Walsh[185], Dai Wang[86], Qiang Wang[117], Bradley T. Webb[22], Mark Weiser[54], Dieter B. Wildenauer[186], Nigel M. Williams[6], Stephanie Williams[51], Stephanie H. Witt[77], Aaron R. Wolen1[64], Emily H. M. Wong[43], Brandon K. Wormley[22], Hualin Simon Xi[187], Clement C. Zai[105,106], Xuebin Zheng[188], Fritz Zimprich[177], Naomi R. Wray[87], Kari Stefansson[176], Peter M. Visscher[87], Wellcome Trust Case-Control Consortium [2189], Rolf Adolfsson[150], Ole A. Andreassen[14,133], Douglas H. R. Blackwood[132], Elvira Bramon[190], Joseph D. Buxbaum[35,36,91,191], Anders D. Børglum[17,58,59,138], Sven Cichon[55,56,95,192], Ariel Darvasi[193], Enrico Domenici[194], Hannelore Ehrenreich[23], Tõnu Esko[3,11,96,135], Pablo V. Gejman[64,65], Michael Gill[5], Hugh Gurling[53], Christina M. Hultman[26], Nakao Iwata[98], Assen V. Jablensky[39,102,186,195], Erik G. Jönsson[12,14], Kenneth S. Kendler[196], George Kirov[6], Jo Knight[105,106,107], ToddLencz[197,198,199], Douglas F. Levinson[19], Qingqin S. Li[86], Jianjun Liu[188,200], Anil K. Malhotra[197,198,199], Steven A. McCarrol[l2,96], Andrew McQuillin[53], Jennifer L. Moran[2], Preben B. Mortensen[15,16,17], Bryan J. Mowry[87,201], Markus M. Nöthen[55,56], Roel A. Ophoff[38,80,34], Michael J. Owen[6,7], Aarno Palotie[2,4,161], Carlos N. Pato[110], Tracey L.

Petryshen[2,128,202], Danielle Posthuma[203,204,205], Marcella Rietsche[l77], Brien P. Riley[196], Dan Rujescu[81,83], Pak C. Sham[43,44,116], Pamela Sklar[82,91,165], David St Clair[206], Daniel R. Weinberger[178,207], Jens R. Wendland[166], Thomas Werge[17,90,208], Mark J. Daly[1,2,3], Patrick F. Sullivan[26,51,160] & Michael C. O'Donovan[6,7]

[1]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [2]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [3]Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [4]Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [5]Neuropsychiatric Genetics Research Group, Department of Psychiatry, Trinity College Dublin, Dublin 8, Ireland. [6]MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff CF24 4HQ,UK. [7]NationalCentre for Mental Health, Cardiff University, Cardiff CF244HQ, UK. [8]Eli Lilly and Company Limited, Erl Wood Manor, Sunninghill Road, Windlesham, Surrey GU20 6PH, UK. [9]Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF,UK. [10]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, DK-2800, Denmark. [11]Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts 02115,USA. [12]Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institutet, SE-17176 Stockholm, Sweden. [13]Department of Psychiatry, Diakonhjemmet Hospital, 0319 Oslo, Norway. [14]NORMENT, KG Jebsen Centre for Psychosis Research, Institute of Clinical

Medicine, University of Oslo, 0424 Oslo, Norway. [15]Centre for Integrative Register-based Research, CIRRAU, Aarhus University, DK-8210 Aarhus, Denmark. [16]National Centre for Register-based Research, Aarhus University, DK-8210Aarhus, Denmark. [17]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark. [18]State Mental Hospital, 85540 Haar, Germany. [19]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California 94305, USA. [20]Department of Psychiatry and Behavioral Sciences, Atlanta Veterans Affairs Medical Center, Atlanta, Georgia 30033, USA. [21]Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, Georgia,30322, USA. [22]Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [23]Clinical Neuroscience, Max Planck Institute of Experimental Medicine, Goettingen 37075, Germany. [24]Department of Medical Genetics, University of Pe´cs, Pe´cs H-7624, Hungary. [25]Szentagothai Research Center, University of Pe´cs, Pe´cs H-7624, Hungary. [26]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm SE-17177, Sweden. [27]Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA. [28]University Medical Center Groningen, Department of Psychiatry, University of Groningen NL-9700 RB, The Netherlands. [29]School of Nursing, Louisiana State University Health Sciences Center, New Orleans, Louisiana 70112, USA. [30]Athinoula A. Martinos Center, Massachusetts General Hospital, Boston, Massachusetts 02129, USA. [31]Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA. [32]Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [33]Department of Psychiatry, University of California at San Francisco, San Francisco, California 94143, USA. [34]University Medical Center Utrecht, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, 3584 Utrecht, The

Netherlands. [35]Department of Human Genetics, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [36]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [37]Centre Hospitalier du Rouvray and INSERM U1079 Faculty of Medicine, 76301 Rouen, France. [38]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. [39]Schizophrenia Research Institute, Sydney NSW 2010, Australia. [40]School of Psychiatry, University of New South Wales, Sydney NSW 2031, Australia. [41]Royal Brisbane and Women's Hospital, University of Queensland, Brisbane, St Lucia QLD 4072, Australia. [42]Institute of Psychology, Chinese Academy of Science, Beijing 100101, China. [43]Department of Psychiatry, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. [44]State Key Laboratory for Brain and Cognitive Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. [45]Department of Computer Science, University of North Carolina, Chapel Hill, North Carolina 27514, USA. [46]Castle Peak Hospital, Hong Kong, China. [47]Institute of Mental Health, Singapore 539747, Singapore. [48]Department of Psychiatry, Washington University, St. Louis, Missouri 63110, USA. [49]Department of Child and Adolescent Psychiatry, Assistance Publique Hopitaux de Paris, Pierre and Marie Curie Faculty of Medicine and Institute for Intelligent Systems and Robotics, Paris 75013, France. [50]Blue Note Biosciences, Princeton, New Jersey 08540, USA [51]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599-7264, USA. [52]Department of Psychological Medicine, Queen Mary University of London, London E1 1BB, UK. [53]Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, LondonWC1E6JJ, UK. [54]ShebaMedicalCenter, Tel Hashomer 52621, Israel. [55]Department of Genomics, Life and Brain Center, D-53127 Bonn, Germany. [56]Institute of Human Genetics, University of Bonn, D-53127 Bonn, Germany.

[57]Applied Molecular Genomics Unit, VIB Department of Molecular Genetics, University of Antwerp, B-2610 Antwerp, Belgium. [58]Centre for Integrative Sequencing, iSEQ, Aarhus University, DK-8000 Aarhus C, Denmark. [59]Department of Biomedicine, Aarhus University, DK-8000 Aarhus C, Denmark. [60]First Department of Psychiatry, University of Athens Medical School, Athens 11528, Greece. [61]Department of Psychiatry, University College Cork, Co. Cork, Ireland. [62]Department of Medical Genetics, Oslo University Hospital, 0424 Oslo, Norway. [63]Cognitive Genetics and Therapy Group, School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Co. Galway, Ireland. [64]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, Illinois 60637, USA. [65]Department of Psychiatry and Behavioral Sciences, North Shore University Health System, Evanston, Illinois 60201, USA. [66]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. [67]Department of Child and Adolescent Psychiatry, University Clinic of Psychiatry, Skopje 1000, Republic of Macedonia. [68]Department of Psychiatry, University of Regensburg, 93053 Regensburg, Germany. [69]Department of General Practice, Helsinki University Central Hospital, University of Helsinki P.O. Box 20, Tukholmankatu 8 B, FI-00014, Helsinki, Finland [70]Folkha¨lsan Research Center, Helsinki, Finland, BiomedicumHelsinki 1, Haartmaninkatu 8, FI-00290, Helsinki, Finland. [71]National Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland. [72]Translational Technologies and Bioinformatics, Pharma Research and Early Development, F. Hoffman-La Roche, CH-4070 Basel, Switzerland. [73]Department of Psychiatry, Georgetown University School of Medicine, Washington DC 20057, USA. [74]Department of Psychiatry, Keck School of Medicine of the University of Southern California, Los Angeles, California 90033, USA. [75]Department of Psychiatry, Virginia Commonwealth University School of Medicine,

Richmond, Virginia 23298, USA. [76]Mental Health Service Line, Washington VA Medical Center, Washington DC 20422, USA. [77]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Heidelberg, D-68159 Mannheim, Germany. [78]Department of Genetics, University of Groningen, University Medical Centre Groningen, 9700 RB Groningen, The Netherlands. [79]Department of Psychiatry, University of Colorado Denver, Aurora, Colorado 80045, USA. [80]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, California 90095, USA. [81]Department of Psychiatry, University of Halle, 06112 Halle, Germany. [82]Division of Psychiatric Genomics, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York, New York 10029, USA. [83]Department of Psychiatry, University of Munich, 80336, Munich, Germany. [84]Departments of Psychiatry and Human and Molecular Genetics, INSERM, Institut de Myologie, Hôpital de la Pitiè-Salpêtrière, Paris 75013, France. [85]Mental Health Research Centre, Russian Academy of Medical Sciences, 115522 Moscow, Russia. [86]Neuroscience Therapeutic Area, Janssen Research and Development, Raritan, New Jersey 08869, USA. [87]Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, QLD 4072, Australia. [88]Academic Medical Centre University of Amsterdam, Department of Psychiatry, 1105 AZ Amsterdam, The Netherlands. [89]Illumina, La Jolla, California, California 92122, USA. [90]Institute of Biological Psychiatry, Mental Health Centre Sct.Hans,Mental Health Services Copenhagen,DK-4000,Denmark. [91]Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [92]J. J. Peters VA Medical Center, Bronx, New York, New York 10468, USA. [93]Priority Research Centre for Health Behaviour, University of Newcastle, Newcastle NSW 2308, Australia. [94]School of Electrical Engineering and Computer Science, University of

Newcastle, Newcastle NSW 2308, Australia. [95]Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel CH-4058, Switzerland. [96]Department of Genetics, Harvard Medical School, Boston, Massachusetts, Massachusetts 02115, USA. [97]Section of Neonatal Screening and Hormones, Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut, Copenhagen DK-2300, Denmark. [98]Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Aichi, 470-1192, Japan. [99]Regional Centre for Clinical Research in Psychosis, Department of Psychiatry, Stavanger University Hospital, 4011 Stavanger, Norway. [100]Rheumatology Research Group, Vall d'Hebron Research Institute, Barcelona 08035, Spain. [101]Centre for Medical Research, The University of Western Australia, Perth WA6009, Australia. [102]The Perkins Institute for Medical Research, The University of Western Australia, Perth WA6009, Australia. [103]Department of Medical Genetics, Medical University, Sofia 1431, Bulgaria. [104]Department of Psychology, University of Colorado Boulder, Boulder, Colorado 80309, USA. [105]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario M5T 1R8, Canada. [106]Department of Psychiatry, University of Toronto, Toronto, Ontario M5T 1R8, Canada. [107]Institute of Medical Science, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [108]Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia. [109]Latvian Biomedical Research and Study Centre, Riga, LV-1067, Latvia. [110]Department of Psychiatry and Zilkha Neurogenetics Institute, Keck School of Medicine at University of Southern California, Los Angeles, California 90089, USA. [111]Faculty of Medicine, Vilnius University, LT-01513 Vilnius, Lithuania. [112]Department of Biology and Medical Genetics, 2nd Faculty of Medicine and University Hospital Motol, 150 06 Prague, Czech Republic. [113]Department of Child and Adolescent Psychiatry, Pierre and Marie Curie Faculty of Medicine, Paris 75013, France.

[114]Duke-NUS Graduate Medical School, Singapore 169857. [115]Department of Psychiatry, Hadassah-Hebrew University Medical Center, Jerusalem 91120, Israel. [116]Centre for Genomic Sciences, The University of Hong Kong, Hong Kong, China. [117]Mental Health Centre and Psychiatric Laboratory, West China Hospital, Sichuan University, Chengdu, 610041 Sichuan, China. [118]Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. [119]Department of Psychiatry, Columbia University, New York, New York 10032, USA. [120]Priority Centre for Translational Neuroscience and Mental Health, University of Newcastle, Newcastle NSW2300, Australia. [121]Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University in Szczecin, 70-453 Szczecin, Poland. [122]Department of Mental Health and Substance Abuse Services; National Institute for Health and Welfare, P.O. BOX30, FI-00271 Helsinki, Finland. [123]Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA. [124]Department of Psychiatry, University of Bonn, D-53127 Bonn, Germany. [125]Centre National de la Recherche Scientifique, Laboratoire de Génétique Moléculaire de la Neurotransmission et des Processus Neurodégénératifs, Hôpital de la Pitié Salpêtrière, 75013 Paris, France. [126]Department of Genomics Mathematics, University of Bonn, D-53127 Bonn, Germany. [127]Research Unit, Sørlandet Hospital, 4604 Kristiansand, Norway. [128]Department of Psychiatry, Harvard Medical School, Boston, Massachusetts 02115, USA. [129]VA Boston Health Care System, Brockton, Massachusetts 02301, USA. [130]Department of Psychiatry, National University of Ireland Galway, Co. Galway, Ireland. [131]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh EH16 4SB, UK. [132]Division of Psychiatry, University of Edinburgh, Edinburgh EH16 4SB, UK. [13]3Division of Mental Health and Addiction, Oslo University Hospital, 0424 Oslo, Norway. [134]Massachusetts

Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Boston, Massachusetts 02114, USA. [135]Estonian Genome Center, University of Tartu, Tartu 50090, Estonia. [136]School of Psychology, University of Newcastle, NewcastleNSW2308, Australia. [137]First Psychiatric Clinic, Medical University, Sofia 1431, Bulgaria. [138]Department P, Aarhus University Hospital, DK-8240 Risskov, Denmark. [139]Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin 2, Ireland. [140]King's College London, London SE5 8AF, UK. [141]MaastrichtUniversity Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, 6229 HX Maastricht, The Netherlands. [142]Institute of Translational Medicine, University of Liverpool, Liverpool L69 3BX, UK. [143]Max Planck Institute of Psychiatry, 80336 Munich, Germany. [144]Munich Cluster for Systems Neurology (SyNergy), 80336 Munich, Germany. [145]Department of Psychiatry and Psychotherapy, Jena University Hospital, 07743 Jena, Germany. [146]Department of Psychiatry, Queensland Brain Institute and Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, St Lucia QLD 4072, Australia. [147]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [148]Department of Psychiatry, Trinity College Dublin, Dublin 2, Ireland. [149]Eli Lilly and Company, Lilly Corporate Center, Indianapolis, 46285 Indiana, USA. [150]Department of Clinical Sciences, Psychiatry, Umea˚ University, SE-901 87 Umea˚, Sweden. [15]1DETECT Early Intervention Service for Psychosis, Blackrock, Co. Dublin, Ireland. [152]Centre for Public Health, Institute of Clinical Sciences, Queen's University Belfast, Belfast BT12 6AB, UK. [153]Lawrence Berkeley National Laboratory, University of California at Berkeley, Berkeley, California 94720, USA. [154]Institute of Psychiatry, King's College London, London SE5 8AF, UK. [156]Melbourne Neuropsychiatry Centre, University of Melbourne & Melbourne Health, Melbourne, Vic 3053,

Australia. [157]Department of Psychiatry, University of Helsinki, P.O. Box 590, FI-00029 HUS, Helsinki, Finland. [158]Public Health Genomics Unit, National Institute for Health and Welfare, P.O. BOX 30, FI-00271 Helsinki, Finland [159]Medical Faculty, University of Belgrade, 11000 Belgrade, Serbia. [160]Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina 27599-7160, USA. [161]Institute for Molecular Medicine Finland, FIMM, University of Helsinki, P.O. Box 20FI-00014, Helsinki, Finland [162]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA. [163]Department of Psychiatry, University of Oxford, Oxford, OX3 7JX, UK. [164]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [165]Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [166]PharmaTherapeutics Clinical Research, Pfizer Worldwide Research and Development, Cambridge, Massachusetts 02139, USA. [167]Department of Psychiatry and Psychotherapy, University of Gottingen, 37073 Göttingen, Germany. [168]Psychiatry and Psychotherapy Clinic, University of Erlangen, 91054 Erlangen, Germany. [169]Hunter New England Health Service, Newcastle NSW 2308, Australia. [170]School of Biomedical Sciences, University of Newcastle, Newcastle NSW 2308, Australia. [171]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. [172]University of Iceland, Landspitali, National University Hospital, 101 Reykjavik, Iceland. [173]Department of Psychiatry and Drug Addiction, Tbilisi State Medical University (TSMU), N33, 0177 Tbilisi, Georgia. [174]Research and Development, Bronx Veterans Affairs Medical Center, New York, New York 10468, USA. [175]Wellcome Trust Centre for Human Genetics, Oxford OX3 7BN, UK. [176]deCODE Genetics, 101 Reykjavik, Iceland. [177]Department of Clinical Neurology, Medical University of Vienna, 1090 Wien, Austria. [178]Lieber Institute for Brain Development, Baltimore, Maryland 21205,

USA. [179]Department of Medical Genetics, University Medical Centre Utrecht, Universiteitsweg 100, 3584 CG, Utrecht, The Netherlands. [180]Berkshire Healthcare NHS Foundation Trust, Bracknell RG12 1BQ, UK. [181]Section of Psychiatry, University of Verona, 37134 Verona, Italy. [182]Department of Psychiatry, University of Oulu, P.O. Box 5000, 90014, Finland. [183]University Hospital of Oulu, P.O. Box 20, 90029 OYS, Finland. [184]Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin 2, Ireland. [185]Health Research Board, Dublin 2, Ireland. [186]School of Psychiatry and Clinical Neurosciences, The University of Western Australia, Perth WA6009, Australia. [187]Computational Sciences CoE, Pfizer Worldwide Research and Development, Cambridge, Massachusetts 02139, USA. [188]Human Genetics, Genome Institute of Singapore, A*STAR, Singapore 138672. [190]University College London, London WC1E 6BT, UK. [191]Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [192]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, 52428 Juelich, Germany. [193]Department of Genetics, The Hebrew University of Jerusalem, 91905Jerusalem, Israel. [194]Neuroscience Discovery and Translational Area, Pharma Research and Early Development, F. Hoffman-La Roche, CH-4070 Basel, Switzerland. [195]Centre for Clinical Research in Neuropsychiatry, School of Psychiatry and Clinical Neurosciences, The University of Western Australia, Medical Research Foundation Building, Perth WA6000, Australia. [196]Virginia Institute for Psychiatric and Behavioral Genetics, Departments of Psychiatry and Human and Molecular Genetics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [197]The Feinstein Institute for Medical Research, Manhasset, New York 11030, USA. [198]The Hofstra NS-LIJ School of Medicine, Hempstead, New York 11549, USA. [199]The Zucker Hillside Hospital, Glen Oaks, New York 11004, USA. [200]Saw Swee Hock School of Public Health, National University of Singapore, Singapore

117597, Singapore. [201]Queensland Centre for Mental Health Research, University of Queensland, Brisbane 4076, Queensland, Australia. [202]Center for Human Genetic Research and Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [203]Department of Child and Adolescent Psychiatry, Erasmus University Medical Centre, Rotterdam 3000, The Netherlands. [204]Department of Complex Trait Genetics, Neuroscience Campus Amsterdam, VU University Medical Center Amsterdam, Amsterdam 1081, The Netherlands. [205]Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam 1081, The Netherlands. [206]University of Aberdeen, Institute of Medical Sciences, Aberdeen AB25 2ZD, UK. [207]Departments of Psychiatry, Neurology, Neuroscience and Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA. [208]Department of Clinical Medicine, University of Copenhagen, Copenhagen 2200, Denmark.

REFERENCE

1. Gusev, A., Lee, S.H., Trynka, G., Finucane, H., Vilhjálmsson, B.J., Xu, H., Zang, C., Ripke, S., Bulik-Sullivan, B., and Stahl, E. (2014). Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. The American Journal of Human Genetics 95, 535-552.
2. Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421-427.
3. Delaneau, O., Marchini, J., and Zagury, J.-F. (2012). A linear complexity phasing method for thousands of genomes. Nature methods 9, 179-181.
4. Howie, B., Marchini, J., and Stephens, M. (2011). Genotype imputation with thousands of genomes. G3 (Bethesda, Md) 1, 457-470.
5. Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Natrue Genetics 45, 984-994.
6. Mehta, D., Tropf, F.C., Gratten, J., Bakshi, A., Zhu, Z., Bacanu, S.-A., Hemani, G., Magnusson, P.K.E., Barban, N., Esko, T., et al. (2016). Evidence for genetic overlap between schizophrenia and age at first birth in women. JAMA Psychiatry 73, 497-505.
7. Collins, R. (2012). What makes UK Biobank special? The Lancet 379, 1173-1174.
8. The Genomes Project Consortium. (2015). A global reference for human genetic variation. Nature 526, 68-74.
9. International Stroke Genetics Consortium, and Wellcome Trust Case Control Consortium. (2012). Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nature genetics 44, 328-333.
10. The International Multiple Sclerosis Genetics, C., Wellcome Trust Case Control, C., Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C.A., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., et al. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214-219.
11. Mells, G.F., Floyd, J.A.B., Morley, K.I., Cordell, H.J., Franklin, C.S., Shin, S.-Y., Heneghan, M.A., Neuberger, J.M., Donaldson, P.T., Day, D.B., et al. (2011). Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nature genetics 43, 329-332.
12. Tsoi, L.C., Spain, S.L., Knight, J., Ellinghaus, E., Stuart, P.E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J.E., et al. (2012). Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat Genet 44, 1341-1348.
13. Lee, S.H., Weerasinghe, W.M.S.P., Wray, N.R., Goddard, M.E., and van der Werf, J.H.J. (2017). Using information of relatives in genomic prediction to apply effective stratified medicine. Scientific Reports 7, 42091.
14. Banda, Y., Kvale, M.N., Hoffmann, T.J., Hesselson, S.E., Ranatunga, D., Tang, H., Sabatti, C., Croen, L.A., Dispensa, B.P., Henderson, M., et al. (2015). Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics 200, 1285-1295.
15. Lee, S.H., and van der Werf, J. (2016). MTG2: An efficient algorithm for multivariate linear mixed model analysis based on genomic information. Bioinformatics 32, 1420-1422.
16. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: A tool for genome-wide complex trait analysis. American Journal of Human Genetics 88, 76-82.
17. Bulik-Sullivan, B.K., Loh, P.-r., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. 47, 291-295.

18. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia of the Wellcome Trust Consortium, Duncan, L., et al. (2015). An Atlas of Genetic Correlations across Human Diseases and Traits. Natrure genetics 47, 1236-1241.
19. Zheng, J., Erzurumluoglu, A.M., Elsworth, B.L., Kemp, J.P., Howe, L., Haycock, P.C., Hemani, G., Tansey, K., Laurin, C., Pourcain, B.S., et al. (2017). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 33, 272-279.
20. Maier, R., Moser, G., Chen, G.-B., Ripke, S., Absher, D., Agartz, I., Akil, H., Amin, F., Andreassen, Ole A., Anjorin, A., et al. (2015). Joint Analysis of Psychiatric Disorders Increases Accuracy of Risk Prediction for Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. The American Journal of Human Genetics 96, 283-294.
21. Finucane, H.K., Bulik-sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-r., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nature genetics 47, 1228-1235.
22. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., and Yang, J. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature 518, 197-206.
23. Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J.a., Kutalik, Z., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46, 1173-1186.