Supplementary Figures

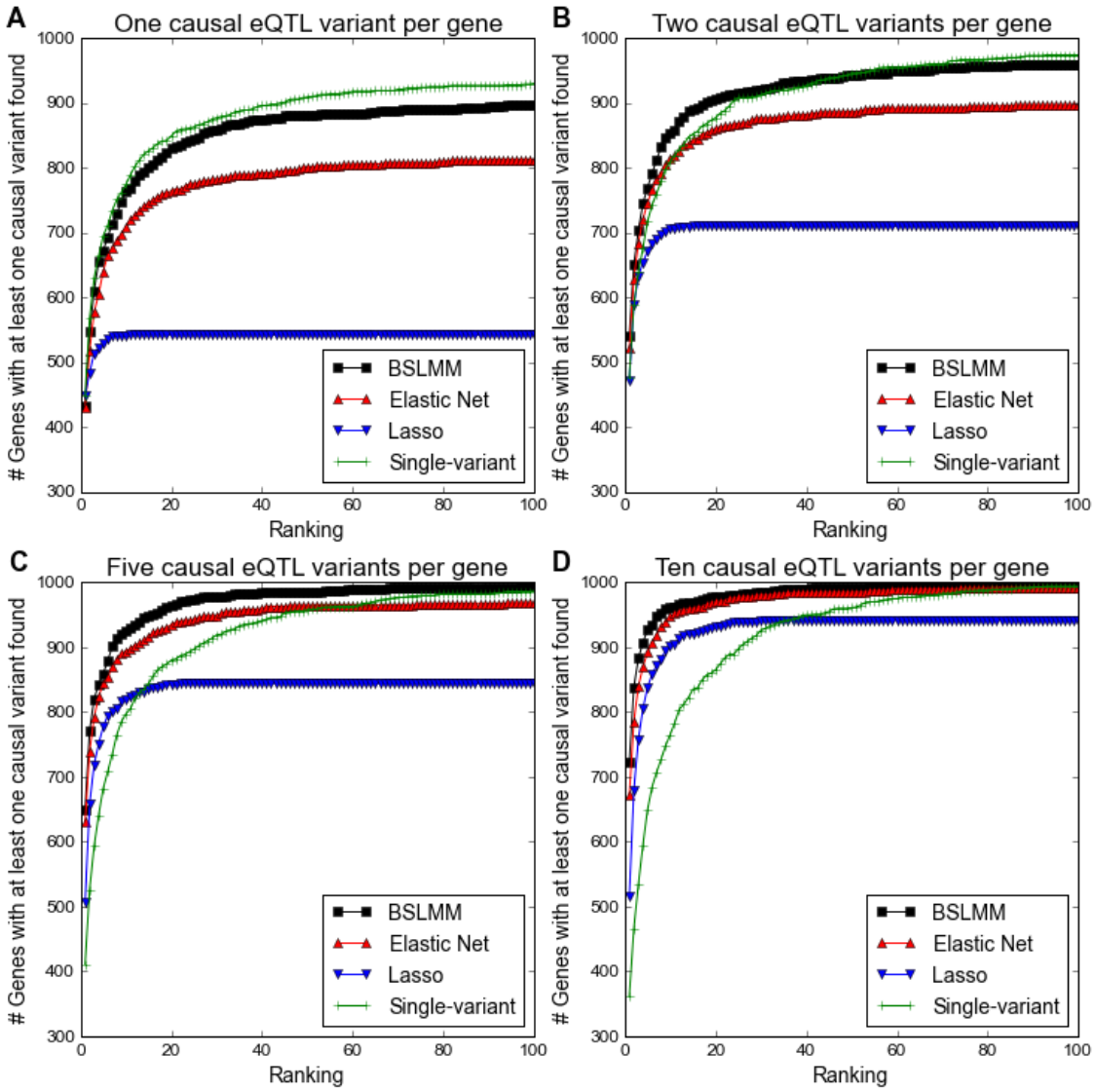

Figure S1. The number of genes with at least one causal eQTL variant found from the simulation data of 503 samples with 60% heritability. The x-axis shows the ranking cutoff and the y-axis shows the number of genes with at least one causal eQTL variant found at the ranking cutoff. (A) One causal eQTL variant per gene. (B) Two causal eQTL variants per gene. (C) Five causal eQTL variants per gene. (D) Ten causal eQTL variants per gene.

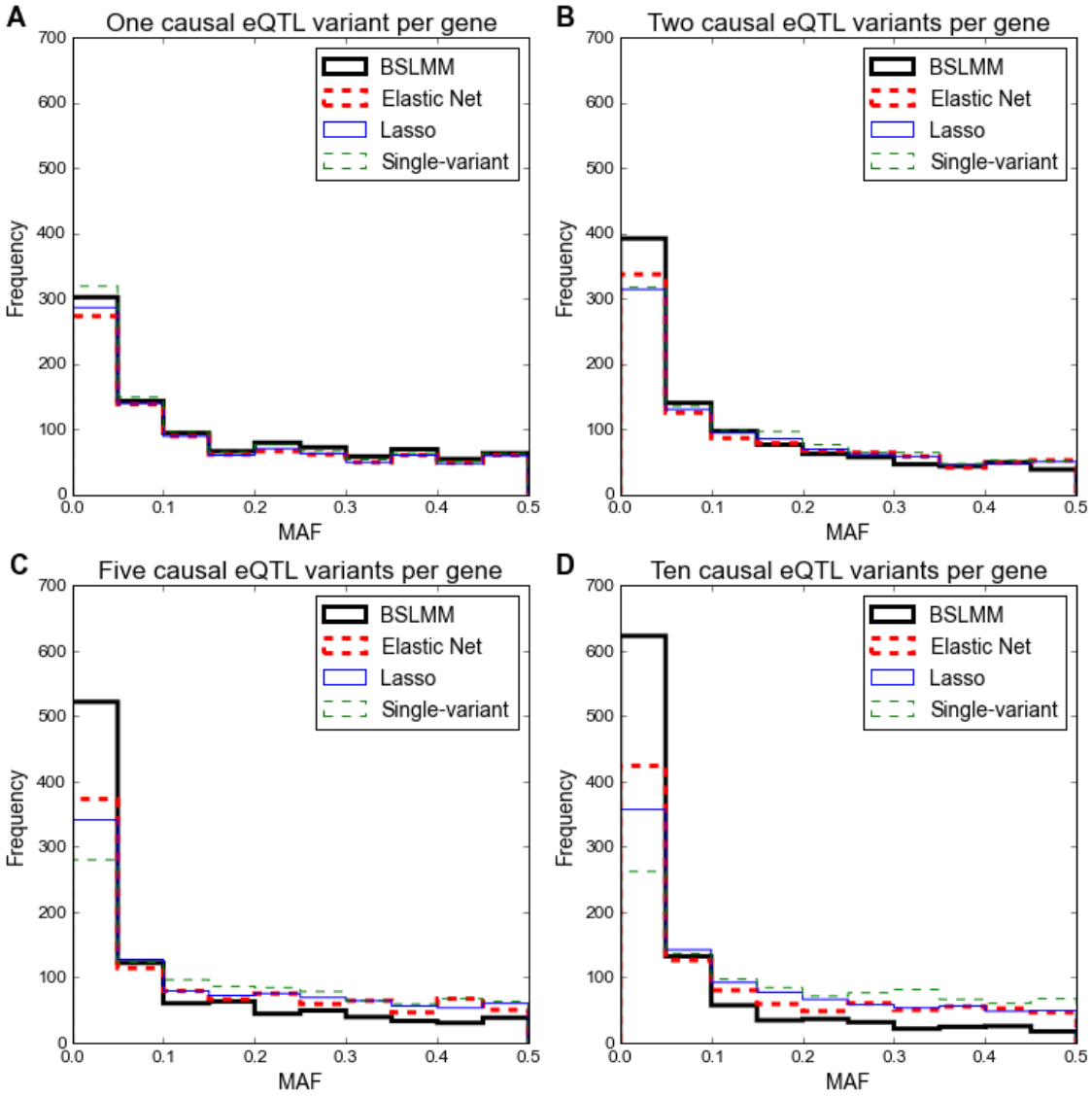

Figure S2. Minor allele frequency of eQTL SNPs identified with each method. The x-axis shows minor allele frequency (MAF) and the y-axis shows the number of top eQTL SNPs within the specified MAF bin identified by each method. (A) One causal eQTL variant per gene. (B) Two causal eQTL variants per gene. (C) Five causal eQTL variants per gene. (D) Ten causal eQTL variants per gene.

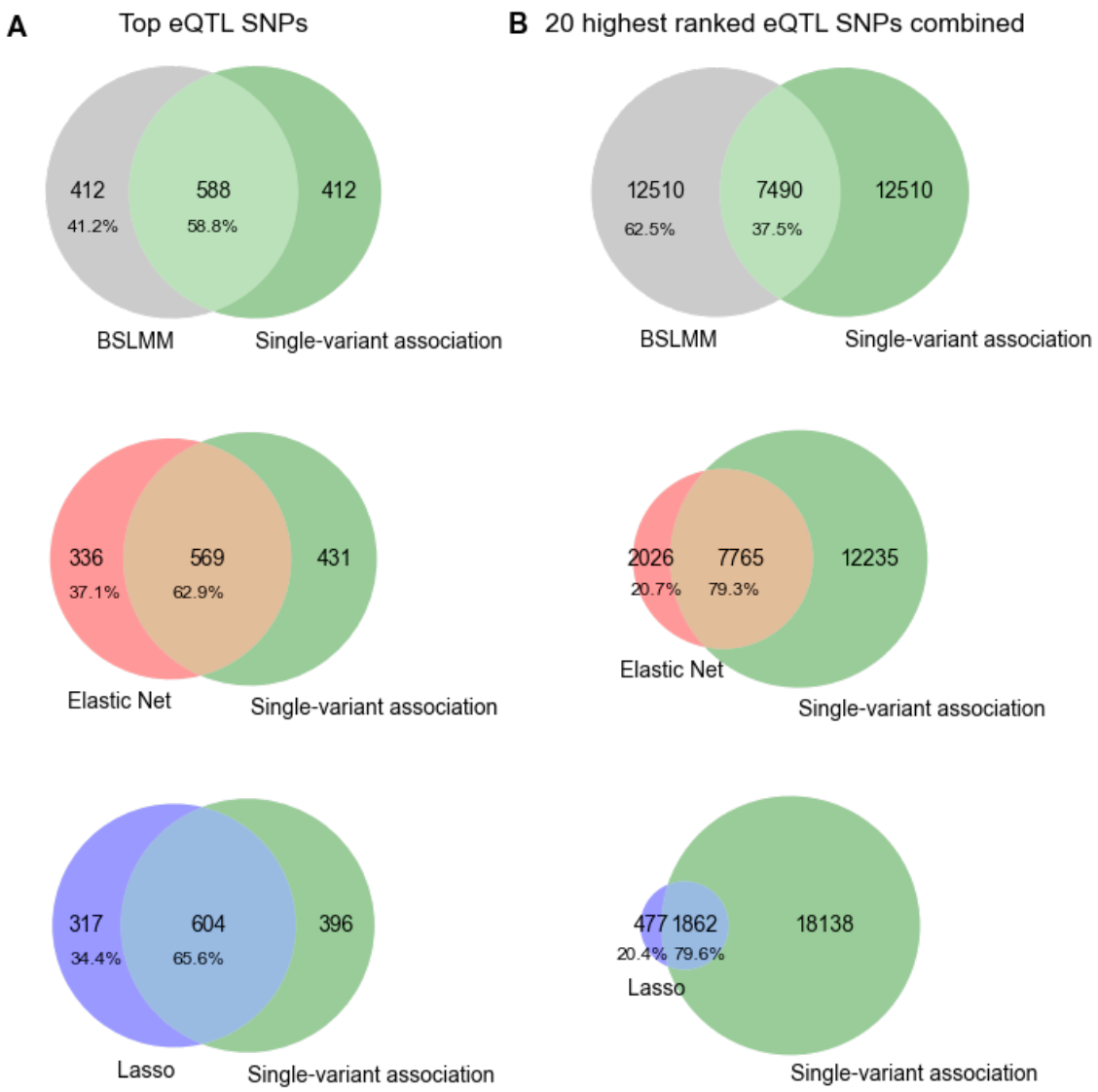

Figure S3. Overlap of the identified eQTL variants from the simulation data of 503 samples with 60% heritability and one causal variant per gene. (A) Overlap of the top eQTL SNPs between each sparse polygenic model (top: BSLMM, middle: Elastic Net, bottom: Lasso) and single-variant association analysis. (B) Overlap of the 20 highest ranked eQTL SNPs combined between each sparse polygenic model (top: BSLMM, middle: Elastic Net, bottom: Lasso) and single-variant association analysis. The percentages shown indicate the fraction of eQTL SNPs identified by the sparse polygenic model that overlap with those identified by single-variant association analysis.

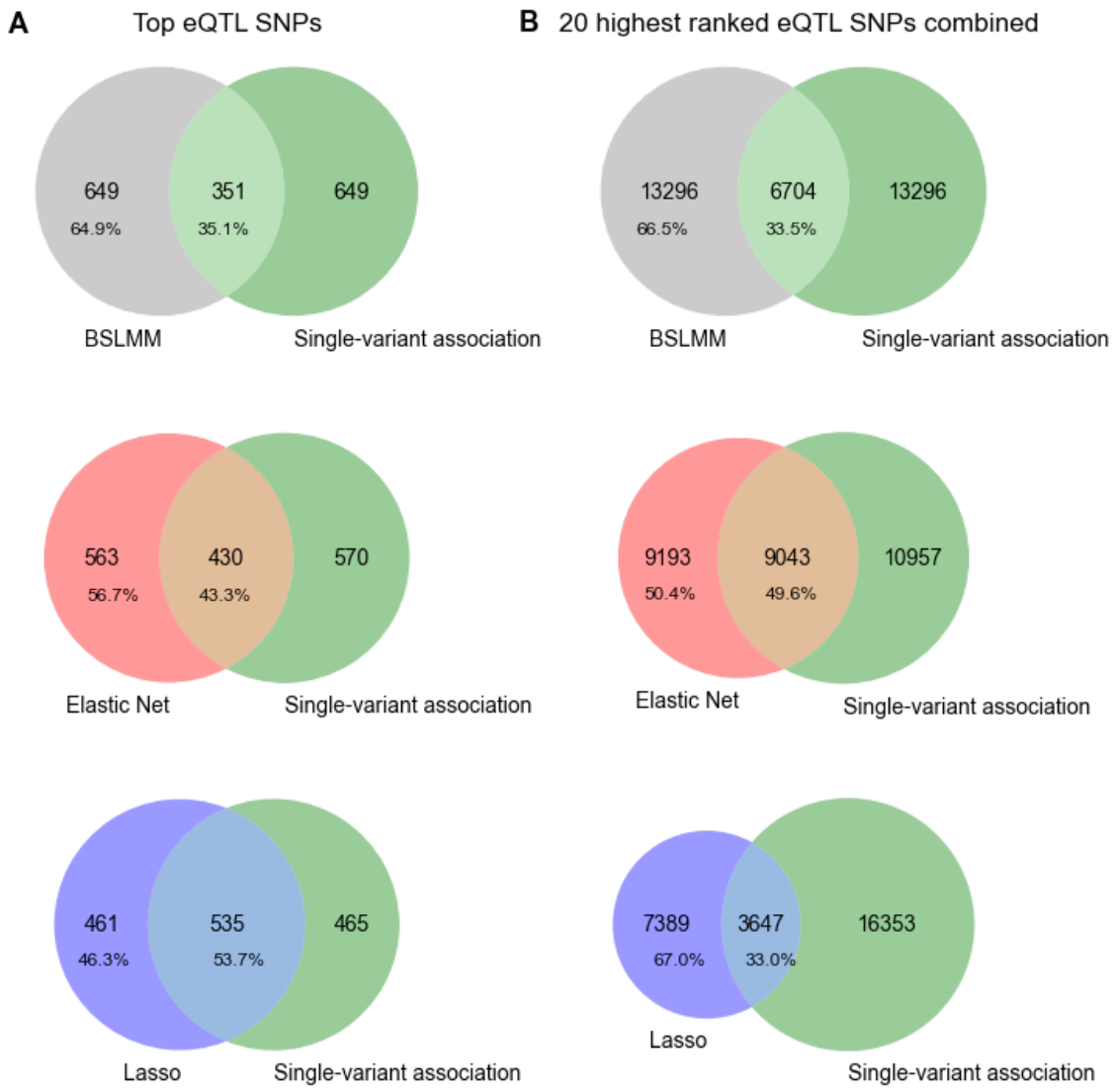

Figure S4. Overlap of the identified eQTL variants from the simulation data of 503 samples with 60% heritability and five causal variants per gene. (A) Overlap of the top eQTL SNPs between each sparse polygenic model (top: BSLMM, middle: Elastic Net, bottom: Lasso) and single-variant association analysis. (B) Overlap of the 20 highest ranked eQTL SNPs combined between each sparse polygenic model (top: BSLMM, middle: Elastic Net, bottom: Lasso) and single-variant association analysis. The percentages shown indicate the fraction of eQTL SNPs identified by the sparse polygenic model that overlap with those identified by single-variant association analysis.

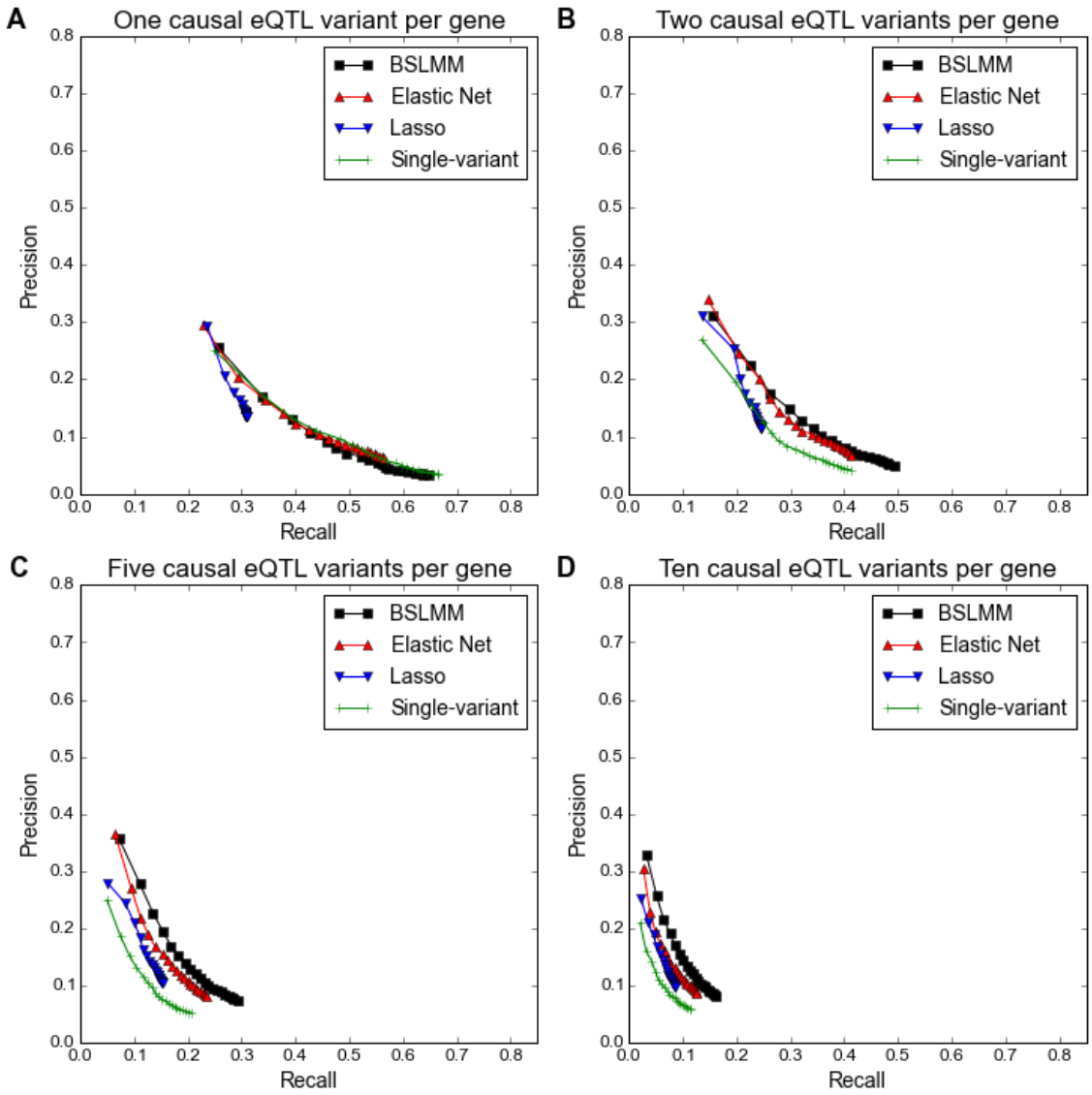

Figure S5. Prediction performance for identifying causal eQTL variants from the simulation data of 100 samples with 60% heritability. The x-axis shows recall and the y-axis shows precision for the 20 highest ranked SNPs of 1,000 randomly selected genes. (A) One causal eQTL variant per gene. (B) Two causal eQTL variants per gene. (C) Five causal eQTL variants per gene. (D) Ten causal eQTL variants per gene.

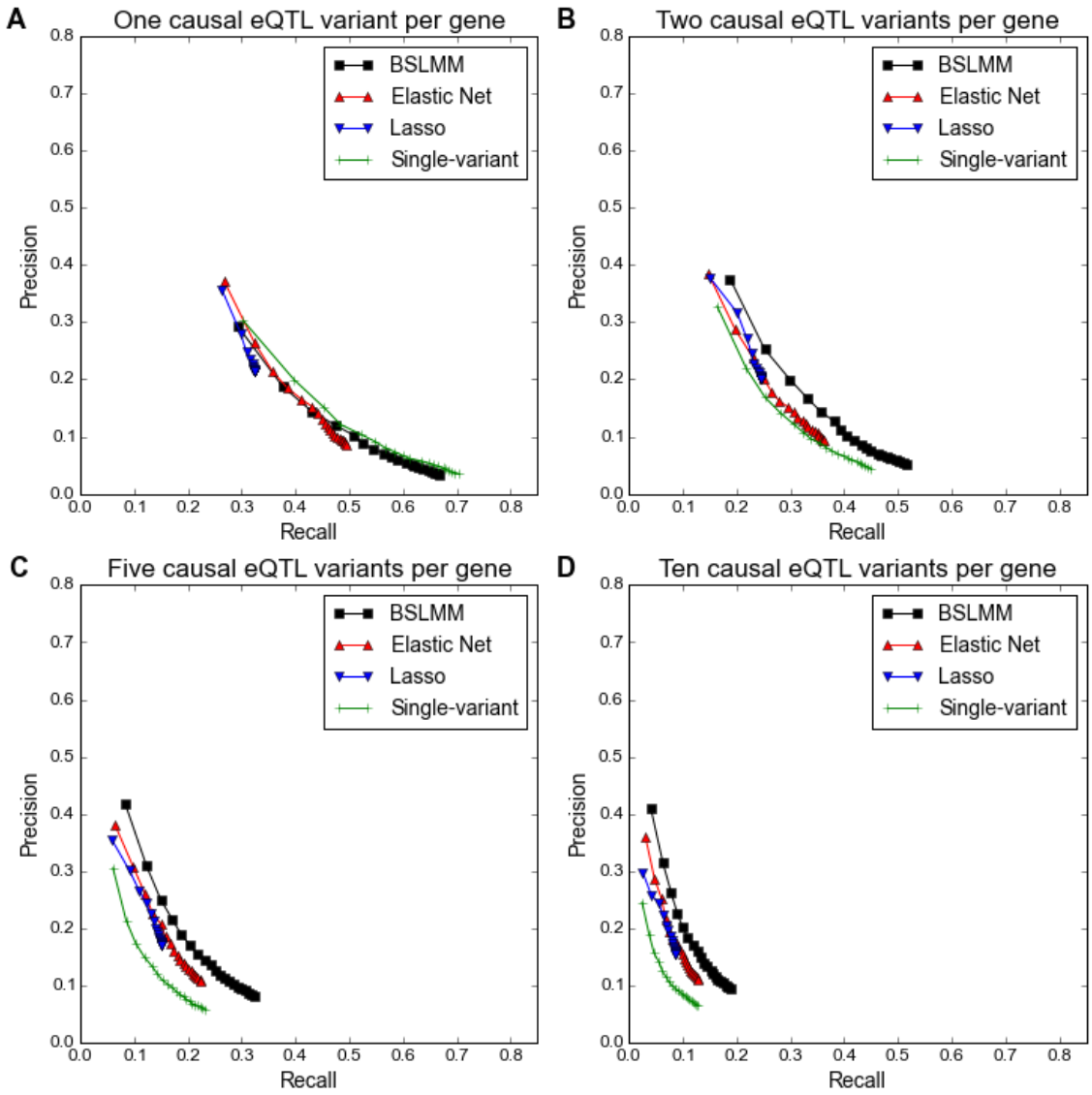

Figure S6. Prediction performance for identifying causal eQTL variants from the simulation data of 503 samples with 20% heritability. The x-axis shows recall and the y-axis shows precision for the 20 highest ranked SNPs of 1,000 randomly selected genes. (A) One causal eQTL variant per gene. (B) Two causal eQTL variants per gene. (C) Five causal eQTL variants per gene. (D) Ten causal eQTL variants per gene.

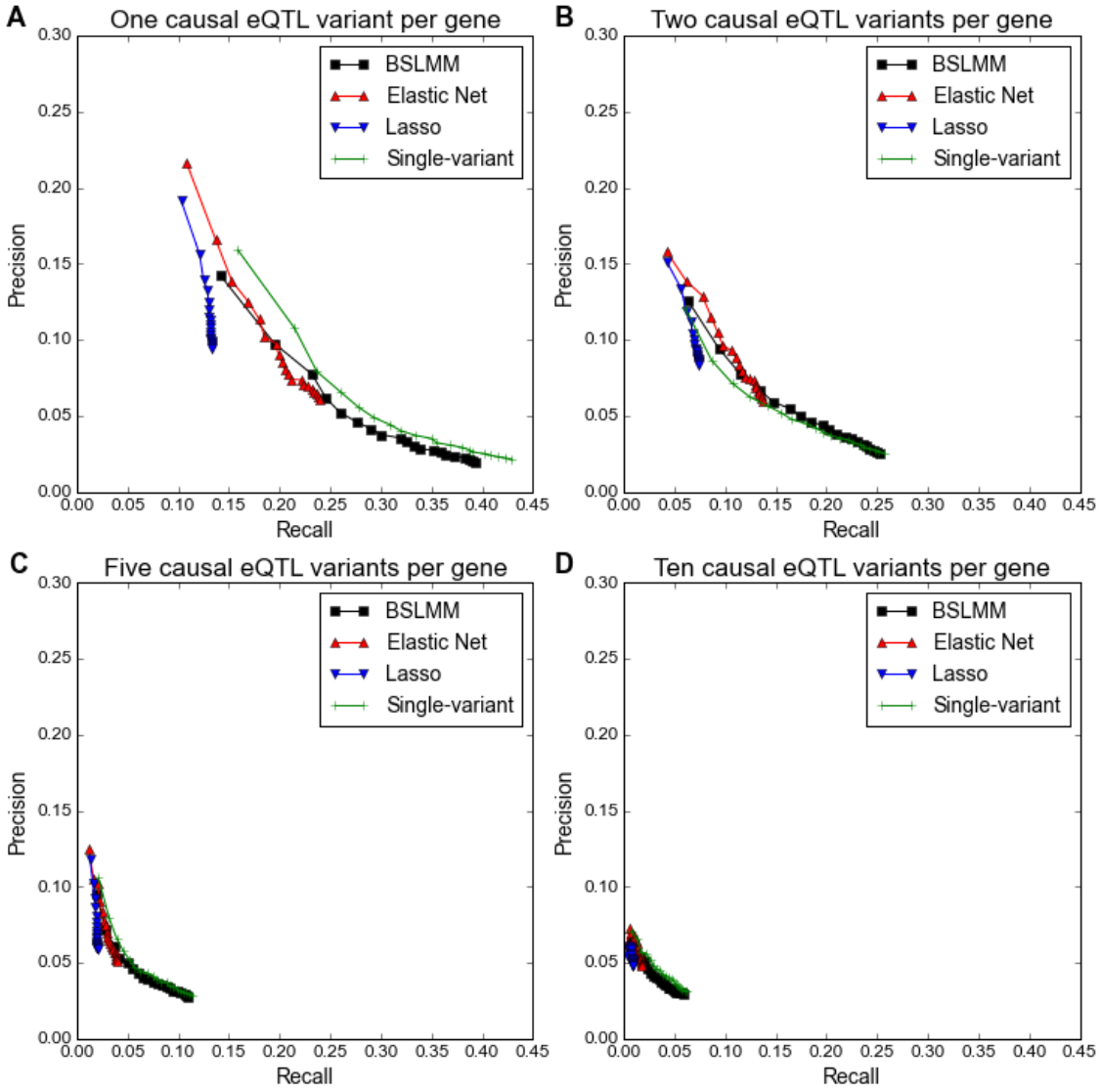

Figure S7. Prediction performance for identifying causal eQTL variants from the simulation data of 100 samples with 20% heritability. The x-axis shows recall and the y-axis shows precision for the 20 highest ranked SNPs of 1,000 randomly selected genes. (A) One causal eQTL variant per gene. (B) Two causal eQTL variants per gene. (C) Five causal eQTL variants per gene. (D) Ten causal eQTL variants per gene.

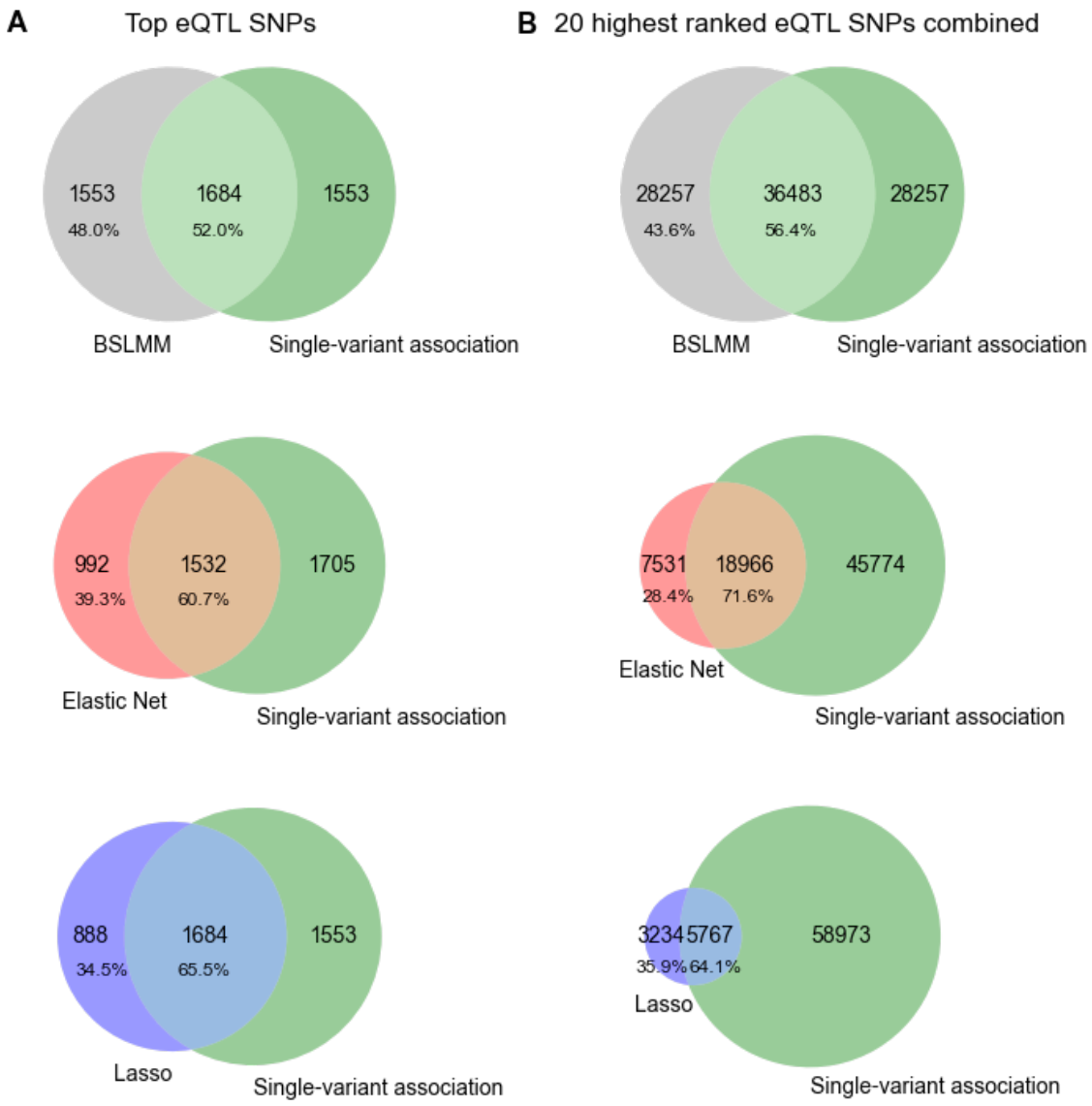

Figure S8. Overlap of eQTL variants of eGenes with one independent eQTL. A) Overlap of the top eQTL SNPs between BSLMM and single-variant association analysis (top), Elastic Net and single-variant association analysis (middle), and Lasso and single-variant association analysis (bottom). (B) Overlap of the 20 highest ranked eQTL SNPs combined between BSLMM and single-variant association analysis (top), Elastic Net and single-variant association analysis (middle), and Lasso and single-variant association analysis (bottom). The percentages shown indicate the fraction of eQTL SNPs identified by the sparse polygenic model that overlap with those identified by single-variant association analysis.

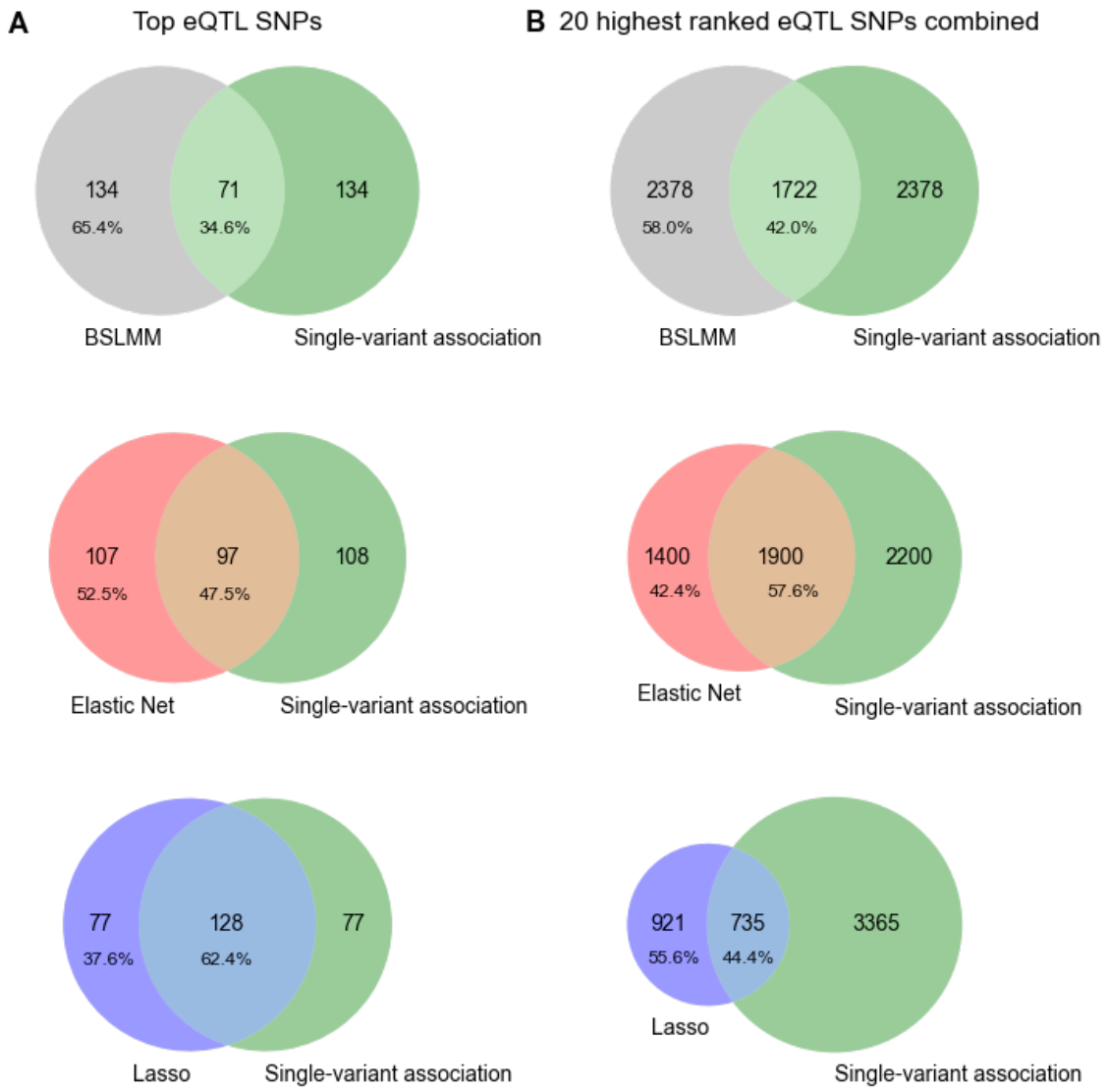

Figure S9. Overlap of eQTL variants of eGenes with more than one independent eQTLs. (A) Overlap of the top eQTL SNPs between BSLMM and single-variant association analysis (top), Elastic Net and single-variant association analysis (middle), and Lasso and single-variant association analysis (bottom). (B) Overlap of the 20 highest eQTL SNPs combined between BSLMM and single-variant association analysis (top), Elastic Net and single-variant association analysis (middle), and Lasso and single-variant association analysis (bottom). The percentages shown indicate the fraction of eQTL SNPs identified by the sparse polygenic model that overlap with those identified by single-variant association analysis.

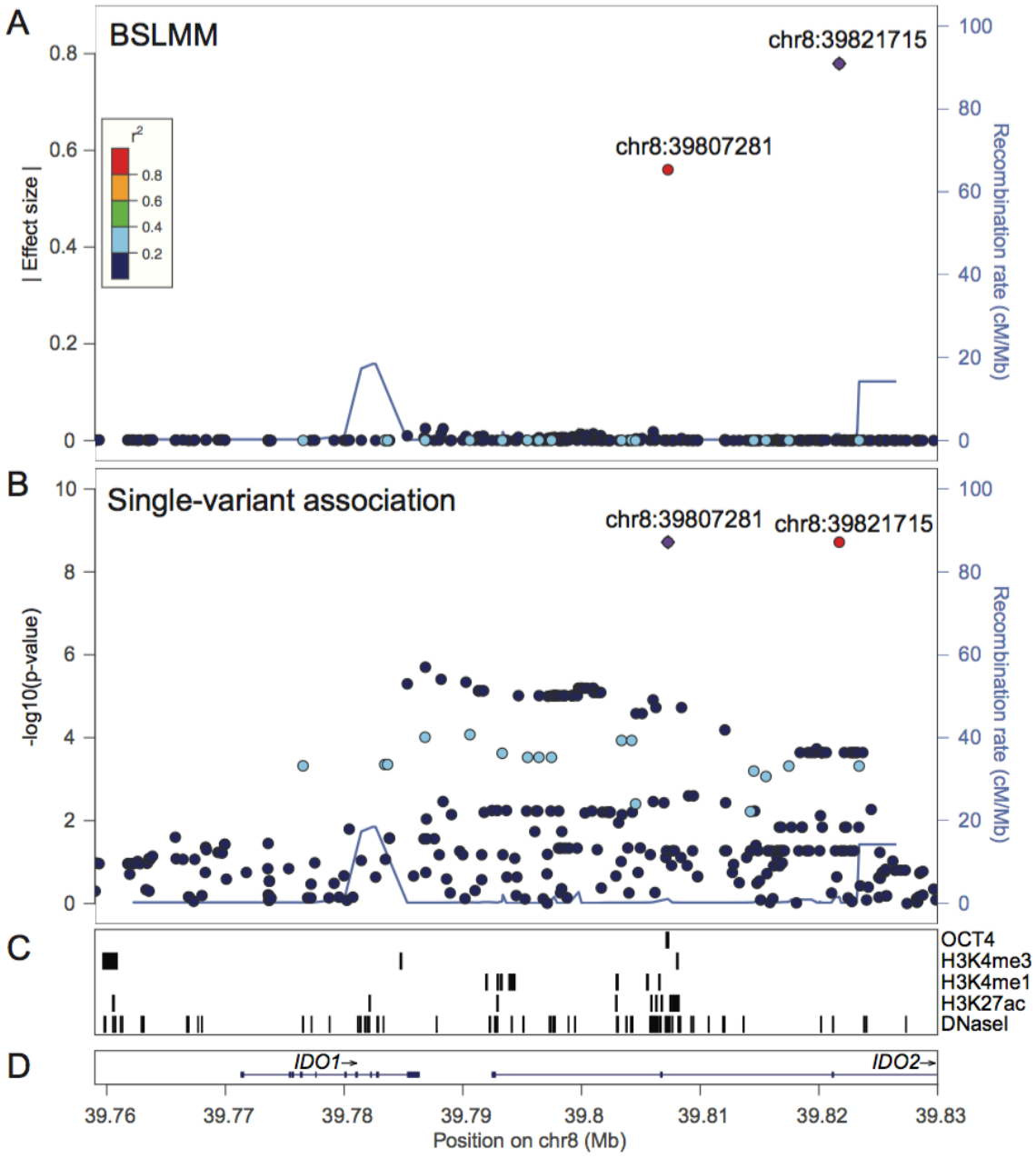

Figure S10. eQTL variants identified as associated with *IDO1* expression. Variants are color-coded based on the strength of LD with the most highly associated eQTL (purple diamond). (A) BSLMM identified two eQTL SNPs as the most highly ranked eQTL variants, located on chr8:39807281 and chr8:39821715, which were also identified as significant by single-variant association analysis. (C) Genomic regions annotated with H1-hESC OCT4, iPSC histone marks (H3K4me3, H3K4me1, and H3K27ac), and iPSC DNase I hypersensitive sites (DHSs). (D) Genomic coordinates of *IDO1* and surrounding genes are based on hg19.