HGGA, Volume *3*

Supplemental information

Stability of polygenic scores across discovery genome-wide association studies

Laura M. Schultz, Alison K. Merikangas, Kosha Ruparel, Sébastien Jacquemont, David C. Glahn, Raquel E. Gur, Ran Barzilay, and Laura Almasy

Supplemental Methods

Pre-imputation quality control (QC) and imputation were done separately for each of the fifteen genotyping array batches that comprise the Philadelphia Neurodevelopmental Cohort (PNC) dataset. Due to the substantial variation in SNPs contained on the arrays and the numbers of samples genotyped on each array (Table S1), we imputed each array batch separately to the 1000 Genomes Mixed/Other reference panel rather than assigning ancestry prior to imputation. The fifteen batches were merged by chromosome after imputation, and post-imputation QC was run on the merged chromosome files.

Genetic ancestry was inferred by KING[1] from the principal components (PCs) derived using multi-dimensional scaling (MDS) of the hard-call dataset that was produced with PLINK 1.9 after concatenating the post-imputation-QC chromosome files. Each array batch included samples from more than one ancestry group (Table S2), thus validating our decision to not impute the array batches to specific ancestry panels.

After splitting the dataset into European-American (EUR) and African-American (AFR) cohorts, we ran a second round of unprojected MDS for each cohort separately. The first ten PCs were later regressed out of the standardized polygenic scores (PGS) to correct for both population structure and array batch effects. Batch effects, which were captured by PC2, were especially pronounced for the AFR samples that were genotyped on array_01 and array_07 (Figure S1). There were no obvious batch effects visible in the second-round PC plot for the EUR samples (Figure S2).

We further analyzed the PNC array batch effects by running a series of logistic-regression GWAS with a single array as a dummy "case" and the other arrays as dummy "controls" within the

EUR and AFR subgroups. We only used arrays that had been used to genotype at least 100 samples as "cases." For the AFR subgroup, these arrays included array_01, array_03, array_04, array_05, array_06, and array_08; the arrays for the EUR subgroup that were run as "cases" were array_03, array_04, array_05, array_06, and array_09. The GWAS were run in PLINK 1.9 both including and not including the first 10 second-round ancestry PCs as covariates so that we could confirm that including the PCs would be an adequate control for array batch effects. *P*-values were generated using Fisher's exact test. We used the R package `qqman`[2] to produce Manhattan and Q-Q plots from the Bonferroni-adjusted *p*-values.

As expected, the most dramatic GWAS results were observed for AFR array_01 (Figure S3). The logistic association model without ancestry PC covariates had many "significant" SNPs, seen as many tall peaks on the Manhattan plot (Figure S3-A) and dramatic curvature on the Q-Q plot (Figure S3-B). When the first 10 PCs were included as covariates, no significant peaks remained in the Manhattan plot (Figure S3-C), and the Q-Q plot did not deviate substantially from the straight line (Figure S3-D). The GWAS results from the other arrays yielded similar Manhattan and Q-Q plots when 10 within-ancestry PCs were included as covariates. Taken together, these results indicated that regressing out the first 10 within-ancestry PCs from our polygenic scores (PGS) would adequately control for any array batch effects.

The ABCD dataset was genotyped exclusively on the Affymetrix NIDA SmokeScreen array. As such, QC and imputation were done on a single dataset. Table S3 shows the ABCD SNP and sample counts before and after the pre-imputation QC. Within-ancestry PCs were computed for the AFR (Figure S4) and EUR (Figure S5) subsets of the ABCD dataset.

PRS-CS[3] requires a single GWAS sample size as an input. Given that most of our discovery GWAS were meta-GWAS that were comprised of individual studies that varied in terms of their sample sizes and the SNPs they included, the effective sample size often varied considerably between SNPs. To account for this reality, we examined the distribution of SNP sample sizes in R and excluded SNPs that had sample sizes that were less than half of the maximum SNP sample size. Of the remaining SNPs, the median SNP sample size was used as the PRS-CS GWAS sample size input.

As an example, consider the Freeze 2 EUR PTSD GWAS produced by the Psychiatric Genomics Consortium (PGC).[4] This meta-GWAS includes 9,766,174 SNPs with effective sample sizes that range from 17,559.4 to 70,237.5 (Figure S6-A). Given that PRS-CS uses only those SNPs that overlap with both the relevant LD panel and the dataset, we started by retaining only the 1,116,862 SNPs that were present in the EUR LD panel. These SNPs also had effective sample sizes ranging from 17,559.4 to 70,237.5 (Figure S6-B). After we removed SNPs with effective sample sizes that were less than 35,000, the remaining 1,113,044 SNPs had effective sample sizes that ranged from 38,250.5 to 70,237.5 (Figure S6-C), with a median of 70,237.5. This median value was truncated to 70,237 and used as the GWAS sample size when we ran PRS-CS. We made similar sample size determinations for the other discovery GWAS.

UK Biobank Experiment

To explore the impact of GWAS sample size on our PGS results, we ran an experiment using imputed genotype data for the UK Biobank that we obtained under data use application 40980. As illustrated in Figure S7, we identified 276,107 unrelated white British subjects who had both imputed genotype data and also a measured standing height phenotype (Data-Field 50), and we

then randomly assigned these samples into seven groups as shown. The non-overlapping Groups A and B were each used to produce a "medium-large" height GWAS (both *N* = 134,000). Groups C and D were sub-sampled from Groups A and B, respectively, and each used to produce a "medium" height GWAS (both *N* = 75,000). Groups E and F were sub-sampled from Groups C and D, respectively, and each used to produce a "small" height GWAS (both *N* = 10,000). Groups A and B were merged to form Group AB, which was used to produce a "large" height GWAS (*N* = 268,000). The remaining 8,107 subjects with height phenotypes comprised the test set for whom we computed PGS using all seven GWAS. All GWAS were computed using the PLINK 2 `--linear` function with sex, age at height measurement, and the first 20 ancestry PCs supplied by the UK Biobank as covariates. The subject characteristics for each GWAS group and the test set are summarized in Table S16, and the mean χ^2 computed by LDSC for each height GWAS is provided in Table S17. Table S18 provides the number of genome-wide significant SNPs ($P < 5 \times 10^{-8}$) for each of the seven discovery GWAS that were included among the 1,113,490 SNPs that were used for PRS-CS computations (i.e., the set of SNPs that were jointly present in the test data set, the discovery GWAS, and the PRS-CS EUR LD panel). We also used LDSC to compute the genetic correlation between each pair of height GWAS (Table S19).

We used PRS-CS to compute seven height PGS for each sample in the test set (i.e., a PGS was computed from each of the seven discovery GWAS) as described in the main Methods section of our paper. We standardized the PGS and then corrected for batch effects and stratification by regressing the first 20 UKBB-supplied ancestry PCs out of the standardized PGS. We then calculated the correlation between each pair of corrected height PGS (Table S20). We also used each set of standardized PGS to predict height in an additive linear regression model that included

sex, age at height measurement, and the first 20 ancestry PCs as covariates. We calculated the coefficient of determination (R^2) for each model as a measure of how well the PGS predicted height, and we also ran a partial F-test for each model to assess the effect of adding the standardized PGS to a base model that included sex, age at height measurement, and the first 20 ancestry PCs as predictors of height (Table S21).

Supplemental Data

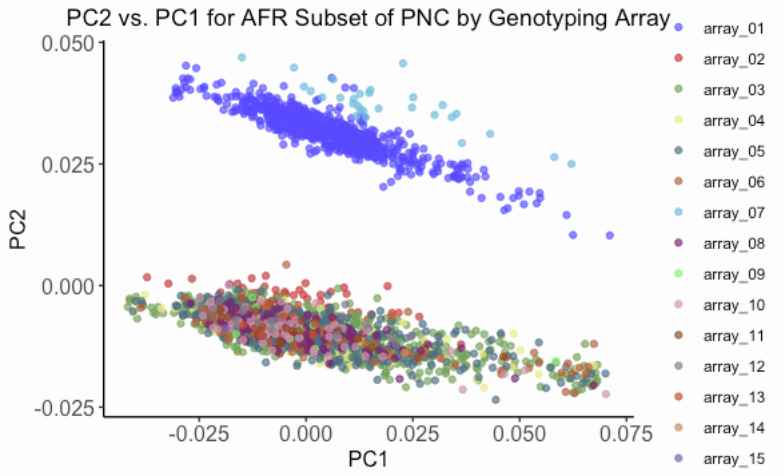

Figure S1. Within-ancestry PC2 vs. PC1 for the AFR subset of PNC with samples color coded by their genotyping array batch. PC2 captures an array batch effect that is most pronounced for array_01 and array_07.

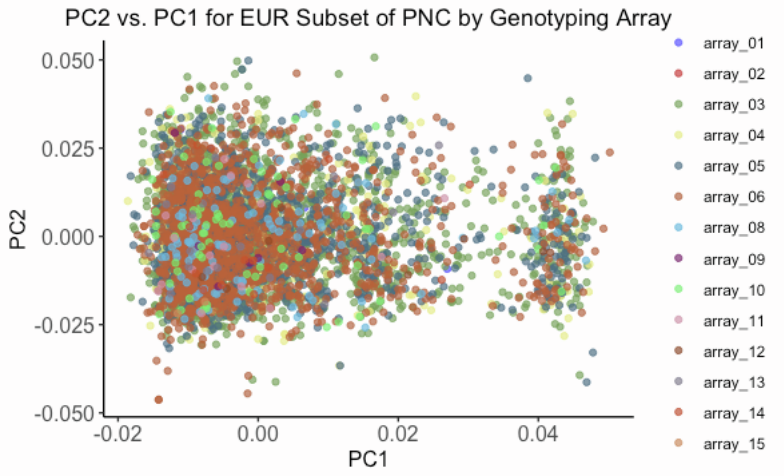

Figure S2. Within-ancestry PC2 vs. PC1 for the EUR subset of PNC with samples color coded by their genotyping array batch.

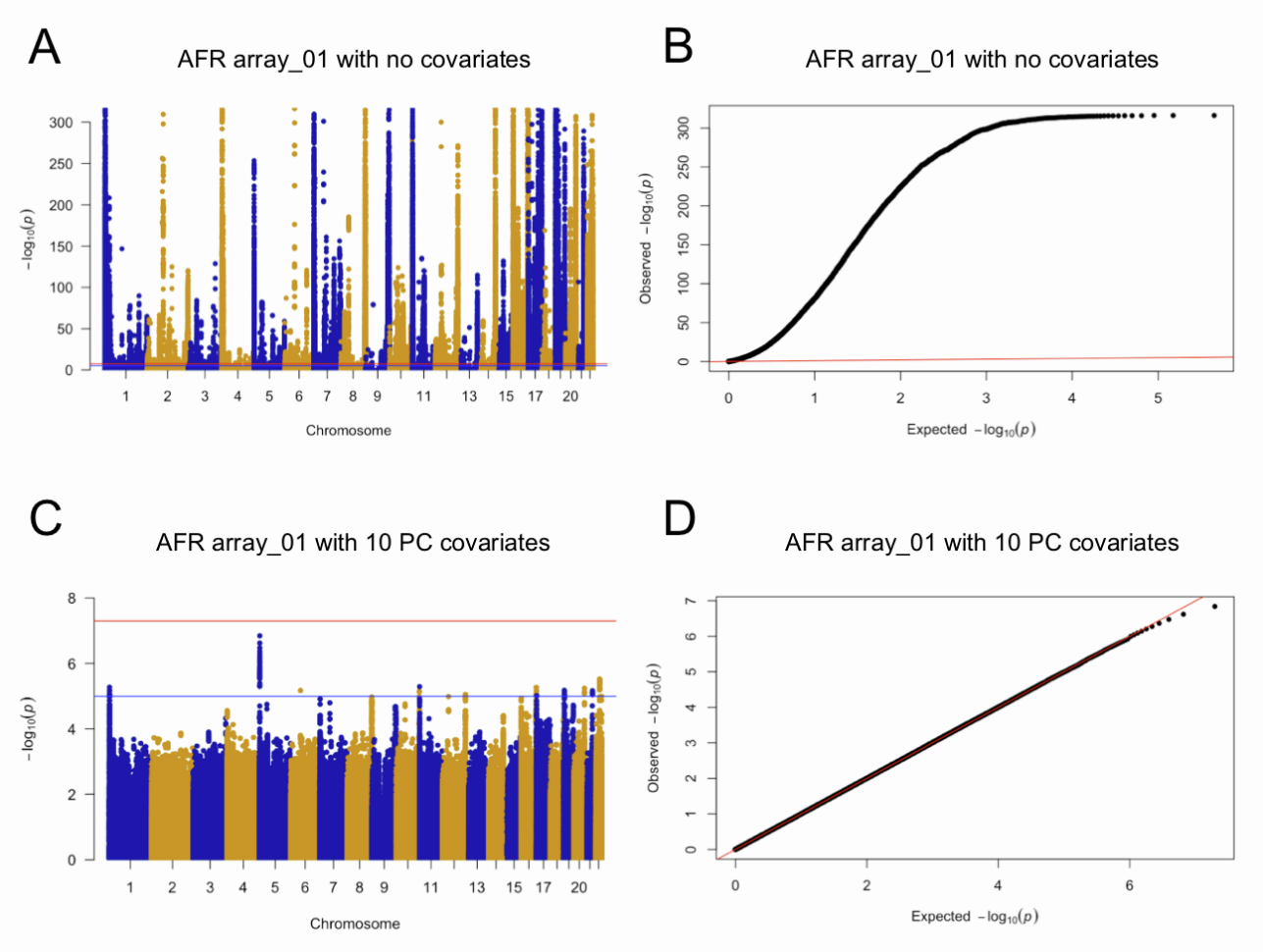

Figure S3. Illustration of PNC batch effects for AFR array_01. The plots are limited to 9,809,388 biallelic SNPs. (A) The Manhattan plot showed many highly significant SNPs when the ancestry PCs were not included as covariates. (B) When PCs were not included as covariates, the Q-Q plot deviated substantially from the expected straight line. (C) When 10 PCs were included as covariates, no significant peaks remained in the Manhattan plot. (D) With 10 PCs included as covariates, the Q-Q plot of observed $-\log_{10}(p)$ versus expected $-\log_{10}(p)$ largely followed the expected straight line.

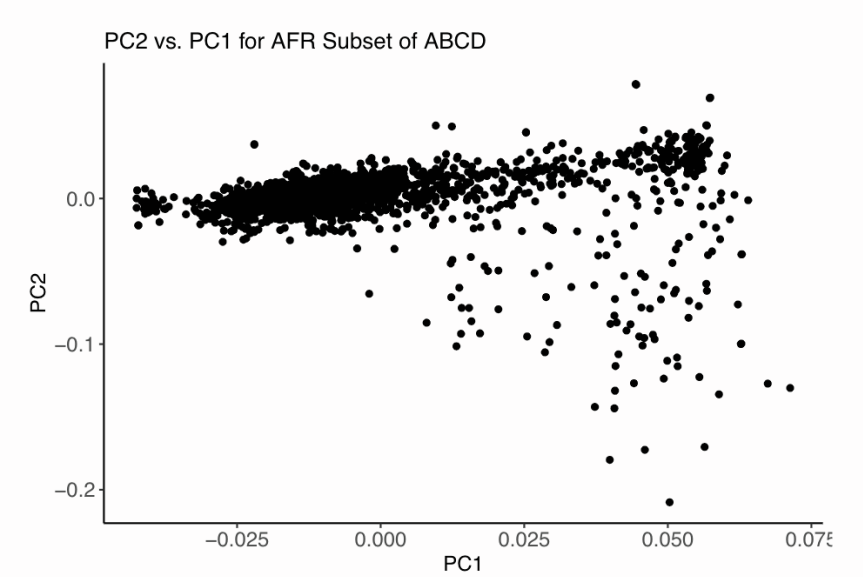

Figure S4. Within-ancestry PC2 vs. PC1 for the AFR subset of the ABCD dataset.

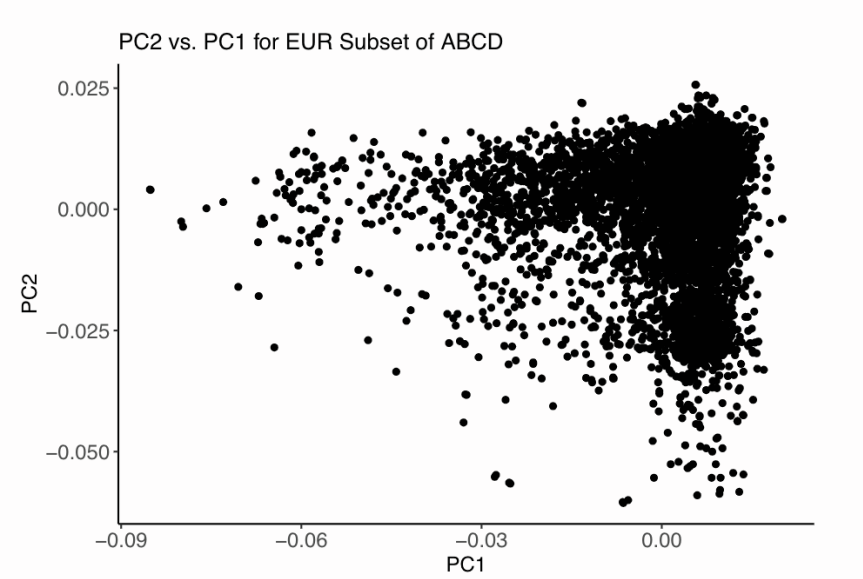

Figure S5. Within-ancestry PC2 vs. PC1 for the EUR subset of the ABCD dataset.

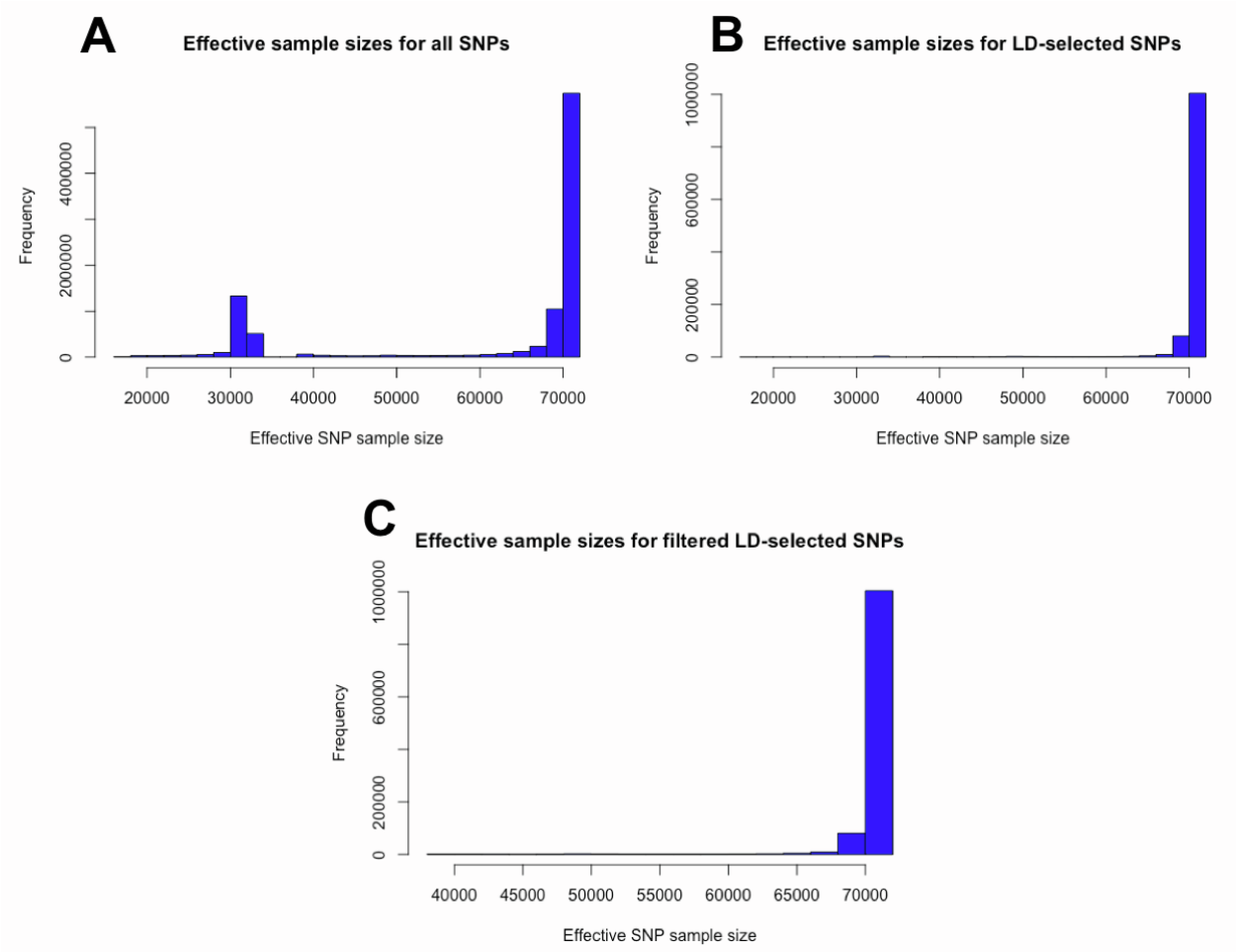

Figure S6. GWAS sample size determination for the PTSD Freeze 2 EUR meta-GWAS. (A) This meta-GWAS contained 9,766,174 SNPs with effective sample sizes that ranged from 17,559.4 to 70,237.5. (B) The 1,116,862 SNPs that were present in the PRS-CS EUR LD panel had this same range of effective sample sizes. (C) Filtering to retain only LD-selected SNPs with effective sample sizes of at least 35,000 resulted in 1,113,044 SNPs with effective sample sizes between 38,250.5 and 70,237.5. The median effective SNP sample size of 70,237.5 for these filtered SNPs was truncated to 70,237 and used as the GWAS sample size in PRS-CS.

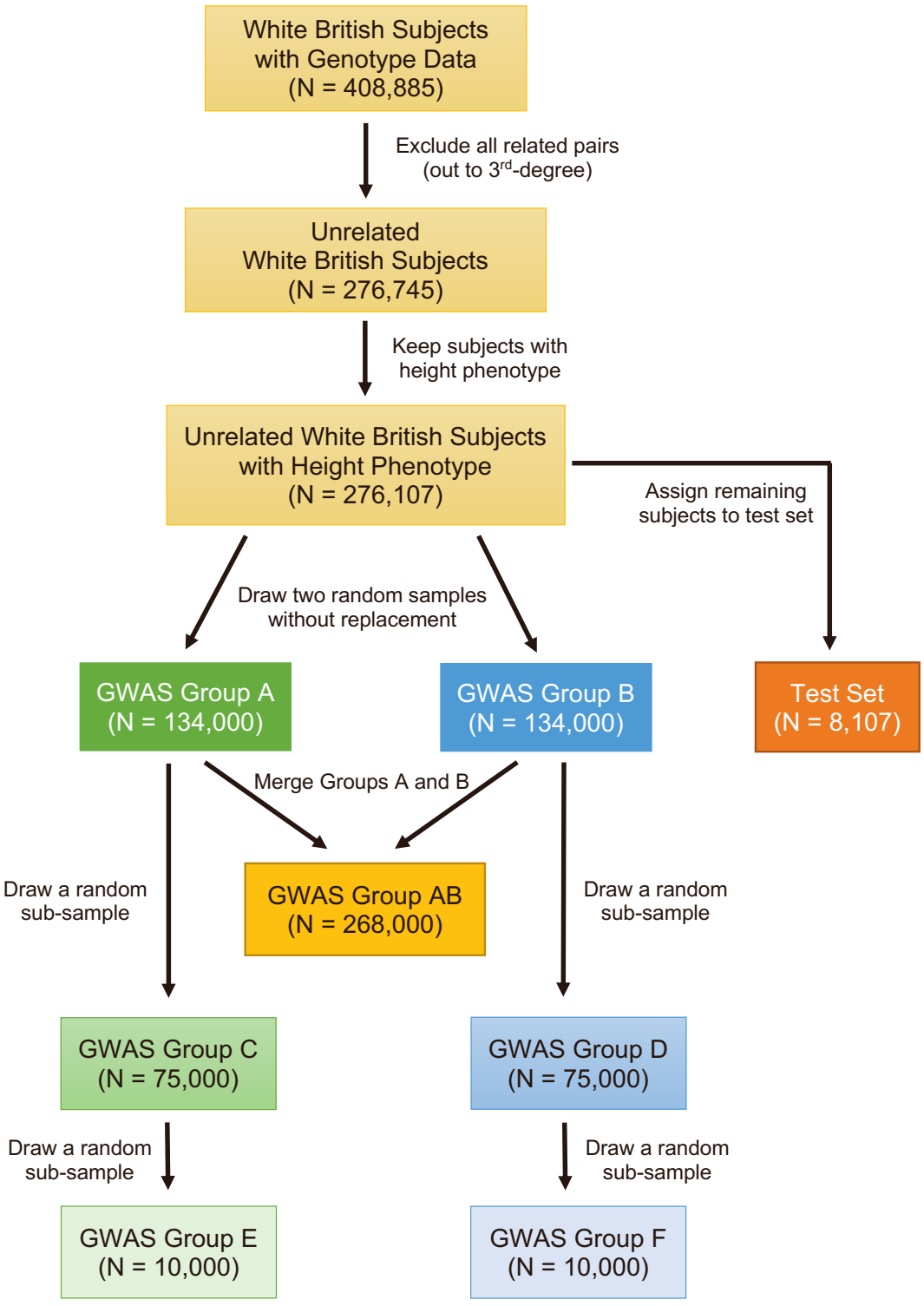

Figure S7. Protocol for assigning unrelated white British subjects from the UK Biobank to seven height GWAS groups and a test set.

Table S1. PNC sample and SNP counts by genotyping array before and after pre-imputation QC.

Array Code	dbGaP Filename	No. of Samples Pre-QC	No. of Samples Post-QC	No. of SNPs Pre-QC	No. of SNPs Post-QC
array_01	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.Axiom.c1.GRU-NPU/GO_Axiom	722	719	567,096	472,217
array_02	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.Genome-Wide_Human_SNP_Array_6.0.c1.GRU-NPU/GO_Affy60	66	66	909,622	725,897
array_03	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.Human610-Quadv1_B.c1.GRU-NPU/GO_Quad_5removed	3,802	3,789	620,901	573,487
array_04	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.HumanHap550_v1.c1.GRU-NPU/GO_v1_1removed	555	552	555,352	533,783
array_05	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.HumanHap550_v3.c1.GRU-NPU/GO_v3_1removed	1,913	1,893	561,466	541,643
array_06	phg000381.v2.NIMH_NeurodevelopmentalGenomics.genotype-calls-matrixfmt.HumanOmniExpress.c1.GRU-NPU/GO_Omni	1,657	1,654	733,202	693,213
array_07	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.Axiom.c1.GRU-NPU/GO_Axiom_set2	40	32	567,096	298,921
array_08	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.Axiom.c1.GRU-NPU/GO_AxiomTx	225	218	767,203	616,881
array_09	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.BDCHP-1X10-HUMANHAP550.c1.GRU-NPU/GO_v1set2	17	17	555,352	490,740
array_10	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.Human1M-Duov3_B.c1.GRU-NPU/GO_1MDuo	141	141	1,199,187	1,040,603
array_11	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.Human610-Quadv1_B.c1.GRU-NPU/GO_Quadset2	40	40	620,901	564,518
array_12	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.HumanHap550_v3.c1.GRU-NPU/GO_v3set2	31	31	561,466	516,726
array_13	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.HumanOmniExpress-12v1_A.c1.GRU-NPU/GO_Omniset2	37	35	733,202	674,803
array_14	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.HumanOmniExpress-12v1_B.c1.GRU-NPU/GO_OMNI12v11	18	18	719,665	578,397
array_15	phg000661.v1.NIMH_NeurodevelopmentalGenomics_v2.genotype-calls-matrixfmt.HumanOmniExpressExome-8v1_A.c1.GRU-NPU/GO_OEE	3	3	951,117	400,300

Table S2. PNC ancestry by genotyping array.

Array	AFR	EUR	Other	Total
array_01	693	8	18	719
array_02	63	2	1	66
array_03	1,341	2,157	291	3,789
array_04	177	341	34	552
array_05	623	1,137	133	1,893
array_06	112	1,354	188	1,654
array_07	29	0	3	32
array_08	101	102	15	218
array_09	9	7	1	17
array_10	62	69	10	141
array_11	20	18	2	40
array_12	19	8	2	29
array_13	9	18	8	35
array_14	1	16	1	18
array_15	1	2	0	3
Total	**3,260**	**5,239**	**707**	**9,206**

Table S3. ABCD sample and SNP counts before and after pre-imputation QC.

Number of Samples Pre-QC	Number of Samples Post-QC	Number of SNPs Pre-QC	Number of SNPs Post-QC
10,461	10,318	517,724	483,017

Table S4. PTSD PGS correlations for PNC EUR cohort when limited to one sample per family versus including all samples.

Comparison	One sample per family *n* = 4928	All EUR samples *n* = 5239
PRS-CS replication using same discovery GWAS (PGC Freeze 2)[4]	*r* = 0.9994 (*t* = 2007 , *P* < 2e-16)	*r* = 0.9994 (*t* = 2057, *P* < 2e-16)
Different discovery GWAS, same ancestry (PGC Freezes 1 and 2)[4; 5]	*r* = 0.388 (*t* = 29.55, *P* < 2e-16)	*r* = 0.392 (*t* = 30.86, *P* < 2e-16)
Different discovery GWAS, different ancestry (PGC Freeze 1)[5]	*r* = -0.00265 (*t* = -0.186, *P* = 0.852)	*r* = 0.00136 (*t* = 0.098, P = 0.922)
Different discovery GWAS, different ancestry (PGC Freeze 2)[4]	*r* = 0.0341 (*t* =2.391, *P* = 0.0168)	*r* = 0.0379 (*t* = 2.746, *P* = 0.00605)

PNC, Philadelphia Neurodevelopmental Cohort; PTSD, post-traumatic stress disorder; PGS, polygenic score; EUR, European-American ancestry; GWAS, genome-wide association study; PGC; Psychiatric Genomics Consortium; *r*, Pearson correlation coefficient; *t*, linear association *t*-test statistic; *P*, two-tailed *P*-value on 4926 (or 5237) degrees of freedom.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS: [4]Nievergelt et al. (2019), [5]Duncan et al. (2018).

Table S5. PTSD PGS correlations for PNC AFR cohort when limited to one sample per family versus including all samples.

Comparison	One sample per family *n* = 2954	All AFR samples *n* = 3260
PRS-CS replication using same discovery GWAS (PGC Freeze 2)[4]	*r* =0.9997 (*t* = 2055, *P* < 2e-16)	*r* = 0.9997 (*t* = 2162, *P* < 2e-16)
Different discovery GWAS, same ancestry (PGC Freezes 1 and 2)[4; 5]	*r* = 0.636 (*t* = 44.76, *P* < 2e-16)	*r* = 0.696 (*t* = 55.26, *P* < 2e-16)
Different discovery GWAS, different ancestry (PGC Freeze 1)[5]	*r* = 0.0399 (*t* = 2.171, *P* = 0.03)	*r* = 0.0417 (*t* = 2.379, *P* = 0.0174)
Different discovery GWAS, different ancestry (PGC Freeze 2)[4]	*r* = 0.000732 (*t* = 0.04, *P* = 0.968)	*r* = 0.00356 (*t* = 0.203, *P* = 0.839)

PNC, Philadelphia Neurodevelopmental Cohort; PTSD, post-traumatic stress disorder; PGS, polygenic score; AFR, African-American ancestry; GWAS, genome-wide association study; PGC; Psychiatric Genomics Consortium; *r*, Pearson correlation coefficient; linear association *t*-test statistic, *P*, two-tailed *P*-value on 2952 (or 3258) degrees of freedom.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS: [4]Nievergelt et al. (2019), [5]Duncan et al. (2018).

Table S6. PGS correlations for PNC AFR cohort (*n* = 3260).

Comparison	PTSD	T2D	Height
PRS-CS replication using same discovery GWAS	r = 0.9997[4] (t = 2162, P < 2e-16)	NA	NA
Different discovery GWAS, same ancestry	r = 0.696[4; 5] (t = 55.26, P < 2e-16)	NA	NA
Different discovery GWAS, different ancestry	r = 0.0417[5] (t = 2.379, P = 0.0174)	r = 0.0185[6; 7] (t = 1.055, P = 0.292)	r = 0.287[8] (t = 17.09, P < 2e-16)
	r = 0.00356[4] (t = 0.203, P = 0.839)	r = 0.0432[7; 9] (t = 2.469, P = 0.0136)	r = 0.258[8; 10] (t = 15.22, P < 2e-16)

PNC, Philadelphia Neurodevelopmental Cohort; PGS, polygenic score; AFR, African-American ancestry; GWAS, genome-wide association study; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; *r*, Pearson correlation coefficient; *t*, linear association *t*-test statistic; *P*, two-tailed *P*-value on 3258 degrees of freedom; NA, not applicable (analysis not run).
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S7. PGS correlations for ABCD AFR cohort (*n* = 1741).

Comparison	PTSD	T2D	Height
Different discovery GWAS, same ancestry	r = 0.657[4; 5] (t = 36.34, P < 2e-16)	NA	NA
Different discovery GWAS, different ancestry	r = -0.00320[5] (t = -0.133, P = 0.894)	r = 0.0219[6; 7] (t = 0.912, P = 0.362)	r = 0.306[8] (t = 13.42, P < 2e-16)
	r = 0.00283[4] (t = 0.118, P = 0.906)	r = -0.0458[7; 9] (t = -1.911, P = 0.0562)	r = 0.312[8; 10] (t = 13.68, P < 2e-16)

ABCD, Adolescent Brain and Cognitive Development Study; PGS, polygenic score; AFR, African-American ancestry; GWAS, genome-wide association study; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; *r*, Pearson correlation coefficient; *t*, linear association *t*-test statistic; *P*, two-tailed *P*-value on 1739 degrees of freedom; NA, not applicable (analysis not run).
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S8. PGS correlations for PNC EUR cohort (*n* = 5239).

Comparison	PTSD	T2D	Height
PRS-CS replication using same discovery GWAS	r = 0.9994[4] (t = 2057, P < 2e-16)	NA	NA
Different discovery GWAS, same ancestry	r = 0.392[4; 5] (t = 30.86, P < 2e-16)	r = 0.602[6; 9] (t = 54.54, P < 2e-16)	r = 0.736[8; 10] (t = 78.78, P < 2e-16)
Different discovery GWAS, different ancestry	r = 0.00136[5] (t = 0.098, P = 0.922)	r = 0.0240[6; 7] (t = 1.739, P = 0.082)	r = 0.403[8] (t = 31.82, P < 2e-16)
	r = 0.0379[4] (t = 2.746, P = 0.00605)	r = 0.00528[7; 9] (t = 0.382, P = 0.703)	r = 0.335[8; 10] (t = 25.25, P < 2e-16)

PNC, Philadelphia Neurodevelopmental Cohort; PGS, polygenic score; EUR, European-American ancestry; GWAS, genome-wide association study; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; *r*, Pearson correlation coefficient; *t*, linear association *t*-test statistic; *P*, two-tailed *P*-value on 5237 degrees of freedom; NA, not applicable (analysis not run).

Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S9. PGS correlations for ABCD EUR cohort (*n* = 5815).

Comparison	PTSD	T2D	Height
Different discovery GWAS, same ancestry	r = 0.378[4; 5] (t = 31.14, P < 2e-16)	r = 0.597[6; 9] (t = 56.79, P < 2e-16)	r = 0.734[8; 10] (t = 82.46, P < 2e-16)
Different discovery GWAS, different ancestry	r = -0.00109[5] (t = -0.083, P = 0.934)	r = 0.0224[6; 7] (t = 1.71, P = 0.0872)	r = 0.404[8] (t = 33.69, P < 2e-16)
	r = 0.000867[4] (t = 0.066, P = 0.947)	r = 0.0188[7; 9] (t = 1.431, P = 0.152)	r = 0.327[8; 10] (t = 26.39, P < 2e-16)

ABCD, Adolescent Brain and Cognitive Development Study; PGS, polygenic score; EUR, European-American ancestry; GWAS, genome-wide association study; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; *r*, Pearson correlation coefficient; *t*, linear association *t*-test statistic; *P*, two-tailed *P*-value on 5813 degrees of freedom.

Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S10. Proportional overlap for PNC AFR polygenic scores (*n* = 3260).

Comparison	Top Quintile (≥ 80th Percentile) *n* = 652	Top Decile (≥ 90th Percentile) *n* = 326	Top Ventile (≥ 95th Percentile) *n* = 163
PRS-CS replication using same discovery GWAS	PTSD:[4] 644/652 (0.987)	PTSD:[4] 318/326 (0.975)	PTSD:[4] 161/163 (0.988)
Different discovery GWAS, same ancestry	PTSD:[4; 5] 331/652 (0.508)	PTSD:[4; 5] 134/326 (0.411)	PTSD:[4; 5] 58/163 (0.356)
Different discovery GWAS, different ancestry	PTSD:[4] 143/652 (0.219) PTSD:[5] 138/652 (0.212) T2D:[6; 7] 137/652 (0.210) T2D:[7; 9] 144/652 (0.221) height:[8] 214/652 (0.328) height:[8; 10] 209/652 (0.321)	PTSD:[4] 37/326 (0.113) PTSD:[5] 36/326 (0.110) T2D:[6; 7] 35/326 (0.107) T2D:[7; 9] 41/326 (0.126) height:[8] 77/326 (0.236) height:[8; 10] 72/326 (0.221)	PTSD:[4] 8/163 (0.0491) PTSD:[5] 12/163 (0.0736) T2D:[6; 7] 12/163 (0.0736) T2D:[7; 9] 13/163 (0.0798) height:[8] 27/163 (0.166) height:[8; 10] 31/163 (0.190)

PNC, Philadelphia Neurodevelopmental Cohort; AFR, African-American ancestry; *n*, number of subjects; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019), [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S11. Proportional overlap for ABCD AFR polygenic scores (*n* = 1741).

Comparison	Top Quintile (≥ 80th Percentile) *n* = 349	Top Decile (≥ 90th Percentile) *n* = 175	Top Ventile (≥ 95th Percentile) *n* = 88
Different discovery GWAS, same ancestry	PTSD:[4; 5] 187/349 (0.536)	PTSD:[4; 5] 83/175 (0.475)	PTSD:[4; 5] 32/88 (0.363)
Different discovery GWAS, different ancestry	PTSD:[4] 66/349 (0.189) PTSD:[5] 62/349 (0.178) T2D:[6; 7] 76/349 (0.218) T2D:[7; 9] 69/349 (0.198) height:[8] 115/349 (0.330) height:[8; 10] 119/349 (0.341)	PTSD:[4] 25/175 (0.143) PTSD:[5] 18/175 (0.103) T2D:[6; 7] 19/175 (0.109) T2D:[7; 9] 19/175 (0.109) height:[8] 45/175 (0.257) height:[8; 10] 34/175 (0.194)	PTSD:[4] 7/88 (0.0795) PTSD:[5] 4/88 (0.0455) T2D:[6; 7] 10/88 (0.114) T2D:[7; 9] 6/88 (0.0343) height:[8] 13/88 (0.148) height:[8; 10] 15/88 (0.170)

ABCD, Adolescent Brain and Cognitive Development Study; AFR, African-American ancestry; *n*, number of subjects; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019), [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S12. Proportional overlap for PNC EUR polygenic scores (*n* = 5239).

Comparison	Top Quintile (≥ 80th Percentile) *n* = 1048	Top Decile (≥ 90th Percentile) *n* = 524	Top Ventile (≥ 95th Percentile) *n* = 262
PRS-CS replication using same discovery GWAS	PTSD:[4] 1026/1048 (0.979)	PTSD:[4] 513/524 (0.979)	PTSD:[4] 255/262 (0.973)
Different discovery GWAS, same ancestry	PTSD:[4; 5] 391/1048 (0.373) T2D:[6; 9] 532/1048 (0.508) height:[8; 10] 625/1048 (0.596)	PTSD:[4; 5] 139/524 (0.265) T2D:[6; 9] 228/524 (0.435) height:[8; 10] 253/524 (0.483)	PTSD:[4; 5] 51/262 (0.195) T2D:[6; 9] 90/262 (0.344) height:[8; 10] 109/262 (0.416)
Different discovery GWAS, different ancestry	PTSD:[4] 233/1048 (0.222) PTSD:[5] 209/1048 (0.199) T2D:[6; 7] 221/1048 (0.211) T2D:[7; 9] 204/1048 (0.195) height:[8] 399/1048 (0.381) height:[8; 10] 381/1048 (0.364)	PTSD:[4] 64/524 (0.122) PTSD:[5] 47/524 (0.0897) T2D:[6; 7] 56/524 (0.107) T2D:[7; 9] 50/524 (0.0954) height:[8] 140/524 (0.267) height:[8; 10] 119/524 (0.227)	PTSD:[4] 20/262 (0.0763) PTSD:[5] 14/262 (0.0534) T2D:[6; 7] 15/262 (0.0573) T2D:[7; 9] 10/262 (0.0382) height:[8] 46/262 (0.176) height:[8; 10] 41/262 (0.156)

PNC, Philadelphia Neurodevelopmental Cohort; EUR, European-American ancestry; *n*, number of subjects; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S13. Proportional overlap for ABCD EUR polygenic scores (*n* = 5815).

Comparison	Top Quintile (≥ 80th Percentile) *n* = 1163	Top Decile (≥ 90th Percentile) *n* = 582	Top Ventile (≥ 95th Percentile) *n* = 291
Different discovery GWAS, same ancestry	PTSD:[4; 5] 444/1163 (0.382) T2D:[8; 9] 586/1163 (0.504) height:[8; 10] 674/1163 (0.580)	PTSD:[4; 5] 153/582 (0.263) T2D:[8; 9] 230/582 (0.395) height:[8; 10] 301/582 (0.517)	PTSD:[4; 5] 67/291 (0.230) T2D:[8; 9] 99/291 (0.340) height:[8; 10] 140/291 (0.481)
Different discovery GWAS, different ancestry	PTSD:[4] 248/1163 (0.213) PTSD:[5] 227/1163 (0.195) T2D:[6; 7] 241/1163 (0.207) T2D:[7; 9] 248/1163 (0.213) height:[8] 432/1163 (0.371) height:[8; 10] 379/1163 (0.326)	PTSD:[4] 66/582 (0.113) PTSD:[5] 65/582 (0.112) T2D:[6; 7] 68/582 (0.117) T2D:[7; 9] 66/582 (0.113) height:[8] 155/582 (0.266) height:[8; 10] 130/582 (0.223)	PTSD:[4] 24/291 (0.0825) PTSD:[5] 17/291 (0.0584) T2D:[6; 7] 19/291 (0.0653) T2D:[7; 9] 12/291 (0.0412) height:[8] 64/291 (0.220) height:[8; 10] 55/291 (0.189)

ABCD, Adolescent Brain and Cognitive Development Study; EUR, European-American ancestry; *n*, number of subjects; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes.
Superscripts are the reference numbers for the discovery GWAS used to calculate PGS:
[4]Nievergelt et al. (2019)**,** [5]Duncan et al. (2018), [6]Mahajan et al. (2018), [7]Chen et al. (2019), [8]Marouli et al. (2017), [9]Scott et al. (2017), [10]Wood et al. (2014).

Table S14. LD score regression results for individual EUR-ancestry GWAS.

Trait	Discovery GWAS	Mean χ^2	λ_{GC}	Intercept (SE)
PTSD	Nievergelt et al. (2019)[4]	1.0789	1.0679	1.0217 (0.0066)
	Duncan et al. (2018)[5]	1.0127	1.0165	0.9939 (0.0059)
T2D	Scott et al. (2017)[9]	1.2335	1.1459	0.9997 (0.0085)
	Mahajan et al. (2018)[6]	1.9562	1.6259	1.0835 (0.0144)
Height	Marouli et al. (2017)[8]	6.4544	2.5641	1.6372 (0.0827)
	Wood et al. (2014)[10]	2.9616	2.0007	1.3295 (0.0193)

LD, linkage disequilibrium; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; SE, standard error estimate obtained via block jackknifing; λ_{GC}, genomic control inflation factor

Table S15. LDSC genetic correlations for pairs of EUR-ancestry GWAS.

	GWAS 1	GWAS 2	Genetic Correlation (SE)
PTSD	Duncan et al. (2018)[5]	Nievergelt et al. (2019)[4]	0.9225 (0.1807)
T2D	Scott et al. (2017)[9]	Mahajan et al. (2018)[6]	*1.1265 (0.019)
Height	Wood et al. (2014)[10]	Marouli et al. (2017)[8]	*1.0231 (0.0205)

LDSC, linkage disequilibrium score regression; PTSD, post-traumatic stress disorder; T2D, type 2 diabetes; SE, standard error estimate obtained via block jackknifing

*Genetic correlations >1 are a known issue with LDSC (https://github.com/bulik/ldsc/issues/89).

Table S16. Characteristics of UK Biobank white British height groups.

Group[†]		Count (%)	Mean ± SD Height (cm)	Mean ± SD Age (Years)
GWAS AB	Female	142,904 (53.3%)	162.82 ± 6.23	56.6 ± 7.9
	Male	125,096 (46.7%)	176.04 ± 6.76	57.1 ± 8.1
	Combined	**268,000**	168.99 ± 9.25	56.8 ± 8.0
GWAS A	Female	71,365 (53.3%)	162.81 ± 6.24	56.6 ± 7.9
	Male	62,635 (46.7%)	176.05 ± 6.75	57.1 ± 8.1
	Combined	**134,000**	169.00 ± 9.26	56.8 ± 8.0
GWAS B	Female	71,539 (53.4%)	162.83 ± 6.22	56.6 ± 7.9
	Male	62,461 (46.6%)	176.03 ± 6.76	57.1 ± 8.1
	Combined	**134,000**	168.98 ± 9.24	56.8 ± 8.0
GWAS C	Female	40,082 (53.4%)	162.79 ± 6.25	56.6 ± 7.9
	Male	34,918 (46.6%)	176.05 ± 6.80	57.1 ± 8.1
	Combined	**75,000**	168.96 ± 9.28	56.8 ± 8.0
GWAS D	Female	39,882 (53.2%)	162.83 ± 6.21	56.6 ± 7.8
	Male	35,118 (46.8%)	176.00 ± 6.76	57.1 ± 8.1
	Combined	**75,000**	169.00 ± 9.23	56.8 ± 7.9
GWAS E	Female	5,378 (53.8%)	162.78 ± 6.30	56.6 ± 7.9
	Male	4,622 (46.2%)	175.97 ± 6.63	57.2 ± 8.1
	Combined	**10,000**	168.88 ± 9.21	56.9 ± 8.0
GWAS F	Female	5,316 (53.2%)	162.89 ± 6.29	56.6 ± 7.8
	Male	4,684 (46.8%)	175.95 ± 6.79	57.0 ± 8.1
	Combined	**10,000**	169.01 ± 9.22	56.8 ± 7.9
Test Set	Female	4,335 (53.5%)	162.85 ± 6.29	56.6 ± 8.0
	Male	3,772 (46.5%)	176.14 ± 6.81	57.2 ± 8.1
	Combined	**8,107**	169.03 ± 9.31	56.9 ± 8.0
All Samples	Female	147,239 (53.3%)	162.82 ± 6.23	56.6 ± 7.9
	Male	128,868 (46.7%)	176.04 ± 6.76	57.1 ± 8.1
	Combined	**276,107**	168.99 ± 9.25	56.8 ± 8.0

[†]The 75,000 individuals included in GWAS C were randomly sampled from the 134,000 individuals included in GWAS A, and the 10,000 individuals included in GWAS E were randomly sampled from those included in GWAS C. The same relationships exist for GWAS B, D, and F. GWAS AB was run using the 268,000 individuals who were included in GWAS A or GWAS B. The test set consists of 8,107 individuals who were not included in any GWAS.

Table S17. LD score regression results for UK Biobank height GWAS.

GWAS	GWAS sample size	SNP count for LDSC†	SNP h^2 (SE)	Mean χ^2	λ_{GC}	Intercept (SE)
AB	268,000	1,174,517	0.4263 (0.0182)	3.7259	2.1633	1.3129 (0.0301)
A	134,000	1,174,519	0.4383 (0.0203)	2.3496	1.6831	1.1353 (0.0181)
B	134,000	1,174,516	0.4490 (0.0204)	2.3752	1.6715	1.1400 (0.0195)
C	75,000	1,174,518	0.4456 (0.0232)	1.7443	1.4210	1.0643 (0.0127)
D	75,000	1,174,514	0.4627 (0.0238)	1.7764	1.4316	1.0682 (0.0142)
E	10,000	1,174,518	0.4199 (0.0624)	1.0982	1.0710	1.0151 (0.0069)
F	10,000	1,174,518	0.4148 (0.0539)	1.0913	1.0649	1.0091 (0.0064)

LDSC, linkage disequilibrium score regression; SE, standard error estimate obtained via block jackknifing; h^2, observed scale heritability; λ_{GC}, genomic control inflation factor

†LDSC was run using the SNPs that were jointly present in the GWAS and a EUR-ancestry LD reference panel. Partitioned LD scores with zero variance were excluded from the analysis.

Table S18. Number of genome-wide significant SNPs included in PRS-CS calculations.

GWAS	GWAS sample size	Number of genome-wide significant SNPs[†]
AB	268,000	22,374
A	134,000	8,998
B	134,000	9,001
C	75,000	3,399
D	75,000	3,920
E	10,000	34
F	10,000	36

[†]These counts are the number of genome-wide SNPs present among the 1,113,490 LD-filtered SNPs that entered into the PRS-CS computations. We are defining genome-wide significance as $P < 5 \times 10^{-8}$.

Table S19. LDSC genetic correlations computed for UKBB Height GWAS.

	GWAS AB (*n* = 268,000)	GWAS A (*n* = 134,000)	GWAS B (*n* = 134,000)	GWAS C (*n* = 75,000)	GWAS D (*n* = 75,000)	GWAS E (*n* = 10,000)	GWAS F (*n* = 10,000)
GWAS AB		0.9944 (0.0029)	0.9966 (0.0028)	0.9907 (0.0067)	0.9922 (0.0060)	1.0558* (0.0584)	1.0637* (0.0541)
GWAS A	0.9944 (0.0029)		0.9854 (0.0097)	0.9946 (0.0039)	0.9829 (0.0121)	1.0372* (0.0539)	1.0515* (0.0587)
GWAS B	0.9966 (0.0028)	0.9854 (0.0097)		0.9868 (0.0123)	0.9976 (0.0037)	1.0789* (0.0638)	1.0709* (0.0533)
GWAS C	0.9907 (0.0067)	0.9946 (0.0039)	0.9868 (0.0123)		0.9877 (0.0152)	1.0219* (0.0491)	1.0809* (0.0640)
GWAS D	0.9922 (0.0060)	0.9829 (0.0121)	0.9976 (0.0037)	0.9877 (0.0152)		1.0870* (0.0631)	1.0680* (0.0484)
GWAS E	1.0558* (0.0584)	1.0372* (0.0539)	1.0789* (0.0638)	1.0219* (0.0491)	1.0870* (0.0631)		1.0789* (0.1064)
GWAS F	1.0637* (0.0541)	1.0515* (0.0587)	1.0709* (0.0533)	1.0809* (0.0640)	1.0680* (0.0484)	1.0789* (0.1064)	

LDSC, linkage disequilibrium score regression

Standard errors in parentheses were estimated via block jackknifing.

*Genetic correlations >1 are a known issue with LDSC (https://github.com/bulik/ldsc/issues/89).

Table S20. Correlation between height polygenic scores computed from pairs of GWAS run using varying numbers of unrelated white British individuals from the UK Biobank.

Height PGS Comparison (GWAS sample size)	Number of Overlapping GWAS Samples	% Overlap	Pearson Correlation Coefficient (*r*)	95% Confidence Interval
AB (*n* = 268,000) vs. A (*n* = 134,000)	134,000	50%	0.905	(0.901, 0.909)
AB (*n* = 268,000) vs. B (*n* = 134,000)	134,000	50%	0.907	(0.903, 0.911)
AB (*n* = 268,000) vs. C (*n* = 75,000)	75,000	28%	0.793	(0.784, 0.801)
AB (*n* = 268,000) vs. D (*n* = 75,000)	75,000	28%	0.792	(0.784, 0.800)
AB (*n* = 268,000) vs. E (*n* = 10,000)	10,000	3.7%	0.350	(0.331, 0.369)
AB (*n* = 268,000) vs. F (*n* = 10,000)	10,000	3.7%	0.357	(0.338, 0.376)
A (*n* = 134,000) vs. B (*n* = 134,000)	0	0%	0.649	(0.637, 0.662)
A (*n* = 134,000) vs. C (*n* = 75,000)	75,000	56%	0.879	(0.874, 0.884)
A (*n* = 134,000) vs. D (*n* = 75,000)	0	0%	0.562	(0.547, 0.577)
A (*n* = 134,000) vs. E (*n* = 10,000)	10,000	7.5%	0.396	(0.377, 0.414)
A (*n* = 134,000) vs. F (*n* = 10,000)	0	0%	0.246	(0.226, 0.267)
B (*n* = 134,000) vs. C (*n* = 75,000)	0	0%	0.567	(0.552, 0.581)
B (*n* = 134,000) vs. D (*n* = 75,000)	75,000	56%	0.880	(0.875, 0.885)
B (*n* = 134,000) vs. E (*n* = 10,000)	0	0%	0.242	(0.221. 0.262)
B (*n* = 134,000) vs. F (*n* = 10,000)	10,000	7.5%	0.405	(0.387, 0.423)
C (*n* = 75,000) vs. D (*n* = 75,000)	0	0%	0.485	(0.468, 0.501)
C (*n* = 75,000) vs. E (*n* = 10,000)	10,000	13.3%	0.458	(0.441, 0.475)
C (*n* = 75,000) vs. F (*n* = 10,000)	0	0%	0.213	(0.192, 0.234)
D (*n* = 75,000) vs. E (*n* = 10,000)	0	0%	0.214	(0.194, 0.235)
D (*n* = 75,000) vs. F (*n* = 10,000)	10,000	13.3%	0.456	(0.438, 0.473)
E (*n* = 10,000) vs. F (*n* = 10,000)	0	0%	0.106	(0.0840, 0.127)

Table S21. Variance explained by PGS computed from each height discovery GWAS.

GWAS	GWAS sample size	†R^2	‡‡Partial F-test statistic	‡‡P-value
AB	268,000	0.6286	1984.4	$< 2.2 \times 10^{-16}$
A	134,000	0.6095	1491.4	$< 2.2 \times 10^{-16}$
B	134,000	0.6138	1598.1	$< 2.2 \times 10^{-16}$
C	75,000	0.5927	1096.6	$< 2.2 \times 10^{-16}$
D	75,000	0.5954	1158.9	$< 2.2 \times 10^{-16}$
E	10,000	0.5482	192.16	$< 2.2 \times 10^{-16}$
F	10,000	0.5496	217.43	$< 2.2 \times 10^{-16}$

†This is the coefficient of determination for the additive linear model that includes sex, age at the time of height measurement, 20 ancestry PCs, and the standardized height PGS computed from the specified discovery GWAS as predictors of height. The R^2 for the base model was 0.5374; the base model was significant with $F(22,8084) = 426.9$, $P < 2.2 \times 10^{-16}$.

‡‡We are reporting the results for a partial-F test computed on 1 and 8083 degrees of freedom for the effect of adding the standardized polygenic score (PGS) to a base model that included sex, age, and the first 20 ancestry PCs as predictors.

Supplemental References

1. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics 26, 2867-2873.

2. Turner, S.D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv, 005165.

3. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C.A., and Smoller, J.W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nature Communications 10, 1776.

4. Nievergelt, C.M., Maihofer, A.X., Klengel, T., Atkinson, E.G., Chen, C.-Y., Choi, K.W., Coleman, J.R.I., Dalvie, S., Duncan, L.E., Gelernter, J., et al. (2019). International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. Nature Communications 10, 4558.

5. Duncan, L.E., Ratanatharathorn, A., Aiello, A.E., Almli, L.M., Amstadter, A.B., Ashley-Koch, A.E., Baker, D.G., Beckham, J.C., Bierut, L.J., Bisson, J., et al. (2018). Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. Mol Psychiatry 23, 666-673.

6. Mahajan, A., Wessel, J., Willems, S.M., Zhao, W., Robertson, N.R., Chu, A.Y., Gan, W., Kitajima, H., Taliun, D., Rayner, N.W., et al. (2018). Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nature Genetics 50, 559-571.

7. Chen, J., Sun, M., Adeyemo, A., Pirie, F., Carstensen, T., Pomilla, C., Doumatey, A.P., Chen, G., Young, E.H., Sandhu, M., et al. (2019). Genome-wide association study of type 2 diabetes in Africa. Diabetologia 62, 1204-1211.

8. Marouli, E., Graff, M., Medina-Gomez, C., Lo, K.S., Wood, A.R., Kjaer, T.R., Fine, R.S., Lu, Y., Schurmann, C., Highland, H.M., et al. (2017). Rare and low-frequency coding variants alter human adult height. Nature 542, 186-190.

9. Scott, R.A., Scott, L.J., Mägi, R., Marullo, L., Gaulton, K.J., Kaakinen, M., Pervjakova, N., Pers, T.H., Johnson, A.D., Eicher, J.D., et al. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes 66, 2888.

10. 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. Nature Genetics 46, 1173-1186.