The American Journal of Human Genetics, Volume *110*

Supplemental information

SDPRX: A statistical method for cross-population prediction of complex traits

Geyu Zhou, Tianqi Chen, and Hongyu Zhao

Supplementary Figures

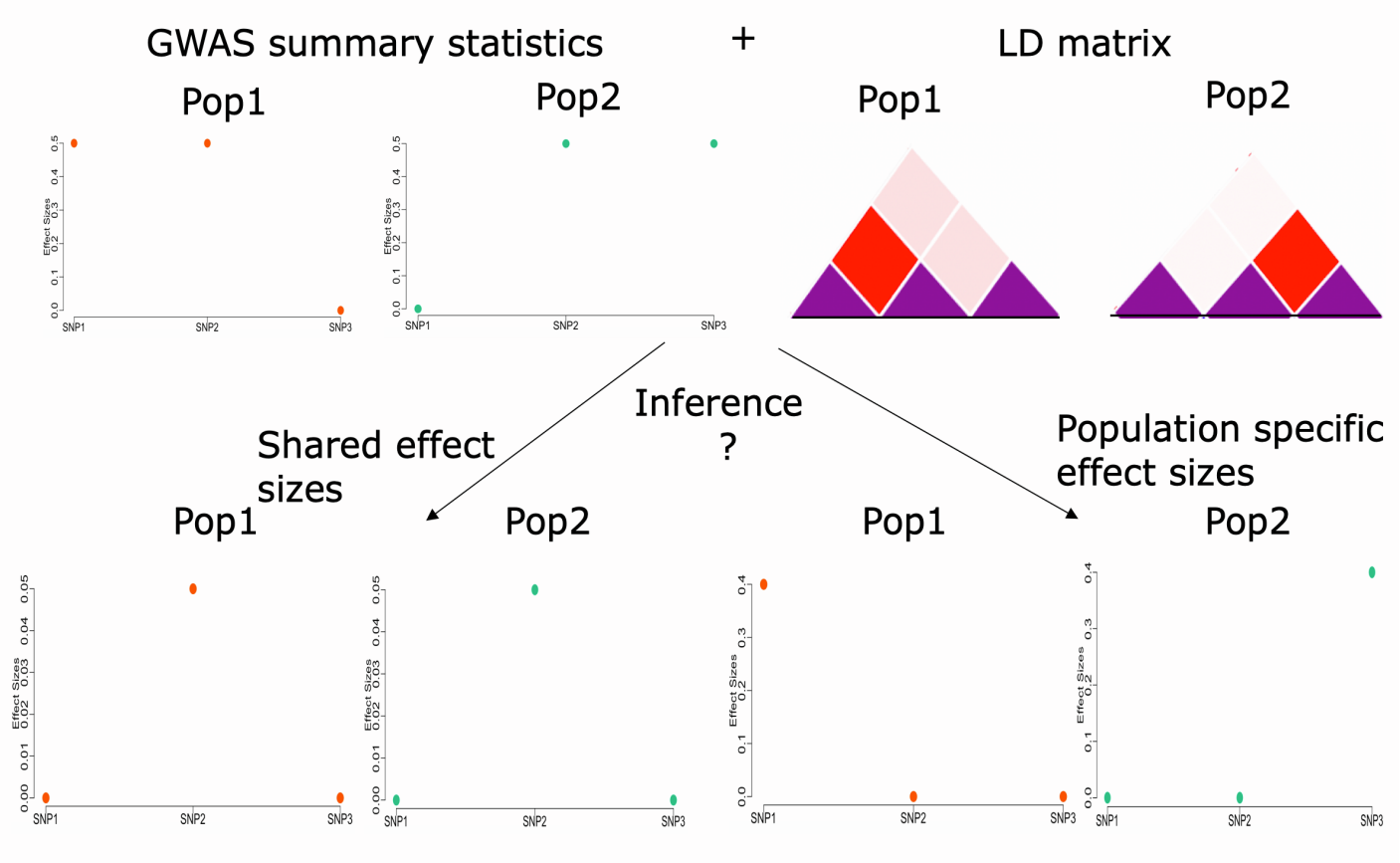

Figure S1. An example of identifiability issue caused by difference of local LD pattern of SNPs in two populations. SNP 1 and SNP 2 are in LD and have similar marginal effect sizes in population 1, while SNP2 and SNP3 are in LD and have similar marginal effect sizes in population 2. Two scenarios are likely to explain the observed GWAS summary statistics, which is not locally identifiable. In the first scenario, SNP 2 is causal and effect sizes are shared between populations. In the second scenario, SNP 1 is causal in population 1 and SNP 3 is causal in population 2. If genetic effects for most SNPs are shared between populations, then inference leading to population specific effect sizes would under-estimate the proportion of SNPs with shared effects and result in the loss of prediction accuracy.

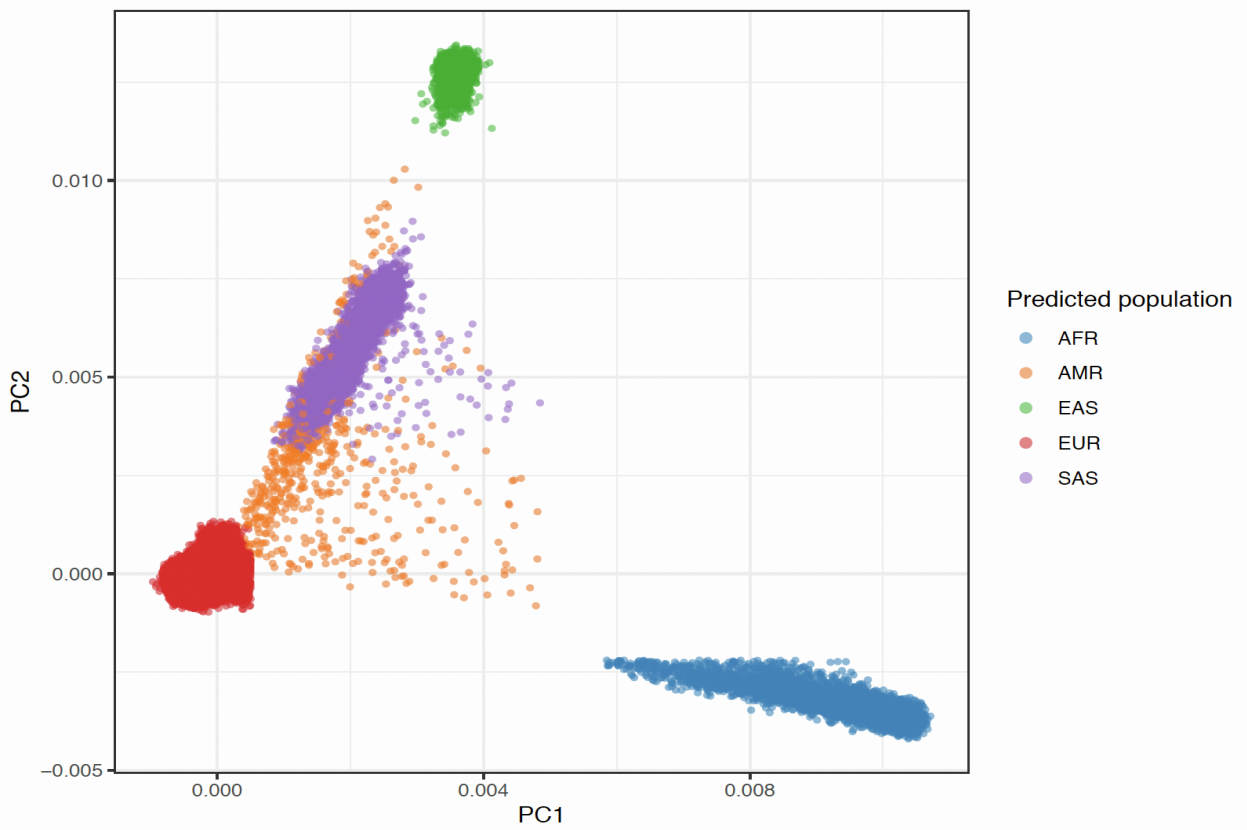

Figure S2. Principal component analysis of UK Biobank individuals after the assignment to one of five super populations (European, East Asian, African, Admixed American, South Asian) by a random forest classifier. We retained 2091 unrelated EAS and 6829 unrelated AFR samples with a predicted probability greater than 0.9 to form the validation and test datasets.

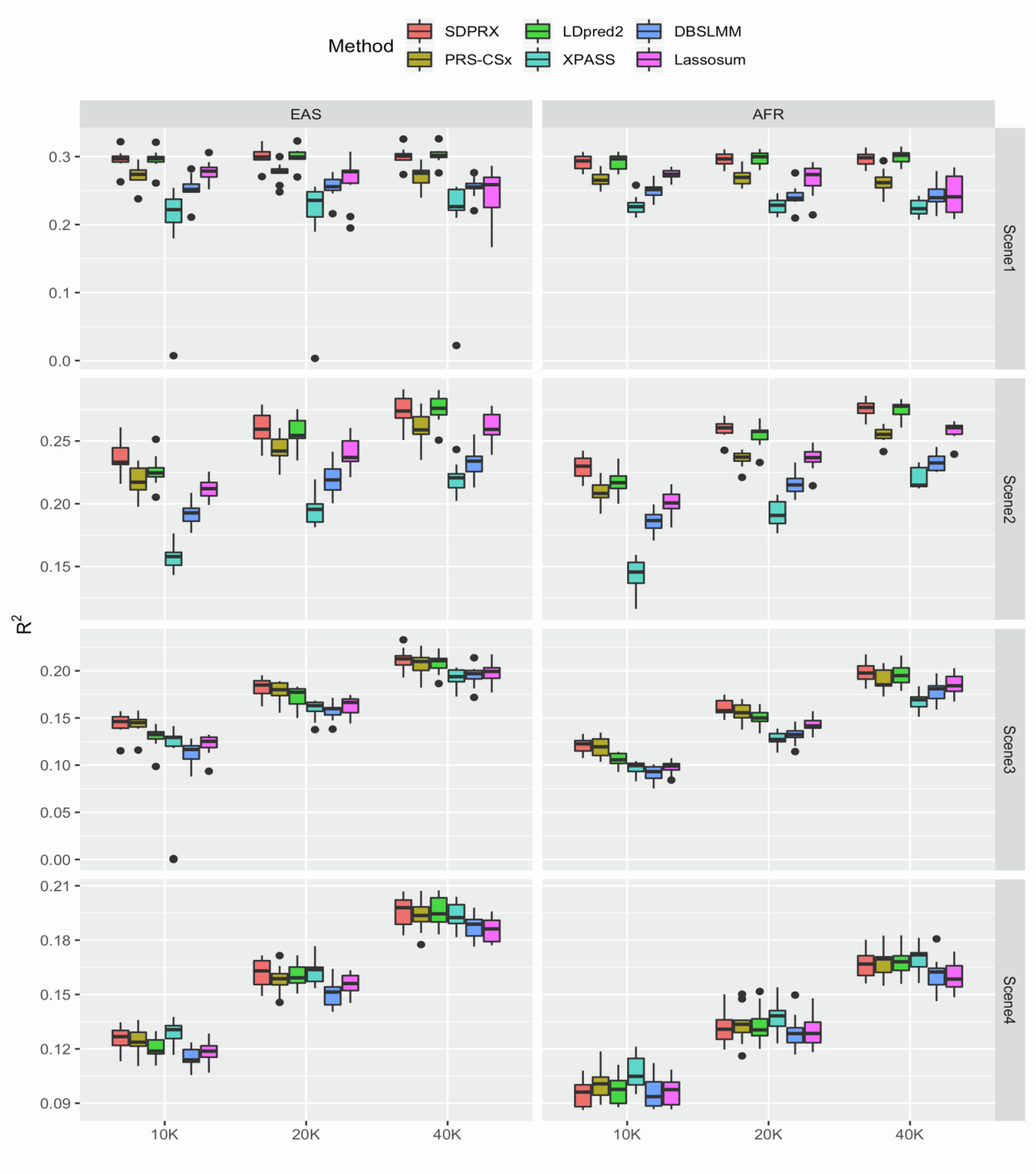

Figure S3. Prediction performance of different methods on simulated data. The proportion of SNPs with population 1 specific, population 2 specific and correlated effect sizes was equally set to be 0.05% (Scenario 1), 0.5% (Scenario 2) and 5% (Scenario 3). For Scenario 4, the proportion

of SNPs with population 1 specific, population 2 specific effect sizes were set to 5% and the proportion of SNPs with shared effect sizes was set to 90%. The cross-population genetic correlation was set to be 0.6 and the heritability was 0.3. Simulation in each scenario was repeated for 10 times. For each boxplot, the central mark is the median and the lower and upper edges represent the 25th and 75th percentiles.

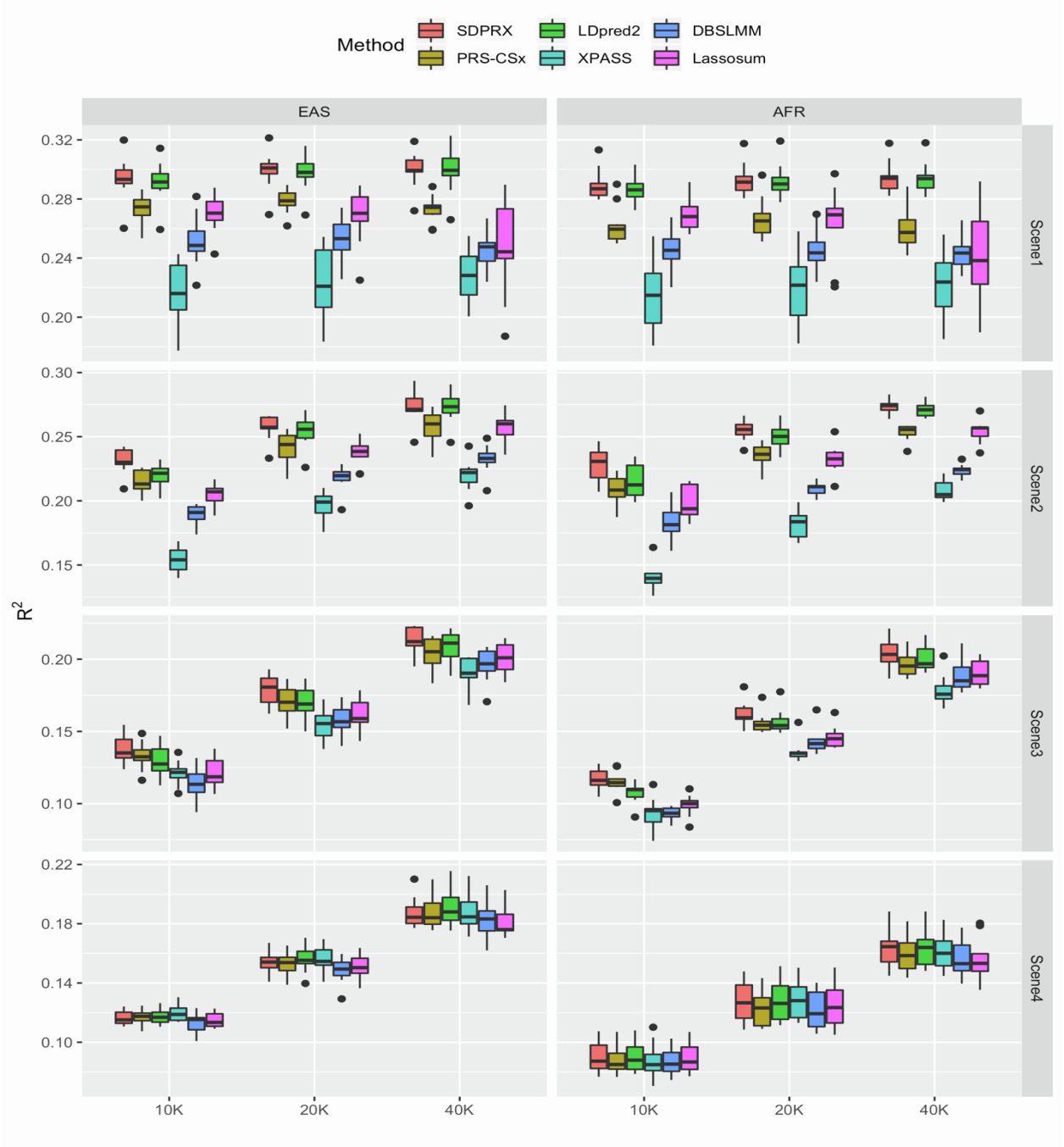

Figure S4. Prediction performance of different methods on simulated data. The proportion of SNPs with population 1 specific, population 2 specific and correlated effect sizes was equally set to be 0.05% (Scenario 1), 0.5% (Scenario 2) and 5% (Scenario 3). For Scenario 4, the proportion of SNPs with population 1 specific, population 2 specific effect sizes were set to 5% and the

proportion of SNPs with shared effect sizes was set to 90%. The cross-population genetic correlation was set to be 0.4 and the heritability was 0.3. Simulation in each scenario was repeated for 10 times. For each boxplot, the central mark is the median and the lower and upper edges represent the 25th and 75th percentiles.

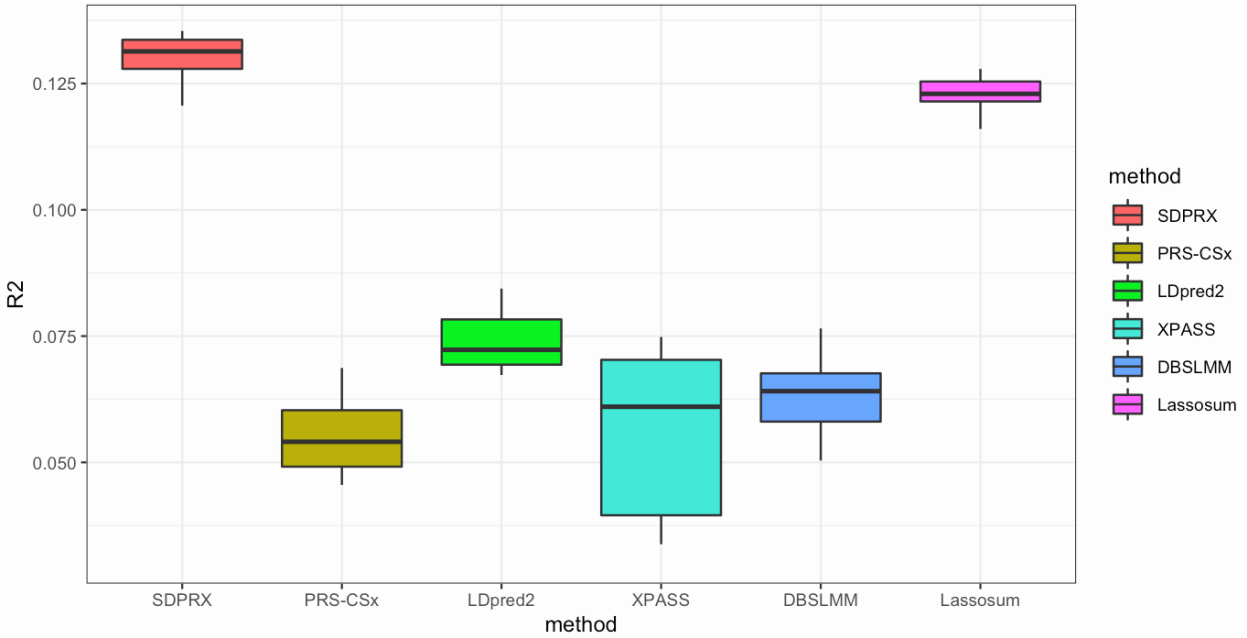

Figure S5. Prediction performance of different methods on simulated data for admixed individuals. The proportion of SNPs with population 1 specific, population 2 specific and correlated effect sizes was equally set to be 5% (Scenario 3). The cross-population genetic correlation was set to be 0.8 and the heritability was 0.4. The sample size was 40K for both EUR and AMR individuals. Simulation in each scenario was repeated for 10 times. For each boxplot, the central mark is the median and the lower and upper edges represent the 25th and 75th percentiles.

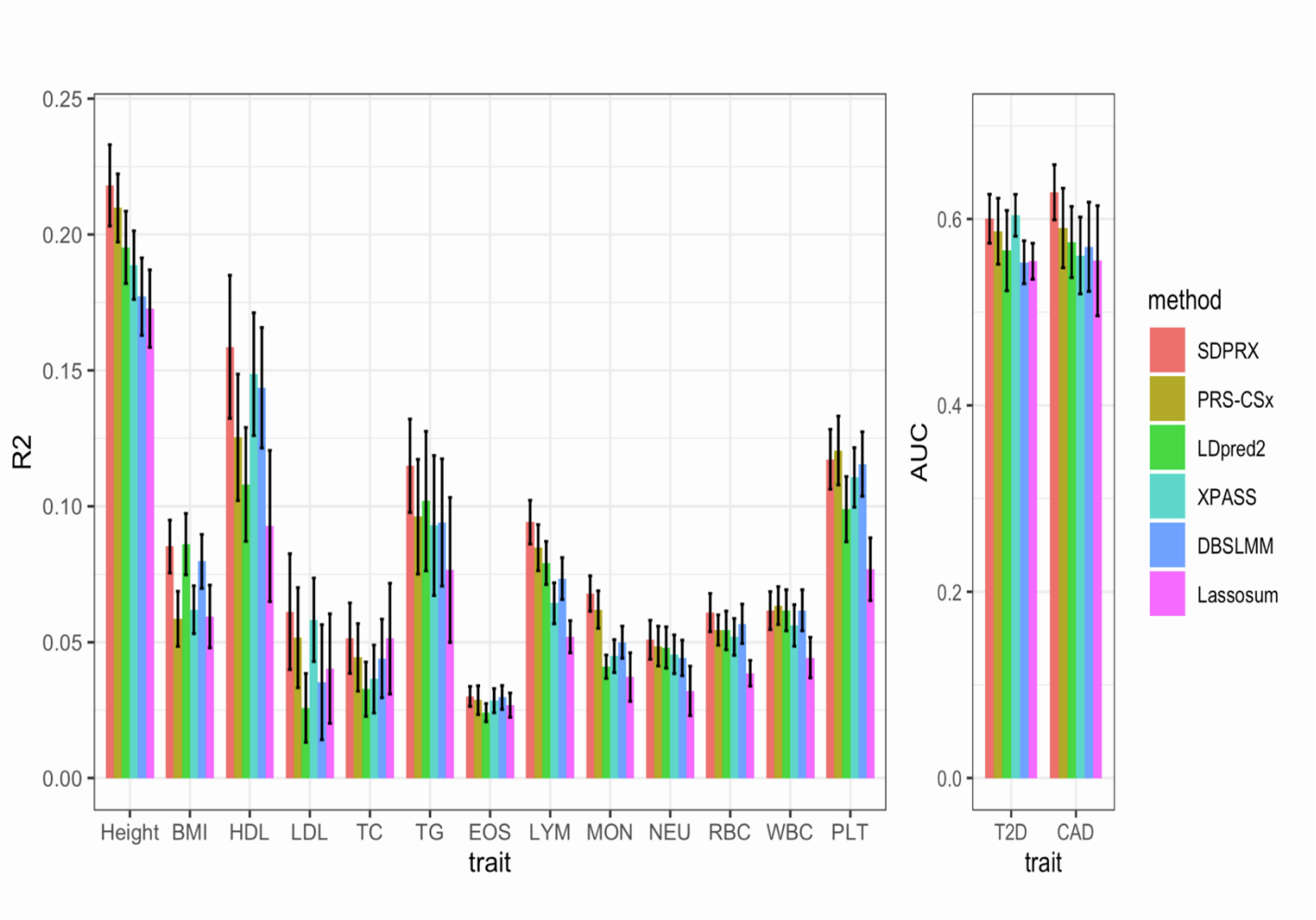

Figure S6. Prediction performance of different methods for 15 quantitative traits and 2 binary traits in EAS samples from UK Biobank with the linear combination of effect sizes. Selected participants with corresponding phenotypes were randomly split to form the validation (1/3) and test datasets (2/3). The mean and standard deviation of R^2 (quantitative trait) and AUC (binary trait) across 20 random splits are showed on the bar plots.

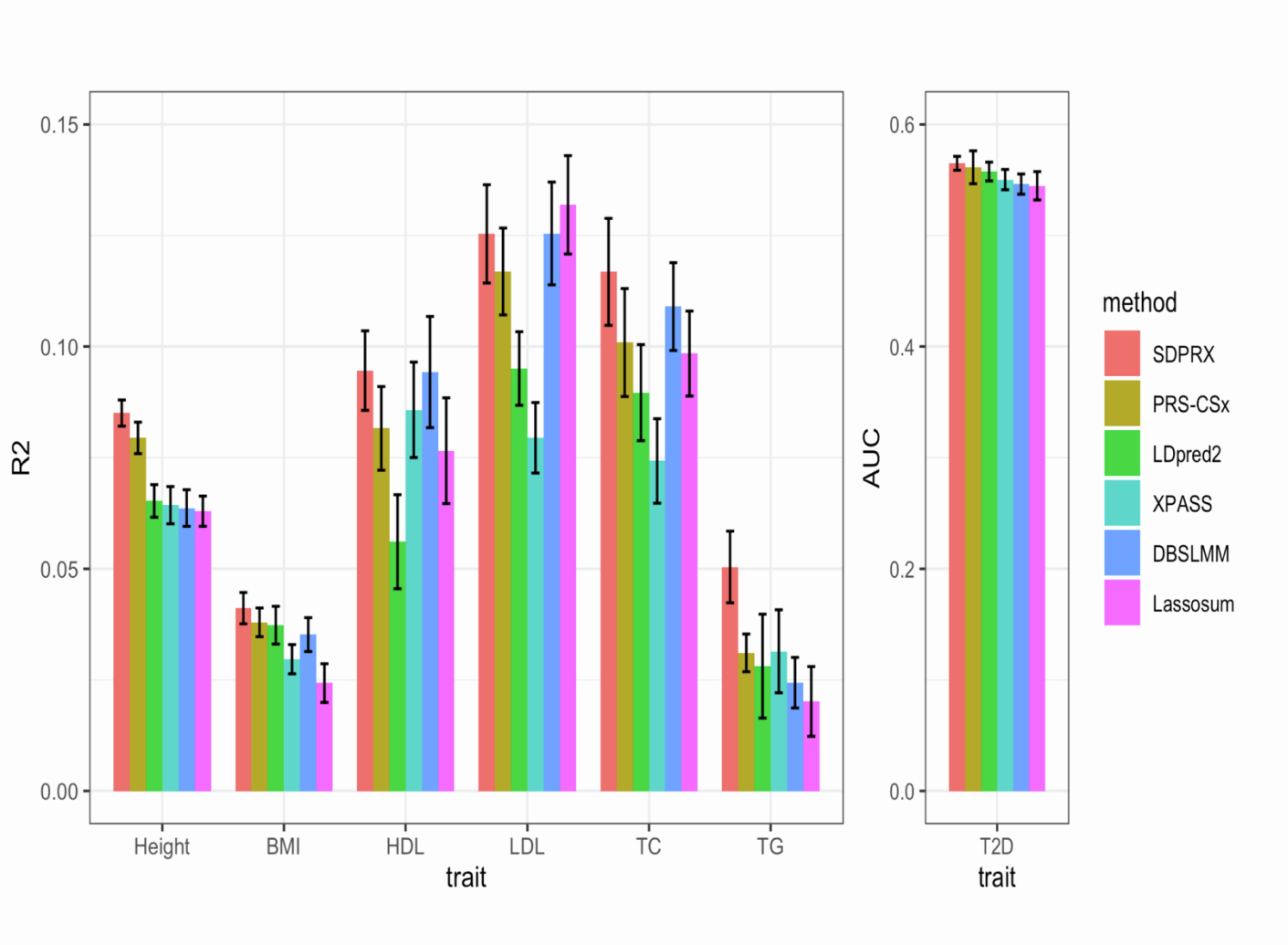

Figure S7. Prediction performance of different methods for 6 quantitative traits and 1 binary trait in AFR samples from UK Biobank with the linear combination of effect sizes. Selected participants with corresponding phenotypes were randomly split to form the validation (1/3) and test dataset (2/3). The mean and standard deviation of R^2 (quantitative trait) and AUC (binary trait) across 20 random splits are showed on the bar plots.

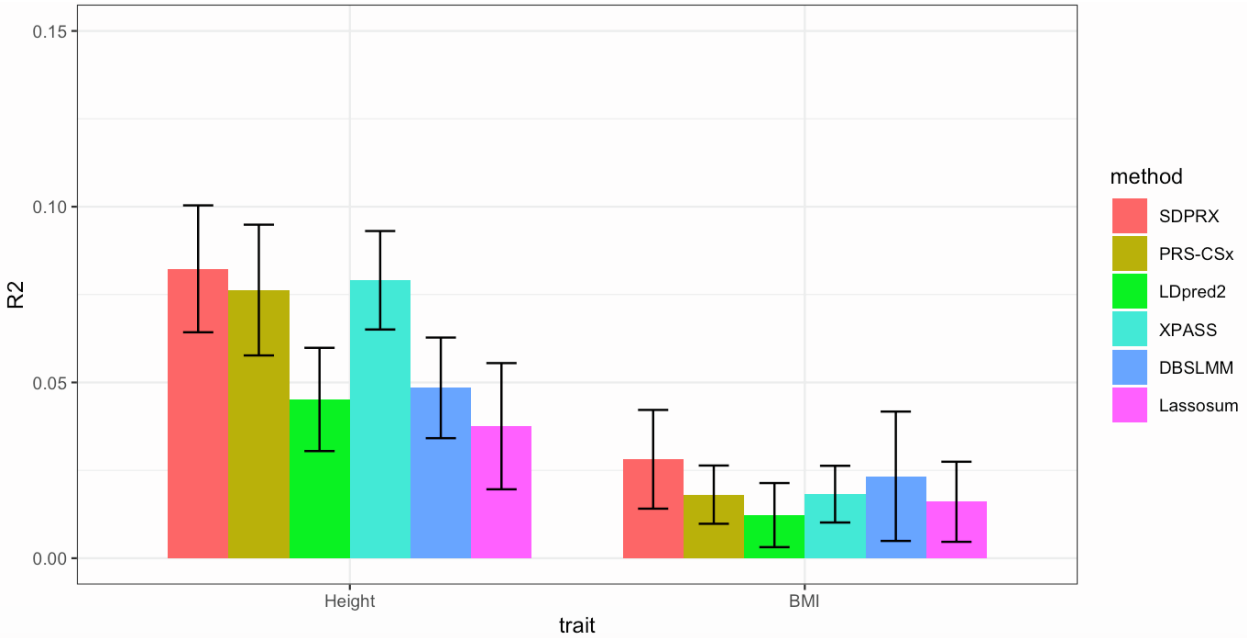

Figure S8. Prediction performance of different methods for 2 quantitative traits in AMR samples from UK Biobank. Selected participants with corresponding phenotypes were randomly split to form the validation (1/3) and test dataset (2/3). The mean and standard deviation of R^2 (quantitative trait) across 20 random splits are showed on the bar plots.

Supplementary Tables

		EAS			AFR		
		10K	20K	40K	10K	20K	40K
Scene 1	SDPRX	0.290	0.292	0.294	0.292	0.294	0.294
	PRS-CSx	0.268	0.274	0.274	0.263	0.268	0.261
	LDpred2	0.290	0.293	0.296	0.292	0.292	0.294
	XPASS	0.224	0.231	0.233	0.225	0.231	0.231
	DBSLMM	0.250	0.256	0.250	0.246	0.244	0.240
	Lassosum	0.270	0.276	0.256	0.269	0.270	0.231
Scene 2	SDPRX	0.230	0.259	0.271	0.229	0.258	0.274
	PRS-CSx	0.219	0.240	0.255	0.210	0.236	0.253
	LDpred2	0.222	0.257	0.269	0.219	0.252	0.273
	XPASS	0.156	0.199	0.221	0.140	0.182	0.213
	DBSLMM	0.189	0.220	0.232	0.181	0.208	0.226
	Lassosum	0.204	0.238	0.253	0.201	0.239	0.262
Scene 3	SDPRX	0.145	0.184	0.214	0.131	0.169	0.208
	PRS-CSx	0.153	0.183	0.209	0.136	0.166	0.196
	LDpred2	0.127	0.171	0.207	0.108	0.154	0.200
	XPASS	0.135	0.164	0.193	0.111	0.140	0.173
	DBSLMM	0.111	0.164	0.199	0.090	0.137	0.183
	Lassosum	0.120	0.161	0.197	0.097	0.144	0.191
Scene 4	SDPRX	0.142	0.174	0.197	0.103	0.135	0.168
	PRS-CSx	0.134	0.167	0.200	0.103	0.135	0.169
	LDpred2	0.122	0.163	0.195	0.092	0.131	0.167
	XPASS	0.131	0.169	0.196	0.107	0.141	0.171
	DBSLMM	0.117	0.149	0.190	0.903	0.128	0.161
	Lassosum	0.122	0.159	0.187	0.089	0.126	0.159

Table S1. The median of square of Pearson correlation across 10 replications when the cross-population genetic correlation was 0.8.

		EAS			AFR		
		10K	20K	40K	10K	20K	40K
Scene 1	SDPRX	0.297	0.299	0.301	0.294	0.297	0.298
	PRS-CSx	0.274	0.278	0.276	0.265	0.269	0.262
	LDpred2	0.297	0.299	0.302	0.297	0.300	0.302
	XPASS	0.222	0.236	0.226	0.226	0.228	0.223
	DBSLMM	0.252	0.256	0.255	0.251	0.239	0.240
	Lassosum	0.279	0.278	0.259	0.274	0.274	0.241
Scene 2	SDPRX	0.233	0.259	0.274	0.230	0.260	0.277
	PRS-CSx	0.217	0.242	0.259	0.208	0.237	0.255
	LDpred2	0.225	0.254	0.276	0.217	0.258	0.278
	XPASS	0.158	0.196	0.221	0.146	0.191	0.215
	DBSLMM	0.193	0.219	0.234	0.187	0.215	0.232
	Lassosum	0.212	0.237	0.259	0.201	0.237	0.261
Scene 3	SDPRX	0.146	0.185	0.213	0.123	0.158	0.198
	PRS-CSx	0.145	0.180	0.210	0.119	0.156	0.185
	LDpred2	0.134	0.178	0.211	0.106	0.150	0.195
	XPASS	0.130	0.163	0.194	0.096	0.127	0.169
	DBSLMM	0.117	0.160	0.197	0.093	0.132	0.181
	Lassosum	0.125	0.167	0.199	0.099	0.142	0.184
Scene 4	SDPRX	0.127	0.163	0.198	0.096	0.131	0.167
	PRS-CSx	0.124	0.159	0.194	0.101	0.134	0.170
	LDpred2	0.119	0.159	0.195	0.098	0.130	0.168
	XPASS	0.131	0.164	0.192	0.105	0.138	0.172
	DBSLMM	0.114	0.151	0.189	0.094	0.128	0.162
	Lassosum	0.119	0.156	0.185	0.098	0.129	0.159

Table S2. The median of square of Pearson correlation across 10 replications when the cross-population genetic correlation was 0.6.

		EAS			AFR		
		10K	20K	40K	10K	20K	40K
Scene 1	SDPRX	0.294	0.301	0.299	0.287	0.291	0.294
	PRS-CSx	0.275	0.279	0.274	0.260	0.265	0.257
	LDpred2	0.292	0.298	0.300	0.286	0.290	0.294
	XPASS	0.216	0.221	0.228	0.215	0.222	0.224
	DBSLMM	0.245	0.254	0.248	0.245	0.244	0.243
	Lassosum	0.270	0.270	0.244	0.268	0.269	0.238
Scene 2	SDPRX	0.230	0.257	0.271	0.231	0.256	0.274
	PRS-CSx	0.213	0.244	0.260	0.208	0.236	0.256
	LDpred2	0.222	0.260	0.273	0.212	0.250	0.271
	XPASS	0.154	0.199	0.222	0.140	0.184	0.205
	DBSLMM	0.191	0.220	0.233	0.181	0.211	0.224
	Lassosum	0.207	0.238	0.260	0.194	0.233	0.257
Scene 3	SDPRX	0.135	0.181	0.212	0.116	0.159	0.203
	PRS-CSx	0.132	0.170	0.205	0.115	0.154	0.195
	LDpred2	0.123	0.169	0.211	0.109	0.154	0.197
	XPASS	0.122	0.155	0.190	0.095	0.135	0.176
	DBSLMM	0.113	0.157	0.197	0.093	0.142	0.185
	Lassosum	0.119	0.159	0.201	0.100	0.145	0.189
Scene 4	SDPRX	0.115	0.154	0.184	0.087	0.127	0.165
	PRS-CSx	0.118	0.154	0.184	0.085	0.123	0.158
	LDpred2	0.117	0.155	0.188	0.088	0.126	0.164
	XPASS	0.119	0.155	0.185	0.085	0.128	0.160
	DBSLMM	0.115	0.149	0.183	0.085	0.193	0.153
	Lassosum	0.113	0.150	0.176	0.087	0.123	0.153

Table S3. The median of square of Pearson correlation across 10 replications when the cross-population genetic correlation was 0.4.

Traits	SDPRX	PRS-CSx	LDpred2	XPASS	DBSLMM	Lassosum
Height	0.213	0.173	0.174	0.171	0.152	0.144
BMI	0.085	0.060	0.071	0.062	0.066	0.071
HDL	0.159	0.117	0.102	0.131	0.106	0.069
LDL	0.066	0.048	0.040	0.061	0.033	0.033
Total cholesterol	0.043	0.039	0.040	0.044	0.035	0.030
Log triglycerides	0.123	0.080	0.080	0.085	0.066	0.051
Eosinophils	0.031	0.027	0.013	0.014	0.028	0.014
Lymphocytes	0.096	0.063	0.039	0.037	0.053	0.025
Monocytes	0.057	0.056	0.039	0.025	0.032	0.035
Neutrophils	0.049	0.044	0.018	0.028	0.041	0.015
Red blood cell	0.061	0.050	0.031	0.036	0.051	0.033
White blood cell	0.062	0.052	0.029	0.035	0.050	0.025
Platelet	0.117	0.103	0.075	0.072	0.097	0.071
Coronary Artery disease	0.626	0.56	0.600	0.589	0.562	0.613
Type 2 diabetes	0.606	0.604	0.573	0.585	0.560	0.536

Table S4. The mean of variance of phenotypes explained by PRS for 13 quantitative traits and AUC for 2 binary traits in EAS across 20 random splits without the linear combination of effect sizes.

Traits	SDPRX	PRS-CSx	LDpred2	XPASS	DBSLMM	Lassosum
Height	0.218	0.210	0.195	0.189	0.177	0.173
BMI	0.085	0.059	0.086	0.062	0.080	0.059
HDL	0.159	0.126	0.108	0.149	0.144	0.093
LDL	0.061	0.052	0.026	0.058	0.035	0.041
Total cholesterol	0.051	0.044	0.033	0.036	0.044	0.051
Log triglycerides	0.115	0.096	0.102	0.093	0.094	0.077
Eosinophils	0.030	0.029	0.024	0.028	0.030	0.027
Lymphocytes	0.094	0.085	0.079	0.064	0.073	0.052
Monocytes	0.068	0.062	0.041	0.045	0.050	0.037
Neutrophils	0.051	0.049	0.048	0.046	0.044	0.032
Red blood cell	0.061	0.054	0.054	0.052	0.057	0.039
White blood cell	0.062	0.064	0.062	0.056	0.062	0.044
Platelet	0.117	0.121	0.099	0.110	0.116	0.077
Coronary Artery disease	0.629	0.590	0.575	0.561	0.570	0.555
Type 2 diabetes	0.600	0.587	0.566	0.604	0.554	0.555

Table S5. The mean of variance of phenotypes explained by PRS for 13 quantitative traits and AUC for 2 binary traits in EAS across 20 random splits with the linear combination of effect sizes.

Traits	SDPRX	PRS-CSx	LDpred2	XPASS	DBSLMM	Lassosum
Height	0.083	0.061	0.051	0.054	0.041	0.038
BMI	0.041	0.025	0.023	0.028	0.021	0.019
HDL	0.095	0.082	0.090	0.083	0.082	0.088
LDL	0.128	0.120	0122	0.079	0.125	0.128
Total cholesterol	0.124	0.102	0.123	0.074	0.108	0.119
Log triglycerides	0.051	0.036	0.038	0.032	0.038	0.025
Type 2 diabetes	0.560	0.551	0.544	0.548	0.539	0.536

Table S6. The mean of variance of phenotypes explained by PRS for six quantitative traits and AUC for one binary trait in AFR across 20 random splits without the linear combination of effect sizes.

Traits	SDPRX	PRS-CSx	LDpred2	XPASS	DBSLMM	Lassosum
Height	0.085	0.079	0.065	0.064	0.064	0.063
BMI	0.041	0.038	0.037	0.030	0.035	0.024
HDL	0.095	0.082	0.056	0.086	0.094	0.077
LDL	0.125	0.117	0.095	0.079	0.125	0.132
Total cholesterol	0.117	0.101	0.090	0.074	0.109	0.098
Log triglycerides	0.050	0.031	0.028	0.031	0.024	0.020
Type 2 diabetes	0.565	0.561	0.558	0.550	0.546	0.545

Table S7. The mean of variance of phenotypes explained by PRS for six quantitative traits and AUC for one binary trait in AFR across 20 random splits with the linear combination of effect sizes.

Traits	SDPRX	PRS-CSx	LDpred2	XPASS	DBSLMM	Lassosum
Height	9.8 (4.4)	6.2 (1.0)	1.1 (26.1)	0.3 (18.0)	0.8 (1.0)	0.1 (2.4)
BMI	8 (4.7)	5.5 (1.0)	1.1 (25.8)	0.3 (22.0)	1.0 (1.0)	0.1 (2.4)
HDL	8.5 (4.7)	5.2 (1.0)	0.9 (23.6)	0.3 (18.3)	0.9 (1.0)	0.1 (2.0)
LDL	8.5 (4.7)	5.3 (1.0)	1.1 (23.5)	0.2 (19.3)	0.9 (1.0)	0.1 (2.3)
TC	5.1 (4.0)	4.3 (1.0)	1.1 (23.6)	0.3 (14.8)	0.9 (1.0)	0.09 (1.6)
TG	8.5 (4.7)	4.3 (1.0)	1.2 (23.6)	0.3 (19.3)	0.6 (0.8)	0.1 (2.3)
EOS	8.4 (4.9)	5.4 (1.0)	1.0 (23.3)	0.3 (19.2)	0.9 (1.0)	0.1 (2.3)
LYM	8.5 (4.9)	5.5 (1.0)	1.3 (23.3)	0.3 (19.2)	0.9 (1.0)	0.1 (2.0)
MON	8.5 (4.9)	4.7 (1.0)	1.3 (23.3)	0.3 (19.2)	1.5 (1.0)	0.1 (2.0)
NEU	8.5 (4.9)	4.3 (1.0)	1.3 (23.3)	0.3 (19.2)	0.9 (1.0)	0.1 (2.0)
RBC	8.6 (4.9)	4.7 (1.0)	1.1 (23.3)	0.3 (19.2)	0.9 (1.0)	0.1 (2.3)
WBC	8.7 (4.8)	5.2 (1.0)	1.1 (23.3)	0.3 (19.2)	0.8 (1.0)	0.1 (2.3)
PLT	8.5 (4.9)	4.3 (1.0)	1.0 (23.3)	0.3 (19.2)	0.8 (1.0)	0.1 (2.0)
T2D	4.5 (4.2)	5.5 (1.0)	1.0 (14.9)	0.3 (15.9)	0.8 (0.9)	0.1 (1.7)
CAD	6.9 (4.5)	4.7 (0.9)	1.1 (22.3)	0.3 (17.4)	0.9 (0.9)	0.1 (2.5)

Table S8. Computational time and memory usage of different methods for 15 traits. The computational time is in hours. Memory usage of each method, as listed in the parenthesis, is measured in the unit of Gigabytes (Gb).

Supplemental Methods

MCMC Algorithm

Here we describe our MCMC algorithm based on Gibbs sampling to obtain the posterior samples. For each SNP j, we introduce a vector $z_j = (m, k), m \in \{0,1,2,3\}, k \in \{1,2,\dots,1000\}$ indicating whether effect sizes are population specific and which variance component it is assigned to. For example, z_j equals (1,4) if the effect sizes of SNP j are population 1 specific and it is assigned to the fourth variance component.

Compute $\boldsymbol{A_1}, \boldsymbol{B_1}, \boldsymbol{A_2}, \boldsymbol{B_2}$: $\boldsymbol{A_1} = (\boldsymbol{R_1} + N_1 a\boldsymbol{I})^{-1}\boldsymbol{R_1}$, $\boldsymbol{B_1} = \boldsymbol{R_1}\boldsymbol{A_1}$. $\boldsymbol{A_2} = (\boldsymbol{R_2} + N_2 a\boldsymbol{I})^{-1}\boldsymbol{R_2}$, $\boldsymbol{B_2} = \boldsymbol{R_2}\boldsymbol{A_2}$.

Sampling z_j: For each LD block, we first integrate out $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ to derive the conditional probability of SNP j whose effect sizes are correlated in two populations and assigned to the k[th] variance component:

$$P(z_j = (3,k)|.)$$

$$\propto \int\int p(\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}|\boldsymbol{\beta_{1j}}, \eta)\, p(\widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}|\boldsymbol{\beta_{2j}}, \eta)\, p(\boldsymbol{\beta_{1j}}, \boldsymbol{\beta_{2j}}|z_j = (3,k), \sigma^2_{3k})\, d\beta_{1j} d\beta_{2j}$$

$$\times\ P(z_j = (3,k))$$

$$\propto \int\int \exp\left\{-\frac{1}{2}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{1}}-\eta \boldsymbol{R}_{\mathbf{1}}\boldsymbol{\beta}_{\mathbf{1}}\right)^{T}\left(\boldsymbol{R}_{\mathbf{1}}/N_1+a\boldsymbol{I}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{1}}-\eta \boldsymbol{R}_{\mathbf{1}}\boldsymbol{\beta}_{\mathbf{1}}\right)\right\}\exp\left\{-\frac{1}{2}\left(\widehat{\boldsymbol{\beta}}_{2}\right.\right.$$

$$\left.-\eta \boldsymbol{R}_{\mathbf{2}}\boldsymbol{\beta}_{\mathbf{2}}\right)^{T}\left(\boldsymbol{R}_{\mathbf{2}}/N_2+a\boldsymbol{I}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{2}}\right.$$

$$\left.\left.-\eta \boldsymbol{R}_{\mathbf{2}}\boldsymbol{\beta}_{\mathbf{2}}\right)\right\}\frac{1}{2\pi\sigma_{3k}^{2}\sqrt{1-\rho^{2}}}\exp\left\{-\frac{1}{2(1-\rho^{2})}\left[\frac{\beta_{1j}^{2}+\beta_{2j}^{2}-2\rho\beta_{1j}\beta_{2j}}{\sigma_{3k}^{2}}\right]\right\}d\beta_{1j}d\beta_{2j}$$

$$\times \pi_{3k}p_3$$

$$\propto \int\int \exp\left\{-\frac{N_1}{2}\eta^{2}\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{T}}\boldsymbol{B}_{\mathbf{1}}\boldsymbol{\beta}_{\mathbf{1}}+N_1\eta\widehat{\boldsymbol{\beta}}_{\mathbf{1}}^{\ T}\boldsymbol{A}_{\mathbf{1}}\boldsymbol{\beta}_{\mathbf{1}}\right\}\exp\left\{-\frac{N_2}{2}\eta^{2}\boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{T}}\boldsymbol{B}_{\mathbf{2}}\boldsymbol{\beta}_{\mathbf{2}}\right.$$

$$\left.+N_2\eta\widehat{\boldsymbol{\beta}}_{\mathbf{2}}^{\ T}\boldsymbol{A}_{\mathbf{2}}\beta_{12}\right\}\frac{1}{2\pi\sigma_{3k}^{2}\sqrt{1-\rho^{2}}}\exp\left\{-\frac{1}{2(1-\rho^{2})}\left[\frac{\beta_{1j}^{2}+\beta_{2j}^{2}-2\rho\beta_{1j}\beta_{2j}}{\sigma_{3k}^{2}}\right]\right\}d\beta_{1j}d\beta_{2j}$$

$$\times \pi_{3k}p_3$$

$$\propto \int\int \exp\left\{-\frac{N_1}{2}\eta^{2}B_{1,jj}\beta_{1j}^{2}-N_1\eta^{2}\sum_{i\neq j}B_{1,ij}\beta_{1i}\beta_{1j}\right.$$

$$\left.+N_1\eta\sum_{i}A_{1,ij}\hat{\beta}_{1i}\beta_{1j}\right\}\exp\left\{-\frac{N_2}{2}\eta^{2}B_{2,jj}\beta_{2j}^{2}-N_2\eta^{2}\sum_{i\neq j}B_{2,ij}\beta_{2i}\beta_{2j}\right.$$

$$\left.+N_2\eta\sum_{i}A_{2,ij}\hat{\beta}_{2i}\beta_{2j}\right\}\frac{1}{2\pi\sigma_{3k}^{2}\sqrt{1-\rho^{2}}}\exp\left\{-\frac{1}{2(1-\rho^{2})}\left[\frac{\beta_{1j}^{2}+\beta_{2j}^{2}-2\rho\beta_{1j}\beta_{2j}}{\sigma_{3k}^{2}}\right]\right\}d\beta_{1j}d\beta_{2j}$$

$$\times \pi_{3k}p_3$$

$$\propto \int\int \exp\left\{-a_{jk1}\beta_{1j}^{2}-a_{jk2}\beta_{2j}^{2}+N_1b_{1j}+N_2b_{2j}+c_k\beta_{1j}\beta_{2j}\right\}\frac{1}{2\pi\sigma_{3k}^{2}\sqrt{1-\rho^{2}}}d\beta_{1j}d\beta_{2j}$$

$$\times \pi_{3k}p_3$$

$$\propto \frac{1}{\left(4a_{jk1}a_{jk2}-c_k^{2}\right)^{\frac{1}{2}}\sigma_{3k}^{2}}\exp\left\{a_{jk1}\mu_{jk1}^{2}+a_{jk2}\mu_{jk2}^{2}-c_k\mu_{jk1}\mu_{jk2}\right\}\times\frac{\pi_{3k}p_3}{\sqrt{1-\rho^{2}}}$$

(1)

where

$$b_{1j} = \eta \sum_i A_{1,ij}\hat{\beta}_{1i} - \eta^2 \sum_{i \neq j} B_{1,ij}\beta_{1i}$$

$$b_{2j} = \eta \sum_i A_{2,ij}\hat{\beta}_{2i} - \eta^2 \sum_{i \neq j} B_{2,ij}\beta_{2i}$$

$$a_{jk1} = \frac{N_1}{2}\eta^2 B_{1,jj} + \frac{1}{2\sigma_{3k}^2(1-\rho^2)}$$

$$a_{jk2} = \frac{N_2}{2}\eta^2 B_{2,jj} + \frac{1}{2\sigma_{3k}^2(1-\rho^2)}$$

$$\mu_{jk1} = \frac{2a_{jk2}N_1 b_{1j} + c_k N_2 b_{2j}}{4a_{jk1}a_{jk2} - c_k^2}$$

$$\mu_{jk2} = \frac{2a_{jk1}N_2 b_{2j} + c_k N_1 b_{1j}}{4a_{jk1}a_{jk2} - c_k^2}$$

$$c_k = \frac{\rho}{(1-\rho^2)\sigma_{3k}^2}$$

We next derive the conditional probability of SNP j whose effect sizes are population specific or null. It can be viewed as the special case to evaluate the last integrand by setting $\rho = 0, \beta_{2j} = 0$ (population 1 specific), $\rho = 0, \beta_{1j} = 0$ (population 2 specific), and $\beta_{1j} = \beta_{2j} = 0$ (both null).

$$P(z_j = (1,k)|.) \propto \frac{1}{\sqrt{N_1\eta^2 B_{1,jj}\sigma_{1k}^2 + 1}} \exp\left\{\frac{N_1^2 b_{1j}^2}{2(N_1\eta^2 B_{1,jj} + \sigma_{1k}^{-2})}\right\} \times \pi_{1k}p_1$$

$$P(z_j = (2,k)|.) \propto \frac{1}{\sqrt{N_2\eta^2 B_{2,jj}\sigma_{2k}^2 + 1}} \exp\left\{\frac{N_2^2 b_{2j}^2}{2(N_2\eta^2 B_{2,jj} + \sigma_{2k}^{-2})}\right\} \times \pi_{2k}p_2$$

$$P(z_j = (0,0)|.) \propto p_0$$

We use log-exp-sum trick to avoid numerical overflow. Note that because SNPs in different LD blocks are approximately independent, we can sample their assignments in parallel. For population 1 specific SNPs, we only need to evaluate $P\left(z_j = (1,k)\middle|.\right)$ and $P\left(z_j = (0,0)\middle|.\right)$.

Sampling $\boldsymbol{\beta_1}, \boldsymbol{\beta_2}$: For SNPs that are non-causal in any populations, we simply set the corresponding entries of $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ as zero. We then jointly sample the effect sizes of causal SNPs in one independent LD block. We introduce two indexes γ_1 and γ_2 such that $\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}$ and $\boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}$ are non-zero. We combine $\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}$ and $\boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}$ into one vector $\boldsymbol{\beta_\gamma}$, which follows a bivariate normal distribution with mean 0 and variance-covariance matrix $\boldsymbol{\Sigma_0}$. The jth diagonal entry of $\boldsymbol{\Sigma_0}$ is $\sigma_{z_j}^2$. If effect sizes of one SNP are non-zero with correlation in two populations, then $\Sigma_{0,ij} = \Sigma_{0,\mathrm{ji}} = \rho\sigma_{z_j}^2$. Other entries of Σ_0 are zero. Note that the special structure of $\boldsymbol{\Sigma_0}$ allows an analytical solution of $\boldsymbol{\Sigma_0^{-1}}$. We next derive the conditional likelihood as:

$$
\begin{aligned}
& p\left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}, \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\middle|.\right) \\
& \propto \exp\left\{-\frac{N_1}{2}\eta^2 \boldsymbol{\beta_1^T B_1 \beta_1}\right. \\
& \left. + \eta\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}{}^{\boldsymbol{T}} \boldsymbol{A_1 \beta_1}\right\} \exp\left\{-\frac{N_2}{2}\eta^2 \boldsymbol{\beta_2^T B_2 \beta_2} + \eta\widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}{}^{\boldsymbol{T}} \boldsymbol{A_2 \beta_2}\right\} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)^{\boldsymbol{T}} \boldsymbol{\Sigma_0^{-1}} \left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)\right\} \\
& \propto \exp\left\{-\frac{1}{2}\eta^2 \left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)^{\boldsymbol{T}} \begin{pmatrix} \boldsymbol{N_1 B_{1,\gamma_1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{N_2 B_{2,\gamma_2}} \end{pmatrix} \left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)\right. \\
& \qquad \left. + \eta\left(N_1 \widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{1},\boldsymbol{\gamma_1}} \quad N_2 \widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)\right\} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)^{\boldsymbol{T}} \boldsymbol{\Sigma_0^{-1}} \left(\boldsymbol{\beta}_{\boldsymbol{1},\boldsymbol{\gamma_1}}\ \boldsymbol{\beta}_{\boldsymbol{2},\boldsymbol{\gamma_2}}\right)\right\} \\
& \propto \exp\left\{-\frac{1}{2}\eta^2 \boldsymbol{\beta_\gamma^T B_\gamma \beta_\gamma} + \eta\widehat{\boldsymbol{\beta}}^{\boldsymbol{T}} \boldsymbol{A_\gamma \beta_\gamma}\right\} \exp\left\{-\frac{1}{2}\boldsymbol{\beta_\gamma^T \Sigma_0^{-1} \beta_\gamma}\right\}
\end{aligned}
$$

$$= MVN(\eta \boldsymbol{\Sigma} \boldsymbol{A}_{\boldsymbol{\gamma}}^{\boldsymbol{T}} \widehat{\boldsymbol{\beta}}_{\boldsymbol{\gamma}},\ \boldsymbol{\Sigma}) \tag{2}$$

where $\boldsymbol{\Sigma} = \left(\eta^2 \boldsymbol{B}_{\boldsymbol{\gamma}} + \boldsymbol{\Sigma}_{\boldsymbol{0}}^{-\boldsymbol{1}}\right)^{-1}, \boldsymbol{A}_{\boldsymbol{\gamma}} = \begin{pmatrix} N_1 \widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{1},\boldsymbol{\gamma}_1} & N_2 \widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{2},\boldsymbol{\gamma}_2} \end{pmatrix}, \boldsymbol{B}_{\boldsymbol{\gamma}} = \begin{pmatrix} N_1 \boldsymbol{B}_{\boldsymbol{1},\boldsymbol{\gamma}_1} & 0 \\ 0 & N_2 \boldsymbol{B}_{\boldsymbol{2},\boldsymbol{\gamma}_2} \end{pmatrix}$. $\boldsymbol{A}_{\boldsymbol{1},\boldsymbol{\gamma}_1}$ is the submatrix by selecting columns from matrices $\boldsymbol{A}_{\boldsymbol{1}}$ based on the index γ_1. $\boldsymbol{B}_{\boldsymbol{1},\boldsymbol{\gamma}_1}$ is the submatrix by selecting rows and columns from matrices $\boldsymbol{B}_{\boldsymbol{1}}$ based on the index γ_1.

Sampling η: The full conditional likelihood is

$$p(\eta|.)$$

$$\propto \exp\left\{-\frac{1}{2} N_1 \eta^2 \sum \boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{1}} \boldsymbol{\beta}_{\boldsymbol{1}}\right.$$

$$\left. + N_1 \eta \sum \widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{1}} \boldsymbol{\beta}_{\boldsymbol{1}} \right\} \exp\left\{-\frac{1}{2} N_2 \eta^2 \sum \boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{2}} \boldsymbol{\beta}_{\boldsymbol{2}} + N_2 \eta \sum \widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{2}} \boldsymbol{\beta}_{\boldsymbol{2}} \right\} \exp\left\{-\frac{\eta^2}{2 \times 10^{-6}}\right\}$$

$$= N\left(\frac{N_1(\sum \widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{1}} \boldsymbol{\beta}_{\boldsymbol{1}}) + N_2(\sum \widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{A}_{\boldsymbol{2}} \boldsymbol{\beta}_{\boldsymbol{2}})}{N_1(\sum \boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{1}} \boldsymbol{\beta}_{\boldsymbol{1}}) + N_2(\sum \boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{2}} \boldsymbol{\beta}_{\boldsymbol{2}}) + 10^{-6}}, \frac{1}{N_1(\sum \boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{1}} \boldsymbol{\beta}_{\boldsymbol{1}}) + N_2(\sum \boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{B}_{\boldsymbol{2}} \boldsymbol{\beta}_{\boldsymbol{2}}) + 10^{-6}}\right)$$

Sampling $\sigma_{1k}^2, \sigma_{2k}^2, \sigma_{3k}^2$: The full conditional likelihood is

$$p(\sigma_{1k}^2|.) \propto \prod_{j:z_j=(1,k)} \frac{1}{\sigma_{1k}} \exp\left\{-\frac{\beta_{1j}^2}{2\sigma_{1k}^2}\right\} \sigma_{1k}^{-2(.5-1)} \exp\left\{-\frac{.5}{\sigma_{1k}^2}\right\}$$

$$= IG\left(\frac{M_{1k}}{2} + .5, \frac{\sum_{j:z_j=(1,k)} \beta_{1j}^2}{2} + .5\right)$$

$$p(\sigma_{2k}^2|.) \propto \prod_{j:z_j=(2,k)} \frac{1}{\sigma_{2k}} \exp\left\{-\frac{\beta_{2j}^2}{2\sigma_{2k}^2}\right\} \sigma_{2k}^{-2(.5-1)} \exp\left\{-\frac{.5}{\sigma_{2k}^2}\right\}$$

$$= IG\left(\frac{M_{2k}}{2} + .5, \frac{\sum_{j:z_j=(2,k)} \beta_{2j}^2}{2} + .5\right)$$

$$p(\sigma_{3k}^2|.) \propto \prod_{j:z_j=(3,k)} \frac{1}{\sigma_{3k}} \exp\left\{-\frac{\beta_{1j}^2 + \beta_{2j}^2 - 2\rho\beta_{1j}\beta_{2j}}{2(1-\rho^2)\sigma_{3k}^2}\right\} \sigma_{3k}^{-2(.5-1)} \exp\left\{-\frac{.5}{\sigma_{3k}^2}\right\}$$

$$= IG\left(\frac{M_{3k}}{2} + .5, \frac{\sum_{j:z_j=(3,k)} \beta_{1j}^2 + \beta_{2j}^2 - 2\rho\beta_{1j}\beta_{2j}}{2(1-\rho^2)} + .5\right)$$

where $M_{1k} = \sum_j I\left(z_j = (1,k)\right), M_{2k} = \sum_j I\left(z_j = (2,k)\right), M_{3k} = \sum_j I\left(z_j = (3,k)\right)$ and I is the indicator function.

Sampling $V_{mk}, m \in \{1,2,3\}, k \in \{1,2,\dots,1000\}$: The full conditional likelihood is

$$p(V_{mk}|.) \propto V_k^{M_{mk}} (1-V_k)^{M_{m(k+1)}+\cdots+M_{1000}+\alpha_M-1}$$

$$= Beta\left(1 + M_{mk}, \alpha + \sum_{l=k+1}^{1000} M_{ml}\right)$$

for j=1,2,3 and k = 1, …, 999. V_{m1000} equals 1 according to the definition of the truncated stick-breaking process.

Computing $\pi_{mk}, m \in \{1,2,3\}$: The prior probability can be computed as

$$\pi_{m1} = V_{m1}$$

$$\pi_{mk} = \prod_{l=1}^{k-1}(1 - V_{ml})\, V_{mk} \;\; (k \geq 2)$$

Sampling p_0, p_1, p_2, p_3: The conditional distribution is:

$$p_0, p_1, p_2, p_3|. \sim Dir(M_0 + 1, M_1 + 1, M_2 + 1, M_3 + 1)$$

where $M_0 = \sum_j I\left(z_j = (0,0)\right), M_1 = \sum_j I\left(z_j = (1,.)\right), M_2 = \sum_j I\left(z_j = (2,.)\right), M_{3k} = \sum_j I\left(z_j = (3,.)\right)$. Note that we exclude population specific variants when computing M_0, M_1, M_2, M_3.

Sampling $\alpha_m, m \in \{1,2,3\}$: The full conditional likelihood is

$$p(\alpha_m|.) \propto \prod_{l=1}^{1000-1} \alpha_m(1 - V_{ml})^{\alpha_m - 1}\alpha_m^{1-1} \exp\{-.1 \times \alpha_m\}$$

$$= Gamma(0.1 + 1000 - 1, 0.1 - \sum_{k=1}^{1000-1} \log(1 - V_{mk}))$$

We record the effect sizes $\eta\boldsymbol{\beta_1}$ and $\eta\boldsymbol{\beta_2}$ together with the heritability $h_1^2 = \boldsymbol{\beta_1^T R_1 \beta_1}$ and $h_2^2 = \boldsymbol{\beta_2^T R_2 \beta_2}$ for each iteration and compute the average of all posterior samples as the final estimator. We note that h_1^2 and h_2^2 together with the maximum of effect sizes can be used to assess whether the algorithm converges.