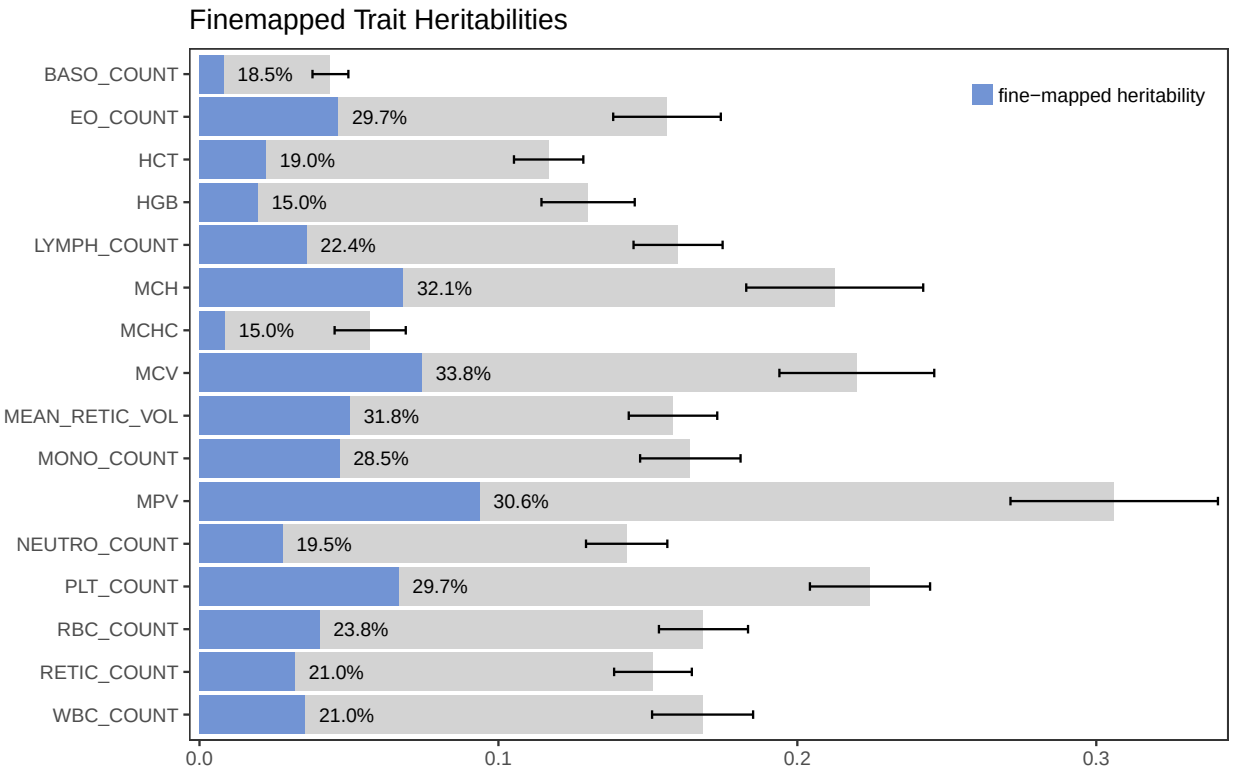
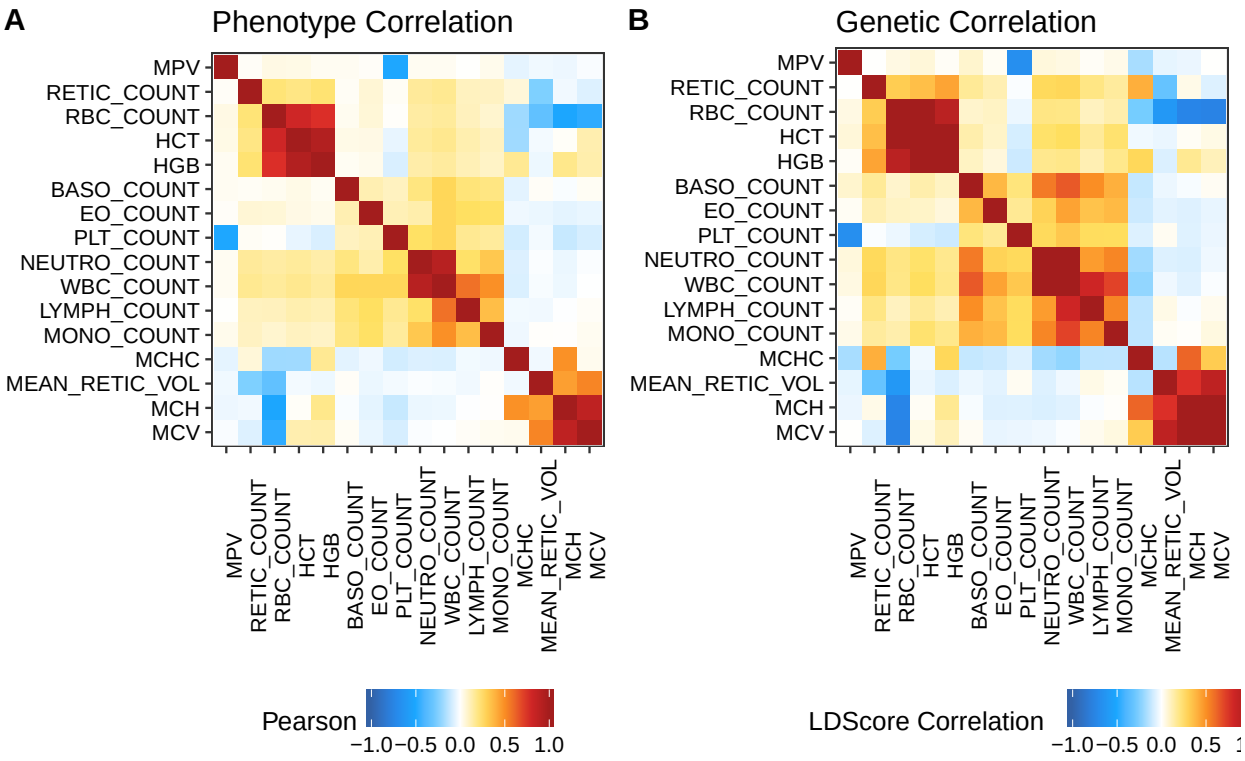
Interrogation of human hematopoiesis at single-cell and single-variant resolution

Supplemental Information

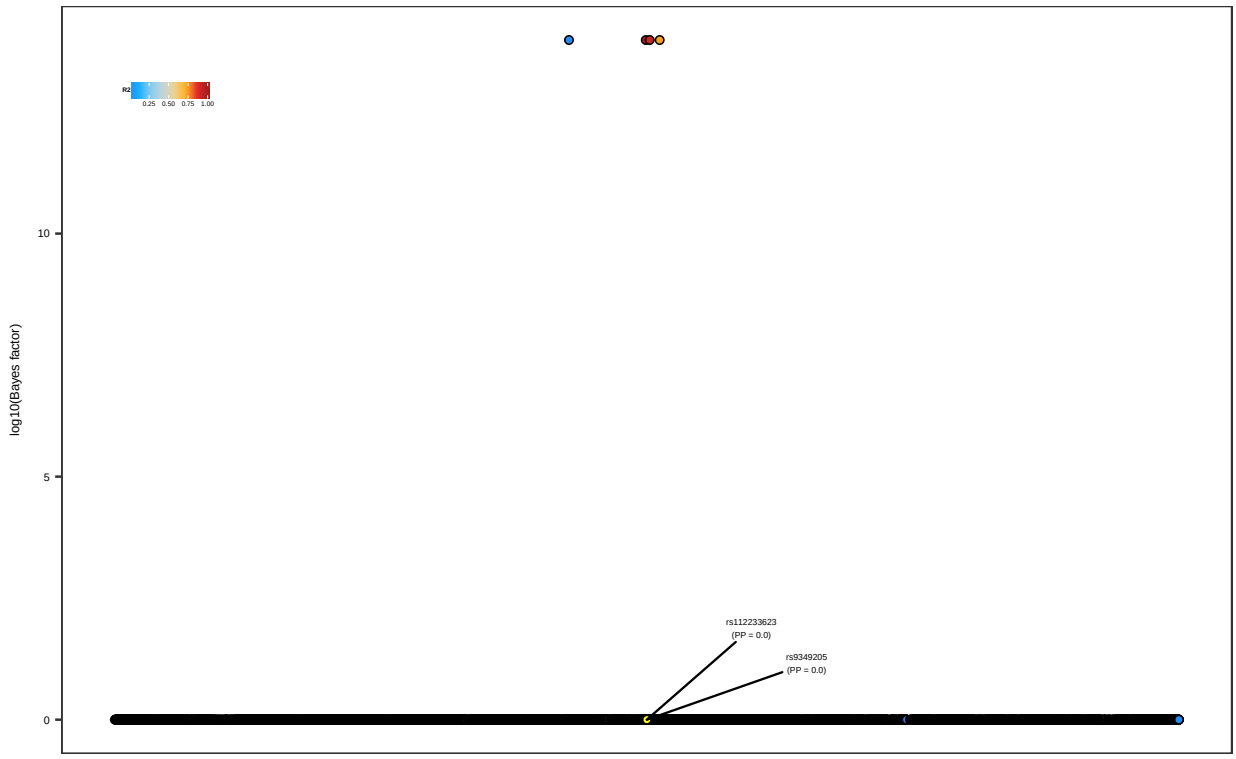
Overview of UK Biobank data



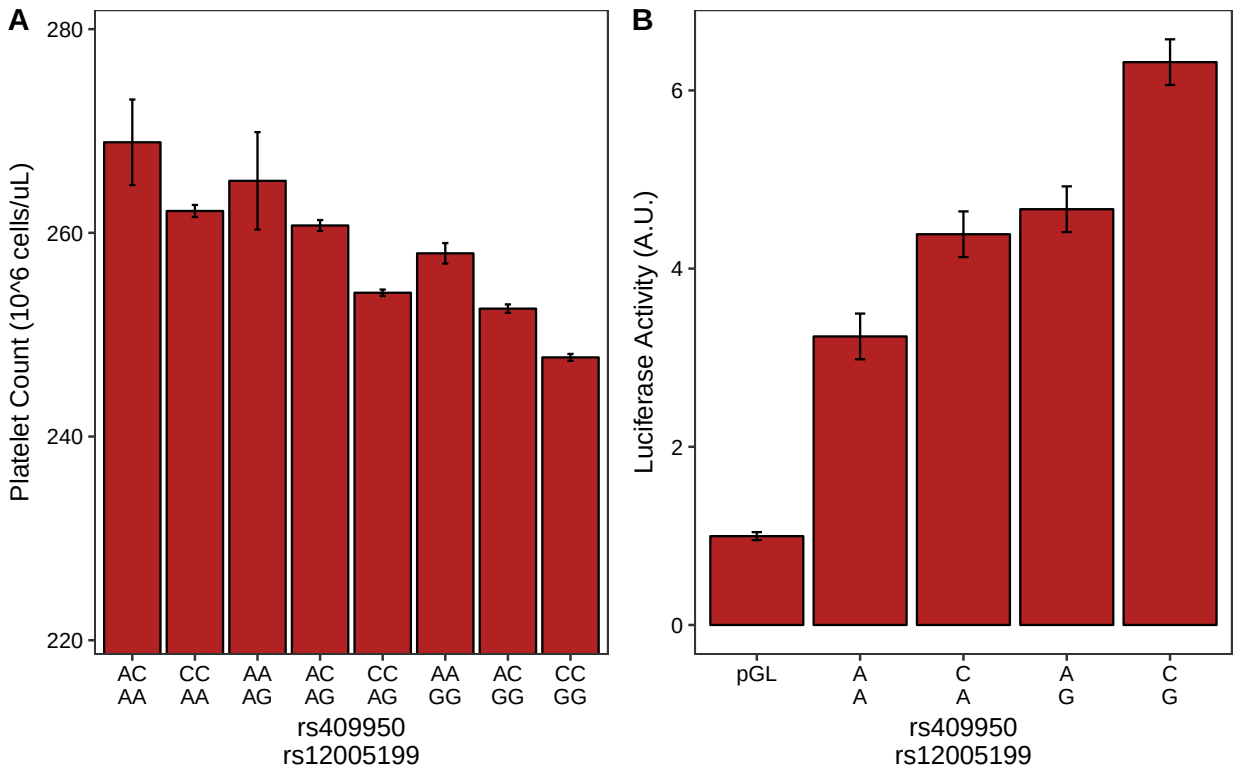
Supplemental Figure 1: Heritability estimates from LD Score Regression across 16 hematopoetic traits. The estimates of the narrow-sense SNP heritabilities are plotted in gray with their corresponding standard erors. Heritability estimates for all variants with fine-mapped posterior probability ¿ 0.001 are plotted in blue for each trait, and the proportions of total narrow-sense heritability captured by these fine-mapped variants (blue bar / gray bar) are indicated by the numbered labels.



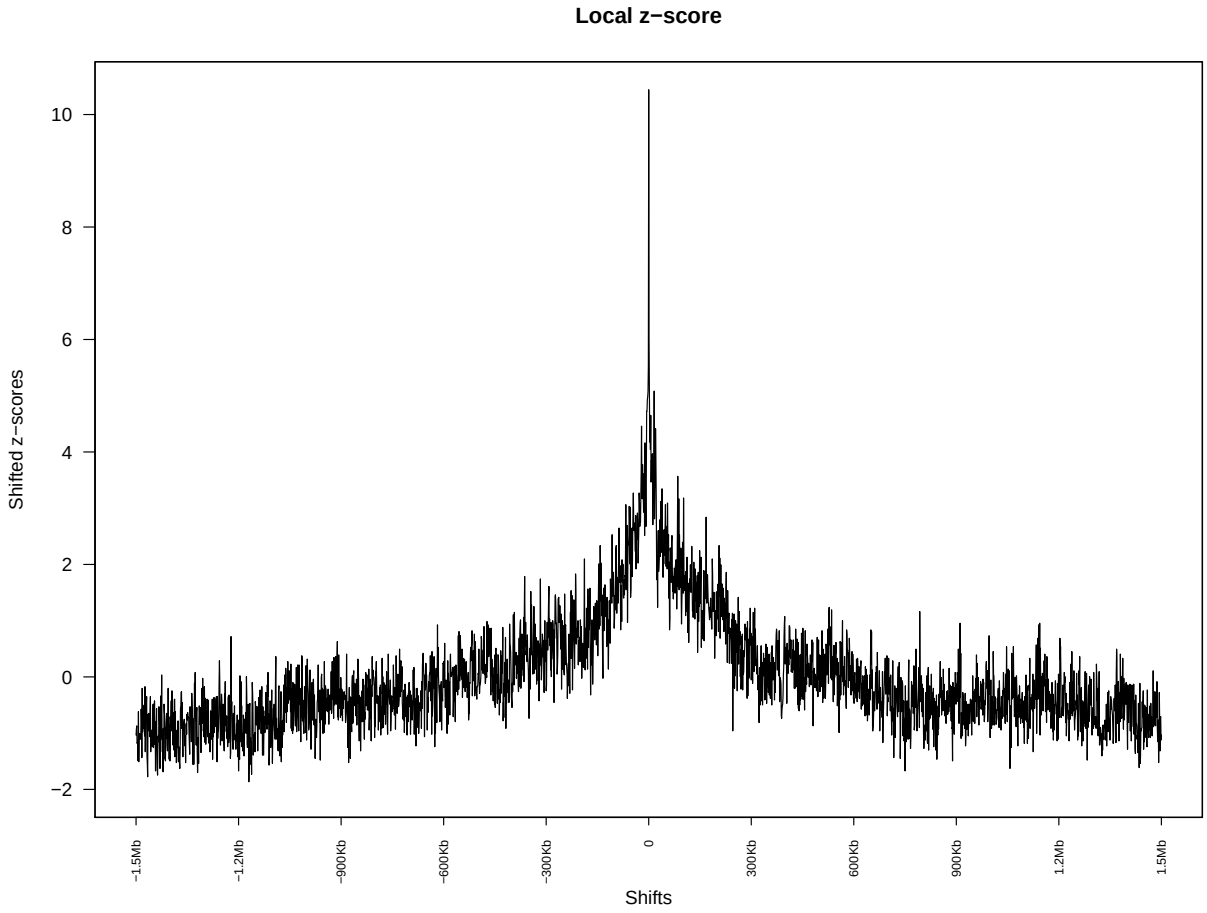
Supplemental Figure 2: Phenotypic and genetic correlations across the 16 traits examined.



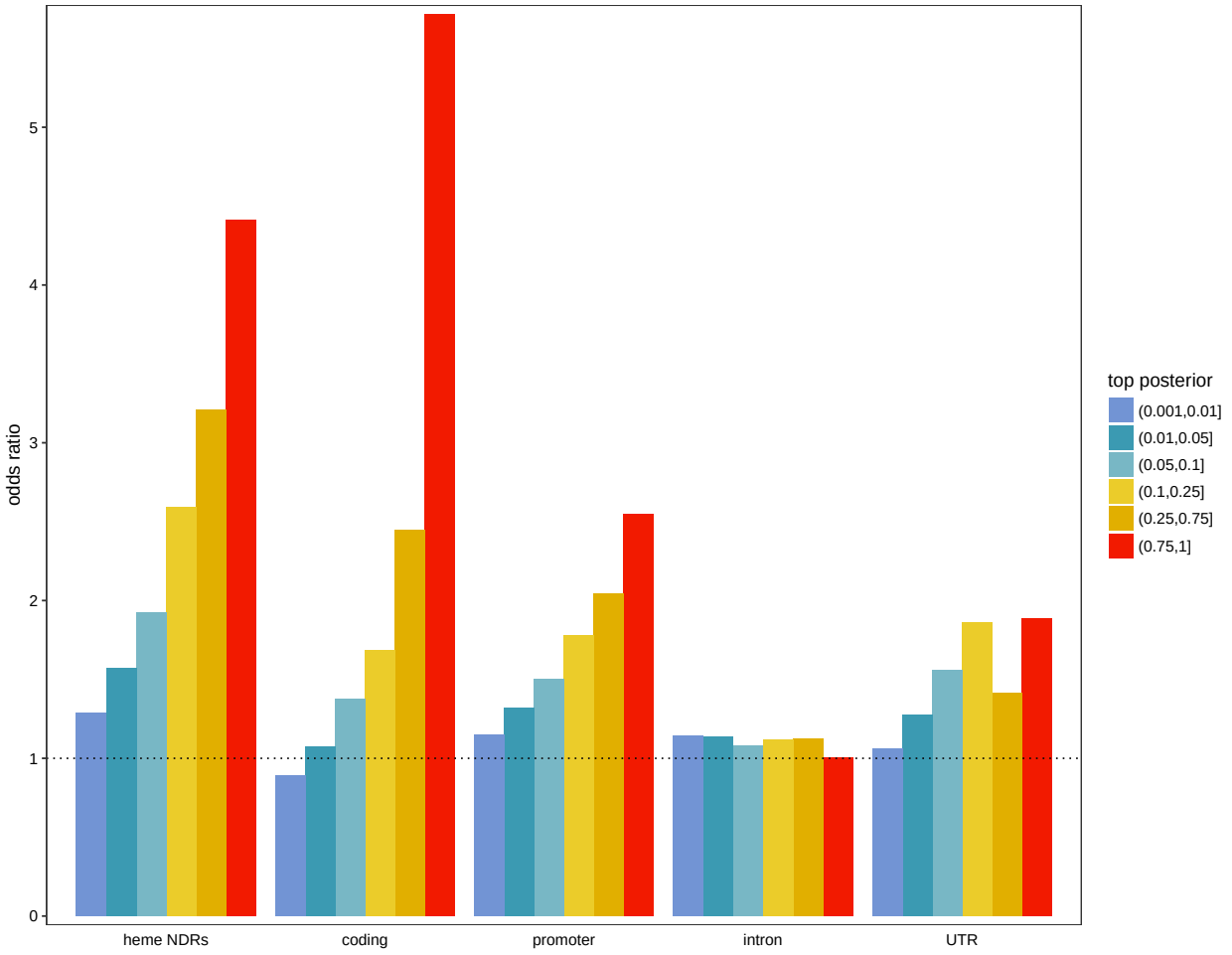
Supplemental Figure 3: Fine-mapped log10(Bayes factor) values for CCND3 variants, with LD estimated from a reference panel of 3,677 individuals from the UK10K cohort.



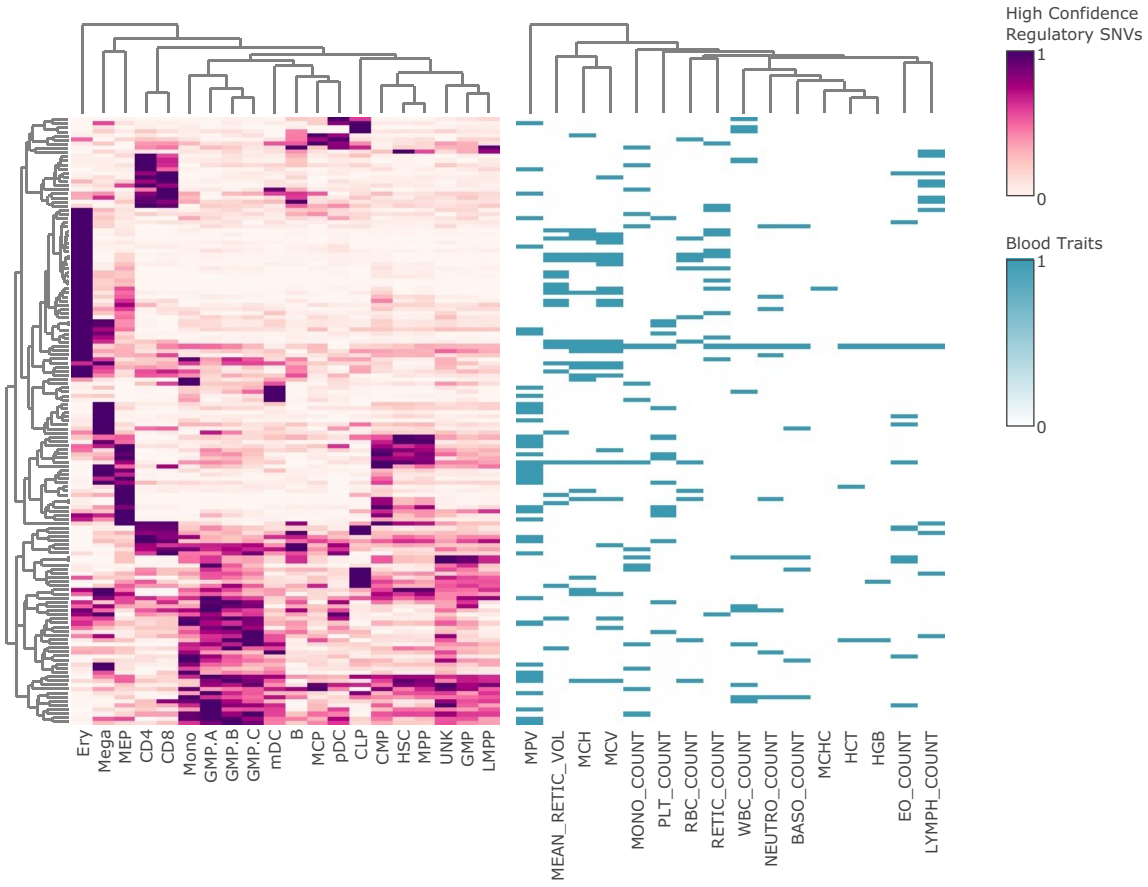
Supplemental Figure 4: Effects of AK3 variants rs409950 and rs12005199 on platelet count. (A) rs409950/rs12005199 haplotypes exert additive effects on platelet count amongst individuals from the UKBB GWAS. (B) Luciferase reporter assay corroborates additive effects of the two SNPs on transcription.



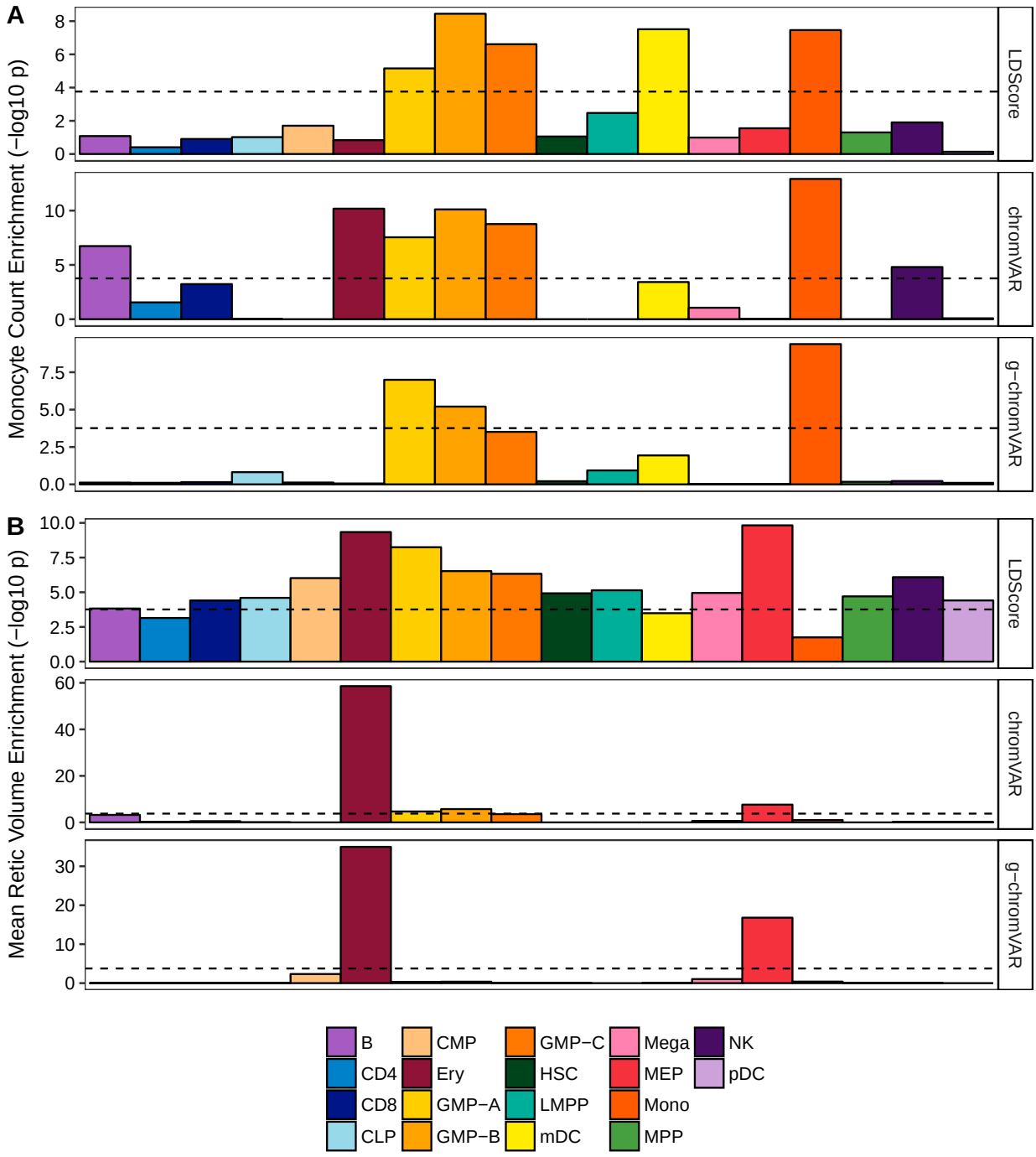
Supplemental Figure 5: Local z-scores for enrichment of hematopoietic nucleosome-depleted regions in the set of fine-mapped variants with posterior probability > 0.10.



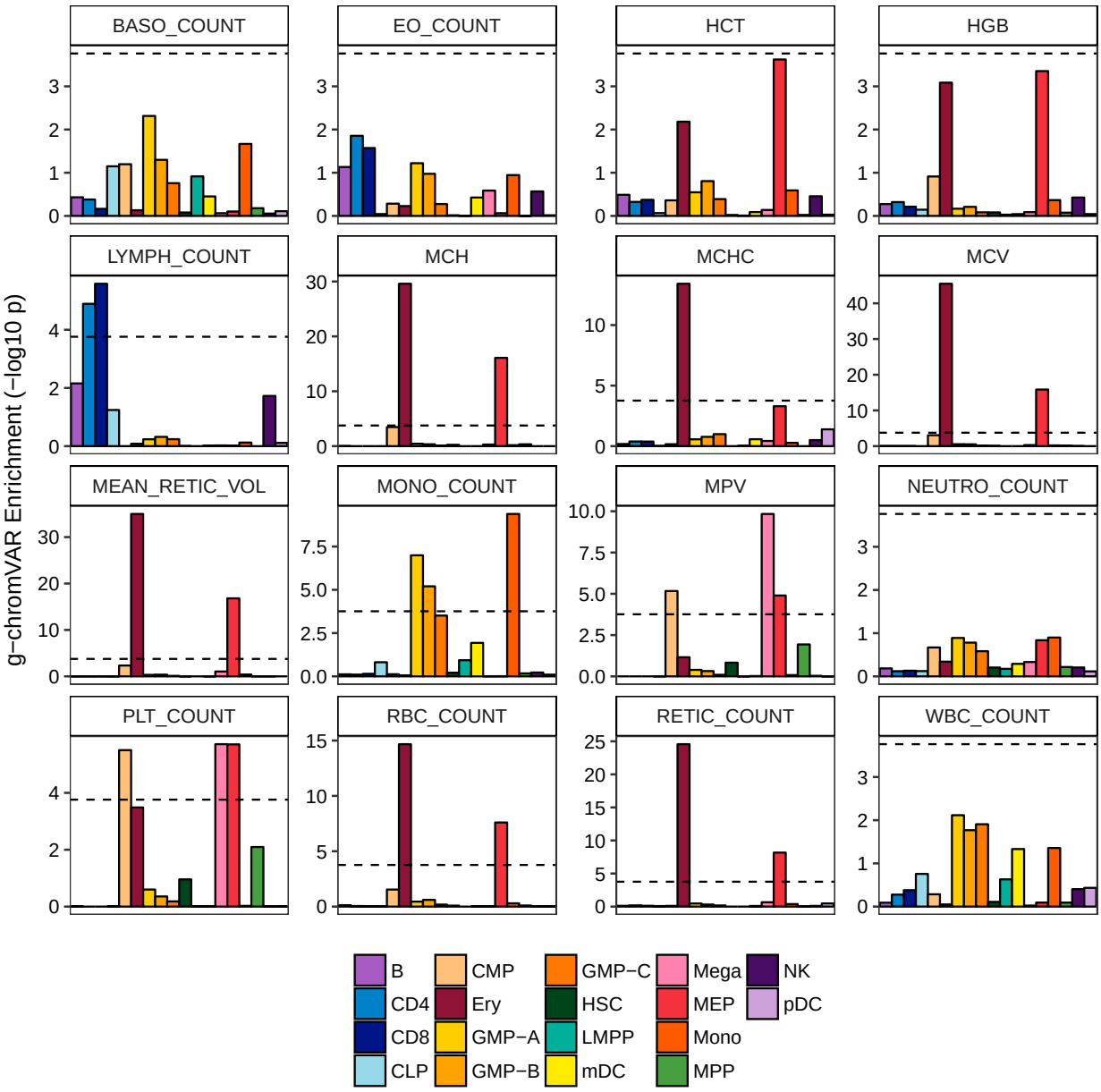
Supplemental Figure 6: Local annotation enrichments for fine-mapped variants, excluding all fine-map variants with $R^2 > 0.80$ to the sentinel variant of any region.



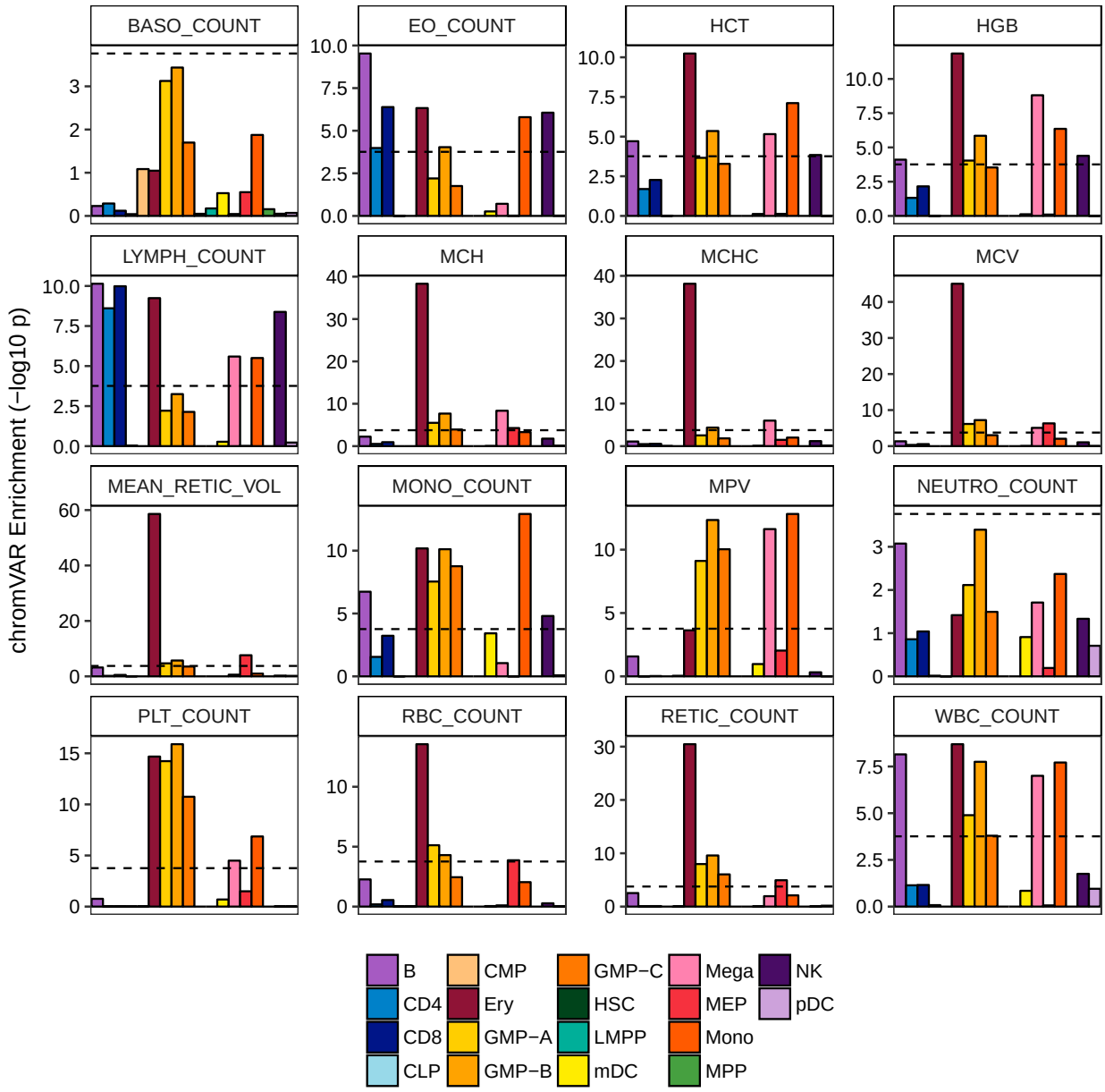
Supplemental Figure 7: Side-by-side heatmaps showing overlap of hematopoietic nucleosome-depleted regions by cell type with fine-mapped variants ($PP > 0.50$) by trait. The two heatmaps share a common y-axis of specific variants.



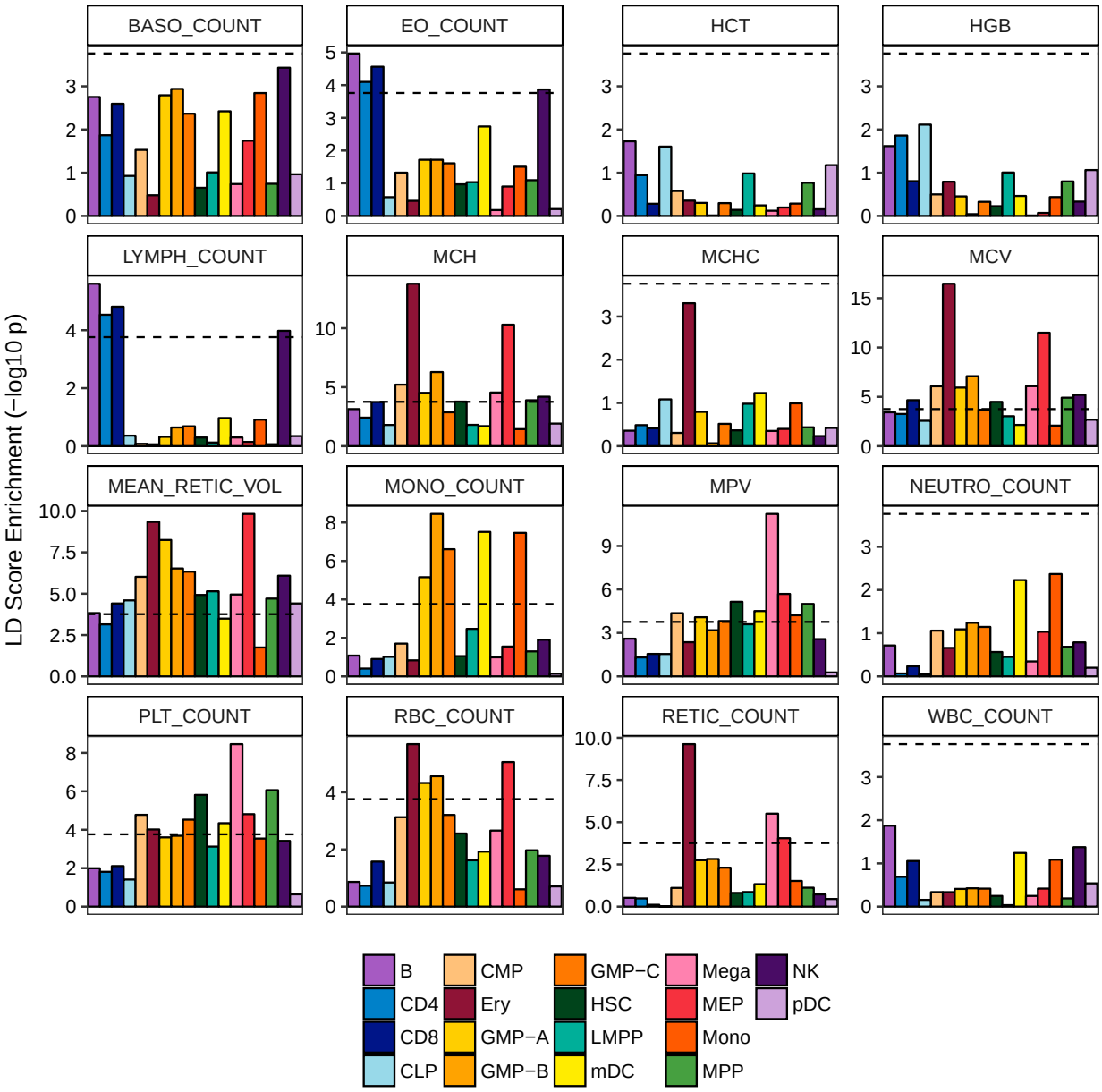
Supplemental Figure 8: Hematopoetic cell type enrichments for Mean Retic Volume and Monocyte count using various methods.



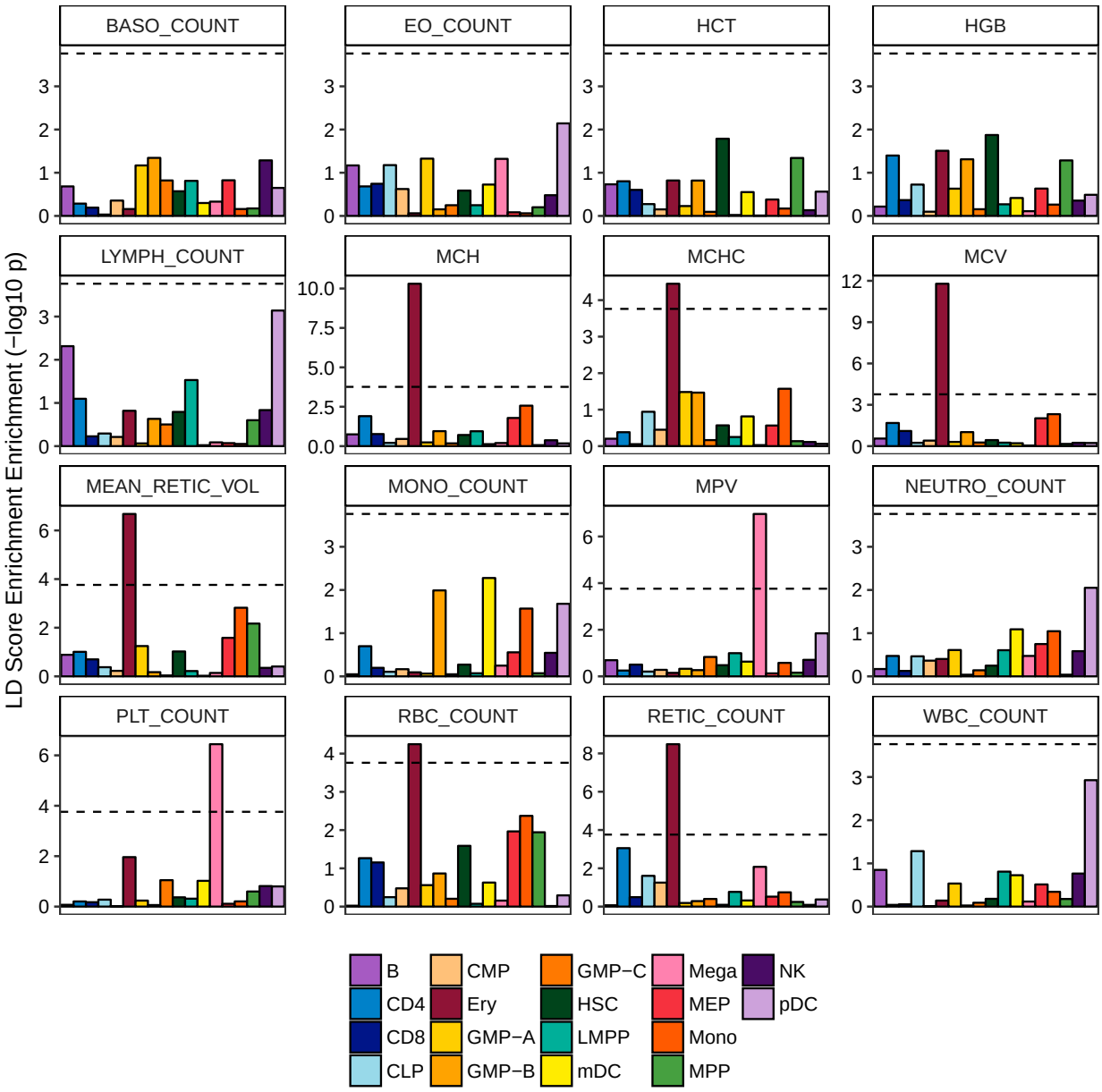
Supplemental Figure 9: All enrichments from g-chromVAR. The horizontal line shows a Bonferonni multiple testing adjusted threshold for statistical significance of enrichment.



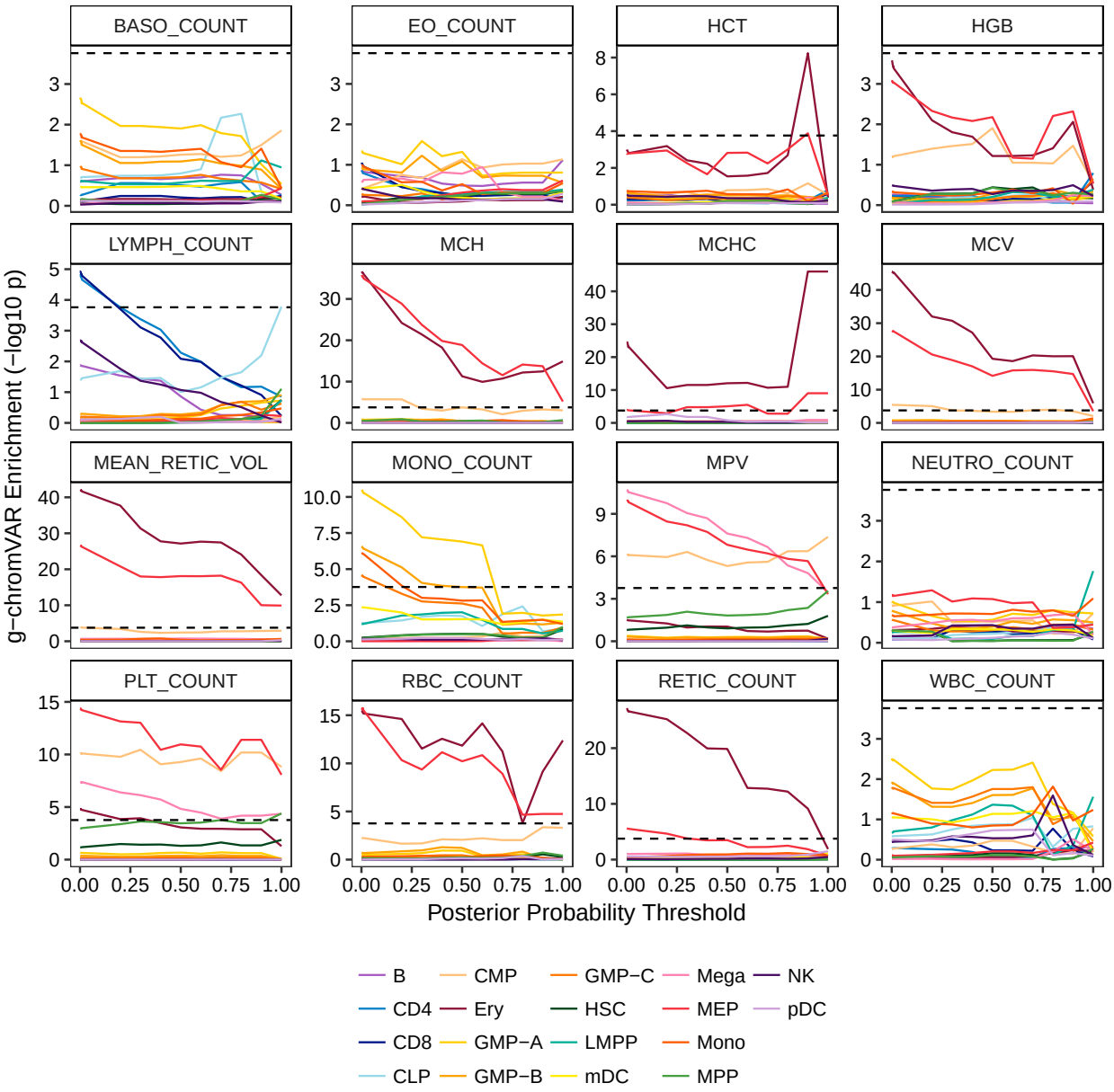
Supplemental Figure 10: All enrichments from chromVAR. The horizontal line shows a Bonferonni multiple testing adjusted threshold for statistical significance of enrichment.



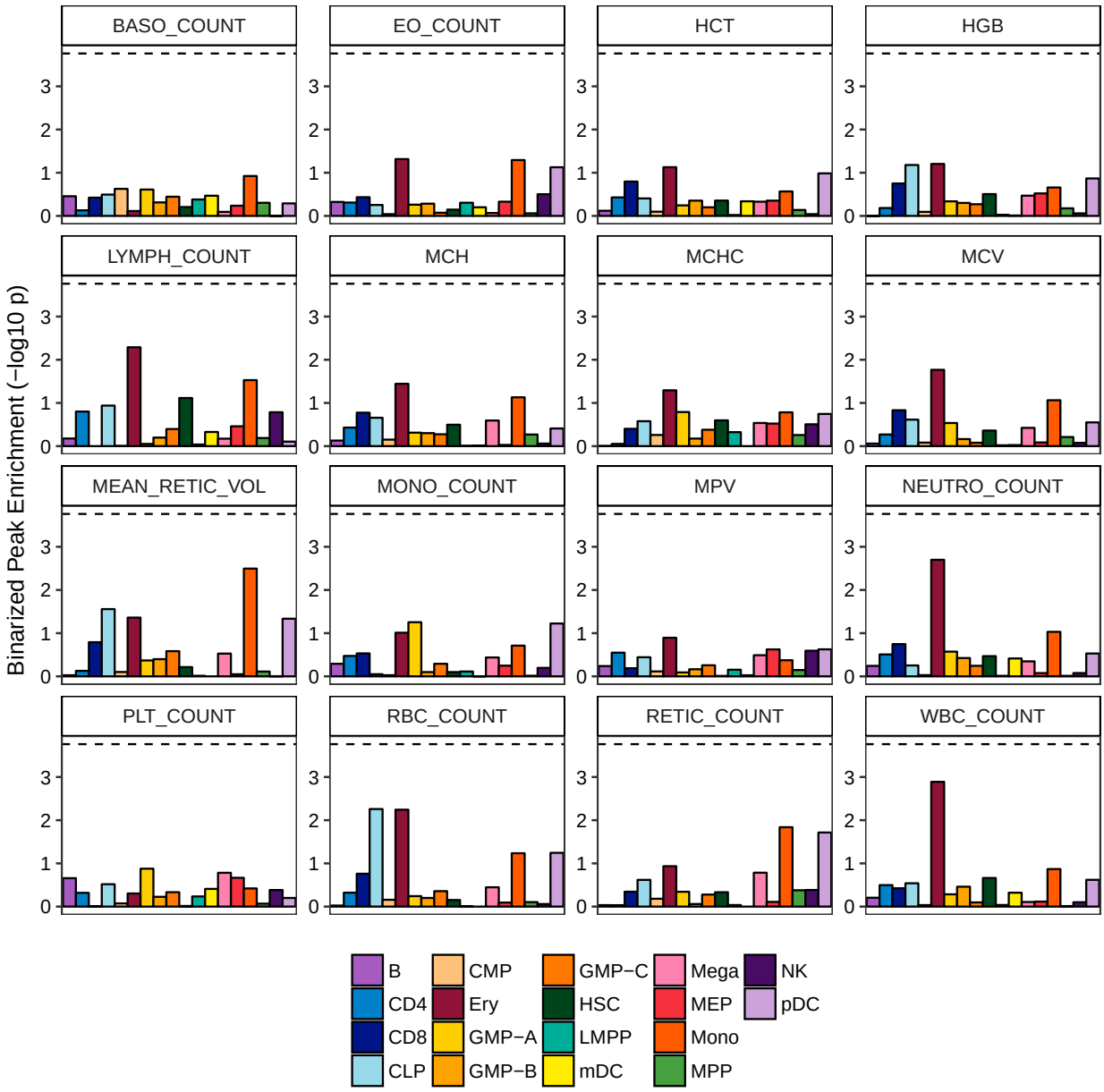
Supplemental Figure 11: All enrichments from LD score regression, calculated from the z-scores of the coefficients for each cell-type-specific annotation added separately to the baseline model. The horizontal line shows a Bonferonni multiple testing adjusted threshold for statistical significance of enrichment.



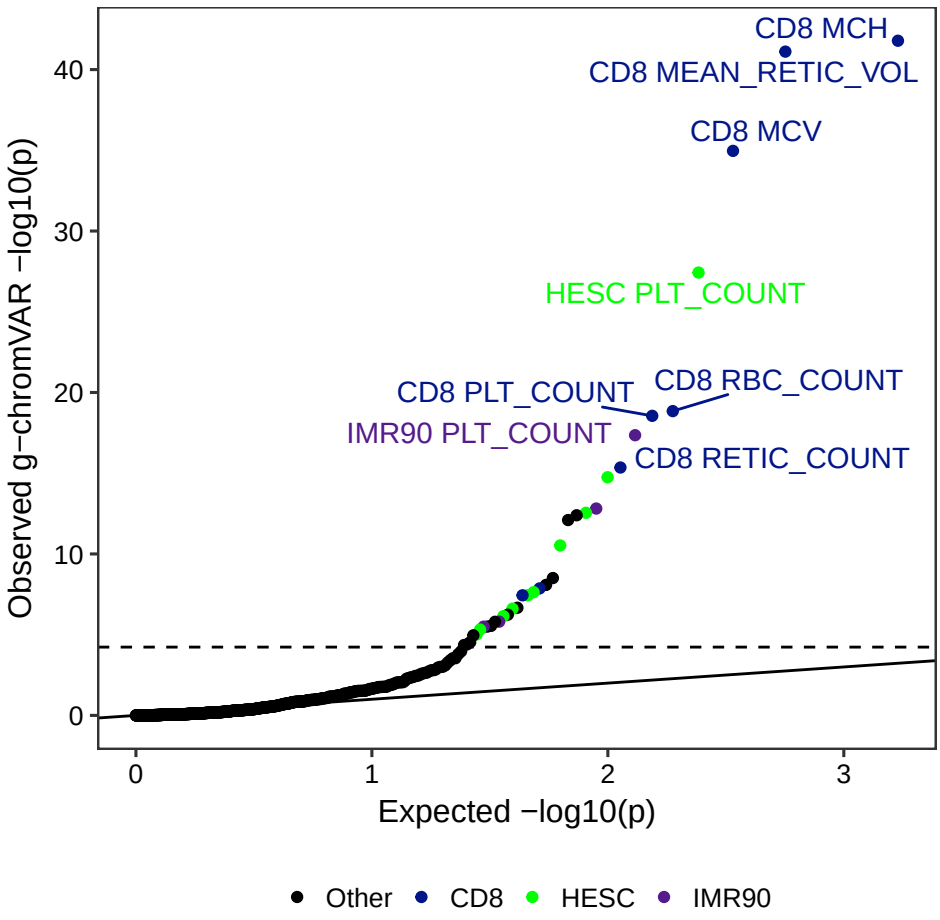
Supplemental Figure 12: All enrichments from LD score regression, calculated from coefficient z-scores after jointly adding all 18 cell-type-specific annotations to the baseline model at once. The horizontal line shows a Bonferonni multiple testing adjusted threshold for statistical significance of enrichment.



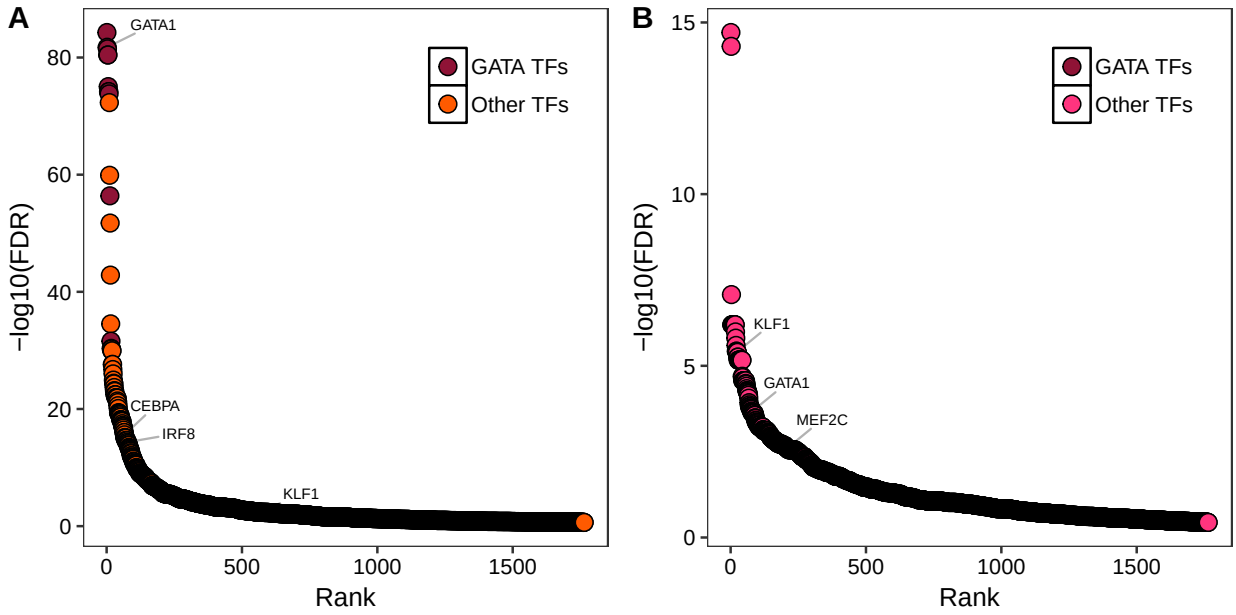
Supplemental Figure 13: Cell type - trait enrichments for g-chromVAR across different finemap variants posterior probability cutoffs. The horizontal line shows a Bonferonni multiple testing adjusted threshold for statistical significance of enrichment.



Supplemental Figure 14: Results of trait/cell type enrichments across hematopoesis using binarized peaks. These results demonstrate the importance of accounting for quantitative chromatin values in peaks to be powered for detecting enriched cell types and traits.



Supplemental Figure 15: QQplot of observed and expected enrichments using DNase hypersensitivity data and identical pre-processing for 53 tissue types from Roadmap. The top eight pairs of cell types and traits are annotated on the plot where 35 total pairs passed Bonferroni-adjusted significance.



Supplemental Figure 16: Two subpopulations of CMP and MEP cells were obtained by k-medoids clustering on ATAC principal components or g-ChromVAR enrichments, respectively. Rank-order plots showing transcription factor binding sites ranked by difference in chromVAR enrichment between the two clusters of (A) CMP and (B) MEP populations