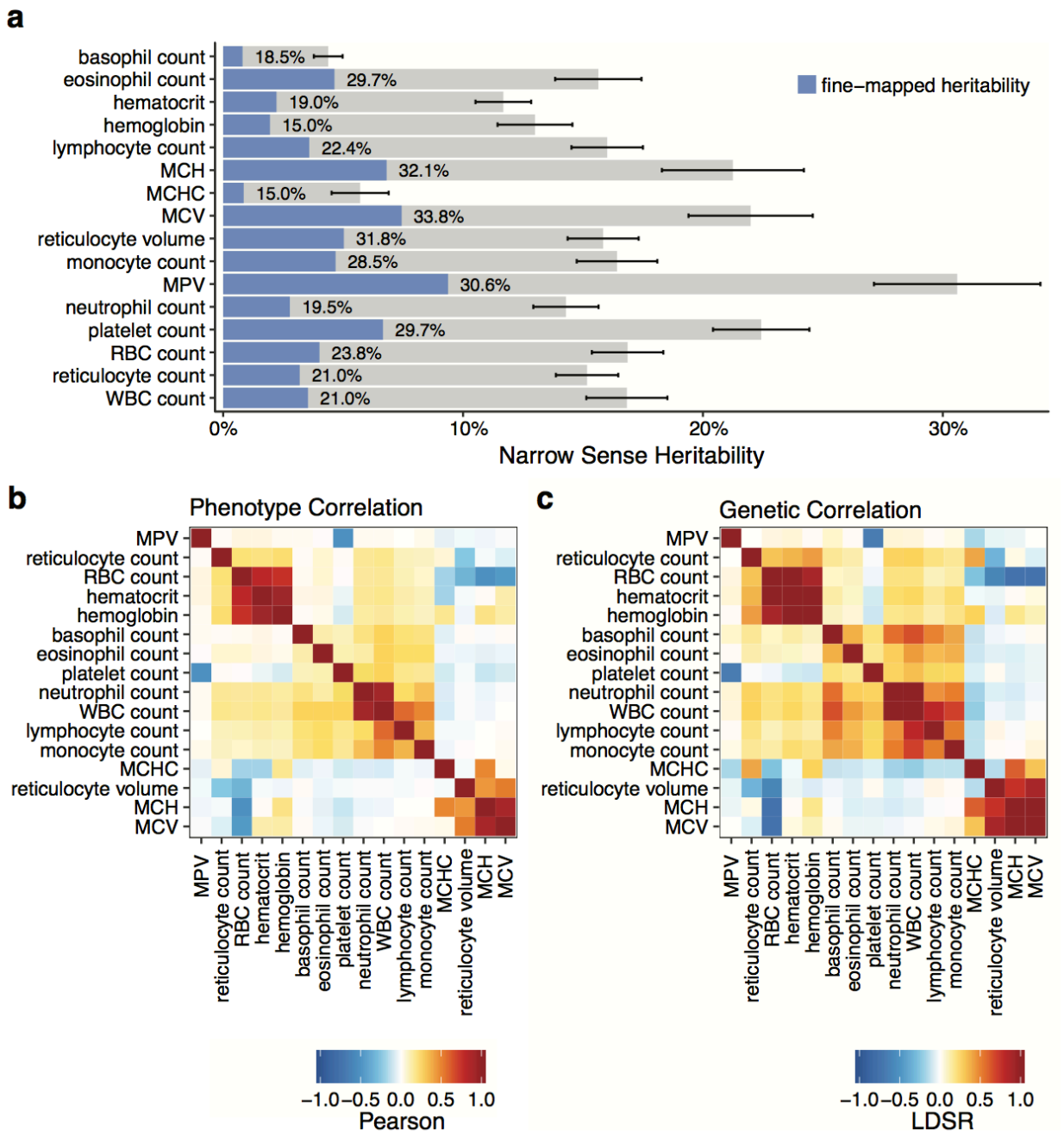


Figure S1. Narrow-sense heritability and genetic/phenotypic correlations between blood traits. (**a**) The estimates of the narrow-sense SNP heritabilities are plotted in gray with their corresponding standard errors. Heritability estimates for all variants with fine-mapped PP > 0.001 are plotted in 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. (**b**) Phenotypic and (**c**) genetic correlations across the 16 traits examined.

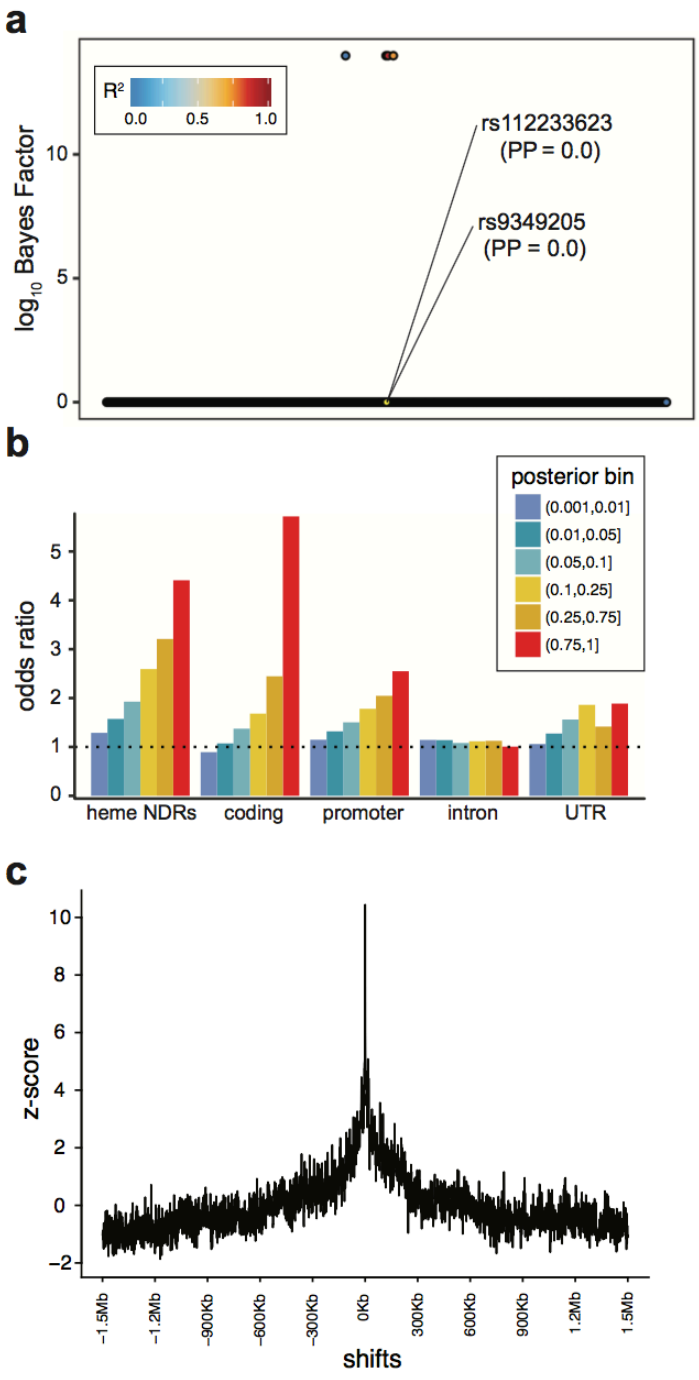


Figure S2. Additional fine mapping diagnostics. (**a**) Fine-mapped log10(Bayes factor) values for CCND3 variants fails to identify known functional variants when LD is estimated from a hard-called UKBB reference panel rather than using imputed genotype probabilities. (**b**) Local shifting enrichments of fine-mapped variants after excluding all variants with high correlation ($R^2 > 0.8$) to the sentinel variants. (**c**) Local z-scores for enrichment of hematopoietic nucleosome-depleted regions in the set of fine-mapped variants with posterior probability > 0.10.

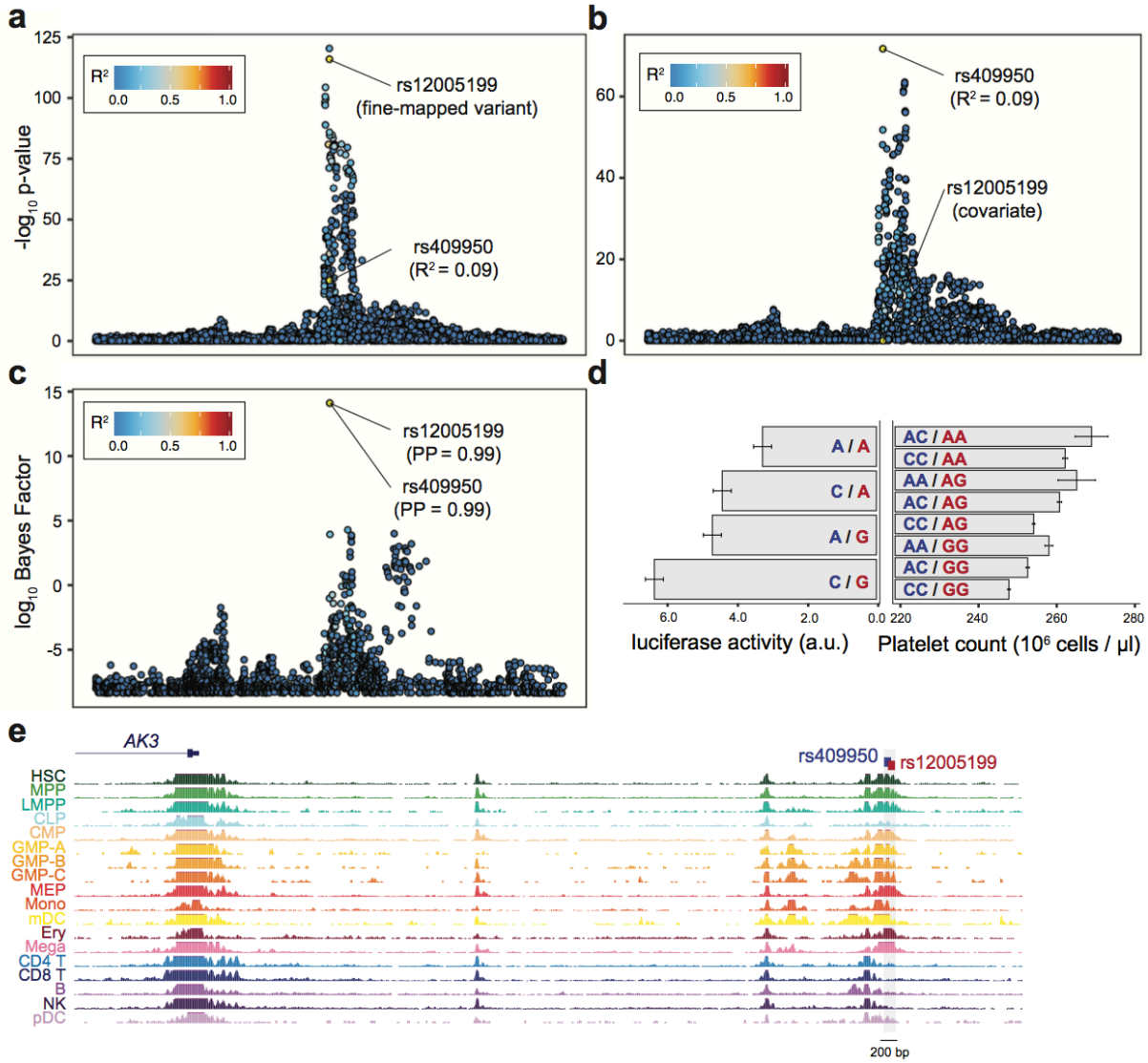
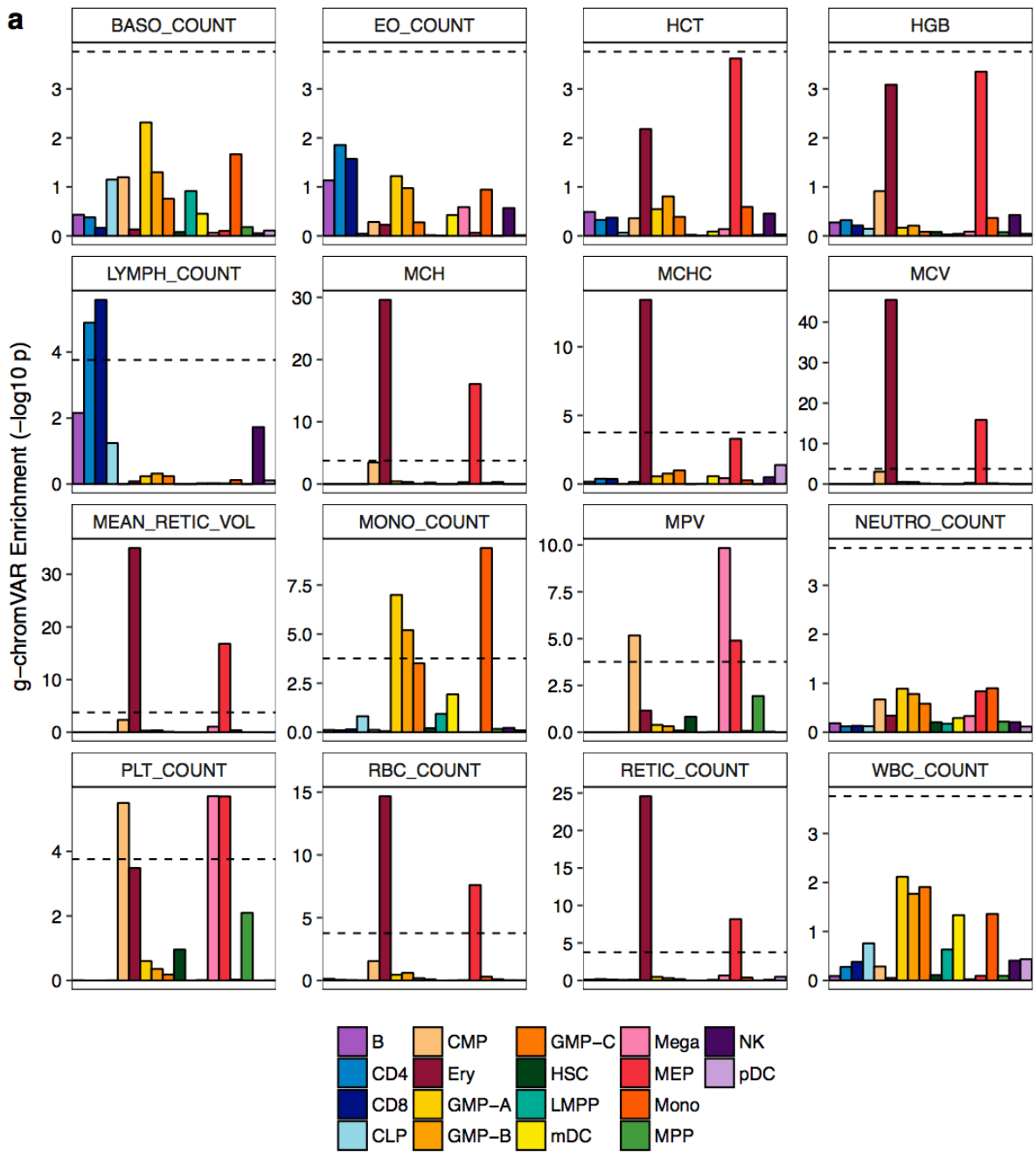
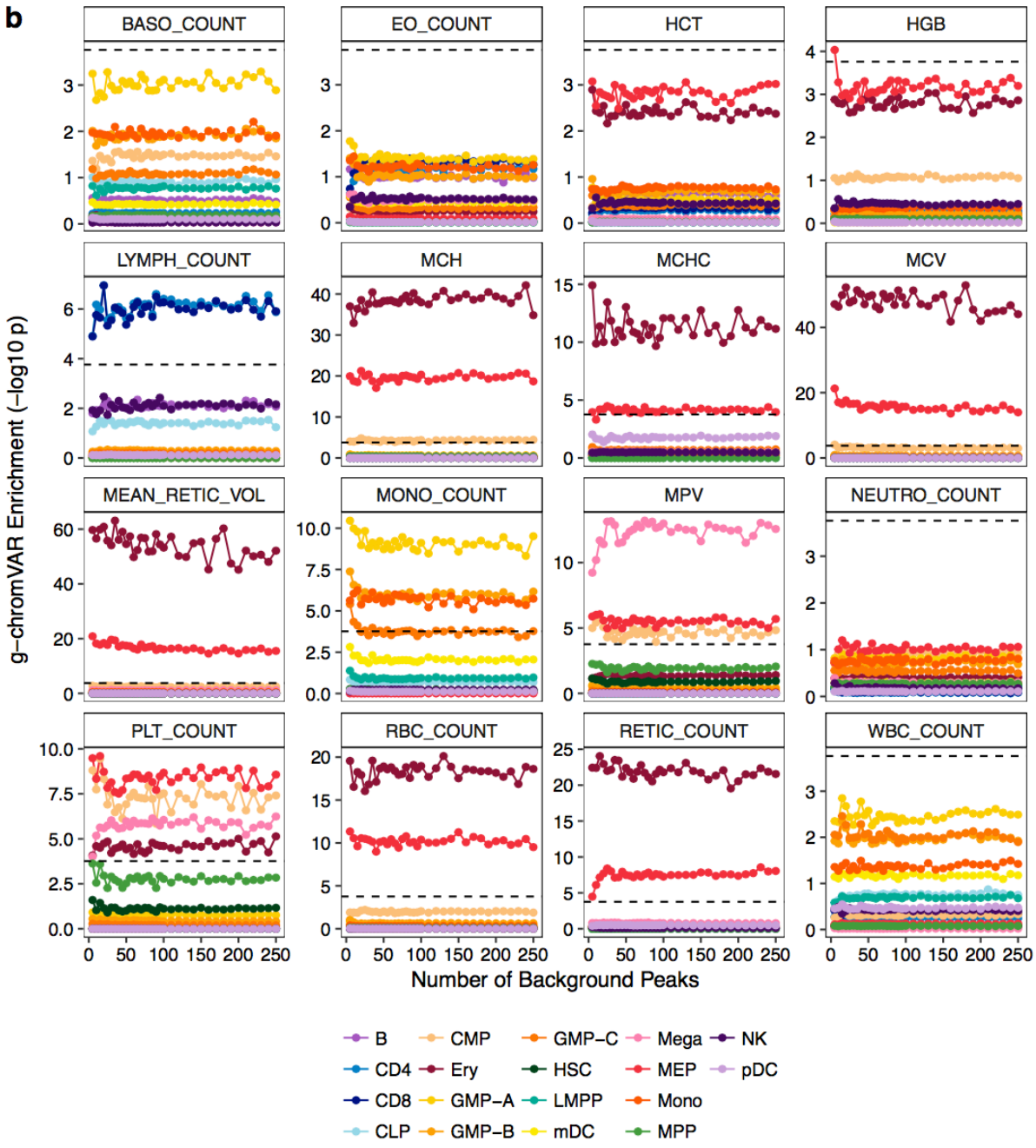
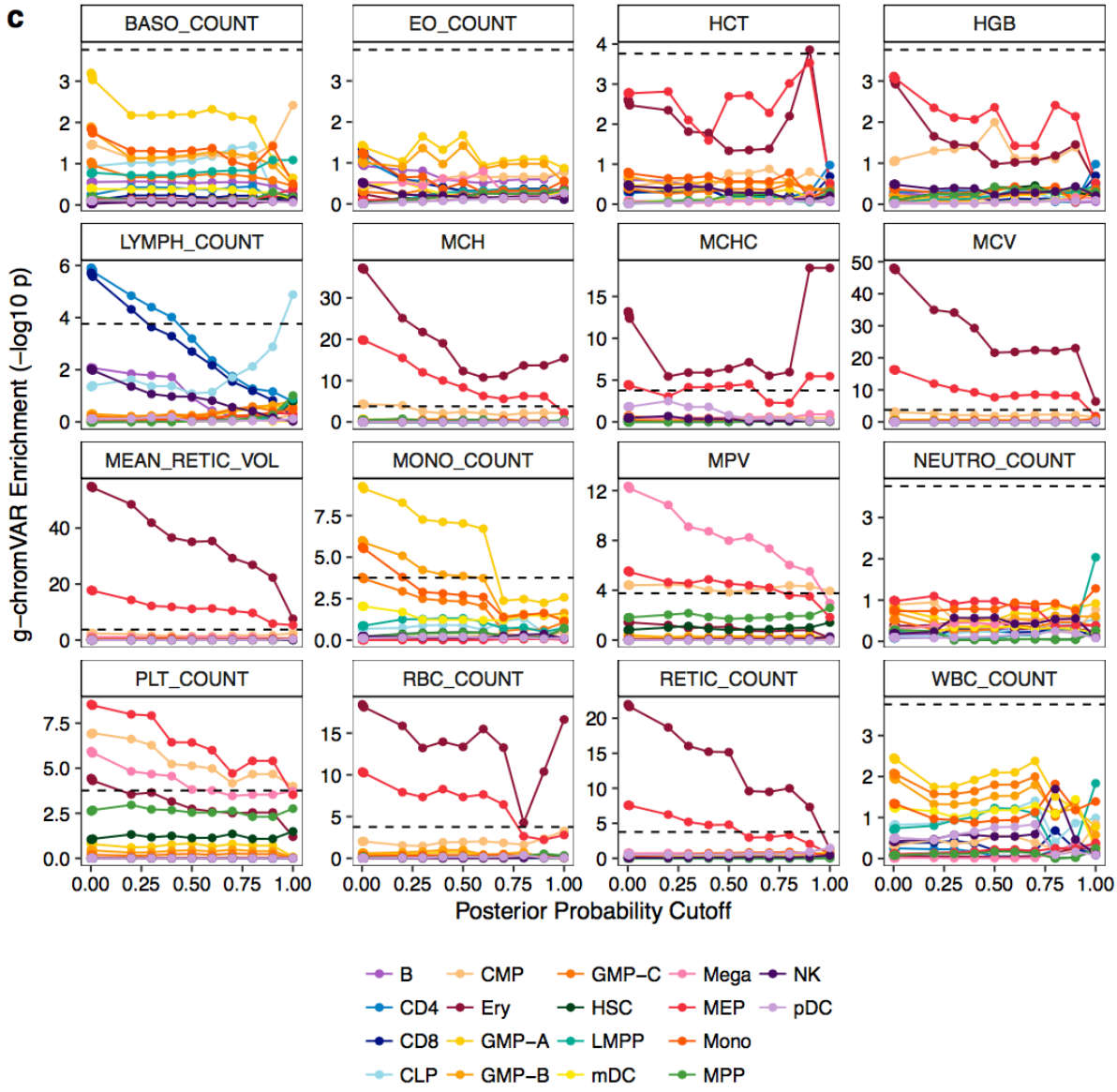


Figure S3. Characterization and validation of *AK3* upstream nucleosome-depleted region (NDR) with multiple causal variants. Regional association plots for RBC count (**a**) from the initial GWAS and (**b**) after conditioning on the top fine-mapped variant rs12005199. (**c**) Fine-mapping identifies two putative causal variants (rs49950 and rs12005199; PP > 0.999; 123 base pairs apart). (**d**) Luciferase reporters for four haplotypes corroborate independent additive effects of rs49950 and rs12005199. (**e**) Both putative causal variants lie within a megakaryocyte NDR, 123 base pairs apart.


a
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
g−chromVAR Enrichment (−log10 p)
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC


b
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
g−chromVAR Enrichment (−log10 p)
Number of Background Peaks
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC


c
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
g−chromVAR Enrichment (−log10 p)
Posterior Probability Cutoff
0.00 0.25 0.50 0.75 1.00
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC

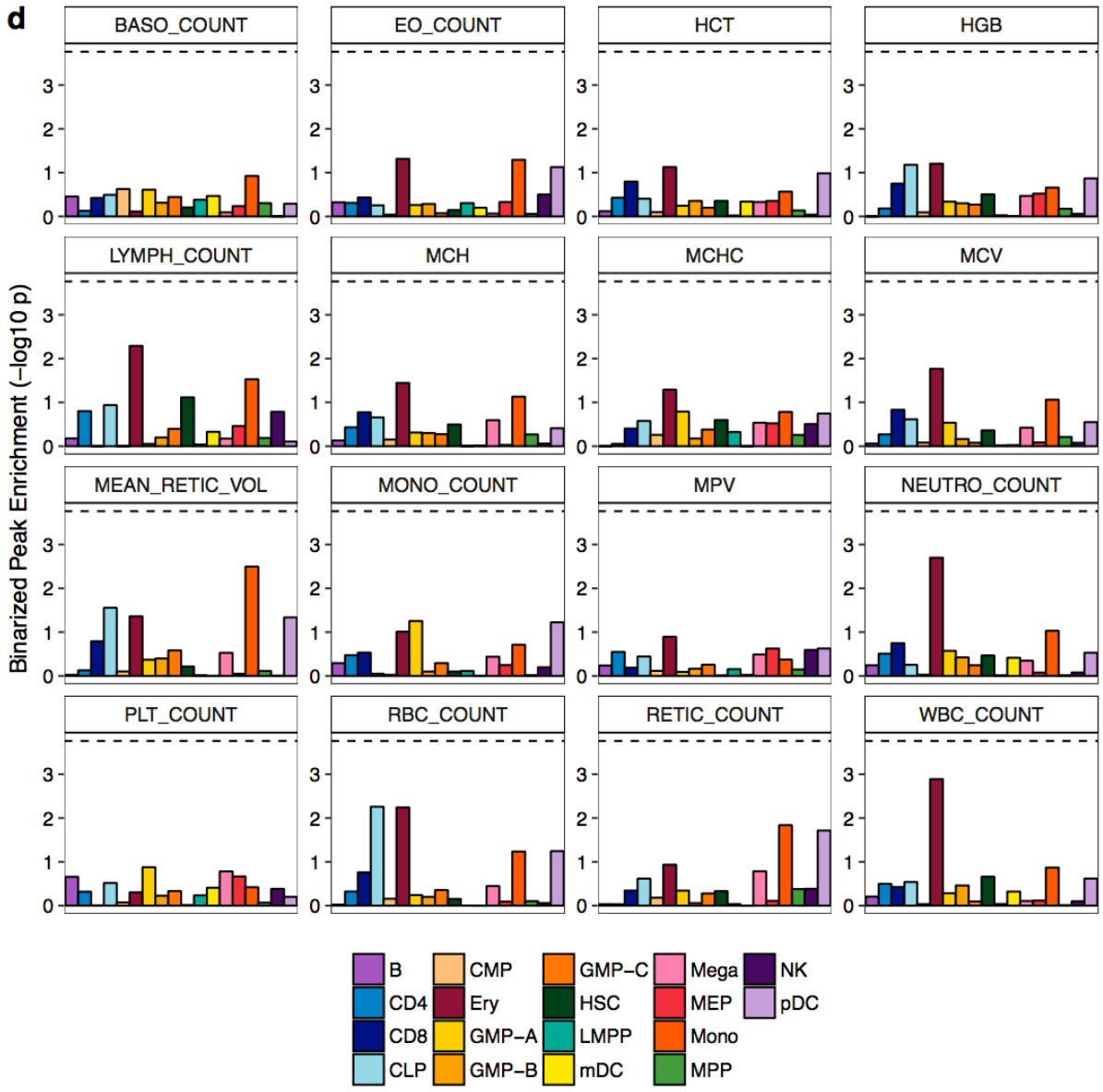


Figure S4. Application of g-chromVAR and parameterization characteristics. (**a**) All trait / cell type pairs scored by g-chromVAR with the Bonferroni-adjusted significance level indicated by the dotted line. (**b**) Characterization of variable background peak numbers in g-chromVAR. By default, 50 background peaks are used per analysis peak. (**c**) Enrichments from g-chromVAR using varied PP cutoffs. Our analysis used PP > 0.001 for all computations. (**d**) Results from running g-chromVAR on binarized peak counts (0/1) show no cell type-trait pairs passing Bonferroni-adjusted significance levels.

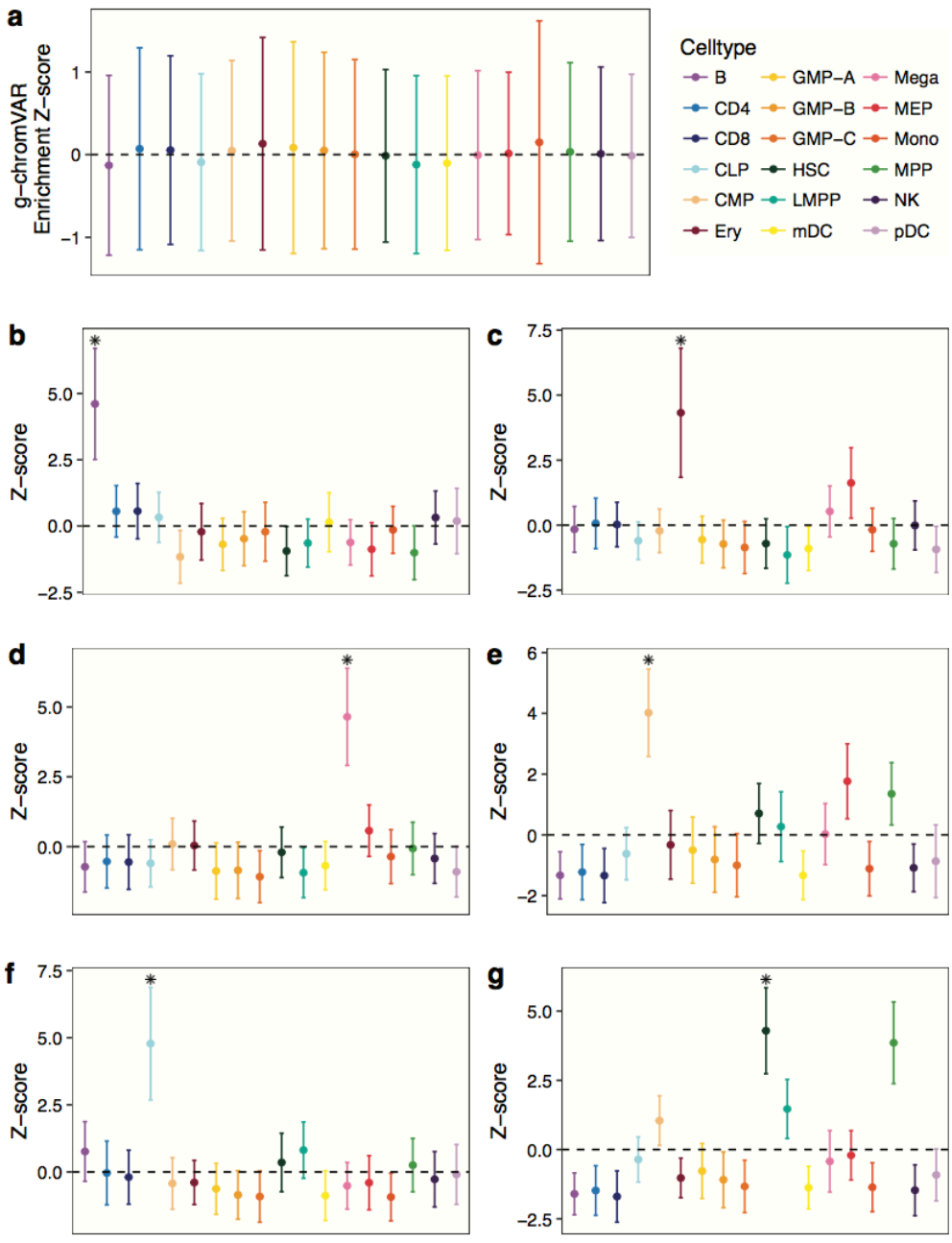
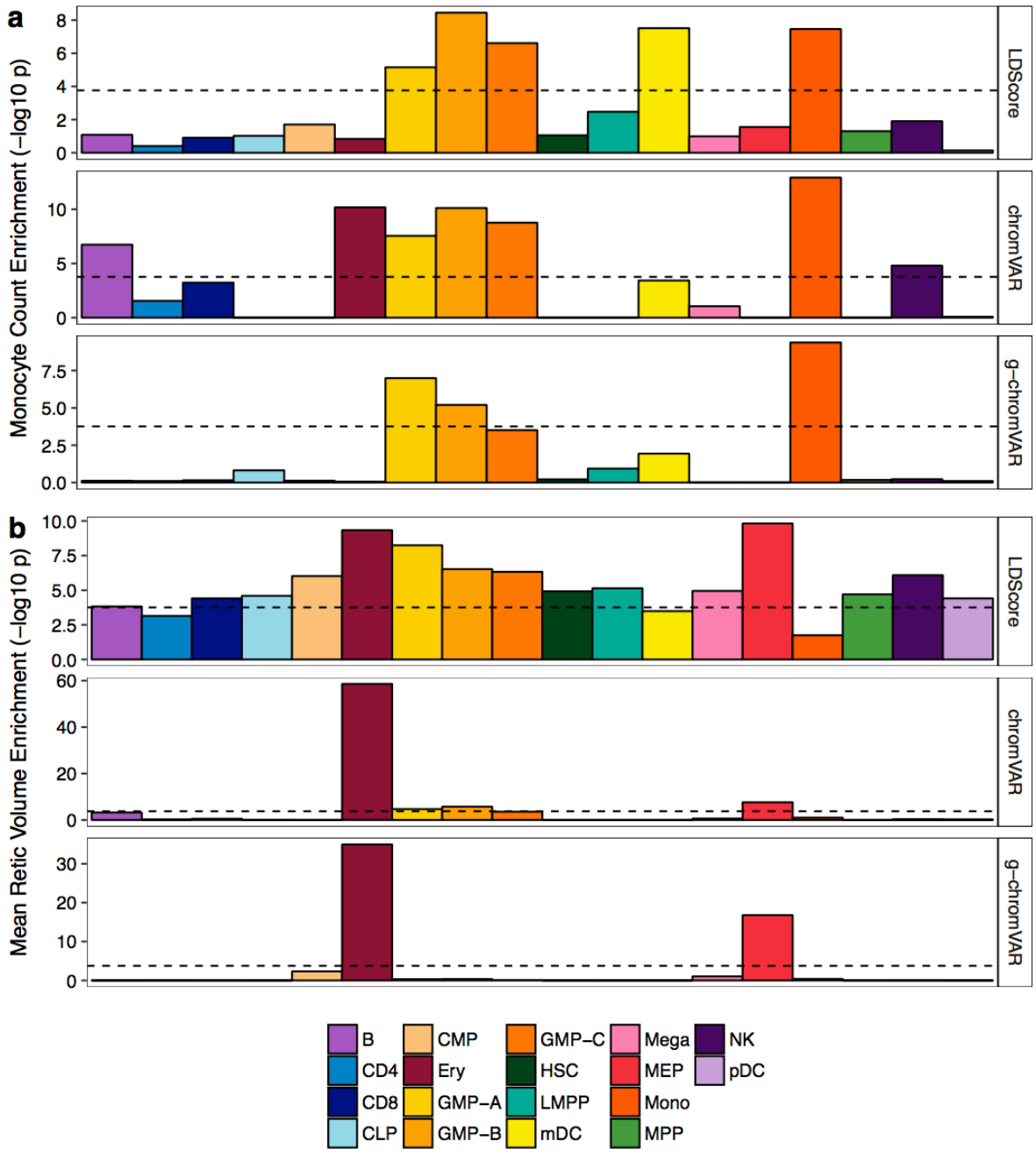
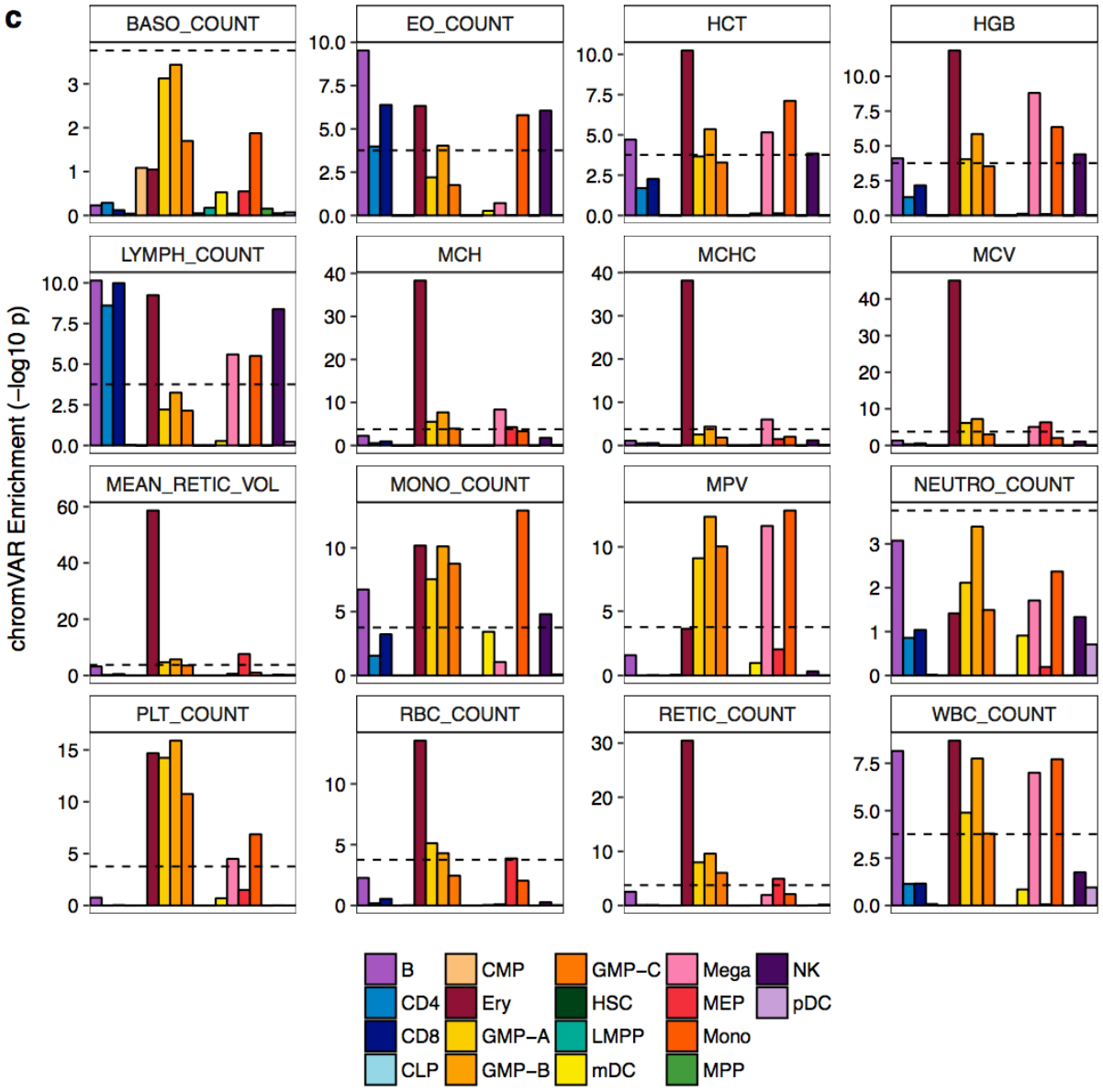
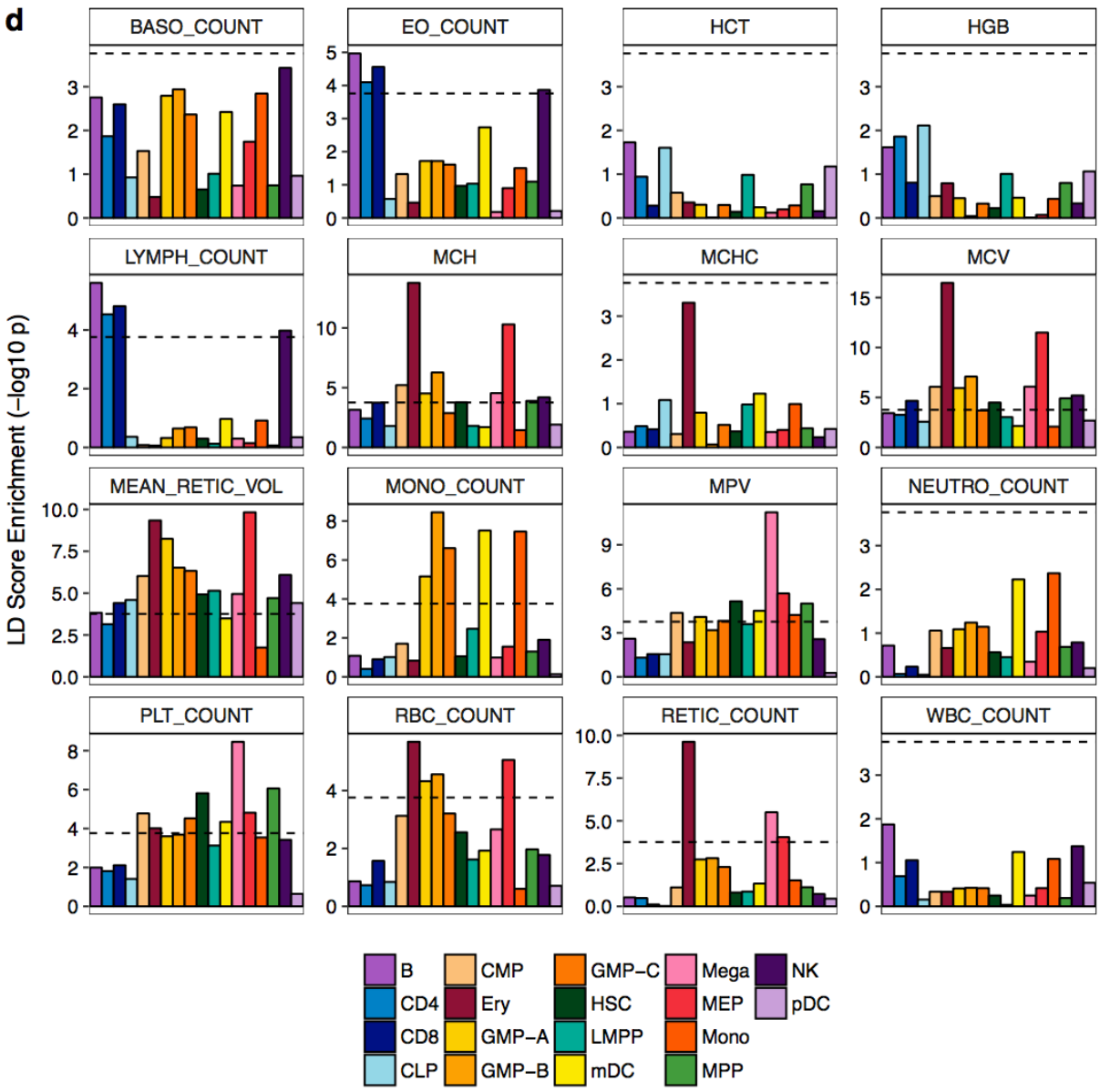
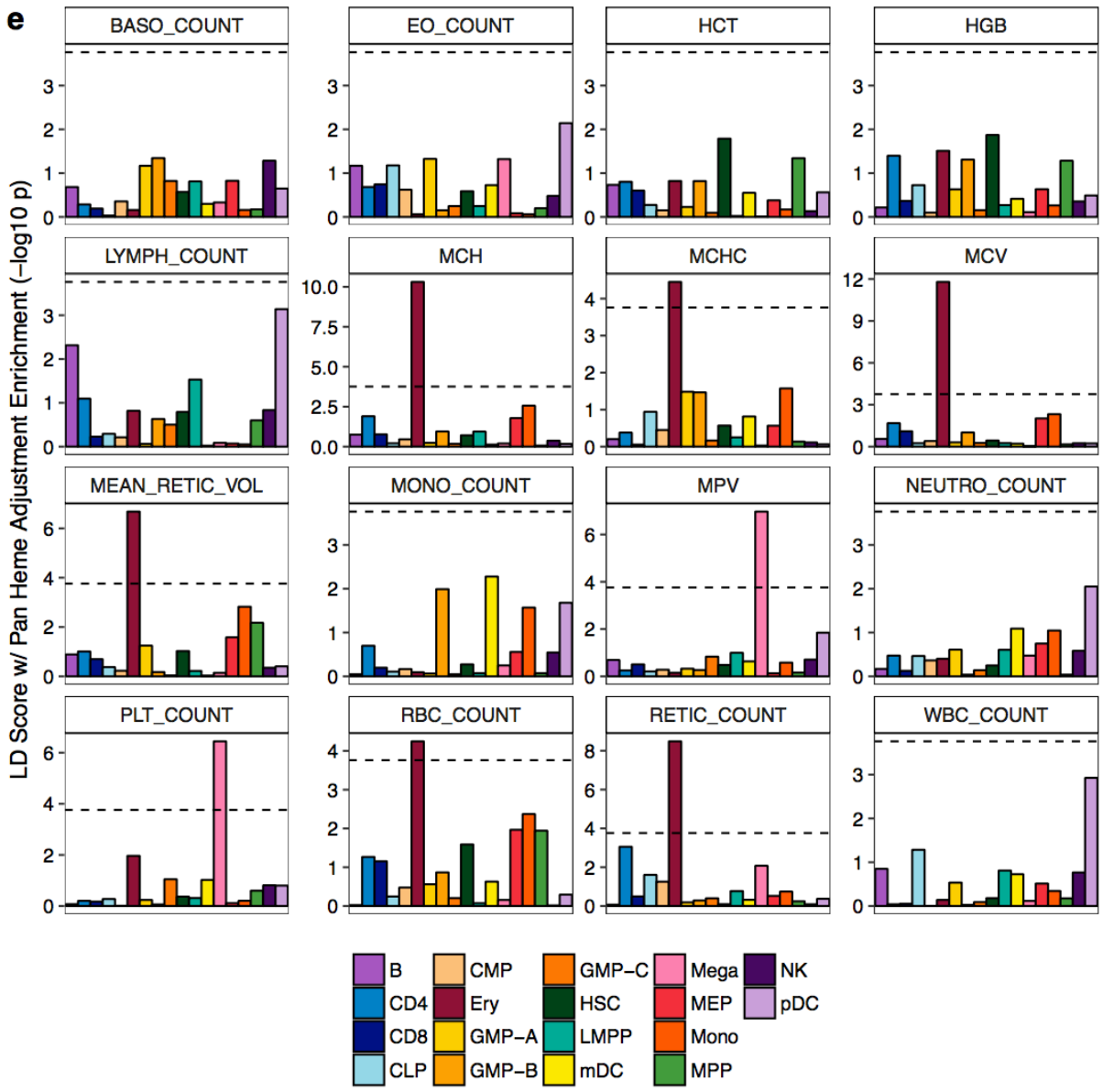


Figure S5. Simulation diagnostics for g-chromVAR. Each plot represents 100 simulations per simulated enrichment. (**a**) Mean and standard deviation of enrichments under a null phenotype simulation in which no cell types are enriched for the simulated phenotype. (**b-g**) Simulation results in which an arbitrary phenotype is enriched for a designated cell type indicated by the asterisk (*). In general, true enrichments for terminally differentiated cell types (**b-d**) and late-stage progenitors (**e,f**) can be reliably detected using g-chromVAR. For early progenitors such as HSCs in (**g**), MPPs also appear enriched due to the high similarity of the chromatin accessibility profiles of these two cell populations.


a
Monocyte Count Enrichment (−log10 p)
LDScore
chromVAR
g−chromVAR
b
Mean Retic Volume Enrichment (−log10 p)
LDScore
chromVAR
g−chromVAR
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC


c
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
chromVAR Enrichment (−log10 p)
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC


d
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
LD Score Enrichment (−log10 p)
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC


e
BASO_COUNT
EO_COUNT
HCT
HGB
LYMPH_COUNT
MCH
MCHC
MCV
MEAN_RETIC_VOL
MONO_COUNT
MPV
NEUTRO_COUNT
PLT_COUNT
RBC_COUNT
RETIC_COUNT
WBC_COUNT
LD Score w/ Pan Heme Adjustment Enrichment (−log10 p)
B
CD4
CD8
CLP
CMP
Ery
GMP−A
GMP−B
GMP−C
HSC
LMPP
mDC
Mega
MEP
Mono
MPP
NK
pDC

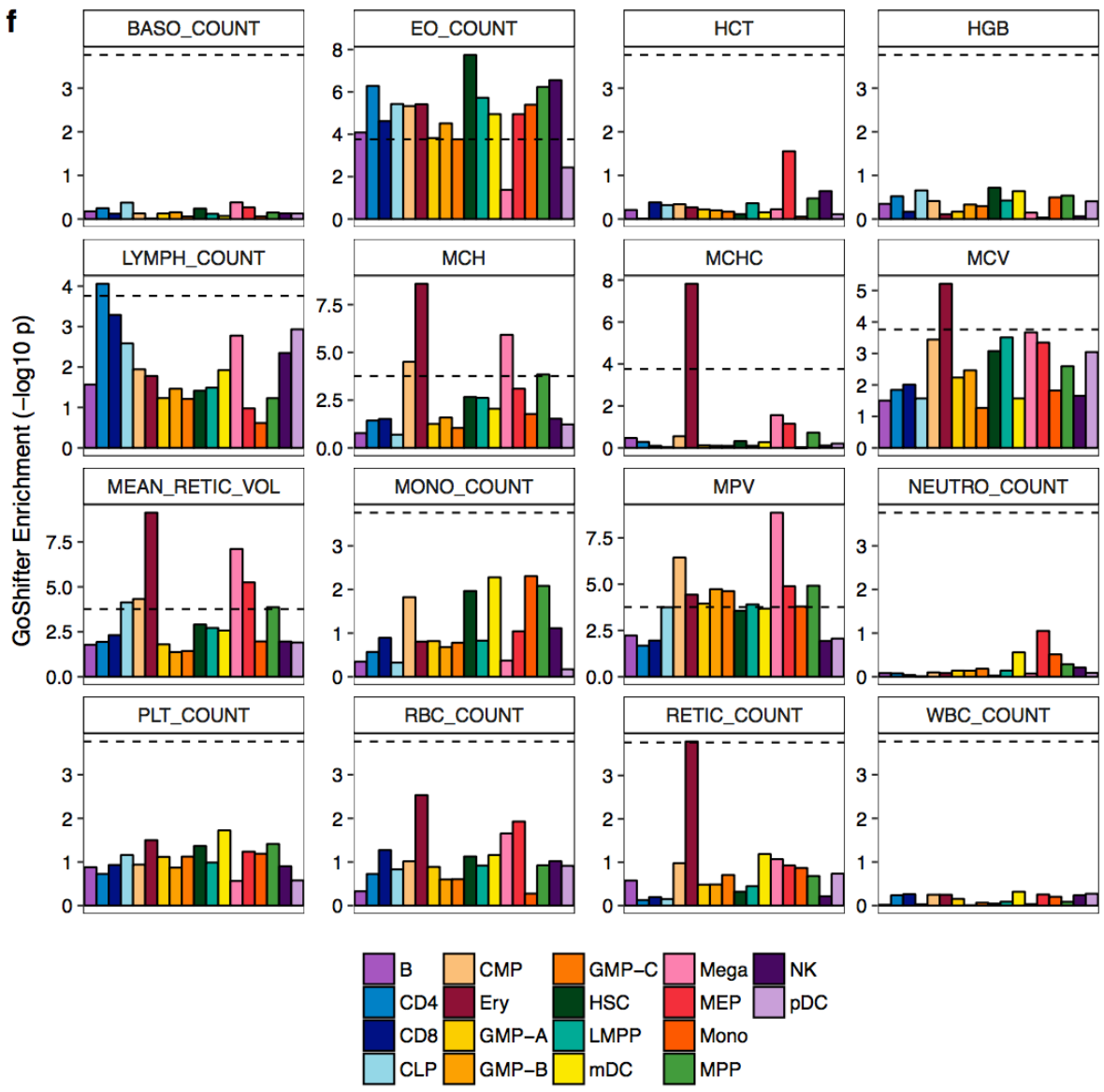


Figure S6. Comparison of g-chromVAR to other methods. (**a,b**) Selected traits (monocyte count and mean reticulocyte respectively) for all hematopoietic enrichments for g-chromVAR, chromVAR, and LD score. Bonferroni-adjusted significance level indicated by the dotted line for all figure panels. Complete visual representation of all trait / cell type pairs scored by (**c**) chromVAR, (**d**) LD score regression, (**e**) LD score regression using an additional covariate for all hematopoietic peaks, and (**f**) GoShifter are shown.

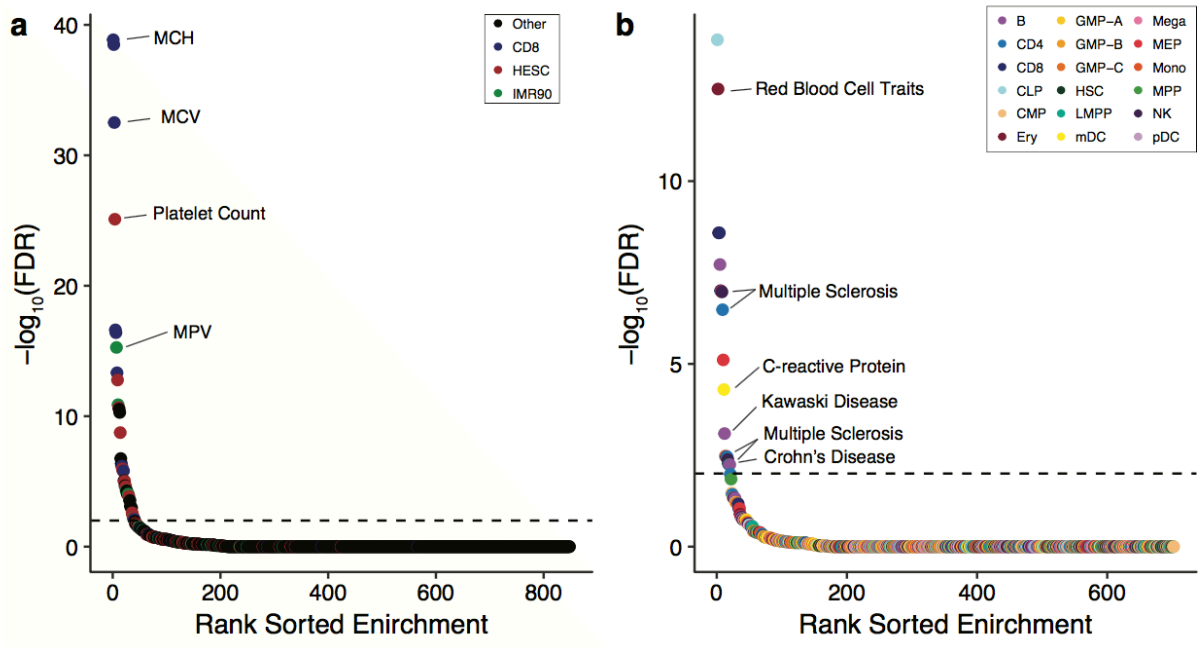


Figure S7. Summary of g-chromVAR application to (**a**) Roadmap DNase Hypersensitivity data with our 16 fine-mapped blood traits and (**b**) our 18 hematopoietic chromatin accessibility profiles with 39 predominately immune-related disorders previously fine-mapped using PICS (allowing for one causal variant per locus). Both plots show the rank ordered enrichment after a false discovery rate (FDR) correction in which the dotted line indicates a 1% FDR.

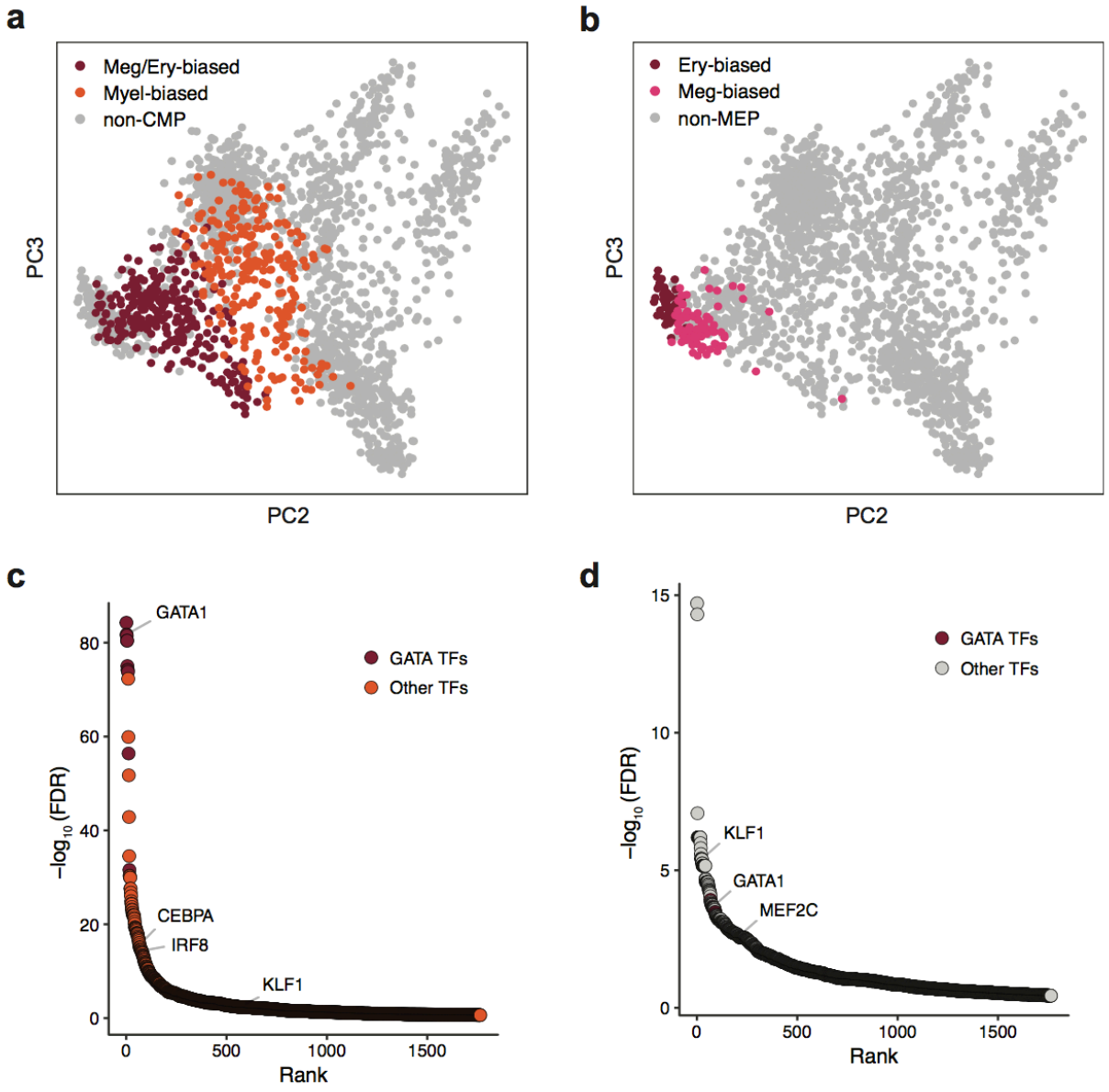
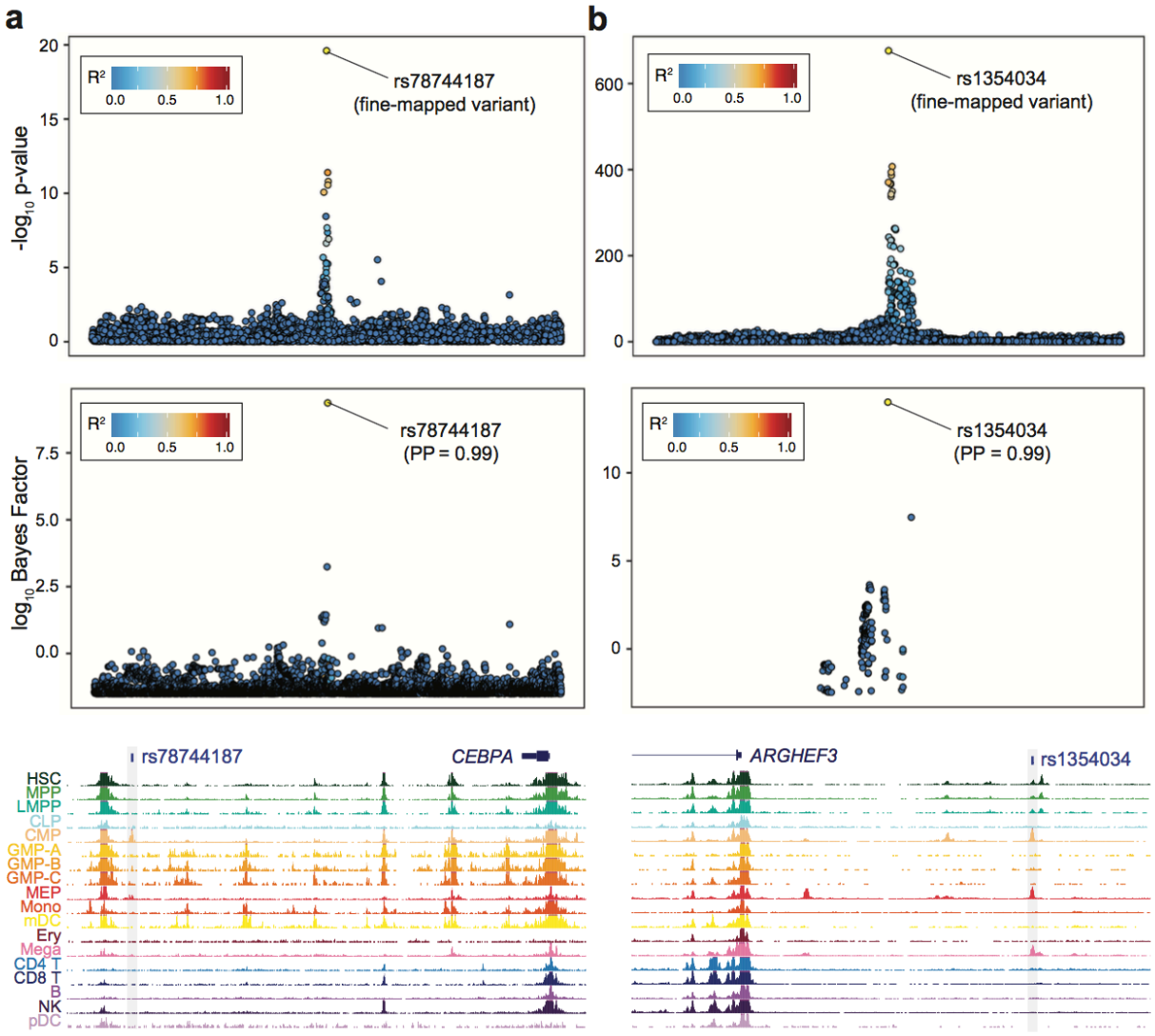
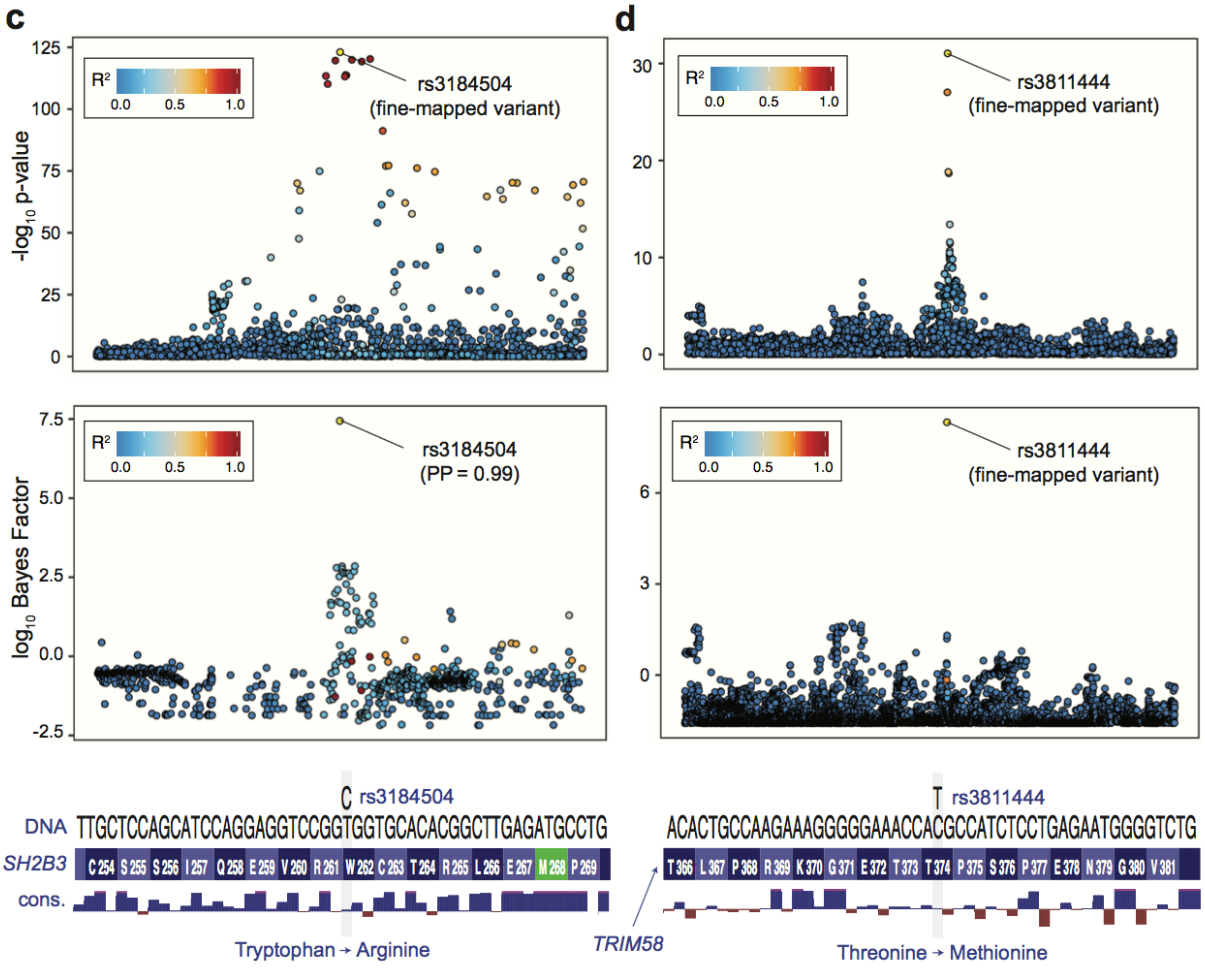
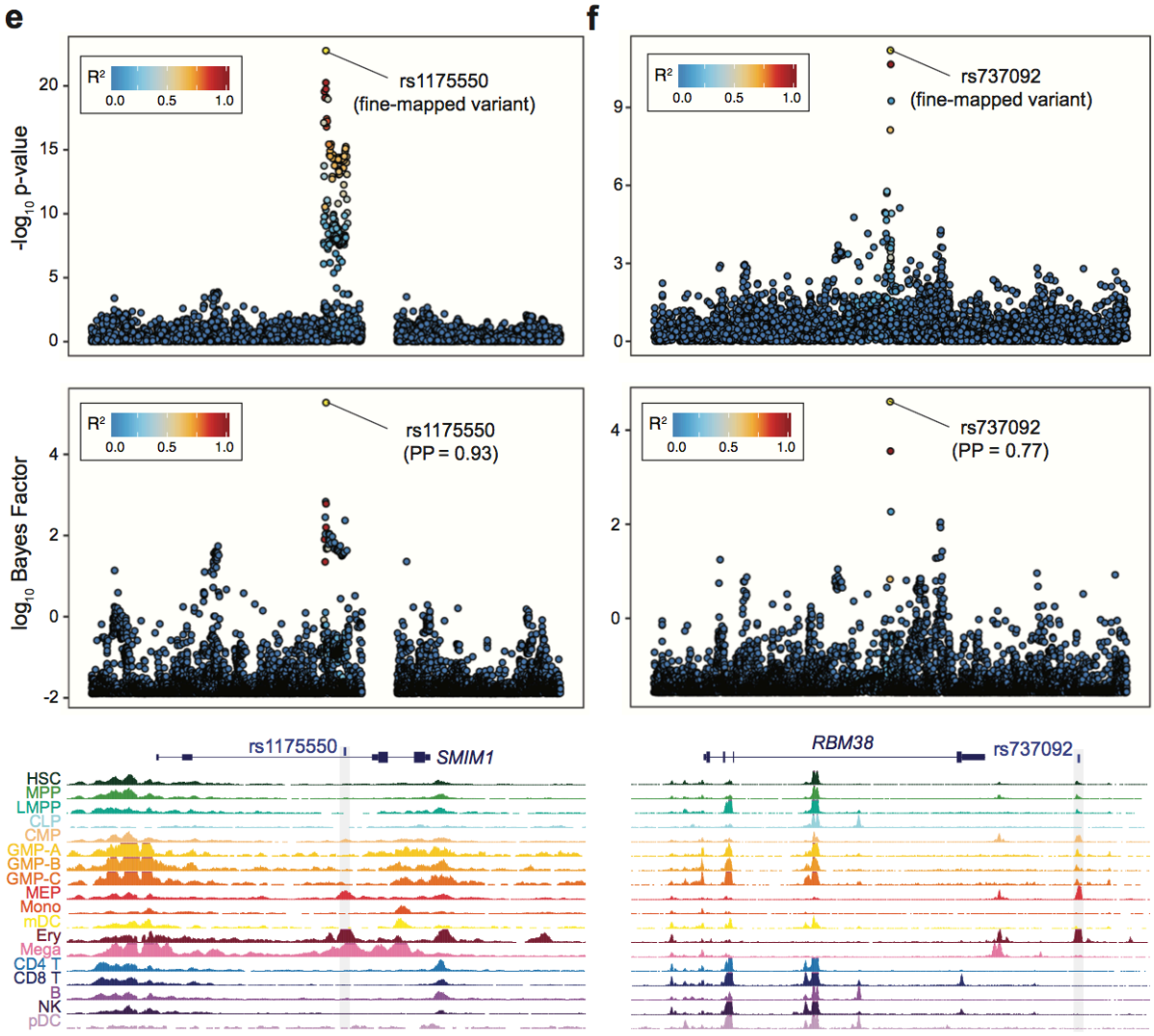


Figure S8. Further diagnostics of scATAC clustering and TF variability. Two-dimensional representation of scATAC samples highlighting defined clusters of (**a**) CMPs and (**b**) MEPs. Additionally, all TFs with variable chromVAR enrichment between scATAC k-medoids clusters in (**c**) CMPs and (**d**) MEPs.


a
b
-log10 p-value
log10 Bayes Factor
R²
0.0 0.5 1.0
rs78744187
(fine-mapped variant)
rs1354034
(fine-mapped variant)
rs78744187
(PP = 0.99)
rs1354034
(PP = 0.99)
rs78744187
CEBPA
ARGHEF3
rs1354034
HSC
MPP
LMPP
CLP
CMP
GMP-A
GMP-B
GMP-C
MEP
Mono
mDC
Ery
Mega
CD4 T
CD8 T
B
NK
pDC


c
d
-log10 p-value
log10 Bayes Factor
R²
0.0 0.5 1.0
rs3184504
(fine-mapped variant)
rs3184504
(PP = 0.99)
rs3811444
(fine-mapped variant)
C rs3184504
T rs3811444
DNA
TTGCTCCAGCATCCAGGAGGTCCGGTGGTGCACACGGCTTGAGATGCCTG
ACACTGCCAAGAAAGGGGGAAACCACGCCATCTCCTGAGAATGGGGTCTG
SH2B3
cons.
TRIM58
Tryptophan → Arginine
Threonine → Methionine


e
f
R²
0.0 0.5 1.0
rs1175550
(fine-mapped variant)
rs737092
(fine-mapped variant)
-log10 p-value
log10 Bayes Factor
rs1175550
(PP = 0.93)
rs737092
(PP = 0.77)
rs1175550
SMIM1
RBM38
rs737092
HSC
MPP
LMPP
CLP
CMP
GMP-A
GMP-B
GMP-C
MEP
Mono
mDC
Ery
Mega
CD4 T
CD8 T
B
NK
pDC

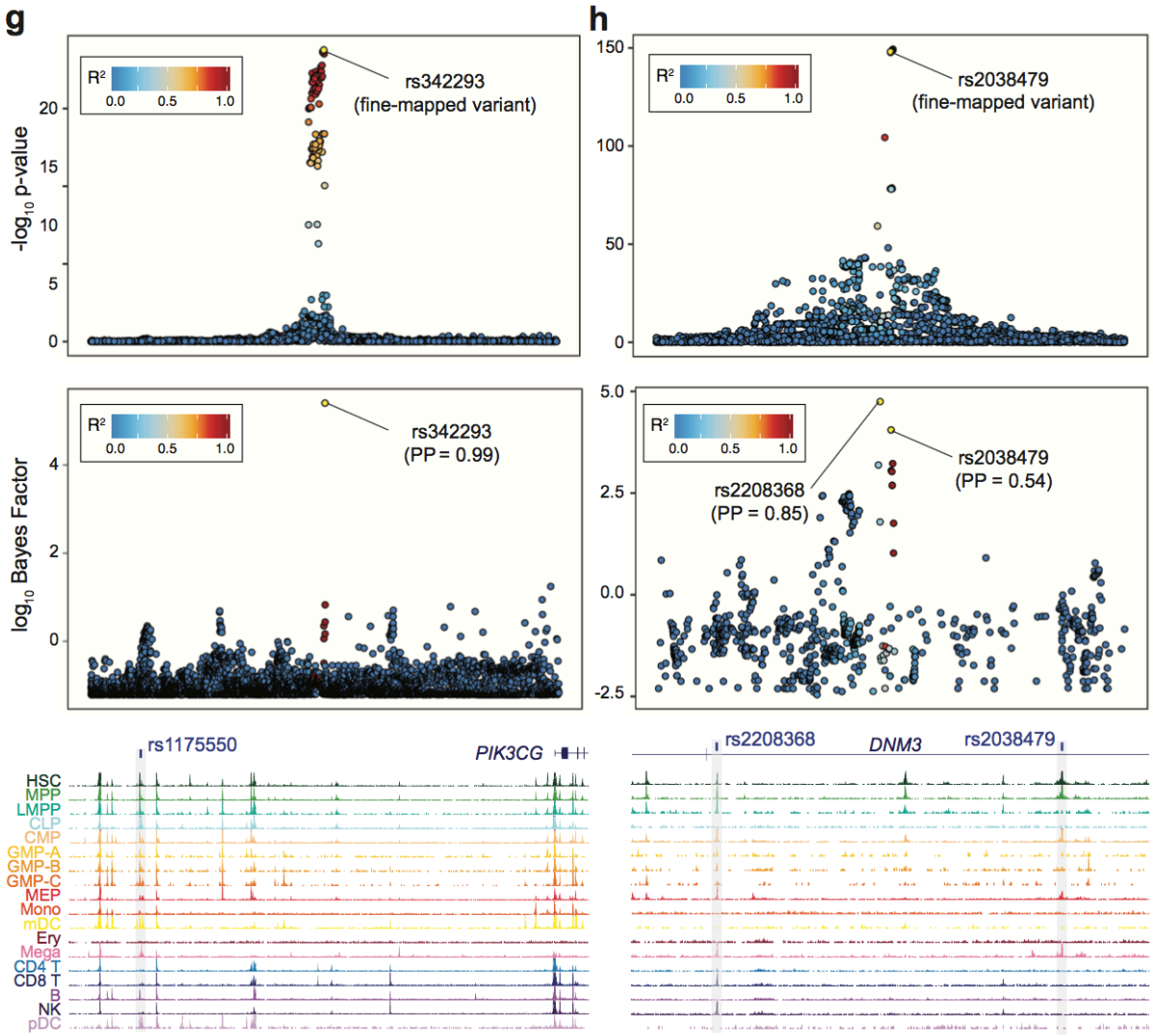


Figure S9. Previously identified causal variants corroborated by fine-mapping results and genomic annotations. (**a**) rs78744187 is associated with basophil count and was shown in Guo et al.[1] to lie in a CMP-specific enhancer that regulates *CEBPA* expression and basophil development. Although this locus is imputed and poorly tagged by genotyped variants, it is readily resolved by our fine-mapping method. (**b**) rs1354034 is associated with platelet traits and was shown in Zou et al.[2] to lie within a megakaryocyte enhancer and is associated (by eQTL) with expression of *ARHGEF3*; *Arhgef3* KO mice were then shown to have larger platelets than normal. This variant is predicted to disrupt a GATA motif and GATA factors are observed to be bound here by ChIP-seq in several hematopoietic cells. (**c**) rs3184504 is a nonsynonymous variant in *SH2B3* associated with multiple blood cell traits; *SH2B3* is a negative regulator of several signaling pathways, and its perturbation by either genome editing or naturally occurring loss-of-function mutations has been shown to have a role in erythroid differentiation.[3] (**d**) rs3811444 is a nonsynonymous variant in *TRIM58* associated with red cell traits; knockdown of *TRIM58* by RNAi showed a role for this gene in erythroblast enucleation through the regulation of dynein degradation.[4] (**e**) rs1175550 is an erythroid enhancer variant associated with red cell

traits that is associated with *SMIM1* expression by genome editing and SMIM1 protein (Vel antigen) by eQTL.[5,6] The Vel antigen is a blood group, and Vel-negative humans often have transfusion reactions to Vel-positive blood. (**f**) rs737092 is an erythroid enhancer variant associated with red cell traits that was found via genome editing to regulate *RBM38* expression.[6] *RBM38* knockdown by RNAi has been shown to regulate RNA splicing and erythroid maturation. (**g**) rs342293 is a megakaryocyte enhancer variant associated with platelet traits that is associated with *PIK3CG* expression (by eQTL); *PIK3CG* knockout mice exhibited dysfunction platelet function.[7] (**h**) rs2038479 is a megakaryocyte enhancer variant associated with platelet traits that is associated with *DNM3* expression by eQTL and exhibited regulatory function in a reporter assay.[8] Inhibition of *DNM3* activity in mouse resulted in reduced platelet formation. In addition to rs2038479, our fine-mapping identified a novel variant rs2208368 that was not marginally associated with platelet traits, but became strongly associated ($p < 10^{-51}$) after conditioning on rs2038479. rs2208368 was similarly localized to a megakaryocyte NDR in an intron of *DNM3*.

Figure S10. Overview of web resource for fine-mapped variants. This interactive resource provides users with an integrated experience, linking variants identified in this study with disease-relevant traits, putative target genes, and plausible molecular functions. (**a**) Resource website header. (**b**) Users can search by variant, region, or gene and view the results in the context of “Manhattan” plots of posterior probabilities, hematopoietic enhancer landscapes, and gene annotations. (**c**) Example of one target gene annotation, based upon promoter capture Hi-C results across hematopoietic cell types. (**d**) Disease relevance for the selected SNP based upon pheWAS results.

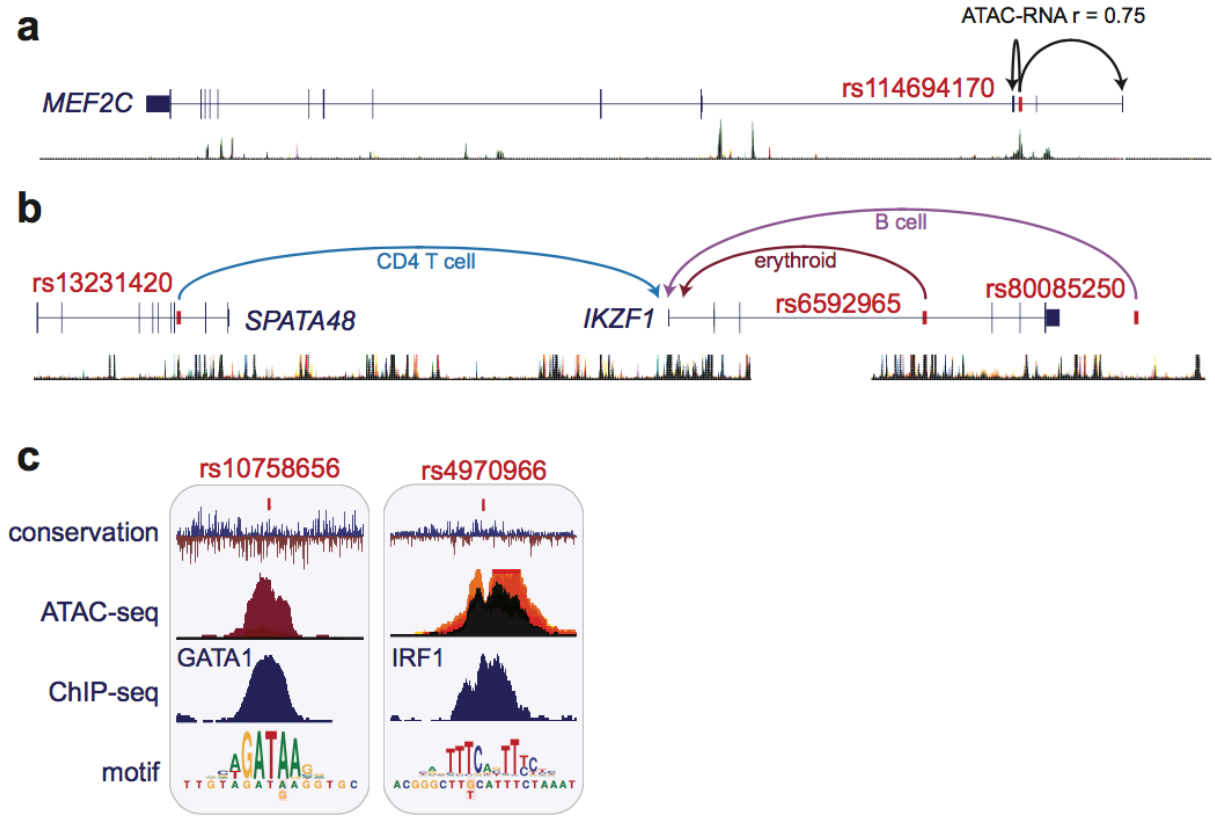


Figure S11. Characterization of additional functional variants. (**a**) Combining genetic fine-mapping for blood cell traits with dense RNA-seq and ATAC-seq profiling of the hematopoietic tree reveals putative causal variants and their target *trans*-acting TFs (rs114694170 is associated with platelet traits). Target genes were identified using ATAC-RNA correlations, and many are additionally supported by promoter-capture Hi-C. (**b**) Three variants near the *IKZF1* locus with PP > 0.75 are associated with multiple blood cell traits and within loops defined by PCHIC. An MCV-associated variant (rs6592965) was within an erythroid-specific loop (red), a lymphocyte count-associated variant (rs13231420) was within a B and T cell-specific loop (blue), and a white blood cell count-associated variant (rs80085250) was specific to a B cell loop (purple). (**c**) Additional putative causal variants that disrupt *cis*-binding of hematopoietic TFs known to be involved in regulating hematopoietic differentiation for various blood cell traits (rs10758656 is associated with red cell traits and rs4970966 is associated with monocyte count). Both IRF4 and IRF1 (IRF8 ChIP-seq was not available) occupied the right panel.

Table S1. Summary statistics and information for all fine-mapped variants PP > 0.001 identified in this work.

Table S2. Per-trait number of fine-mapped variants at various thresholds in addition to variants identified through classical GWAS methods.

Table S3. Summary statistics, including sequencing depth, for bulk ATAC-seq libraries considered in this work.

Table S4. Comprehensive hematopoietic enrichments for each trait/cell type pair analyzed for various methods. The last column indicates whether the trait was defined as lineage-specific or not.

Table S5. Application of g-chromVAR to DNase Roadmap data. Each trait/cell type pair is scored and ranked.

Table S6. Application of g-chromVAR to 39 predominately immune-related disorders previously fine-mapped with PICS to 18 chromatin accessibility profiles. Each trait/cell type pair is scored and ranked.

References

1. Guo, M.H. *et al.* Comprehensive population-based genome sequencing provides insight into hematopoietic regulatory mechanisms. *Proc Natl Acad Sci U S A* **114**, E327-E336 (2017).
2. Zou, S. *et al.* SNP in human ARHGEF3 promoter is associated with DNase hypersensitivity, transcript level and platelet function, and Arhgef3 KO mice have increased mean platelet volume. *PLoS One* **12**, e0178095 (2017).
3. Giani, F.C. *et al.* Targeted Application of Human Genetic Variation Can Improve Red Blood Cell Production from Stem Cells. *Cell Stem Cell* **18**, 73-78 (2016).
4. Thom, C.S. *et al.* Trim58 degrades Dynein and regulates terminal erythropoiesis. *Dev Cell* **30**, 688-700 (2014).
5. Cvejic, A. *et al.* SMIM1 underlies the Vel blood group and influences red blood cell traits. *Nat Genet* **45**, 542-545 (2013).
6. Ulirsch, J.C. *et al.* Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits. *Cell* **165**, 1530-1545 (2016).
7. Paul, D.S. *et al.* Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. *PLoS Genet* **7**, e1002139 (2011).
8. Nurnberg, S.T. *et al.* A GWAS sequence variant for platelet volume marks an alternative DNM3 promoter in megakaryocytes near a MEIS1 binding site. *Blood* **120**, 4859-68 (2012).