Cell Genomics, Volume *4*

Supplemental information

MUSSEL: Enhanced Bayesian polygenic risk prediction leveraging information across multiple ancestry groups

Jin Jin, Jianan Zhan, Jingning Zhang, Ruzhang Zhao, Jared O'Connell, Yunxuan Jiang, 23andMe Research Team, Steven Buyske, Christopher Gignoux, Christopher Haiman, Eimear E. Kenny, Charles Kooperberg, Kari North, Bertram L. Koelsch, Genevieve Wojcik, Haoyu Zhang, and Nilanjan Chatterjee

Supplemental Figures

for

MUSSEL: Enhanced Bayesian Polygenic Risk Prediction Leveraging Information across Multiple Ancestry Groups

This document includes:

Supplementary Figures S1-S19

Supplementary Figure Legends

Supplemental References

Figure S1: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 15,000/45,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestry groups but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

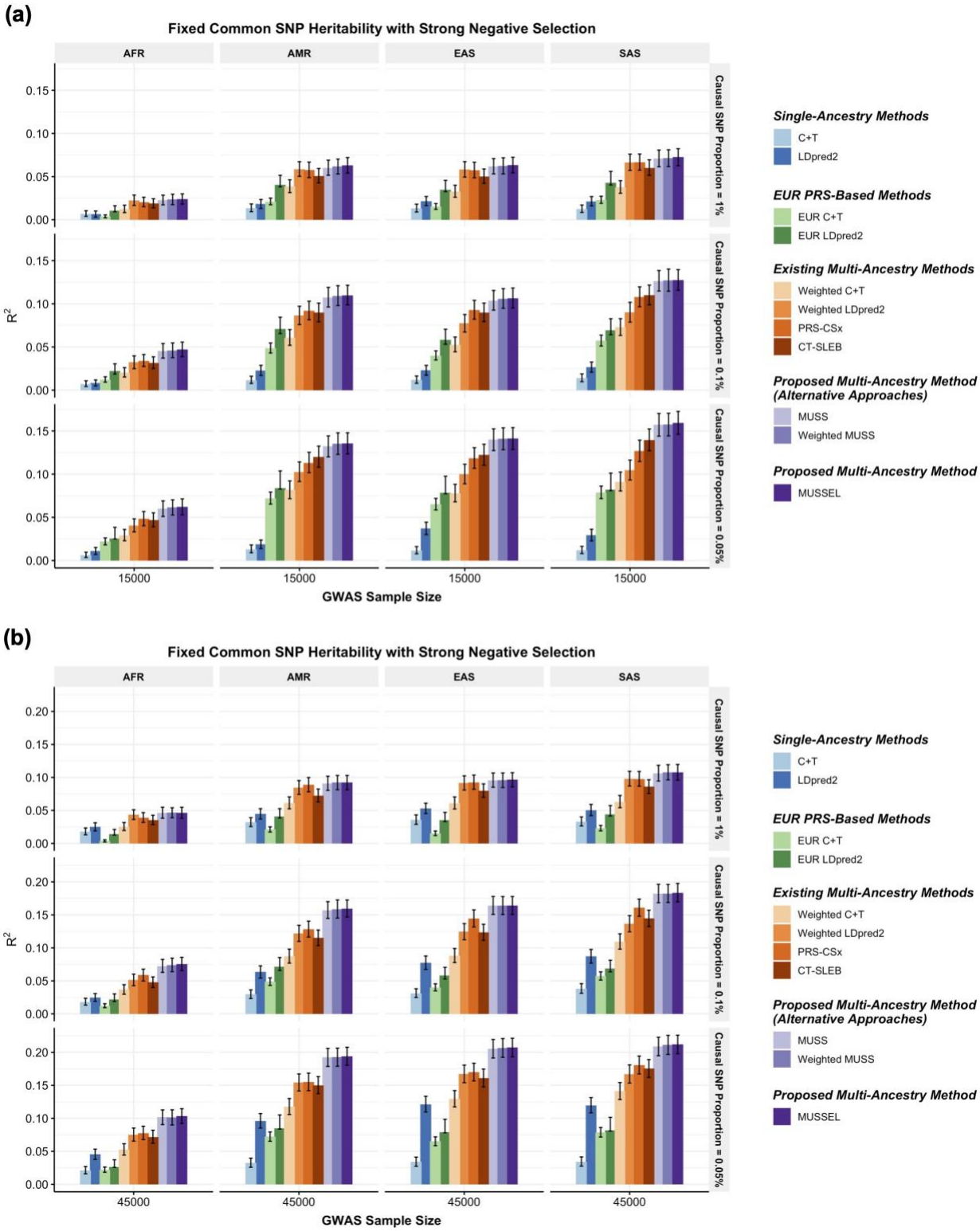

(a)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.00
0.05
0.10
0.15
15000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.00
0.05
0.10
0.15
0.20
45000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S2: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 80,000/100,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

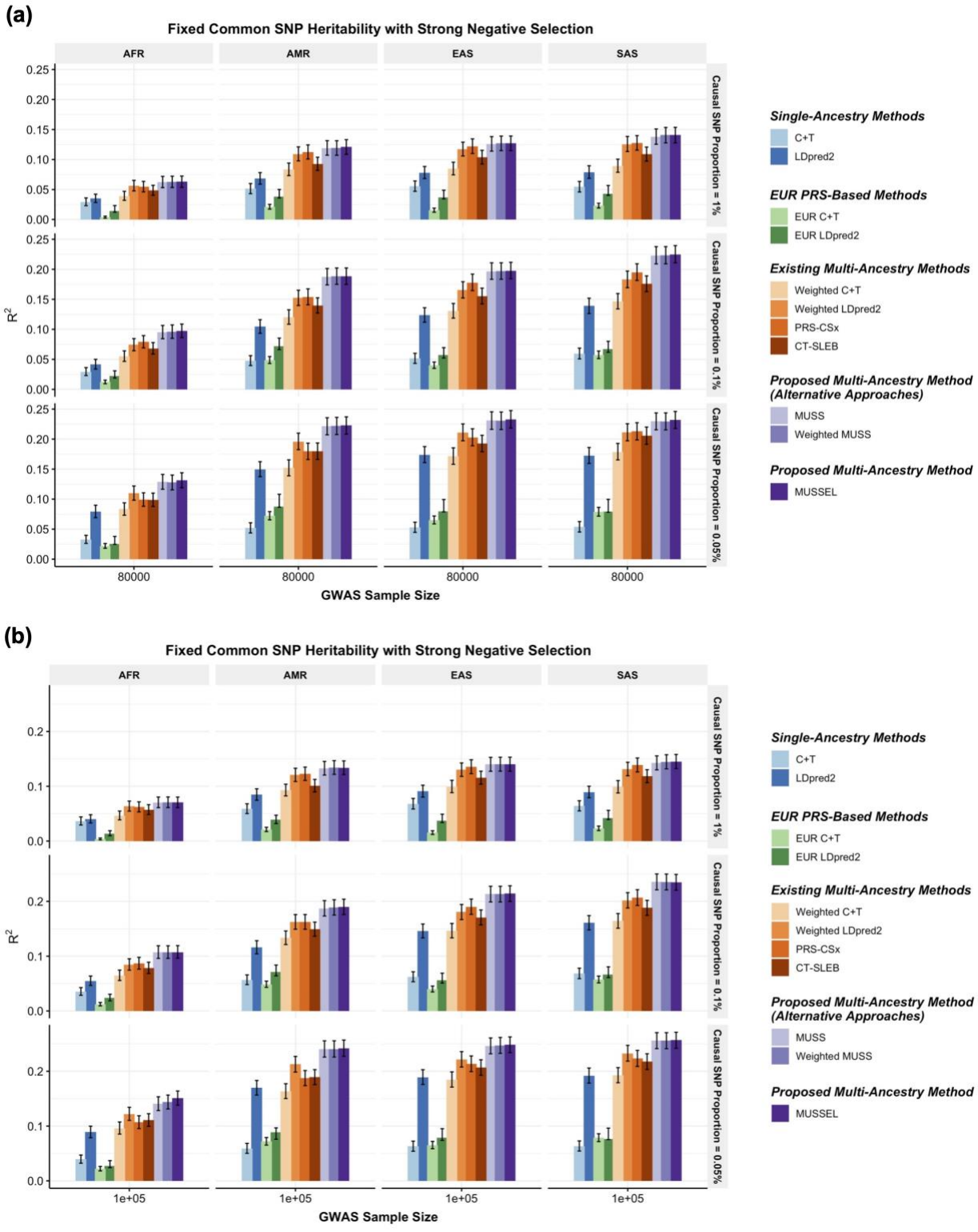

(a)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.25
0.20
0.15
0.10
0.05
0.00
80000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.2
0.1
0.0
1e+05
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S3: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed per-SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 15,000/45,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

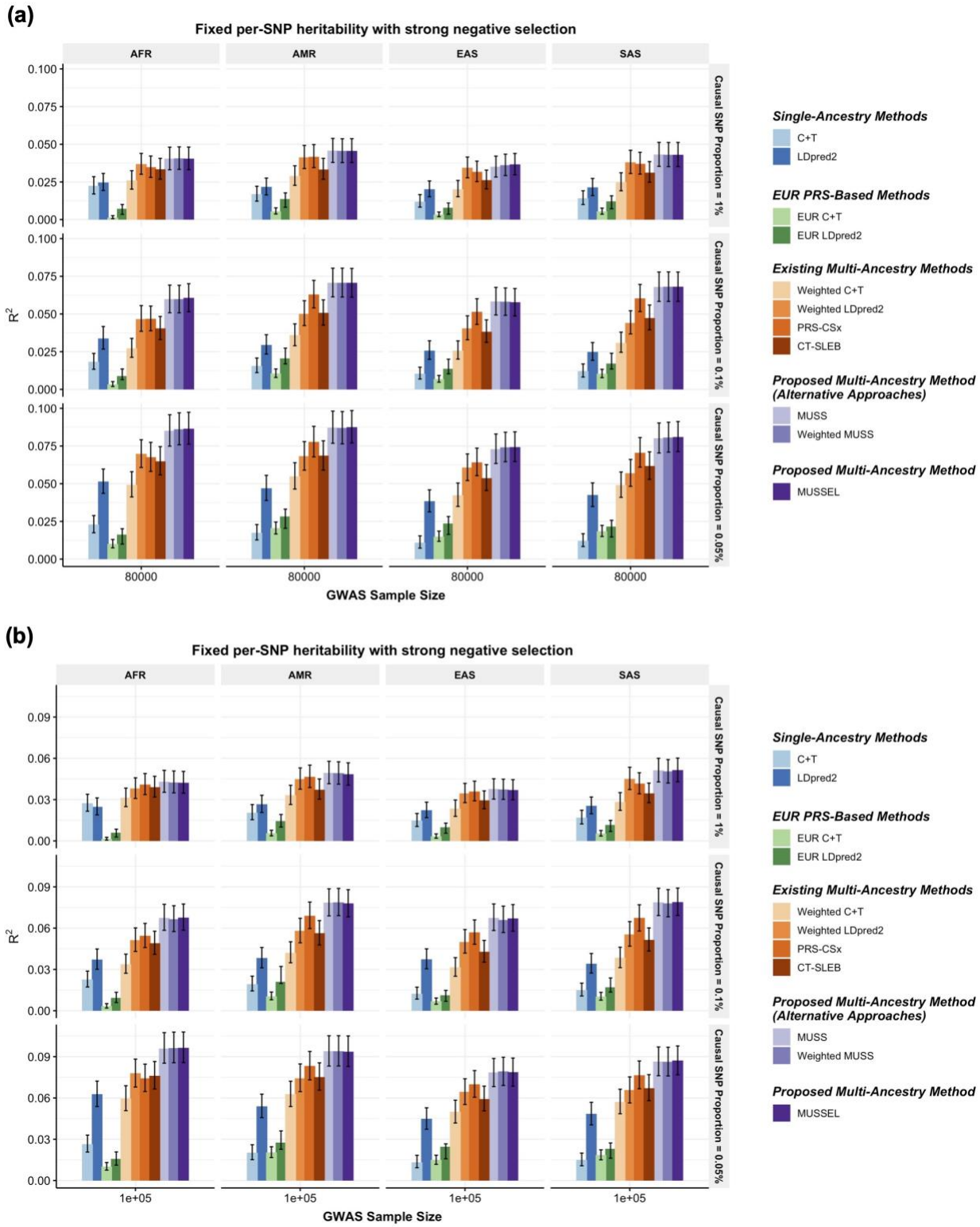

(a)
Fixed per-SNP heritability with strong negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.100
0.075
0.050
0.025
0.000
80000
GWAS Sample Size
(b)
Fixed per-SNP heritability with strong negative selection
0.09
0.06
0.03
0.00
1e+05
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S4: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed per-SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 80,000/100,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

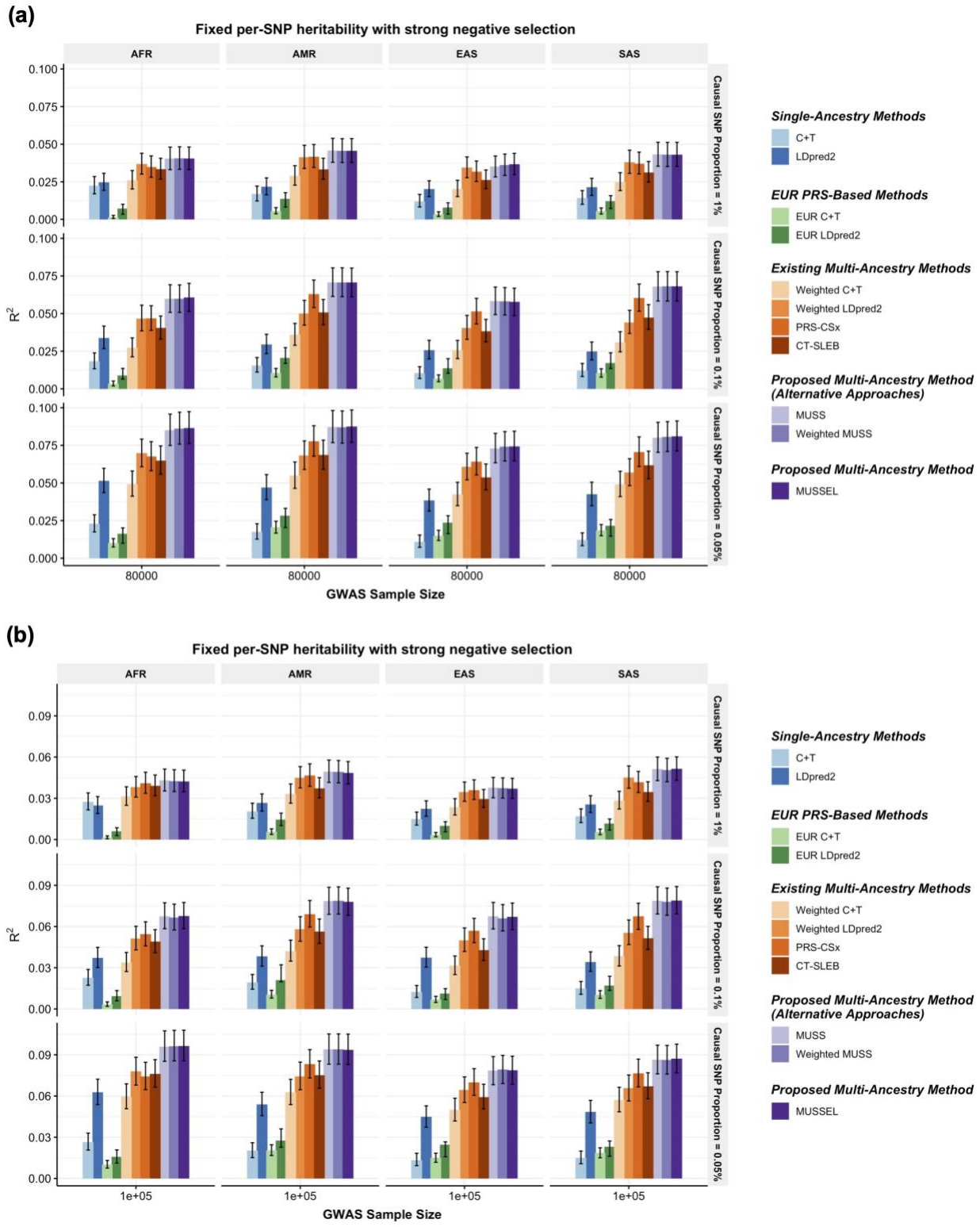

(a)
Fixed per-SNP heritability with strong negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.100
0.075
0.050
0.025
0.000
80000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed per-SNP heritability with strong negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.09
0.06
0.03
0.00
1e+05
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S5: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed per-SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency but with weaker cross-population (0.6 across all pairs of populations), with a GWAS sample size of 15,000/45,000 for each non-EUR population, related to Figure 2. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

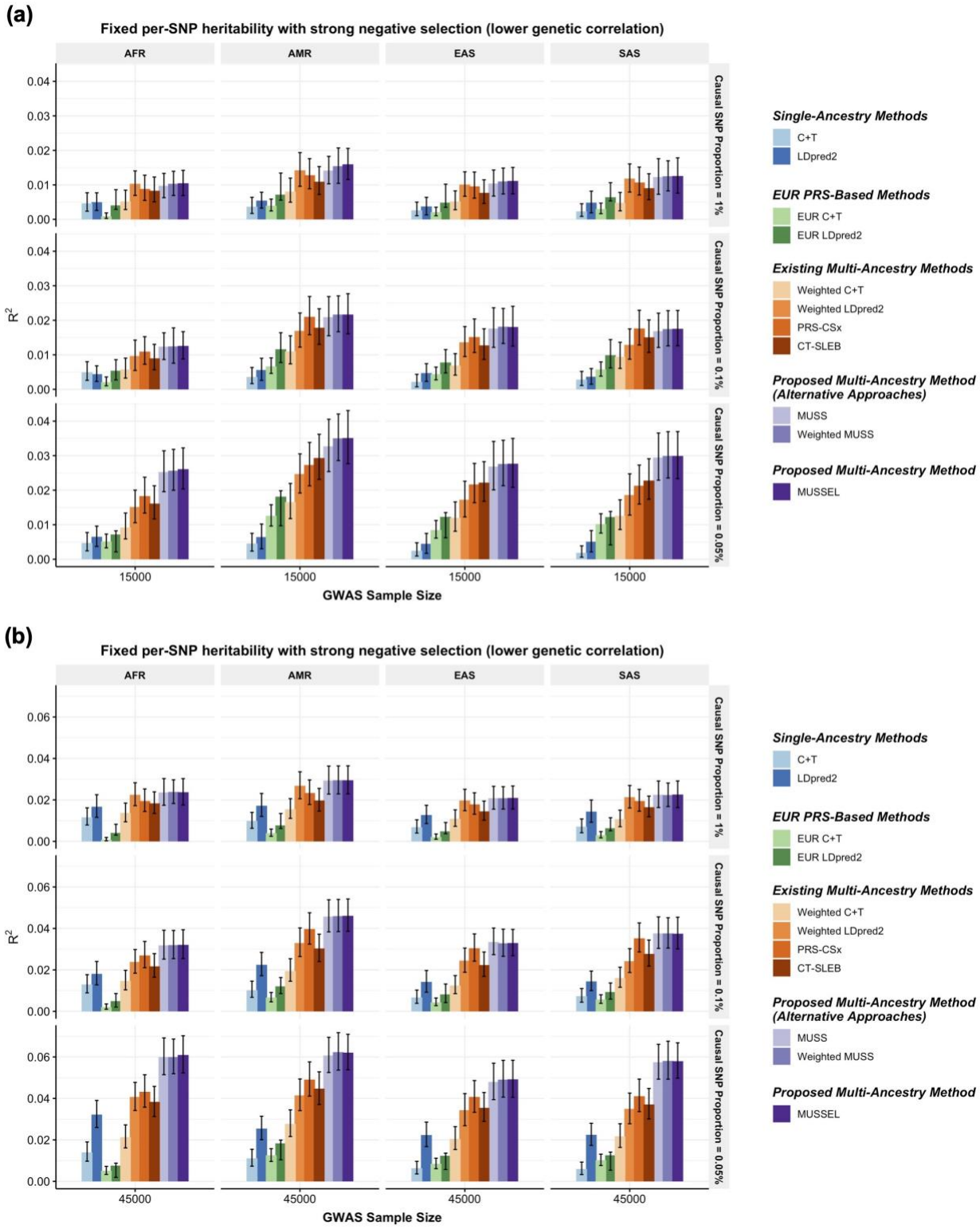

(a)
Fixed per-SNP heritability with strong negative selection (lower genetic correlation)
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R^2
0.04
0.03
0.02
0.01
0.00
15000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed per-SNP heritability with strong negative selection (lower genetic correlation)
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R^2
0.06
0.04
0.02
0.00
45000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S6: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed per-SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency but with weaker cross-population (0.6 across all pairs of populations), with a GWAS sample size of 80,000/100,000 for each non-EUR population, related to Figure 2. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~192K, 19.2K, or 9.6K causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

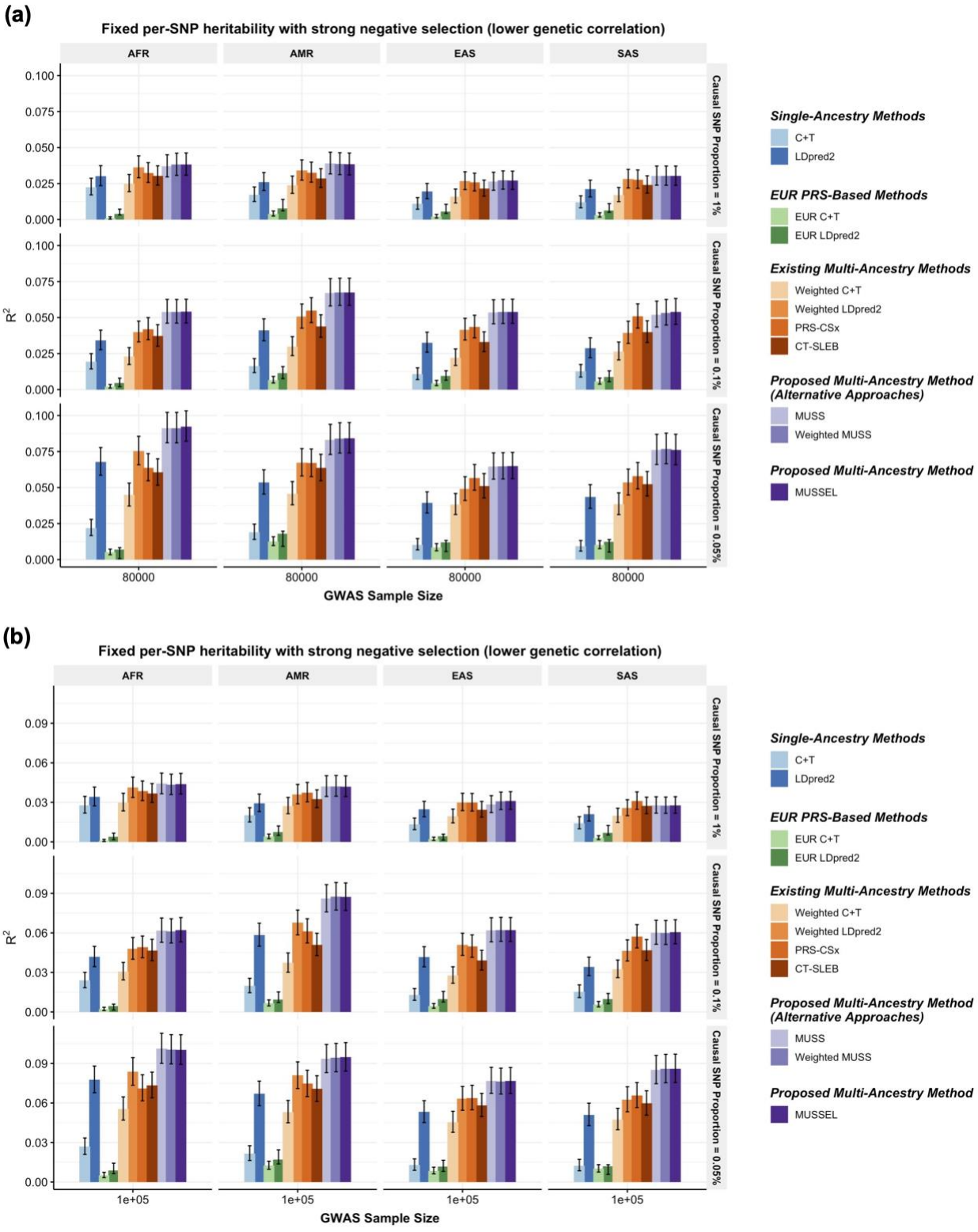

(a)
Fixed per-SNP heritability with strong negative selection (lower genetic correlation)
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.100
0.075
0.050
0.025
0.000
80000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed per-SNP heritability with strong negative selection (lower genetic correlation)
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.09
0.06
0.03
0.00
1e+05
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S7: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries with no negative selection for the relationship between SNP effect size and allele frequency with a GWAS sample size of 15,000/45,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

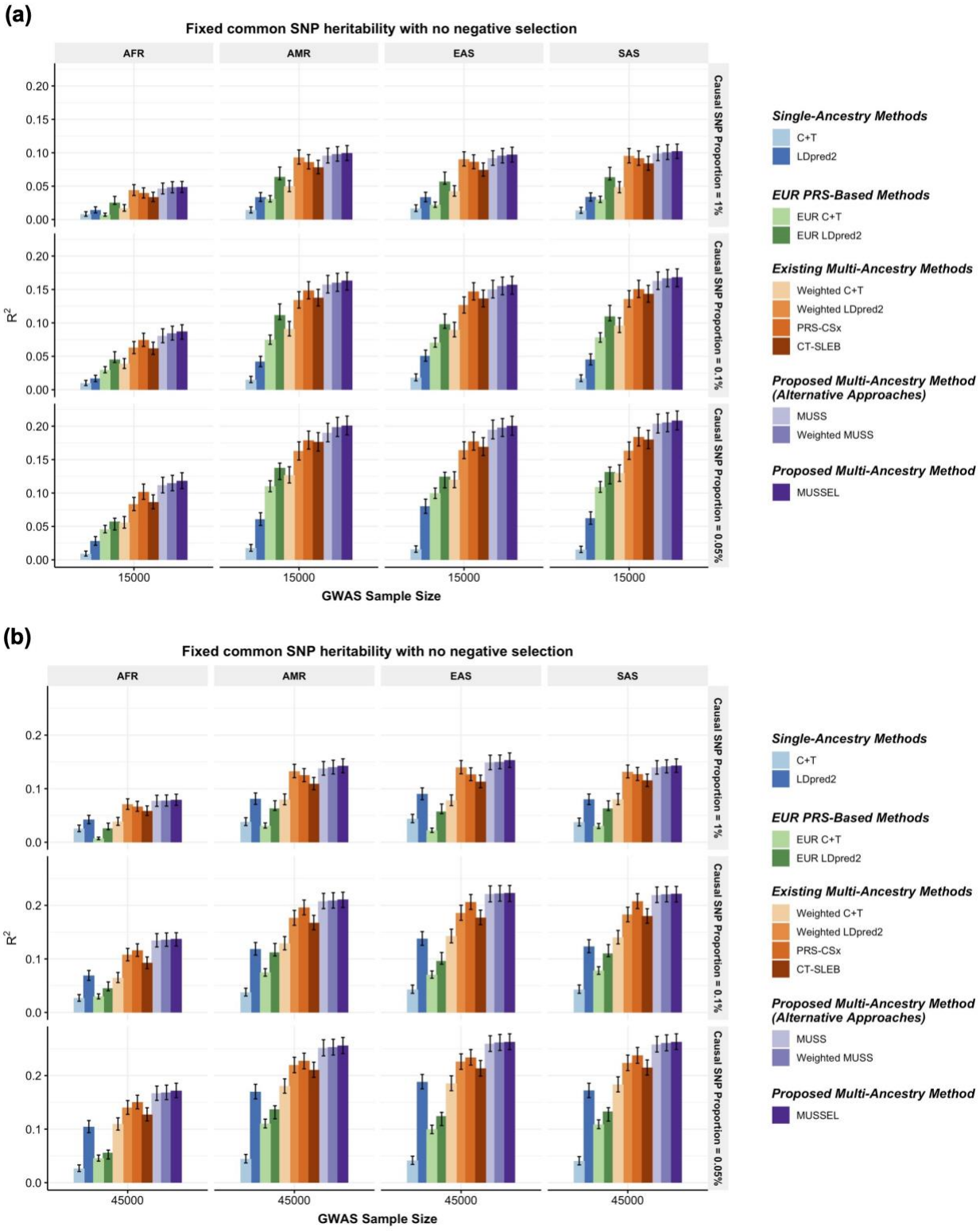

(a)
Fixed common SNP heritability with no negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.00
0.05
0.10
0.15
0.20
15000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed common SNP heritability with no negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
45000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S8: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries with no negative selection for the relationship between SNP effect size and allele frequency with a GWAS sample size of 80,000/100,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

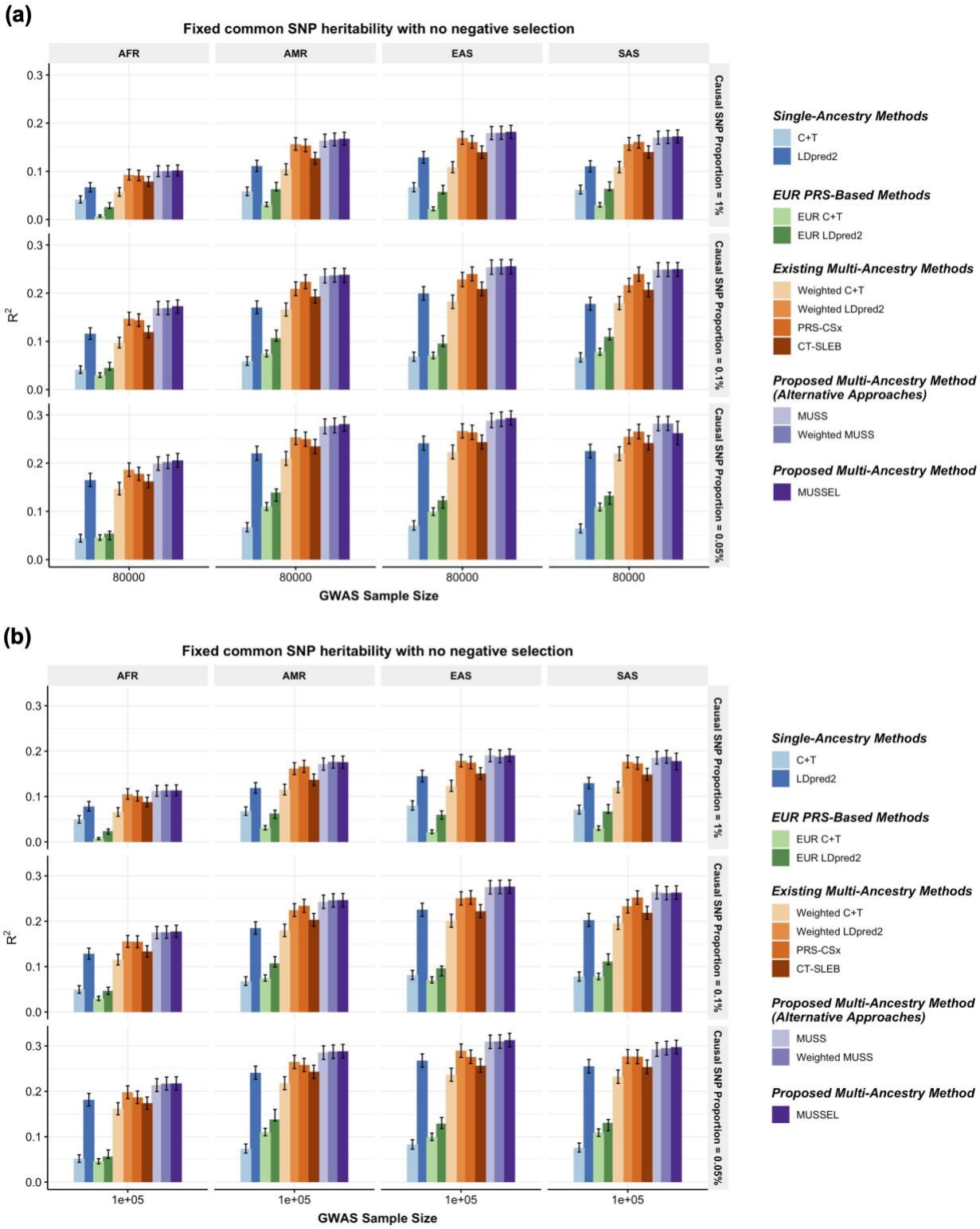

(a)
Fixed common SNP heritability with no negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
0.3
80000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed common SNP heritability with no negative selection
1e+05
GWAS Sample Size

Figure S9: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries under a mild negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 15,000/45,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

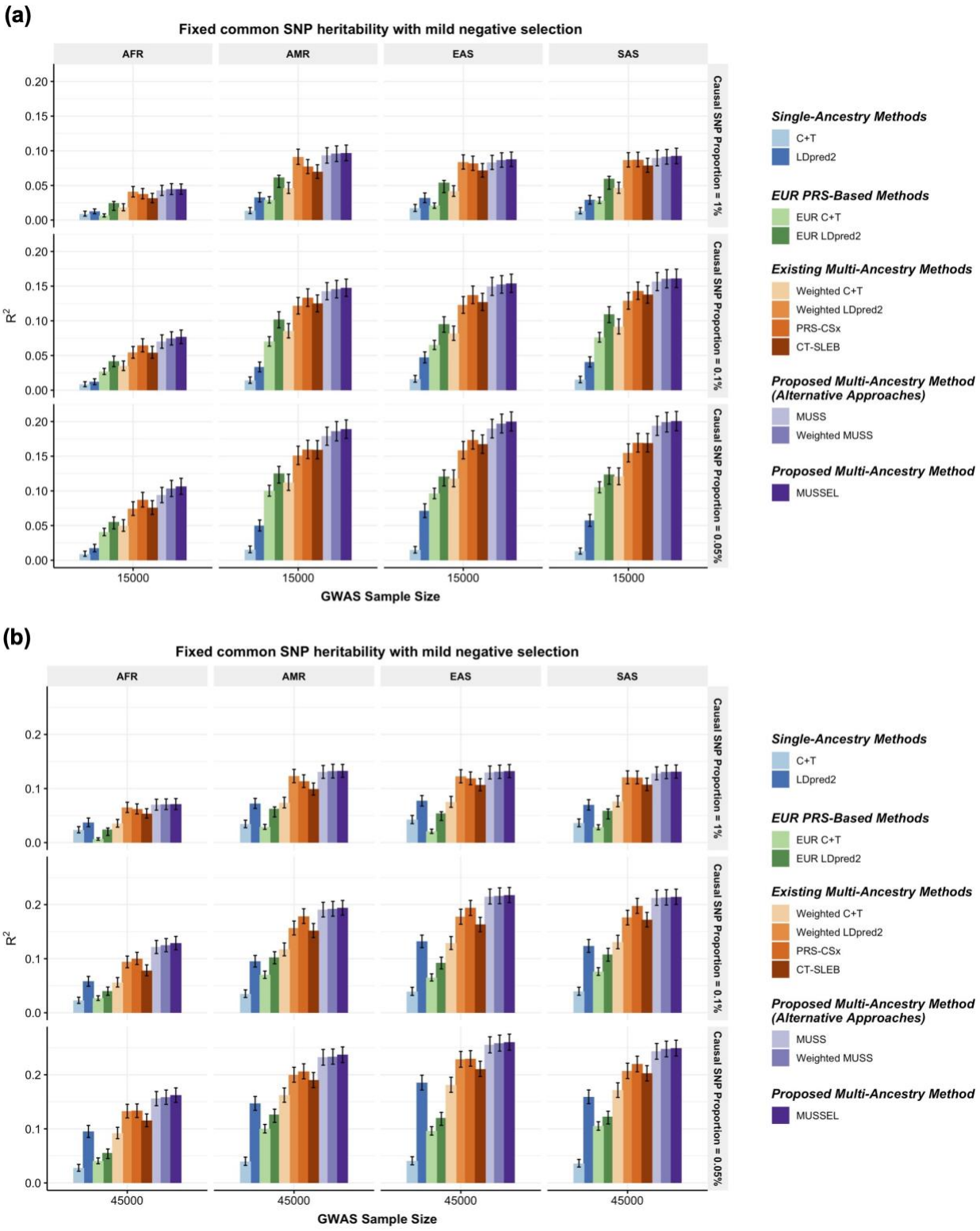

(a)
Fixed common SNP heritability with mild negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.00
0.05
0.10
0.15
0.20
15000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed common SNP heritability with mild negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
45000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S10: Simulation results showing performance of the PRS constructed by MUSSEL and various existing methods, assuming a fixed common SNP heritability (0.4) across ancestries under a mild negative selection model for the relationship between SNP effect size and allele frequency with a GWAS sample size of 80,000/100,000 for each non-EUR population, related to Figure 2. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestries but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group. The corresponding 95% bootstrap CIs are obtained from the same testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot".

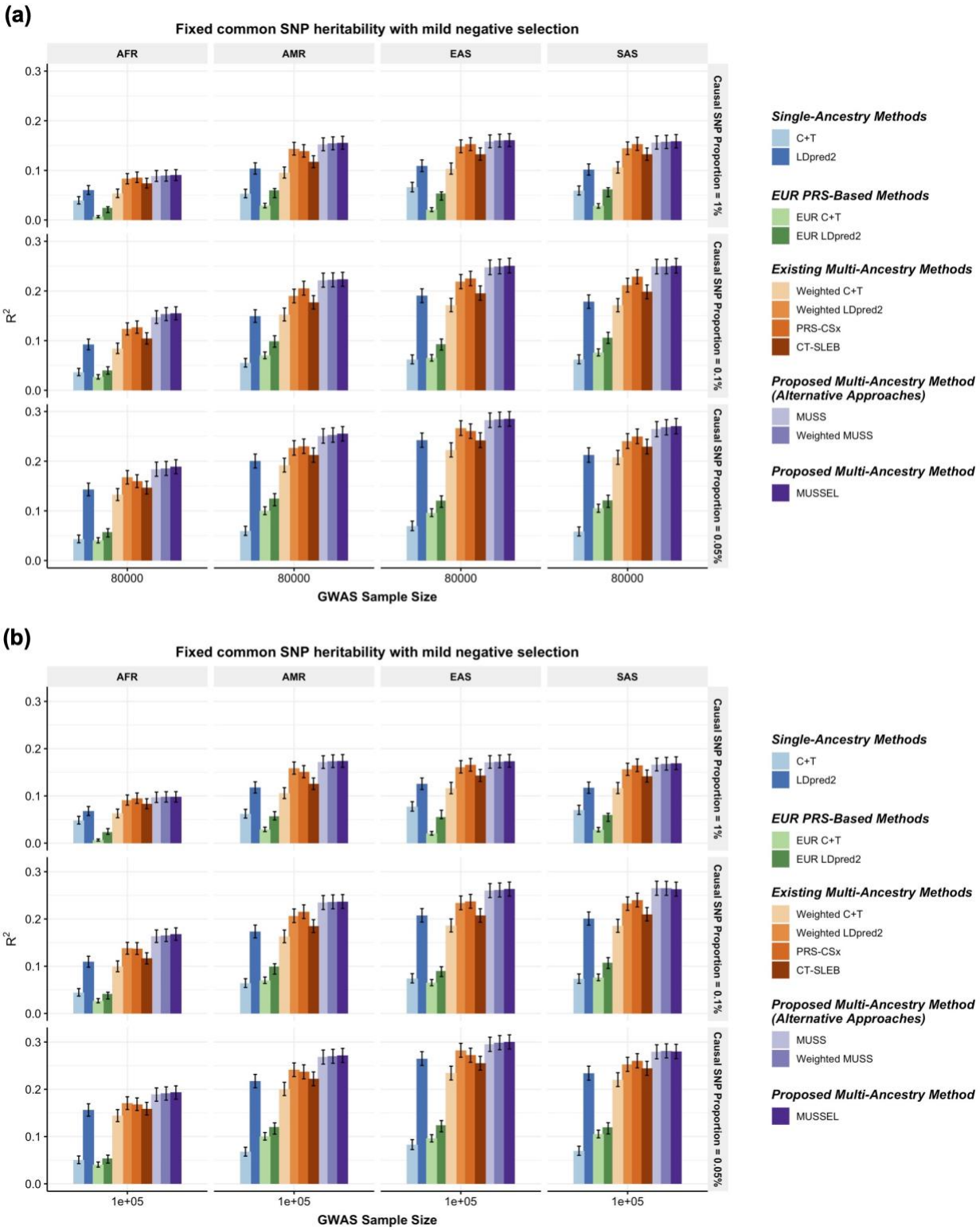

(a)
Fixed common SNP heritability with mild negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
0.3
80000
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL
(b)
Fixed common SNP heritability with mild negative selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
0.3
1e+05
GWAS Sample Size
Single-Ancestry Methods
C+T
LDpred2
EUR PRS-Based Methods
EUR C+T
EUR LDpred2
Existing Multi-Ancestry Methods
Weighted C+T
Weighted LDpred2
PRS-CSx
CT-SLEB
Proposed Multi-Ancestry Method (Alternative Approaches)
MUSS
Weighted MUSS
Proposed Multi-Ancestry Method
MUSSEL

Figure S11: Prediction R^2 with 95% bootstrap CIs on validation individuals of AFR (N=2,015–3,428), EAS (N=2,316-4,647), and AMR ancestries (N=3,479-4,397) in PAGE based on discovery GWAS from PAGE (AFR N_{GWAS}=7,775 – 13,699, AMR N_{GWAS}=13,894 – 17,558), BBJ (EAS N_{GWAS}=70,657 – 158,284), and UKBB (EUR N_{GWAS}=315,133 – 355,983), related to Figure 3. We used genotype data from 1000 Genomes Project (498 EUR, 659 AFR, 347 AMR, 503 EAS, 487 SAS) as the LD reference dataset. All methods were evaluated on the ~2.0 million SNPs that are available in HapMap 3 + MEGA, except for PRS-CSx which is evaluated based on the HapMap 3 SNPs only, as implemented in their software. Ancestry- and trait-specific GWAS sample sizes, number of SNPs included, and validation sample sizes are summarized in Table S7. A random half of the validation individuals is used as the tuning set to tune model parameters, as well as train the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The other half of the validation set is used as the testing set to report R^2 values for PRS on each ancestry, after adjusting for whether or not the sample is from BioMe and the top 10 genetic principal components for BMI, and additionally the age at lipid measurement and sex. The 95% bootstrap CIs of the estimated R^2 are obtained from the testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot". Detailed 95% bootstrap CIs are reported in Table S17.

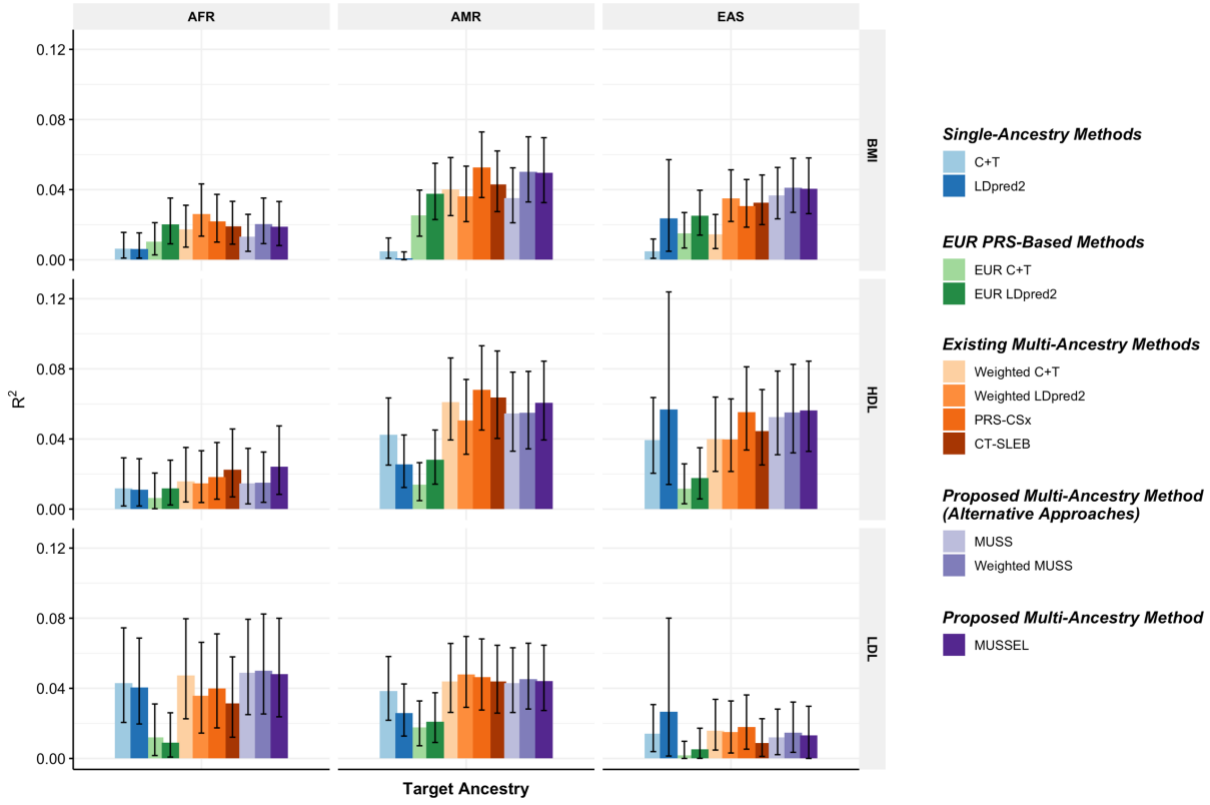


Figure S12: Prediction R^2 with 95% bootstrap CIs on UKBB validation individuals of EUR (17,457 – 19,030), AFR (7,954 – 8,598), EAS (1,752 – 1,921), and SAS (9,385 – 10,288) origin based on discovery GWAS from GLGC on EUR (N_{GWAS} =842,660 – 930,671), AFR or admixed AFR (N_{GWAS} =87,760 – 92,555), Hispanic/Latino (N_{GWAS} =46,040 – 49,582), EAS (N_{GWAS} =82,587 – 146,492), and SAS (N_{GWAS} =33,658 – 34,135). EUR (N_{GWAS} =842,660 – 930,671), AFR or admixed AFR (N_{GWAS} =87,760 – 92,555), Hispanic/Latino (N_{GWAS} =46,040 – 49,582), EAS (N_{GWAS} =82,587 – 146,492), and SAS (N_{GWAS} =33,658 – 34,135), related to Figure 4. The LD reference data is either (a) 1000 Genomes Project (498 EUR, 659 AFR, 347 AMR, 503 EAS, 487 SAS), or (b) UKBB data (PRS-CSx: default UKBB LD reference data which overlap with our testing samples including 375,120 EUR, 7,507 AFR, 687 AMR, 2,181 EAS, and 8,412 SAS; all other methods: UKBB tuning samples including 10,000 EUR, 4,585 AFR, 1,010 EAS, and 5,427 SAS). The ancestry of UKBB individuals were determined by a genetic ancestry prediction approach (Supplementary Notes). Due to the low prediction accuracy of genetic component analysis and extremely small validation sample size of UKBB AMR, prediction R^2 on UKBB AMR is unreliable and thus is not reported here. All methods were evaluated on the ~2.0 million SNPs that are available in HapMap 3 + MEGA, except for PRS-CSx which is evaluated based on the HapMap 3 SNPs only, as implemented in their software. Ancestry- and trait-specific GWAS sample sizes, number of SNPs included, and validation sample sizes are summarized in Table S9. A random half of the validation individuals is used as the tuning set to tune model parameters, as well as train the SL in CT-SLEB and MUSSEL or the linear combination model in weighted LDpred2, PRS-CSx, and weighted MUSS. The other half of the validation set is used as the testing set to report R^2 values for each ancestry. The 95% bootstrap CIs of the estimated R^2 are obtained from the testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot". Detailed 95% bootstrap CIs are reported in Table S17. In (b), PRS-CSx and other methods do not have a fair comparison because the UKBB LD reference data provided by the PRS-CSx software ($UKBB_{PRS\text{-}CSx}$) is much larger than that for other methods, and thus the R^2 of PRS-CSx PRS may be inflated due to a big overlap between $UKBB_{PRS\text{-}CSx}$ and the UKBB testing sample.

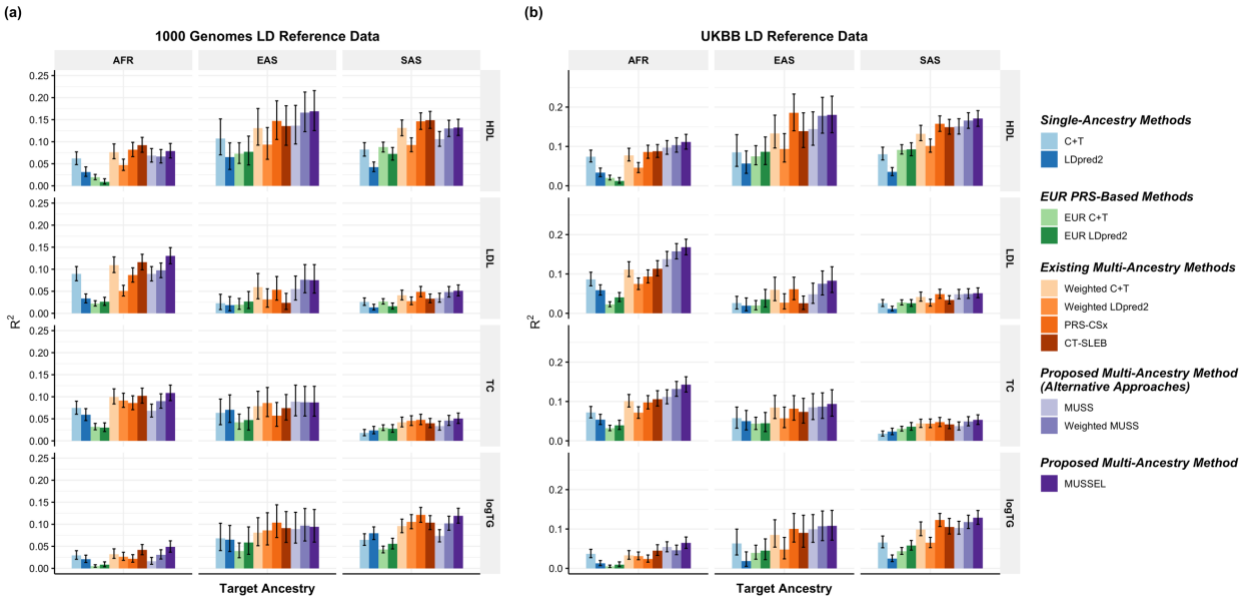


Figure S13: Prediction R^2 with 95% bootstrap CIs on UKBB validation individuals of AFR (N=9,169) origin based on discovery GWAS from AoU on EUR (N_{GWAS} =48,229 – 48,332), AFR (N_{GWAS} =21,514 – 21,550), and Hispanic/Latino (N_{GWAS} =15,364 – 15,413), related to Figure 5. The LD reference data is either **(a)** 1000 Genomes Project (498 EUR, 659 AFR, 347 AMR, 503 EAS, 487 SAS), or **(b)** UKBB data (PRS-CSx: default UKBB LD reference data which overlap with our testing samples including 375,120 EUR, 7,507 AFR, 687 AMR, 2,181 EAS, and 8,412 SAS; all other methods: UKBB tuning samples including 10,000 EUR, 4,585 AFR, 1,010 EAS, and 5,427 SAS). The ancestry of UKBB individuals were determined by a genetic ancestry prediction approach (Supplementary Notes). Due to the low prediction accuracy of genetic component analysis and extremely small validation sample size of UKBB AMR, prediction R^2 on UKBB AMR is unreliable and thus is not reported here. All methods were evaluated on the ~2.0 million SNPs that are available in HapMap3 + MEGA, except for PRS-CSx which is evaluated based on the HapMap 3 SNPs only, as implemented in their software. Ancestry- and trait-specific sample sizes of GWAS, number of SNPs included, and validation sample sizes are summarized in Table S11. A random half of the validation individuals is used as the tuning set to tune model parameters, as well as train the SL in CT-SLEB and MUSSEL or the linear combination model in weighted LDpred2, PRS-CSx, and weighted MUSS. The other half of the validation set is used as the testing set to report R^2 values for each ancestry, after adjusting for age, sex, and the top 10 genetic principal components. The 95% bootstrap CIs of the estimated R^2 are obtained from the testing set based on 10,000 bootstrap samples using the Bca approach[1] implemented in the R package "boot". Detailed 95% bootstrap CIs are reported in Table S17. In (b), PRS-CSx and other methods do not have a fair comparison because the UKBB LD reference data provided by the PRS-CSx software ($UKBB_{PRS\text{-}CSx}$) is much larger than that for other methods, and thus the R^2 of PRS-CSx may be inflated due to a big overlap between $UKBB_{PRS\text{-}CSx}$ and the UKBB testing sample.

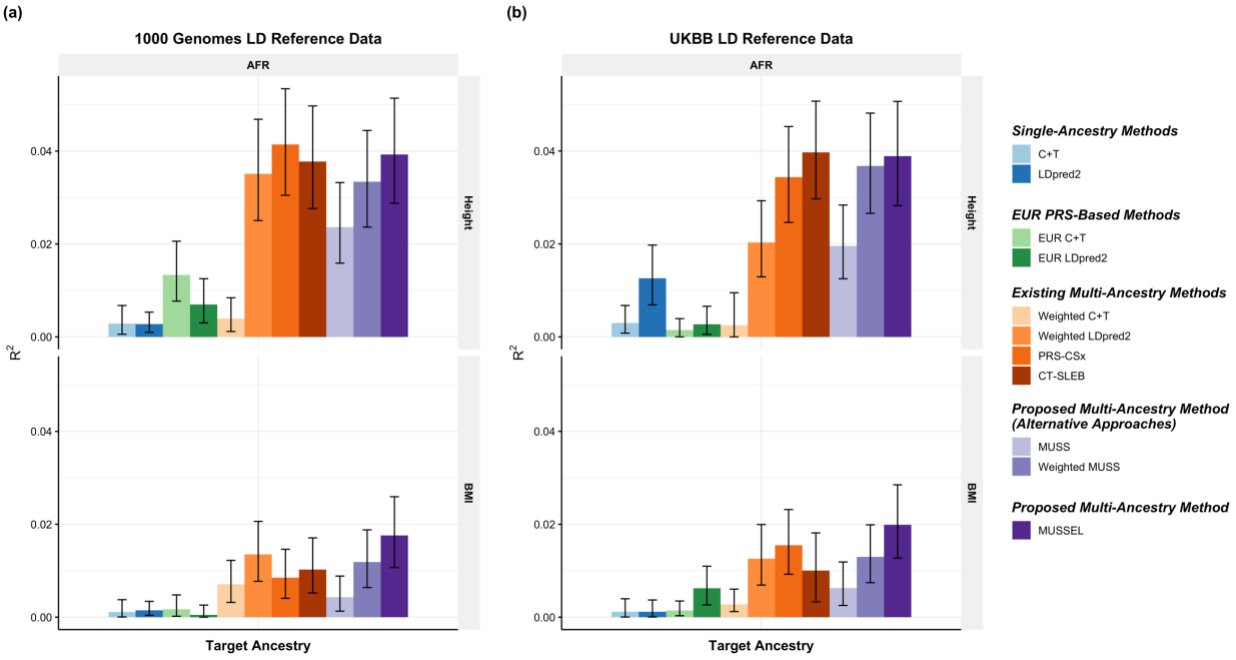


Figure S14: Simulation results with 20% ancestry mis-specification in the LD reference sample, assuming a fixed common SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency, related to Figure 2. The LD matrix for each ancestry group is estimated based on a slightly mis-specified LD reference sample that contains 800 individuals from the same ancestry group and 50 individuals from each of the other four ancestry groups, totaling 200 individuals with ancestry mismatch. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestry groups but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 15,000 or (b) 45,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group.

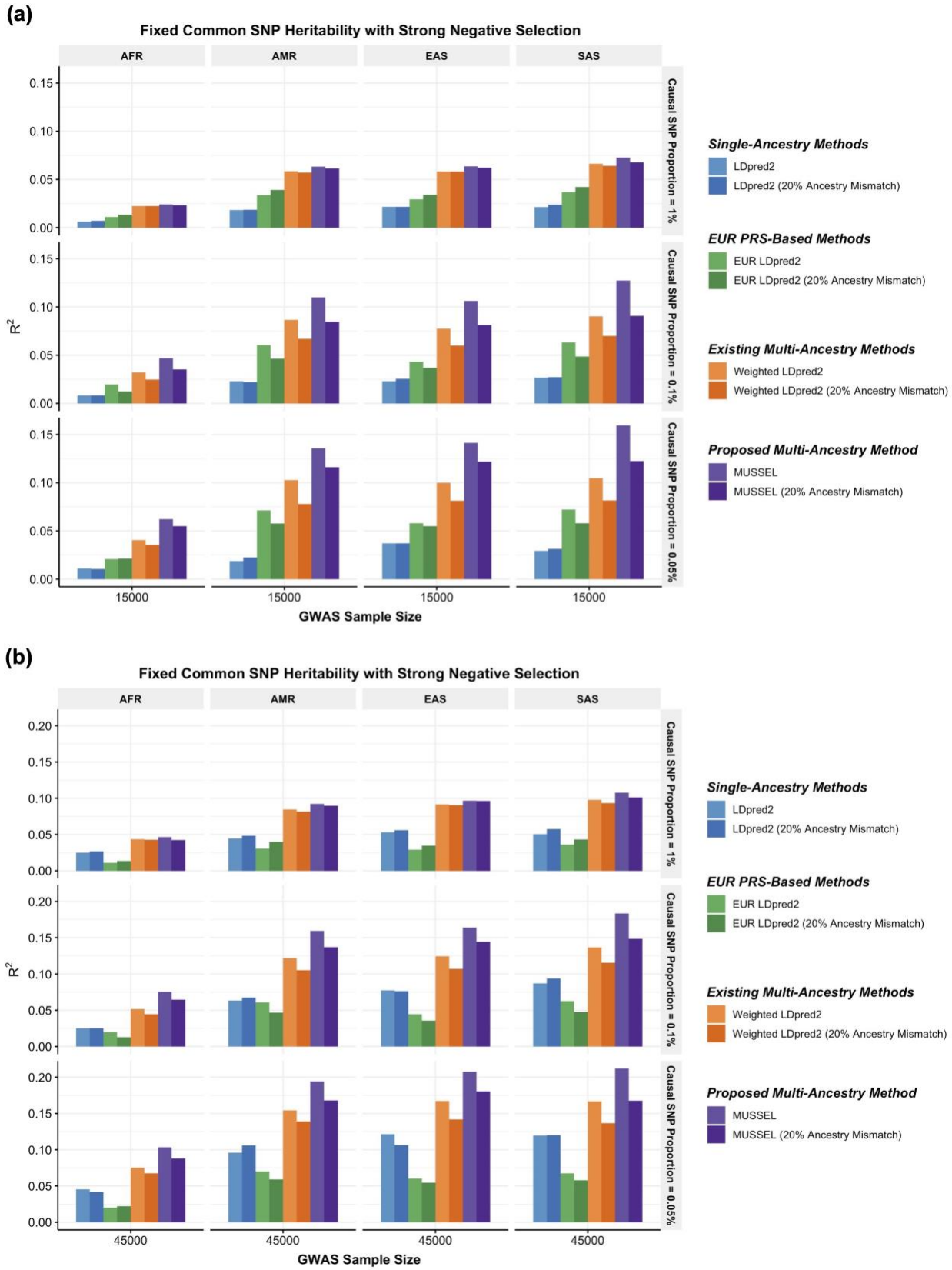
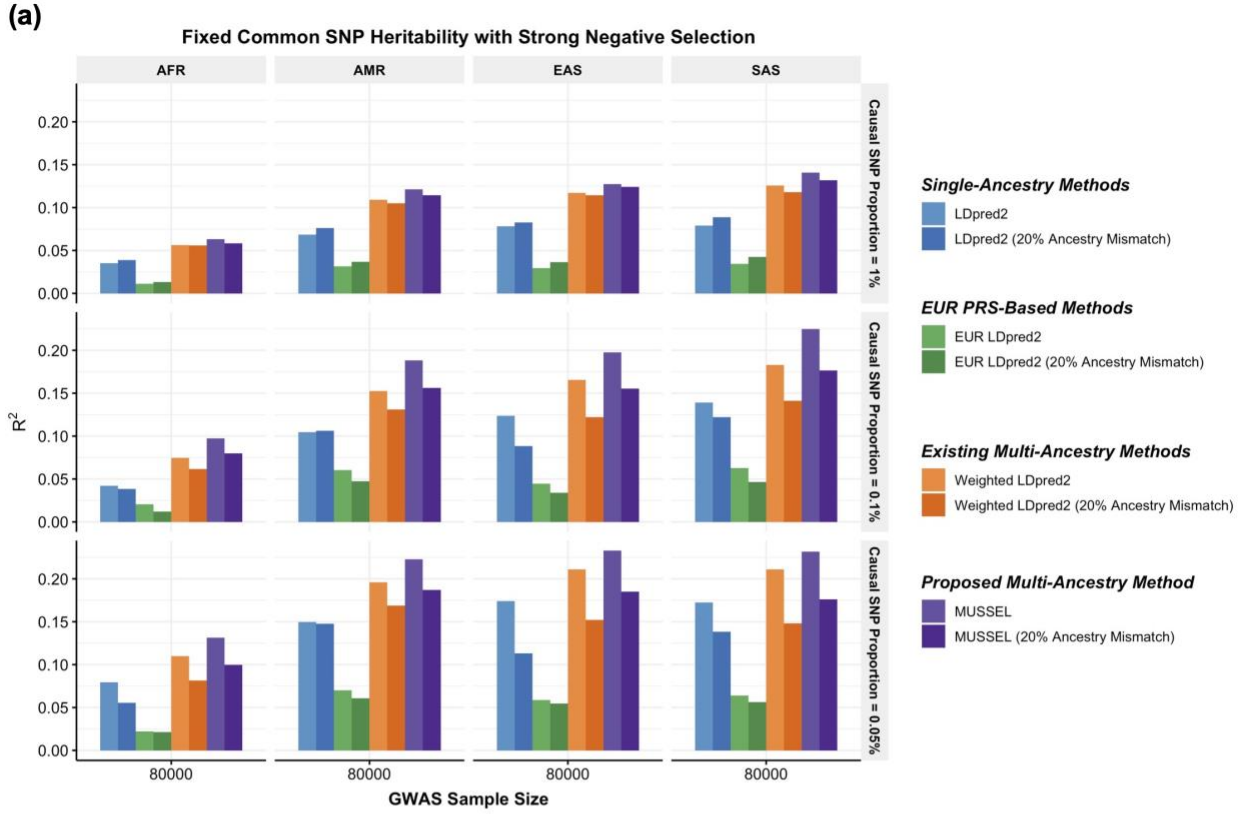

(a)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.15
0.10
0.05
0.00
15000
GWAS Sample Size
Single-Ancestry Methods
LDpred2
LDpred2 (20% Ancestry Mismatch)
EUR PRS-Based Methods
EUR LDpred2
EUR LDpred2 (20% Ancestry Mismatch)
Existing Multi-Ancestry Methods
Weighted LDpred2
Weighted LDpred2 (20% Ancestry Mismatch)
Proposed Multi-Ancestry Method
MUSSEL
MUSSEL (20% Ancestry Mismatch)
(b)
Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.20
0.15
0.10
0.05
0.00
45000
GWAS Sample Size
Single-Ancestry Methods
LDpred2
LDpred2 (20% Ancestry Mismatch)
EUR PRS-Based Methods
EUR LDpred2
EUR LDpred2 (20% Ancestry Mismatch)
Existing Multi-Ancestry Methods
Weighted LDpred2
Weighted LDpred2 (20% Ancestry Mismatch)
Proposed Multi-Ancestry Method
MUSSEL
MUSSEL (20% Ancestry Mismatch)

Figure S15: Simulation results with 20% ancestry mis-specification in the LD reference sample, assuming a fixed common SNP heritability (0.4) across ancestries under a strong negative selection model for the relationship between SNP effect size and allele frequency, related to Figure 2. The LD matrix for each ancestry group is estimated based on a slightly mis-specified LD reference sample that contains 800 individuals from the same ancestry group and 50 individuals from each of the other four ancestry groups, totaling 200 individuals with ancestry mismatch. The genetic correlation in SNP effect size is set to 0.8 across all pairs of populations. The causal SNP proportion (degree of polygenicity) is set to 1.0%, 0.1%, or 0.05% (~$192K$, $19.2K$, or $9.6K$ causal SNPs). We generate data for ~19 million common SNPs (MAF≥1%) across the five ancestry groups but conduct analyses only on the ~2.0 million SNPs in HapMap 3 + MEGA. The discovery GWAS sample size is set to **(a) 80,000 or (b) 100,000** for each non-EUR ancestry, and 100,000 for EUR. A tuning set consisting of 10,000 individuals is used for parameter tuning, as well as training the SL in CT-SLEB and MUSSEL or the linear combination model in weighted C+T, weighted LDpred2, PRS-CSx, and weighted MUSS. The reported R^2 values are calculated on an independent testing set of 10,000 individuals for each ancestry group.

(a)


Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.00
0.05
0.10
0.15
0.20
80000
GWAS Sample Size
Single-Ancestry Methods
LDpred2
LDpred2 (20% Ancestry Mismatch)
EUR PRS-Based Methods
EUR LDpred2
EUR LDpred2 (20% Ancestry Mismatch)
Existing Multi-Ancestry Methods
Weighted LDpred2
Weighted LDpred2 (20% Ancestry Mismatch)
Proposed Multi-Ancestry Method
MUSSEL
MUSSEL (20% Ancestry Mismatch)

(b)

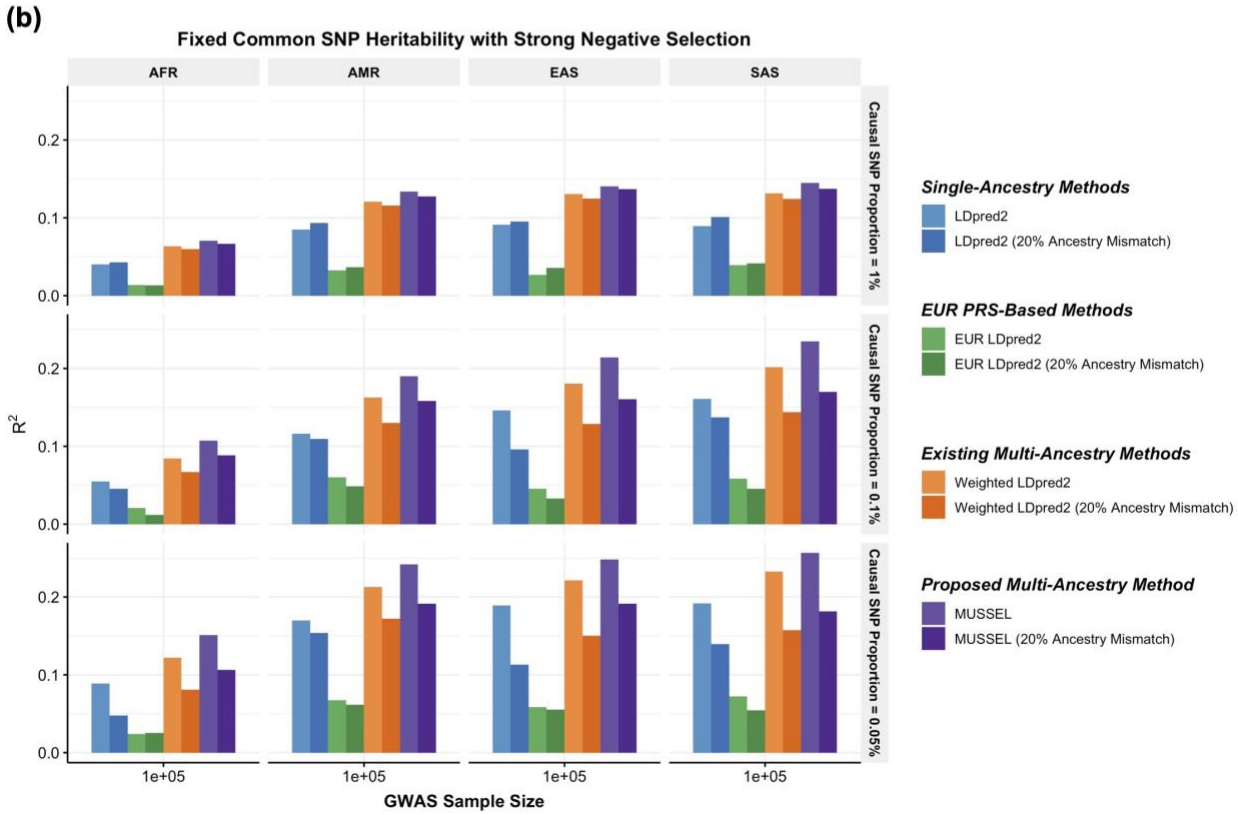

Fixed Common SNP Heritability with Strong Negative Selection
AFR
AMR
EAS
SAS
Causal SNP Proportion = 1%
Causal SNP Proportion = 0.1%
Causal SNP Proportion = 0.05%
R2
0.0
0.1
0.2
1e+05
GWAS Sample Size
Single-Ancestry Methods
LDpred2
LDpred2 (20% Ancestry Mismatch)
EUR PRS-Based Methods
EUR LDpred2
EUR LDpred2 (20% Ancestry Mismatch)
Existing Multi-Ancestry Methods
Weighted LDpred2
Weighted LDpred2 (20% Ancestry Mismatch)
Proposed Multi-Ancestry Method
MUSSEL
MUSSEL (20% Ancestry Mismatch)

Figure S16: Average R^2 and the 95% bootstrap CIs calculated by ancestry group across all traits available in the PAGE + UKBB + BBJ, GLGC, and AoU data analyses, related to Figures 3 - 6. For EUR and AFR, the calculations were conducted across all nine traits in the three data analyses; for EAS and SAS, the calculations were conducted across the seven traits in the PAGE + UKBB + BBJ and GLGC analyses; and for AMR (Hispanic/Latino), the calculations were conducted across the three traits in the PAGE + UKBB + BBJ analysis. To account for the effect of validation sample size on R^2, we calculated the weighted average of R^2, where weights are proportional to the validation sample sizes for different traits. Detailed results are summarized in Table S17.

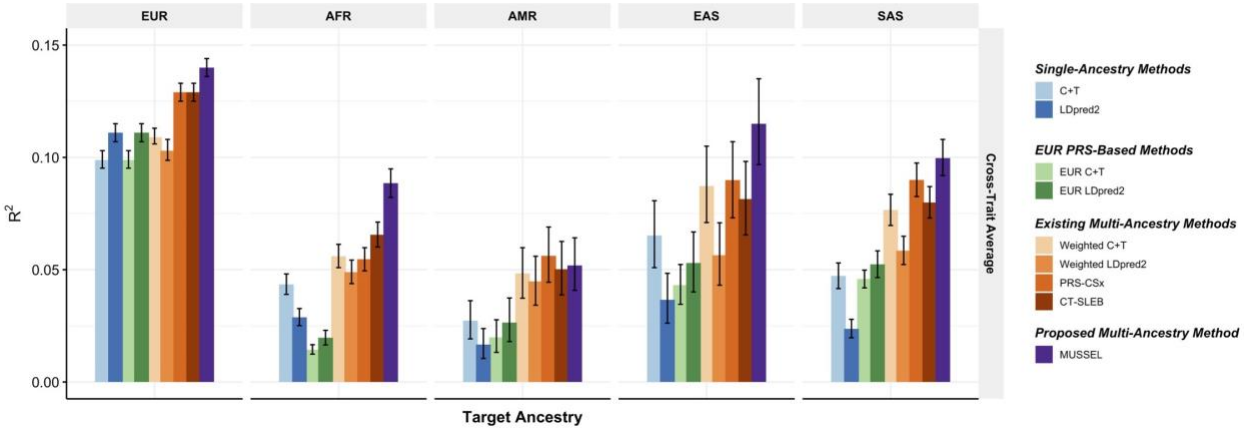
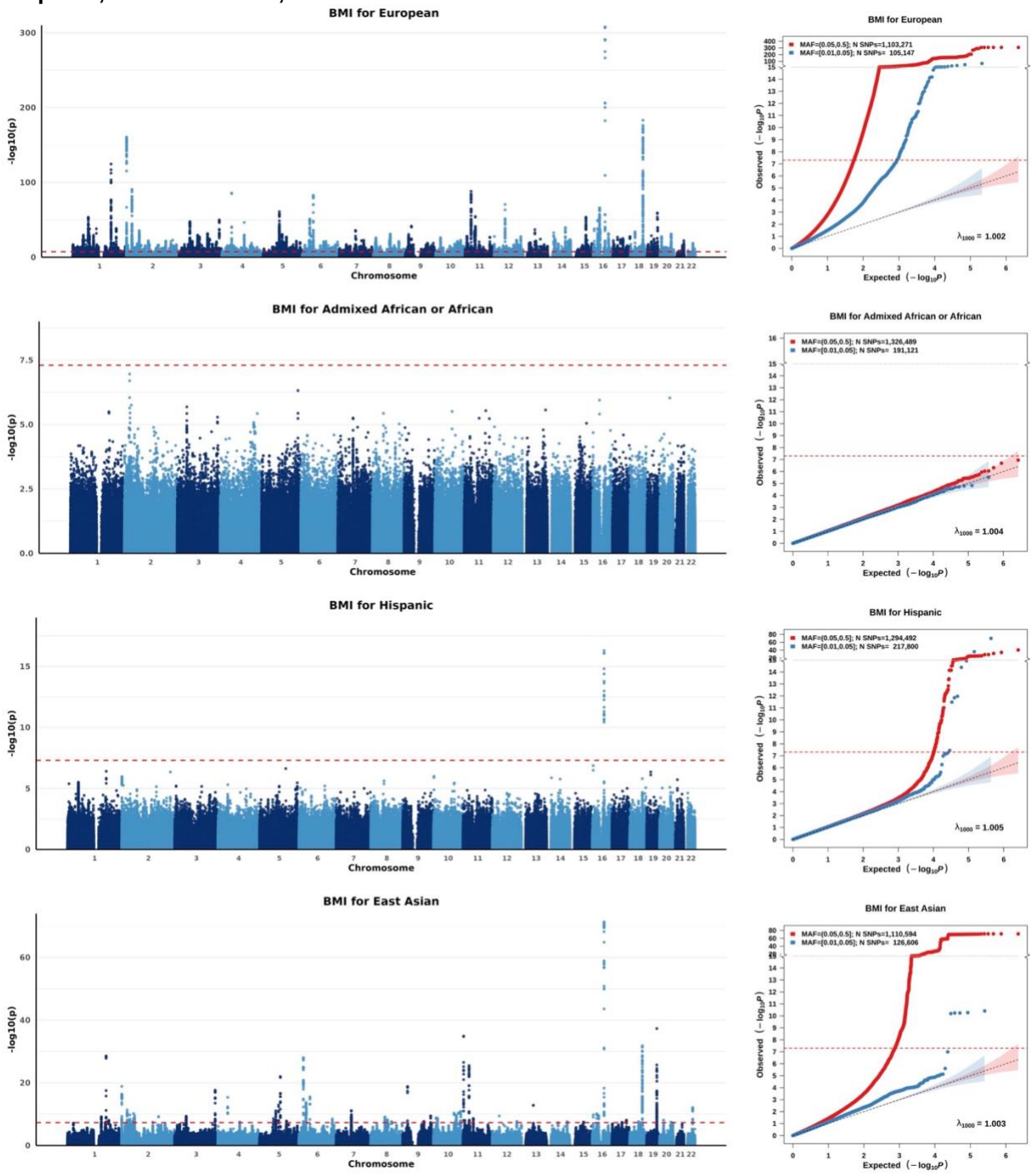
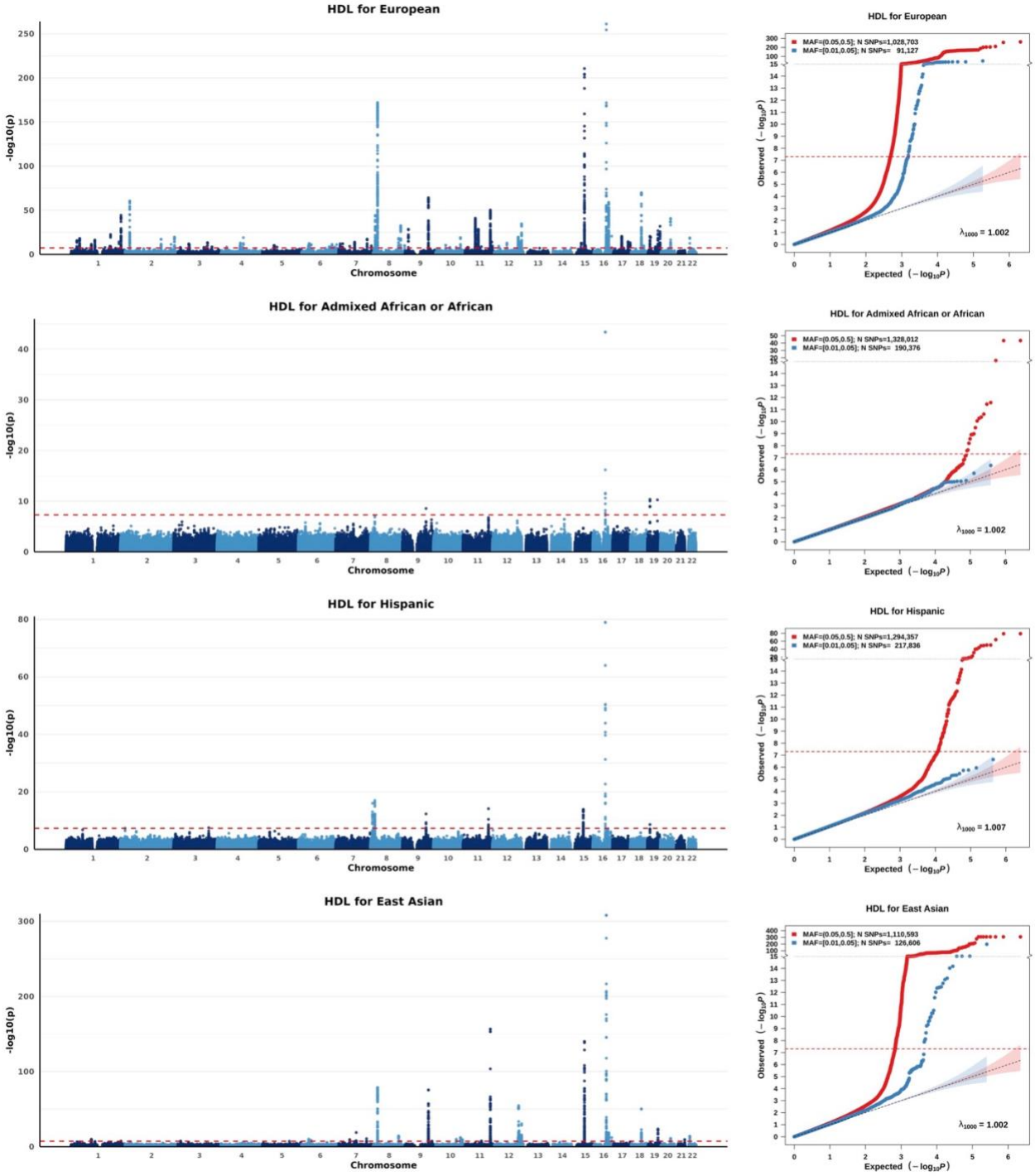


Figure S17: Manhattan plot and QQ plot[1] based on the GWAS summary-level association statistics from PAGE for BMI in four populations: European, Admixed African or African, Hispanic, and East Asian, related to STAR Methods.



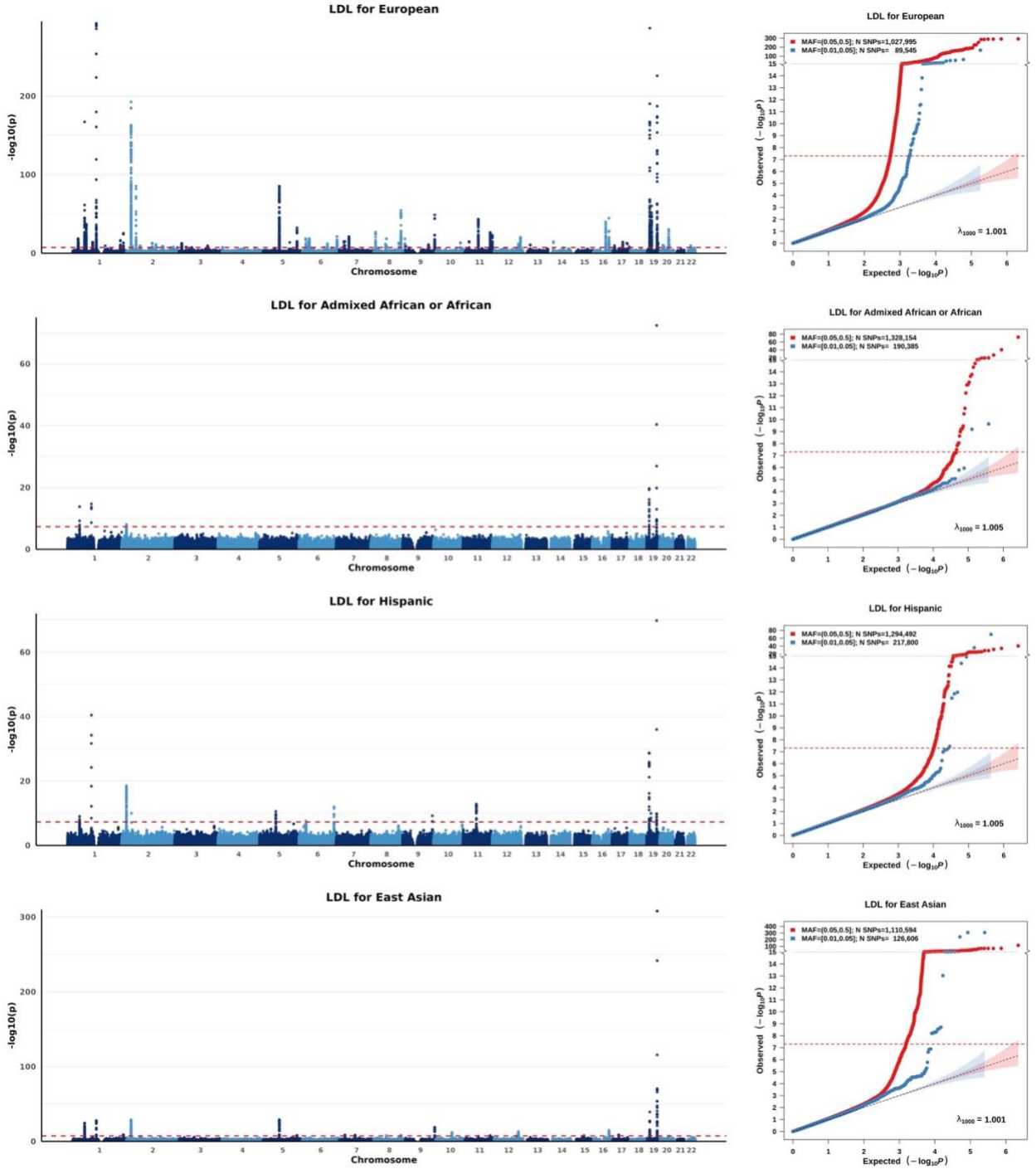
[1] For continuous traits, λ_{1000} scales the genomic inflation factor λ to a study with 1000 subjects using $\lambda_{1000} = 1 + 1000(\lambda - 1)/N$, where N is the total sample size. For binary traits, λ_{1000} scales λ to a study with 1000 cases and 1000 controls using $\lambda_{1000} = 1 + 1000(\lambda - 1)(\frac{1}{N_{case}} + \frac{1}{N_{control}})$.

Figure S18: Manhattan plot and QQ plot[1] based on the GWAS summary-level association statistics from PAGE for high-density lipoprotein (HDL) in four populations: European, Admixed African or African, Hispanic, and East Asian, related to STAR Methods.



[1] For continuous traits, λ_{1000} scales the genomic inflation factor λ to a study with 1000 subjects using $\lambda_{1000} = 1 + 1000(\lambda - 1)/N$, where N is the total sample size. For binary traits, λ_{1000} scales λ to a study with 1000 cases and 1000 controls using $\lambda_{1000} = 1 + 1000(\lambda - 1)(\frac{1}{N_{case}} + \frac{1}{N_{control}})$.

Figure S19: Manhattan plot and QQ plot[1] based on the GWAS summary-level association statistics from PAGE for low-density lipoprotein (LDL) in four populations: European, Admixed African or African, Hispanic, and East Asian, related to STAR Methods.



[1] For continuous traits, λ_{1000} scales the genomic inflation factor λ to a study with 1000 subjects using $\lambda_{1000} = 1 + 1000(\lambda - 1)/N$, where N is the total sample size. For binary traits, λ_{1000} scales λ to a study with 1000 cases and 1000 controls using $\lambda_{1000} = 1 + 1000(\lambda - 1)(\frac{1}{N_{case}} + \frac{1}{N_{control}})$.

Supplemental References

1. DiCiccio TJ, Efron B. Bootstrap Confidence Intervals. Statistical Science 1996;11:189-212.
2. Zhang, H., Zhan, J., Jin, J., Zhang, J., Ahearn, T.U., Yu, Z., O'Connell, J., Jiang, Y., Chen, T., Garcia-Closas, M., et al. (2023). A new method for multiancestry polygenic prediction improves performance across diverse populations. Nature Genetics 55(10), 1757-1768. 10.1038/s41588-023-01501-z.