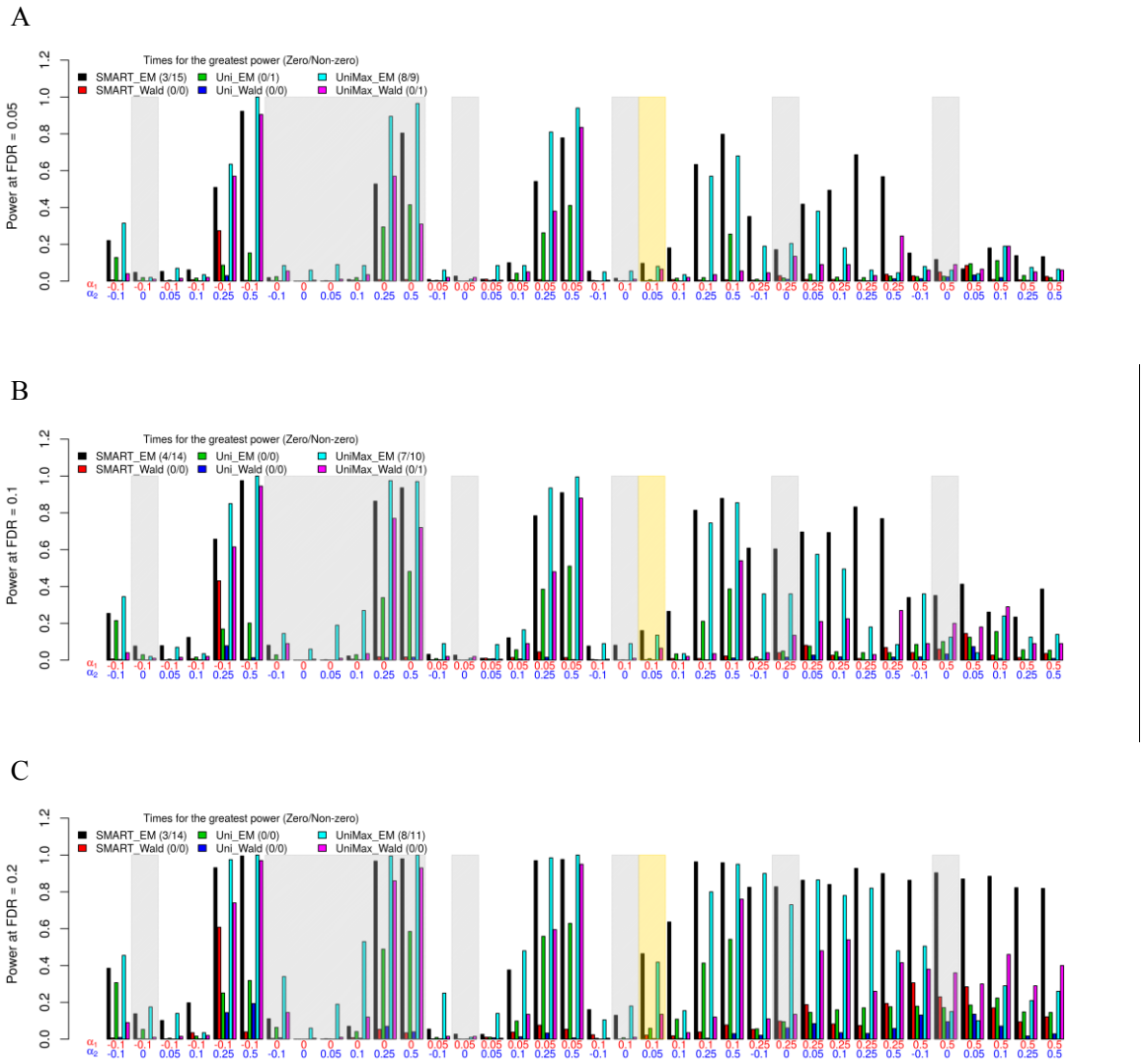


Figure S1 Simulation results for comparing using multiple annotations versus using a single annotation. Power to detect trait-relevant tissues by different approaches in various settings at a fixed FDR of 0.05 (A) or 0.1 (B) or 0.2 (C). The first number for each method in the figure legend represents the number of times each method is ranked as the best in 25 simulation settings where none of the annotations have zero coefficients; while the second number represents the number of times each method is ranked as the best in 11 simulation settings where at least one annotation has a zero coefficient. x-axis shows the values of the two annotation coefficients used in the simulations. Settings where at least one annotation coefficient is zero are shaded in grey. The setting where the annotation coefficients equal to the median estimates from real data (i.e. $\alpha = (0.1, 0.05)$) is shaded in gold.

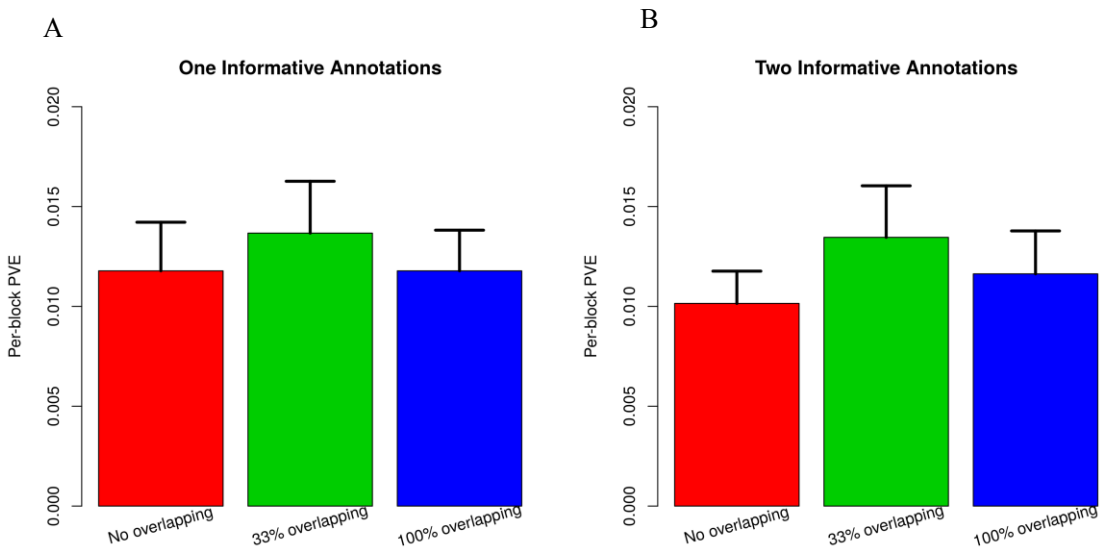


Figure S2 Per-block PVE of the ten causal blocks for the second set simulations. (A) The cases of one informative annotations at $\boldsymbol{\alpha_0 = 0.5}$ and $\boldsymbol{(\alpha_1, \alpha_2) = (0.4, 0)}$; (B) The cases of two informative annotations at $\boldsymbol{\alpha_0 = 0.5}$ and $\boldsymbol{(\alpha_1, \alpha_2) = (0.4, 0.4)}$.The bar indicates the standard error across simulation replicates.

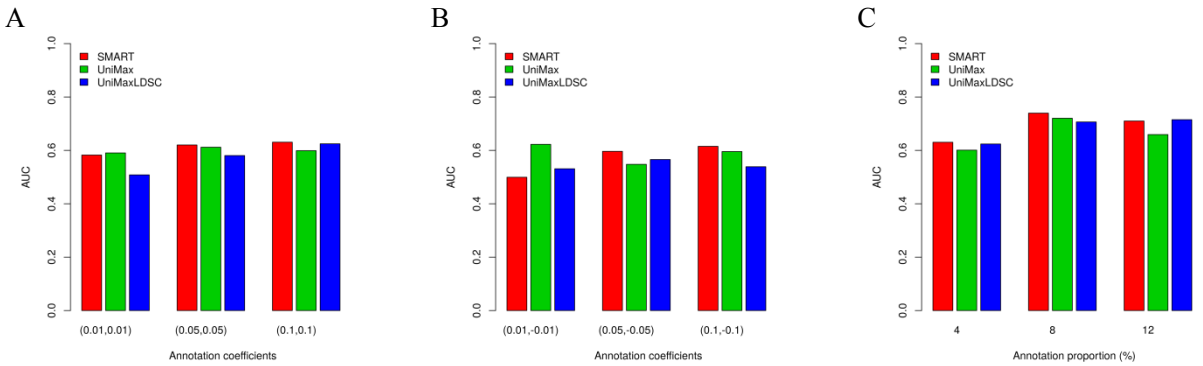


Figure S3 Various annotation characteristics influence the power in identifying trait-relevant tissues in simulations. Methods for comparison include SMART (red), UniMax (green), and UniMaxLDSC (blue). Area under the curve (AUC) is used to measure method performance. (A) Power to identify trait-relevant tissue generally increases with increasingly large annotation coefficients when the two coefficients have the same sign. (B) Power also increases with increasingly large annotation coefficients when the two coefficients have the opposite sign. (C) Power is relatively stable with the genome coverage of the two annotations varied from 4%, 8% to 12%.

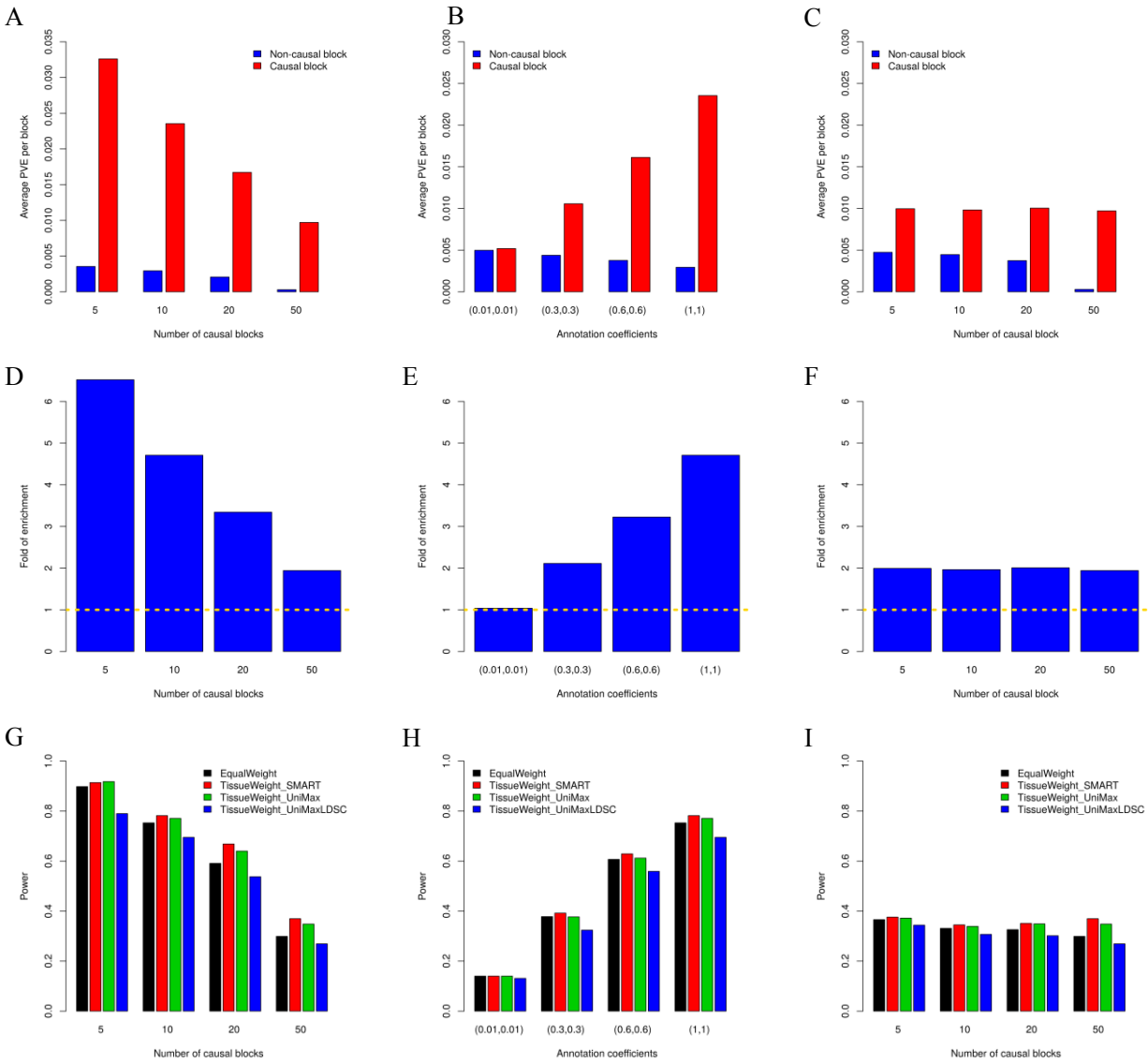
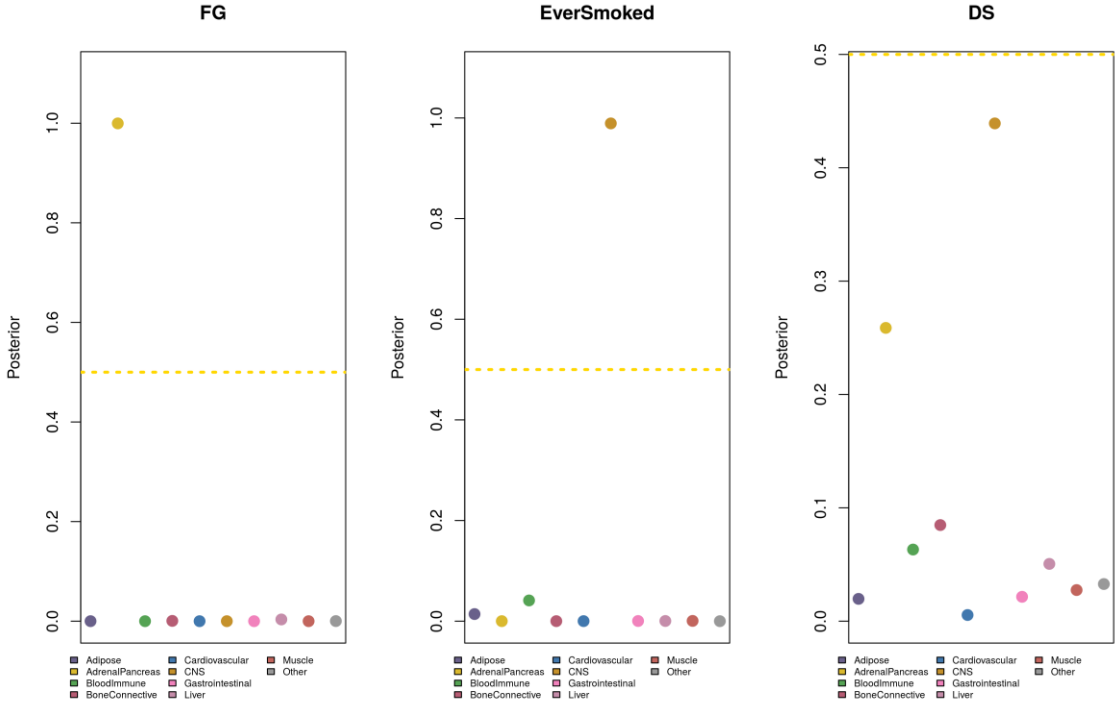
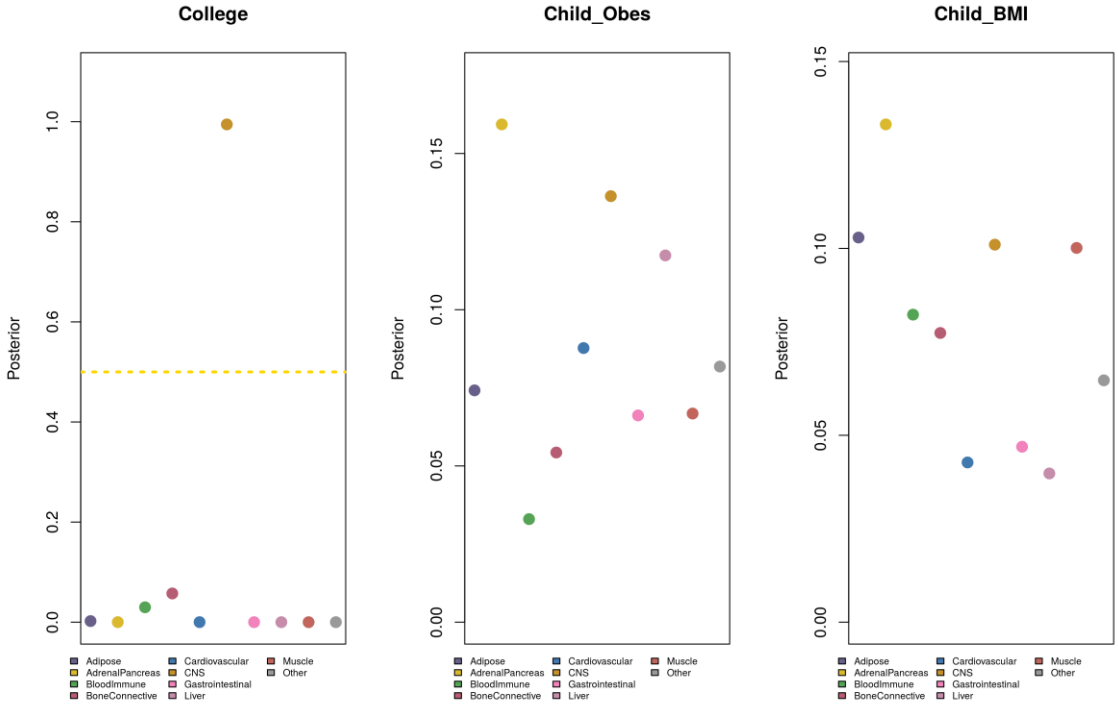
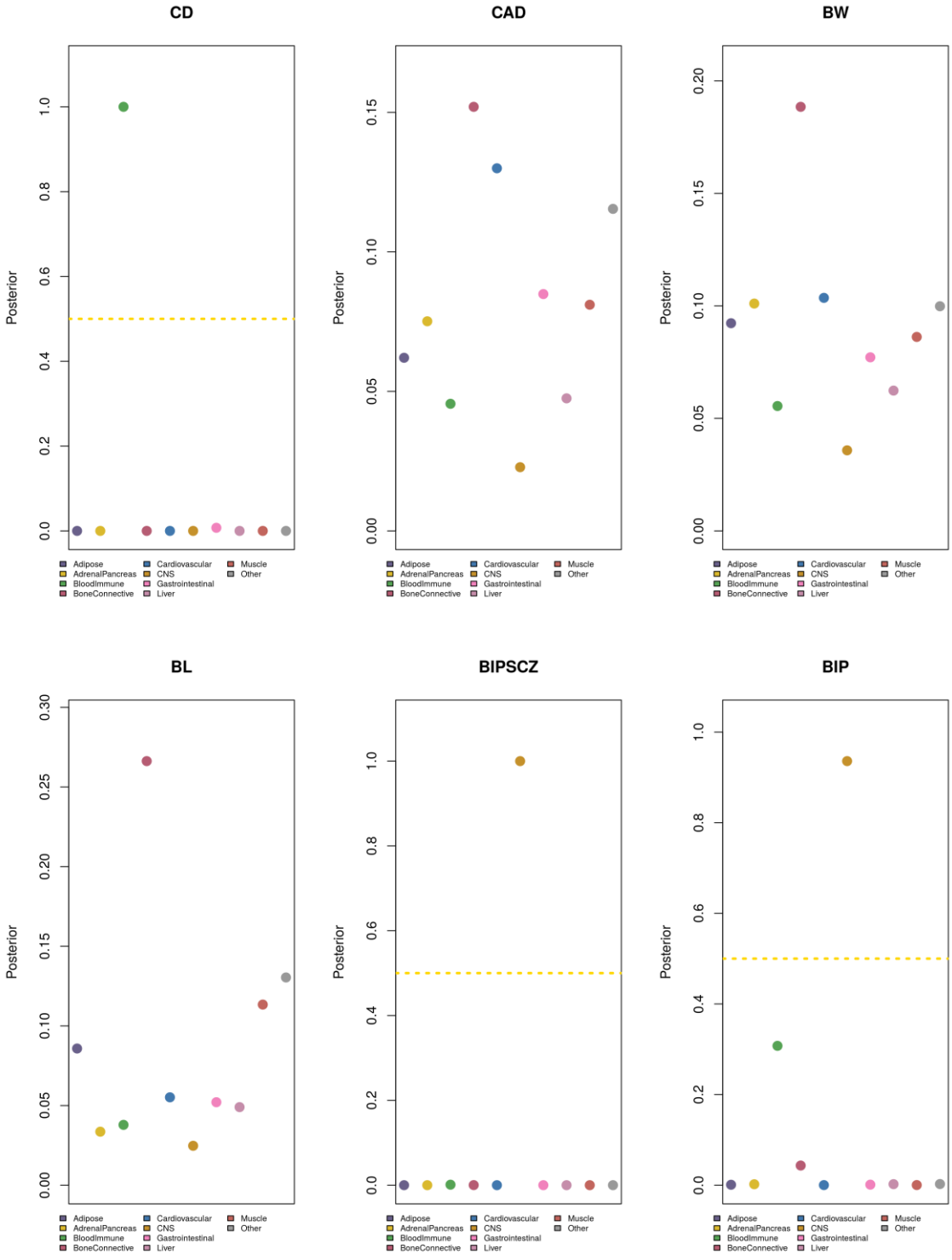
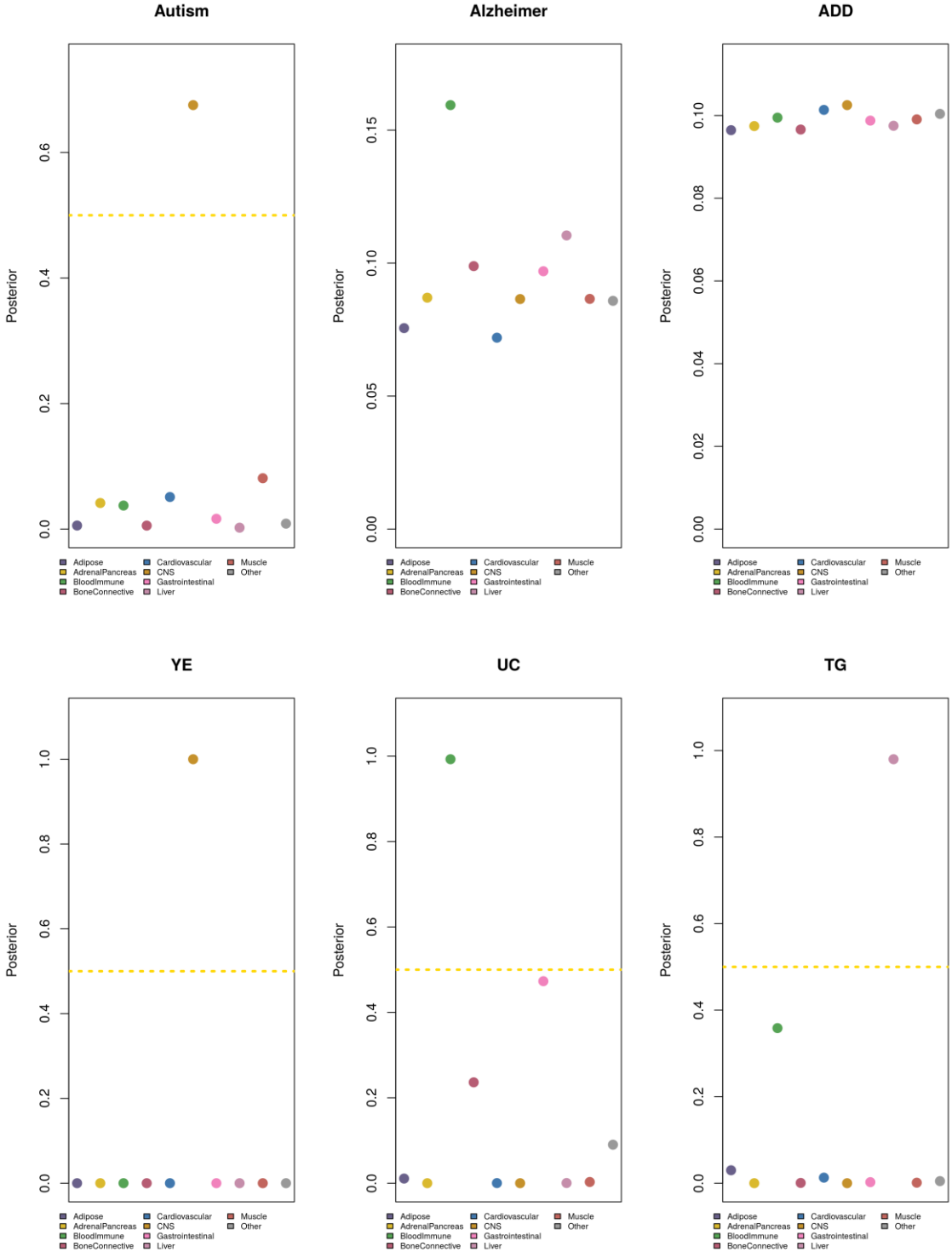


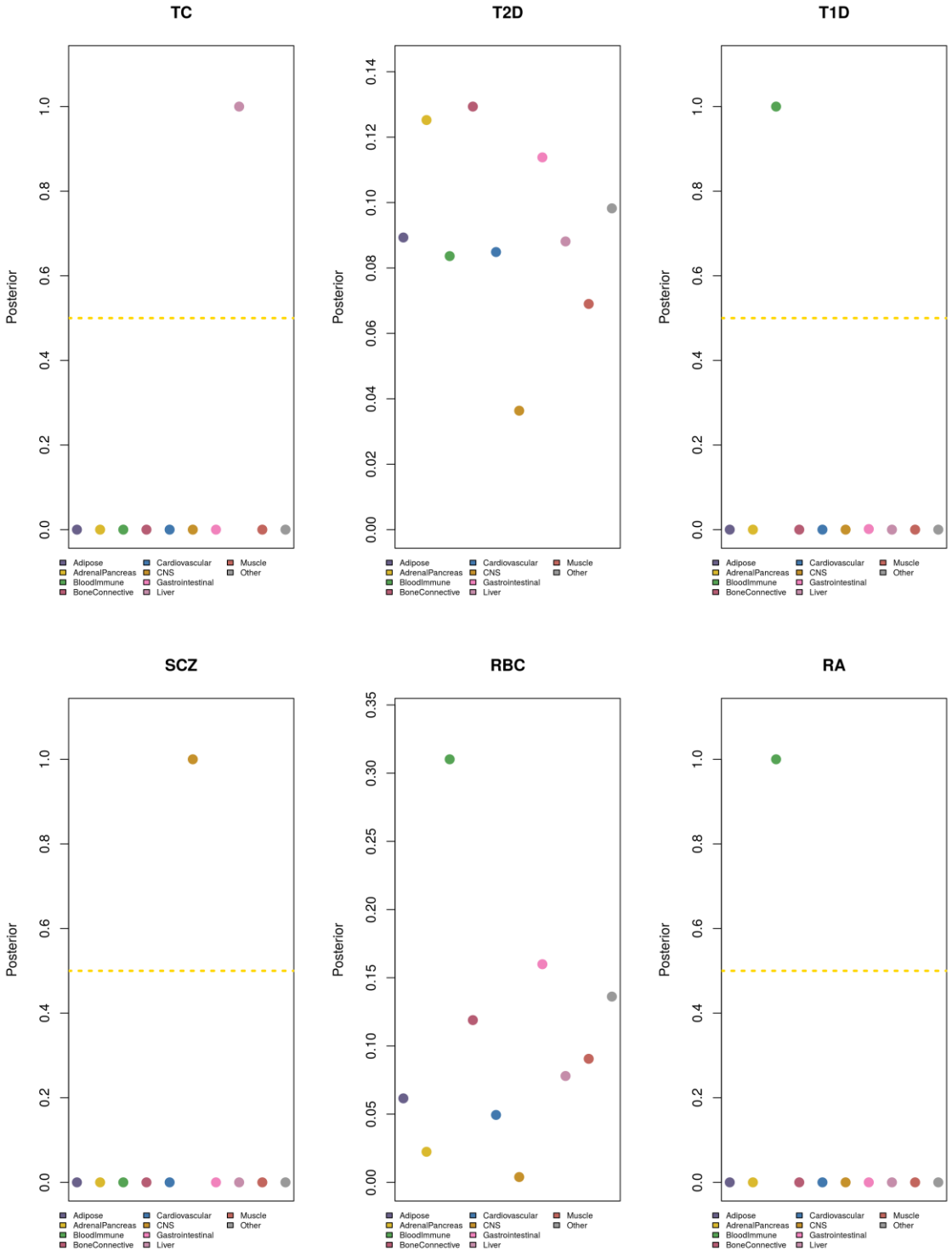
Figure S4 Various factors influence the power of SNP set analysis in simulations. Left columns (A, D, G): annotation coefficients are fixed to be (1, 1) while the number of causal blocks changes from 5, 10, 20 to 50. Middle columns (B, E, H): the number of causal blocks is fixed to be 10 while the annotation coefficients change from (0.01, 0.01), (0.3, 0.3), (0.6, 0.6) to (1, 1). Right columns (C, F, I): per-block PVE are approximately fixed while the number of causal blocks and annotation coefficients vary. Top rows (A, B, C) show the average proportion of phenotype variance explained (PVE) by non-causal or causal blocks. Middle rows (D, E, F) show the fold enrichment. Bottom rows (G, H, I) show SNP set analysis power for various methods.

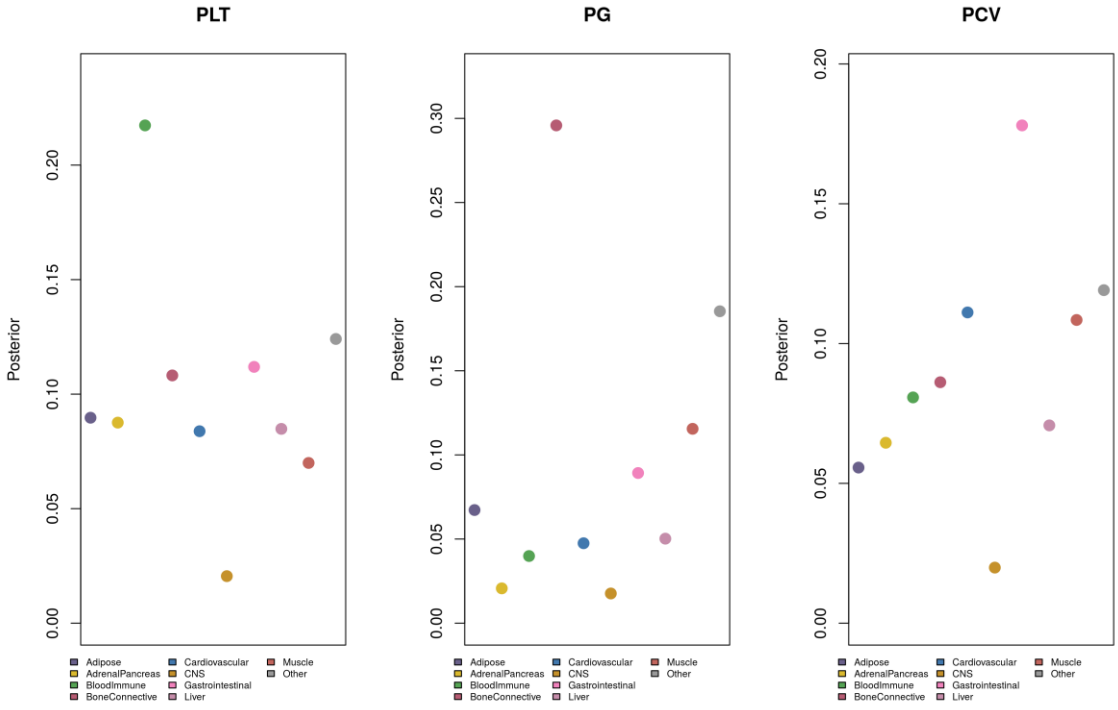

FG
EverSmoked
DS
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other

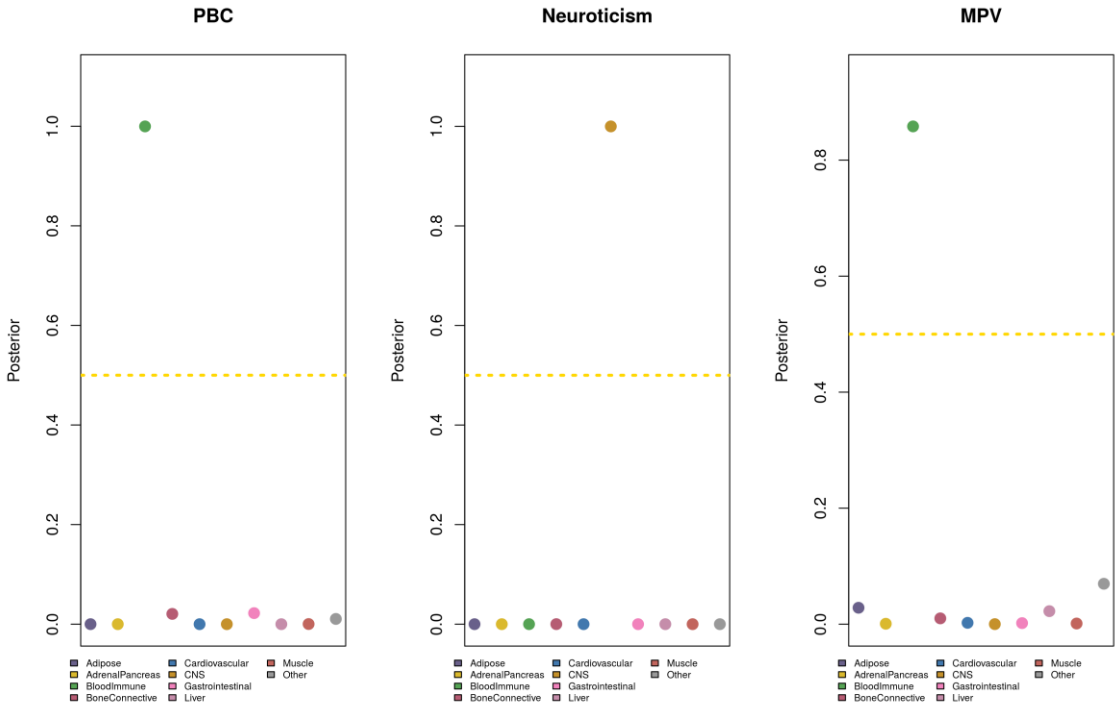

College
Child_Obes
Child_BMI
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other

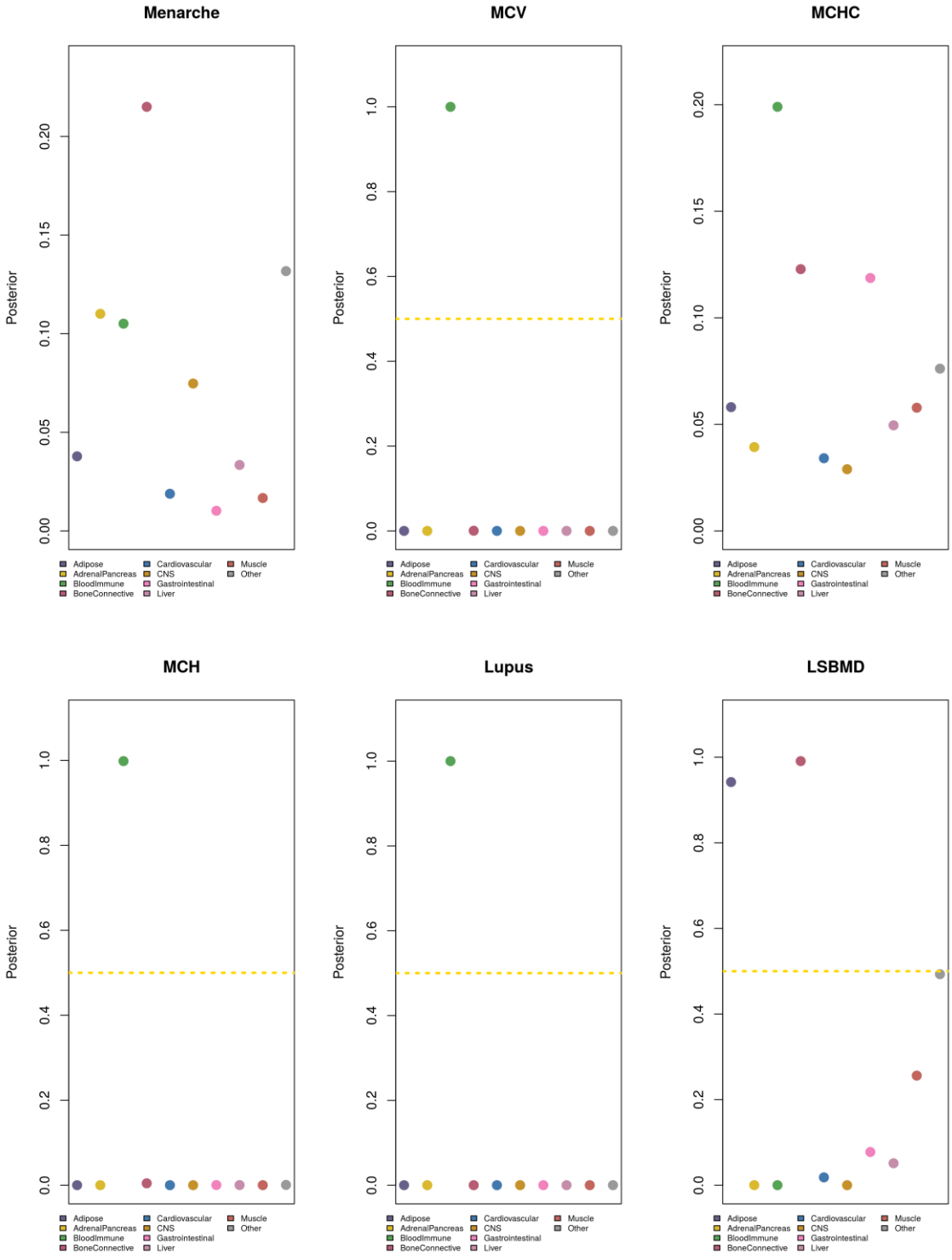

CD
CAD
BW
BL
BIPSCZ
BIP
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other

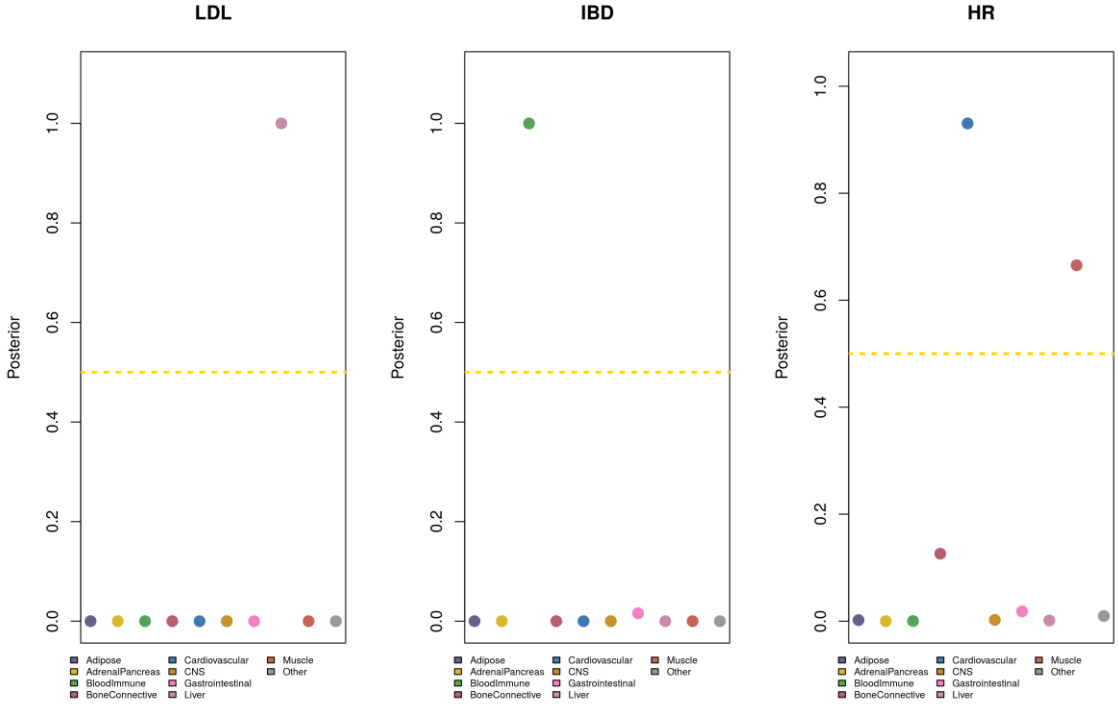

Autism
Alzheimer
ADD
YE
UC
TG
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other

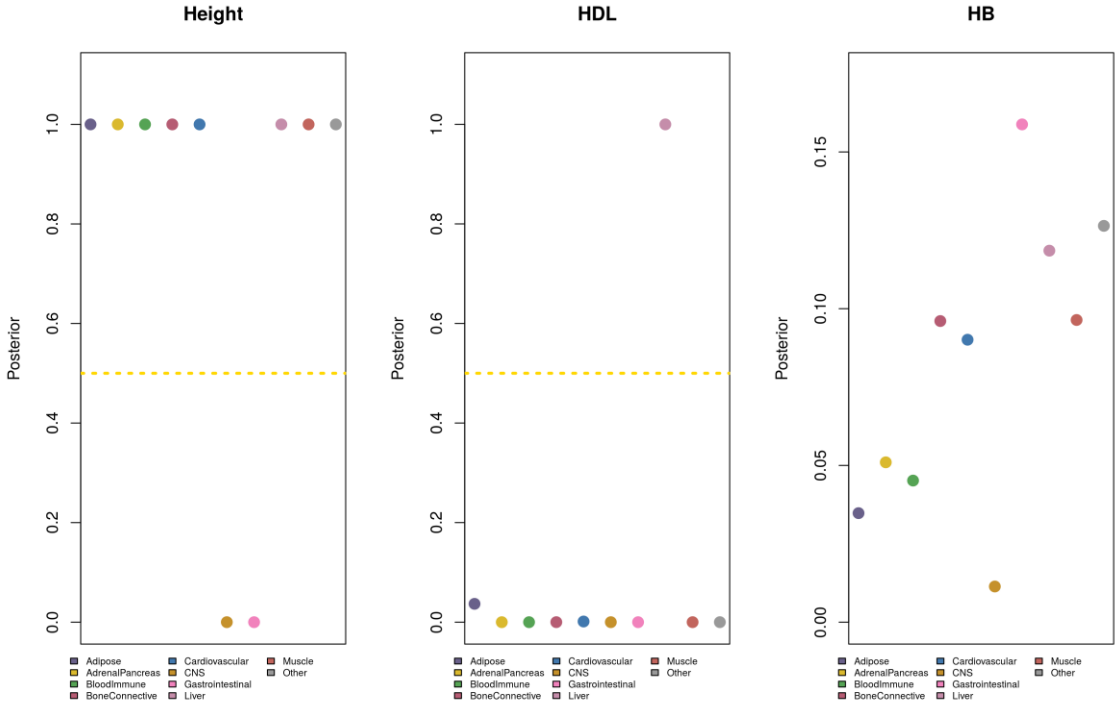

TC
T2D
T1D
SCZ
RBC
RA
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other


PLT
PG
PCV
Posterior
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other


PBC
Neuroticism
MPV
Posterior
0.0
0.2
0.4
0.6
0.8
1.0
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other


Menarche
MCV
MCHC
MCH
Lupus
LSBMD
Posterior
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other


LDL
IBD
HR
Posterior
0.0
0.2
0.4
0.6
0.8
1.0
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other


Height
HDL
HB
Posterior
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
Adipose
AdrenalPancreas
BloodImmune
BoneConnective
Cardiovascular
CNS
Gastrointestinal
Liver
Muscle
Other

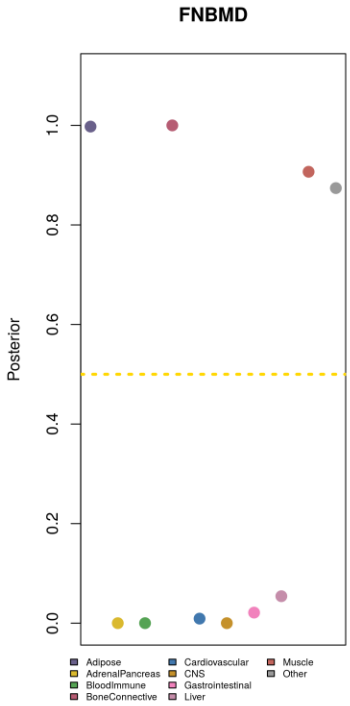
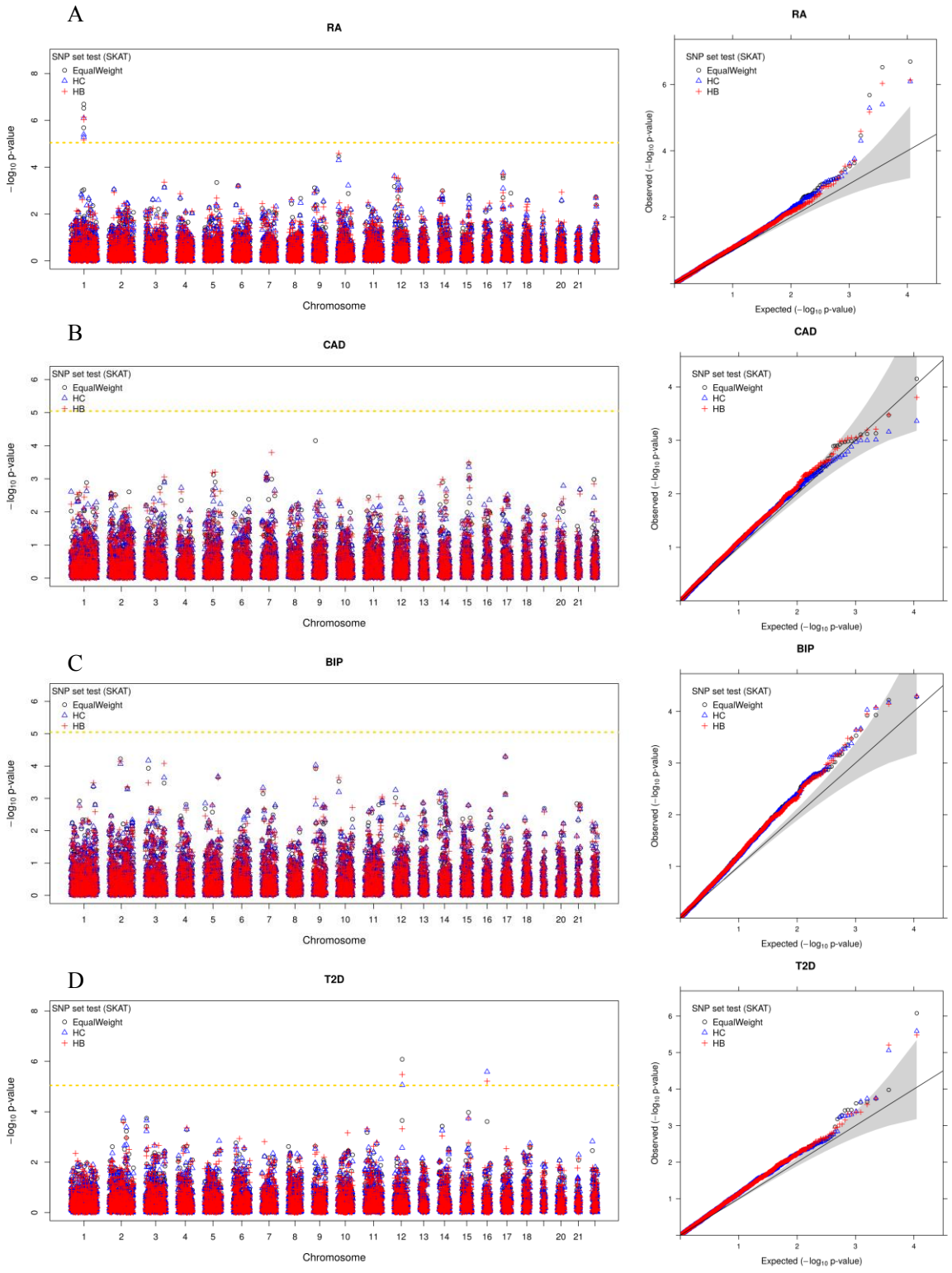
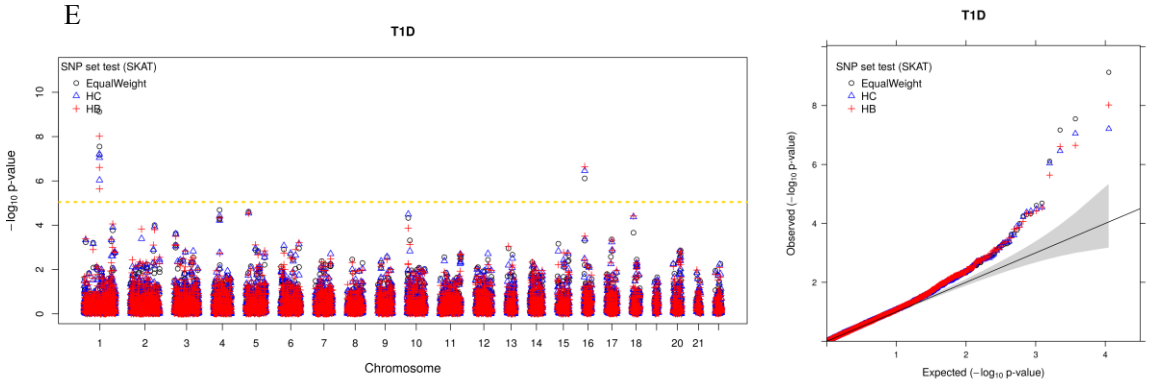


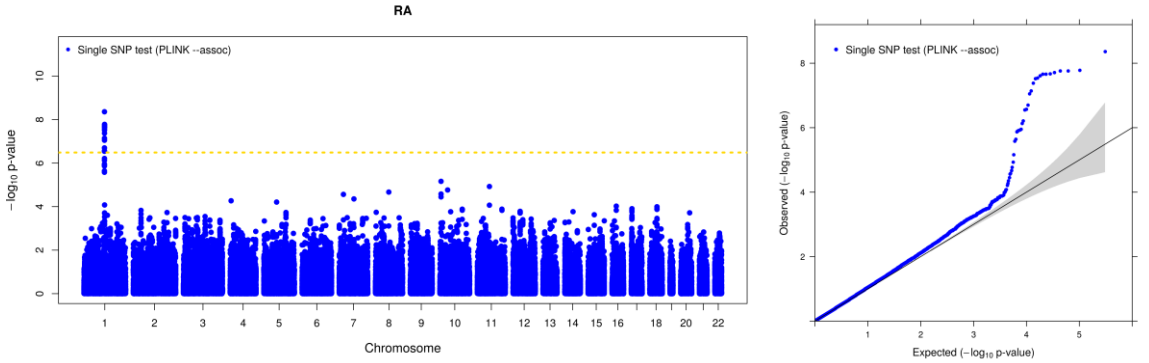
Figure S5 Posterior probabilities of 10 tissue groups for being relevant to each of the 43 GWAS traits by HC. The gold dashed line represents a horizontal line at 0.5.


A
RA
SNP set test (SKAT)
EqualWeight
HC
HB
−log10 p-value
Chromosome
Observed (−log10 p-value)
Expected (−log10 p-value)
B
CAD
C
BIP
D
T2D

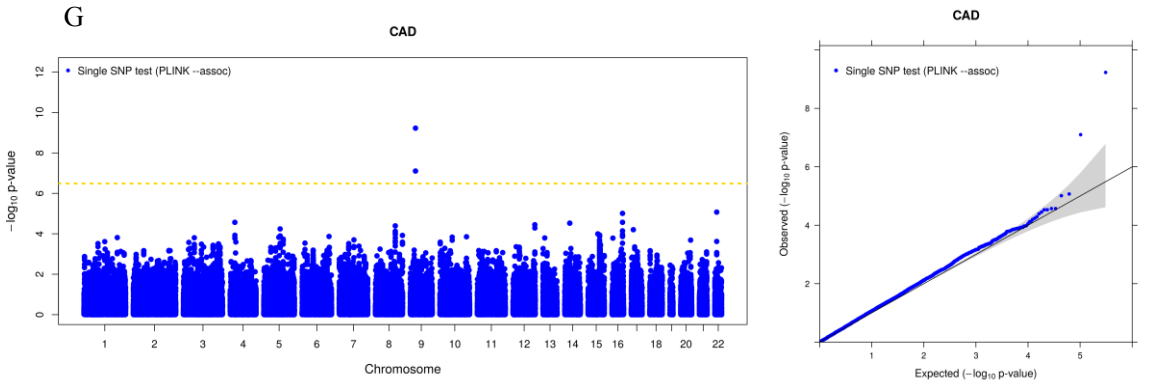
E


T1D
SNP set test (SKAT)
EqualWeight
HC
HB
−log10 p-value
Chromosome
T1D
SNP set test (SKAT)
EqualWeight
HC
HB
Observed (−log10 p-value)
Expected (−log10 p-value)

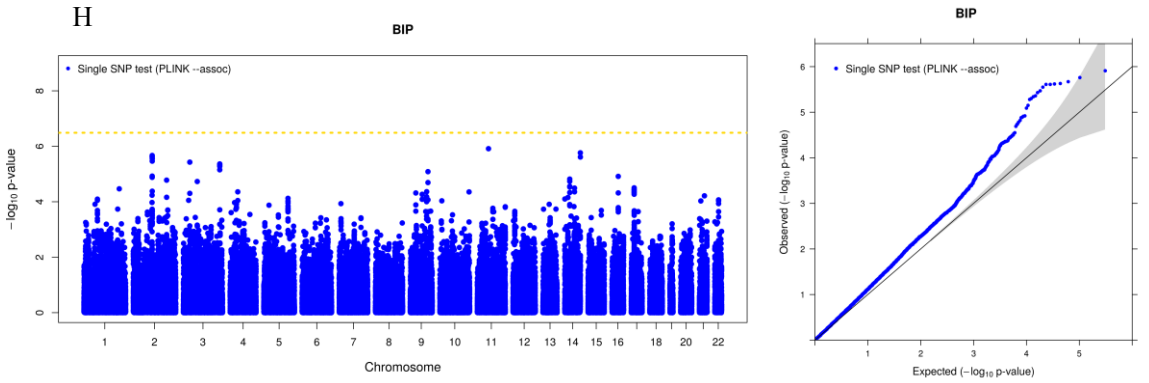
F


RA
Single SNP test (PLINK --assoc)
−log10 p-value
Chromosome
RA
Single SNP test (PLINK --assoc)
Observed (−log10 p-value)
Expected (−log10 p-value)

G


CAD
Single SNP test (PLINK --assoc)
−log10 p-value
Chromosome
CAD
Single SNP test (PLINK --assoc)
Observed (−log10 p-value)
Expected (−log10 p-value)

H


BIP
Single SNP test (PLINK --assoc)
−log10 p-value
Chromosome
BIP
Single SNP test (PLINK --assoc)
Observed (−log10 p-value)
Expected (−log10 p-value)

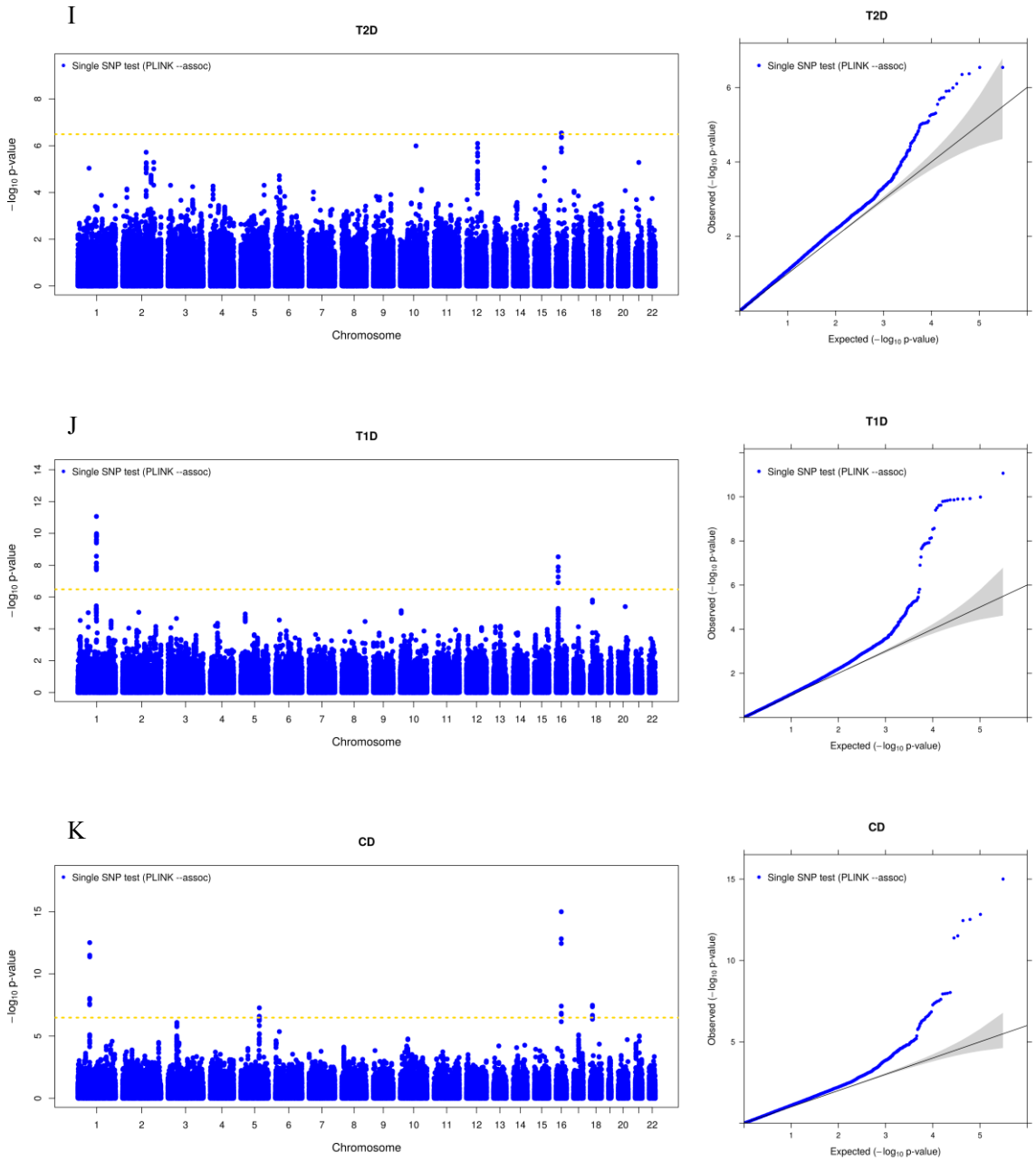


Figure S6 Manhattan and QQ plots display the SNP set test results for five common diseases in WTCCC using different SNP weights. Results are shown for rheumatoid arthritis (RA; A), cardiovascular disease (CAD; B), bipolar disease (BIP; C), type II diabetes (T2D; D), and type I diabetes (T1D; E). For comparison, association results based on univariate SNP tests are also shown in F-K. EqualWeight (black): equal SNP weights. *HC* (blue): SNP weights constructed using the estimated coefficient parameters for continuous histone mark based annotations in a GWAS consortium study. *HB* (red): SNP weights constructed using the estimated coefficient parameters for binary histone mark based annotations in a GWAS consortium study. For Manhattan plots, gold dashed lines represent genome-wide significance thresholds: 0.05/153,813 for univariate tests and 0.05/5,588 for SNP set tests. For QQ plots, grey shaded area represents the 95% point-wise confidence interval.

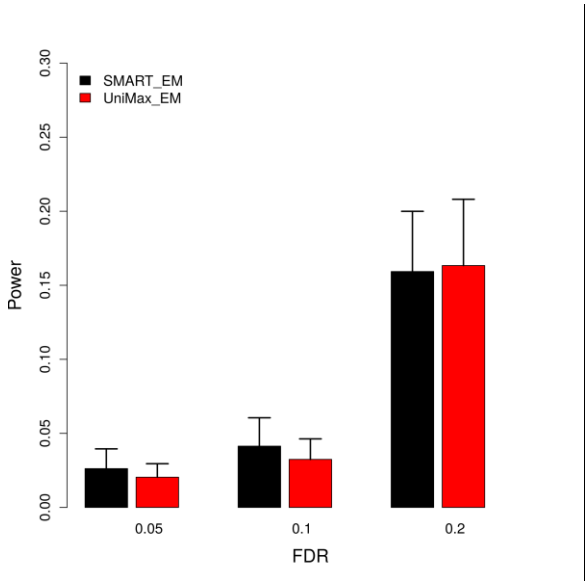


Figure S7 Simulation results when there are two trait relevant tissues. Among the ten tissues, we randomly selected two of them to be trait-relevant. We selected one annotation (H3K4me1) from the first trait-relevant tissue and another annotation (H3K4me3) from the second trait-relevant tissue to simulate SNP effects. The annotation coefficients for these two annotations are set to be 0.1 and 0.05, with $\alpha_0 = 0.1$. We performed 5,000 simulation replicates that were divided into 10 groups. We computed power in each simulation group separately, and obtained the mean and standard deviation of power across 10 groups. The plot shows the mean power and its standard deviation to detect either of the two trait-relevant tissues by different approaches at a fixed FDR value of 0.05, 0.1 or 0.2.