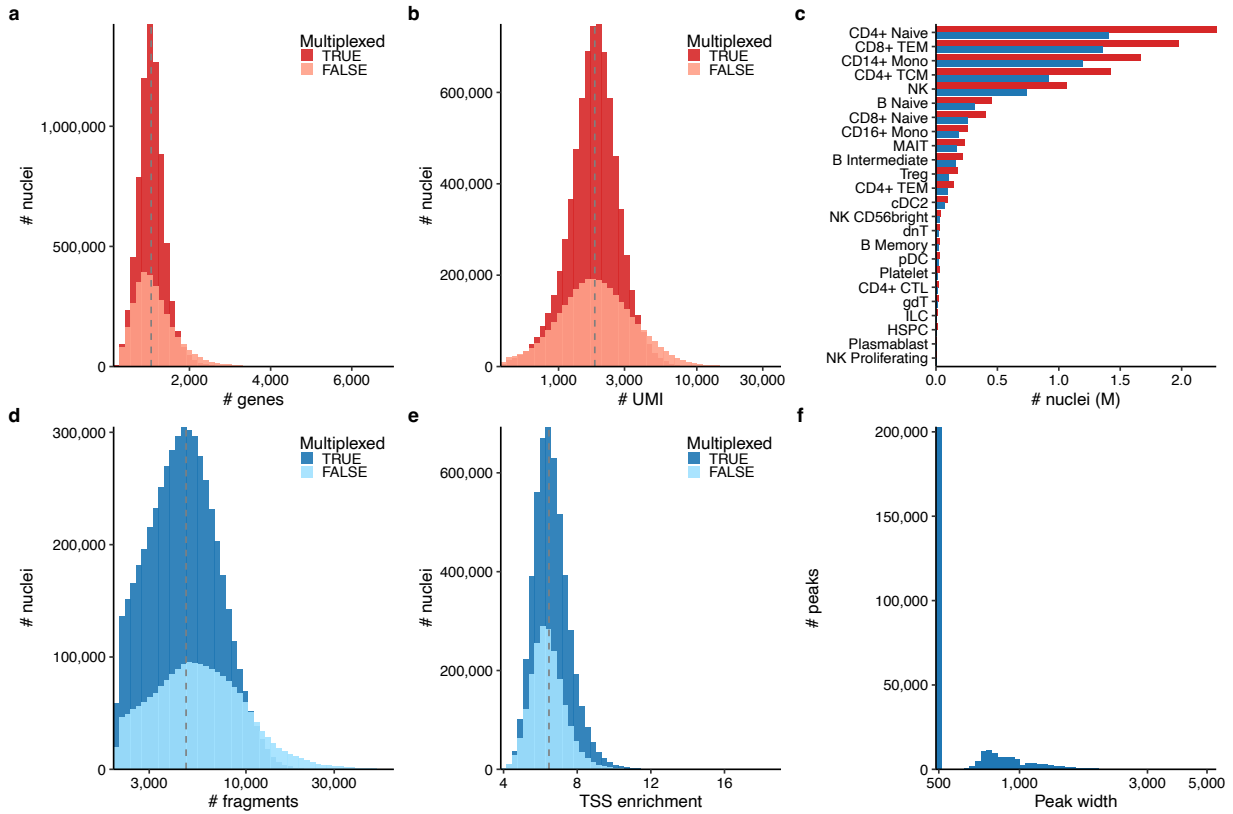

Supplementary information

**Population-scale immune multiome atlas
reveals regulatory disease mechanisms**

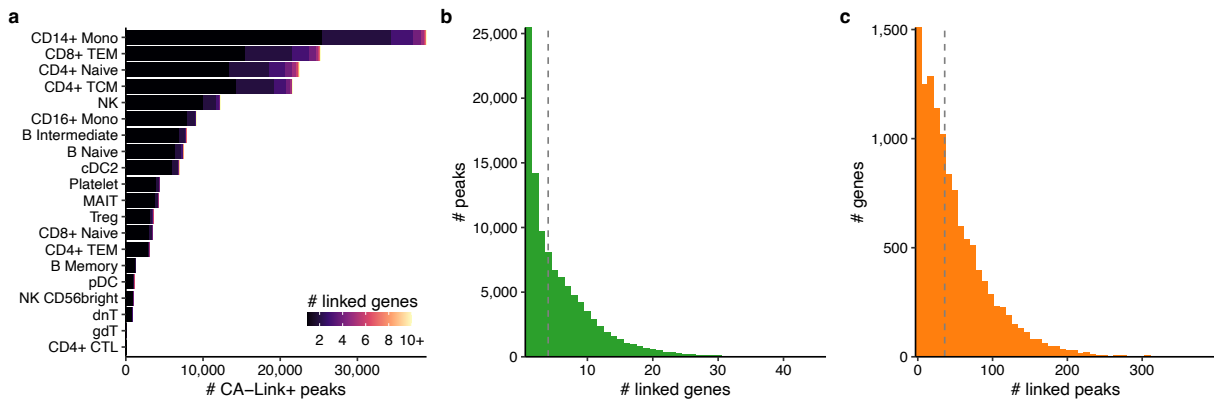
In the format provided by the
authors and unedited

Supplementary Figures



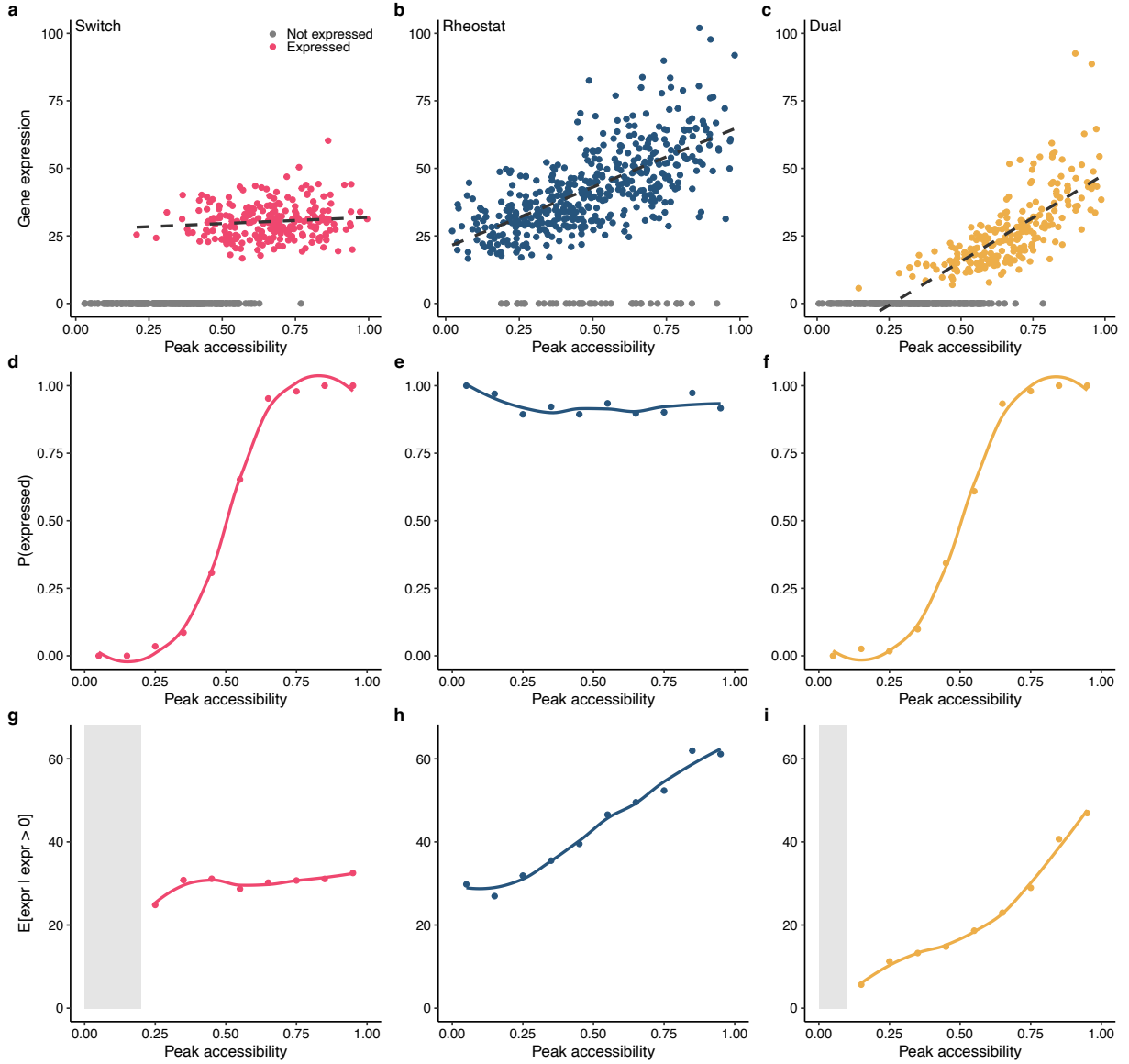
Supplementary Fig. 1 | Quality control metrics and cell type composition of snRNA-seq and snATAC-seq atlases.

a,b, The number of genes (**a**) and the number of UMIs (**b**) per nucleus in the snRNA-seq atlas (10,612,211 nuclei from 1,108 donors), stratified by multiplexing status (see **Methods**). **c**, The number of nuclei for each level 2 (L2) cell types in the snRNA-seq (red) and snATAC-seq (blue). **d,e**, The number of fragments (**d**) and TSS enrichment (**e**) per nucleus in the snATAC-seq atlas (7,100,899 high-quality from 1,093 donors), stratified by multiplexing status. **f**, Distribution of peak widths across 297,024 consensus peaks. L2 cell-type-specific, fixed-width (500bp) peaks called by MACS2 were merged using iterative overlap peak merging (see **Methods**).



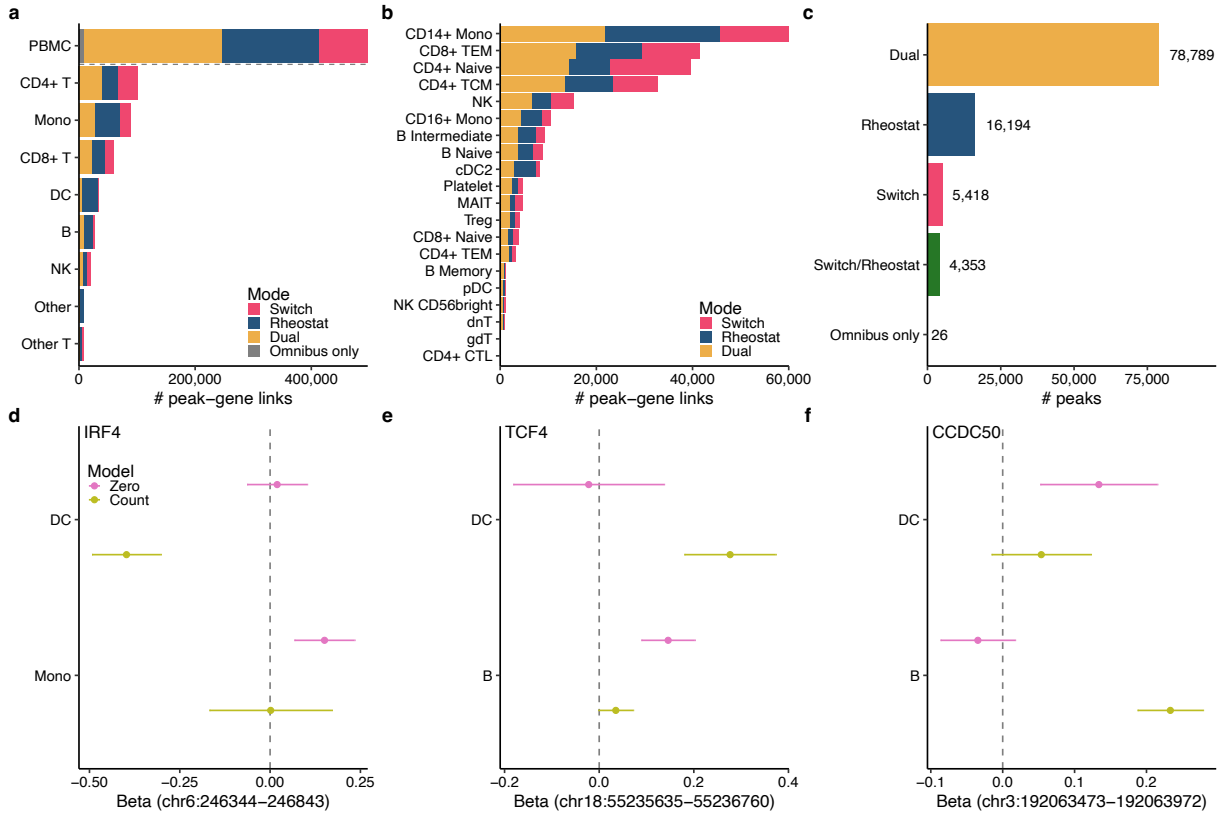
Supplementary Fig. 2 | Peak-gene links across cell types, peaks, and genes.

a, The number of chromatin accessibility (CA) peaks with significant peak-gene links (CA-Link+; see **Methods**) across L2 cell types. Color represents the number of linked genes for each CA peak. **b,c**, The number of linked genes per peak (**b**) and the number of linked peaks per gene (**c**). The dotted vertical line represents the median of the distribution.



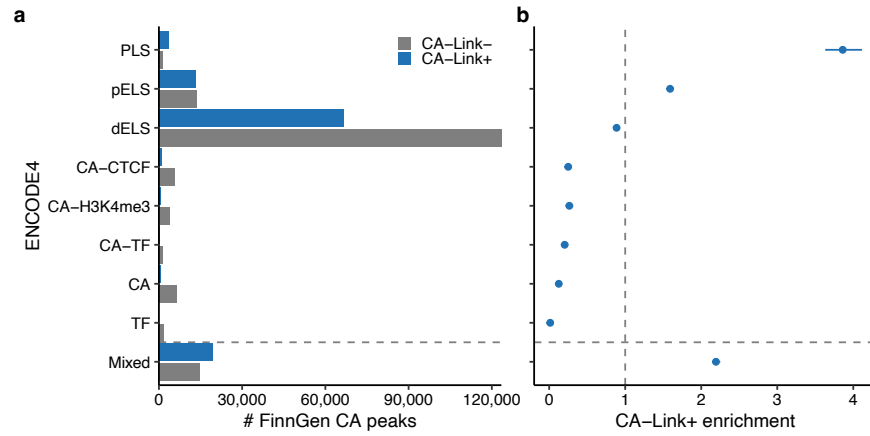
Supplementary Fig. 3 | Schematic overview of the hurdle model classifications for peak-gene link regulatory modes.

a–c, Simulated data illustrating three regulatory modes through which a chromatin accessibility peak can affect gene expression: switch (**a**), rheostat (**b**), and dual (**c**). Dashed lines represent linear fits among expressing nuclei. **d–f**, The probability of expression, $P(\text{expressed})$, across peak accessibility bins for each mode. **g–i**, The mean expression among expressing nuclei, $E[\text{expr} | \text{expr} > 0]$, across peak accessibility bins. Grey shaded regions denote accessibility bins in which no nuclei expressed the gene, leaving the count-model component undefined. In the switch mode (**a,d,g**), peak accessibility governs the probability that the gene is expressed, whereas expression level among expressing nuclei is independent of accessibility. In the rheostat mode (**b,e,h**), nearly all nuclei express the gene and accessibility modulates only the expression level. In the dual mode (**c,f,i**), accessibility drives both on/off switching and dose-response among expressing nuclei. In the peak-gene link analysis, we used a hurdle model with significance determined via a joint χ^2 score test with stage-wise confirmation to classify these regulatory modes (see **Methods**). Because single-cell chromatin accessibility and gene expression are sparse and zero-inflated, the per-nucleus relationship between peak accessibility and gene expression is dominated by zeros and does not visually resolve the three modes even for strongly significant links. We therefore illustrate the modes with simulated data and classify real links statistically with the hurdle model.



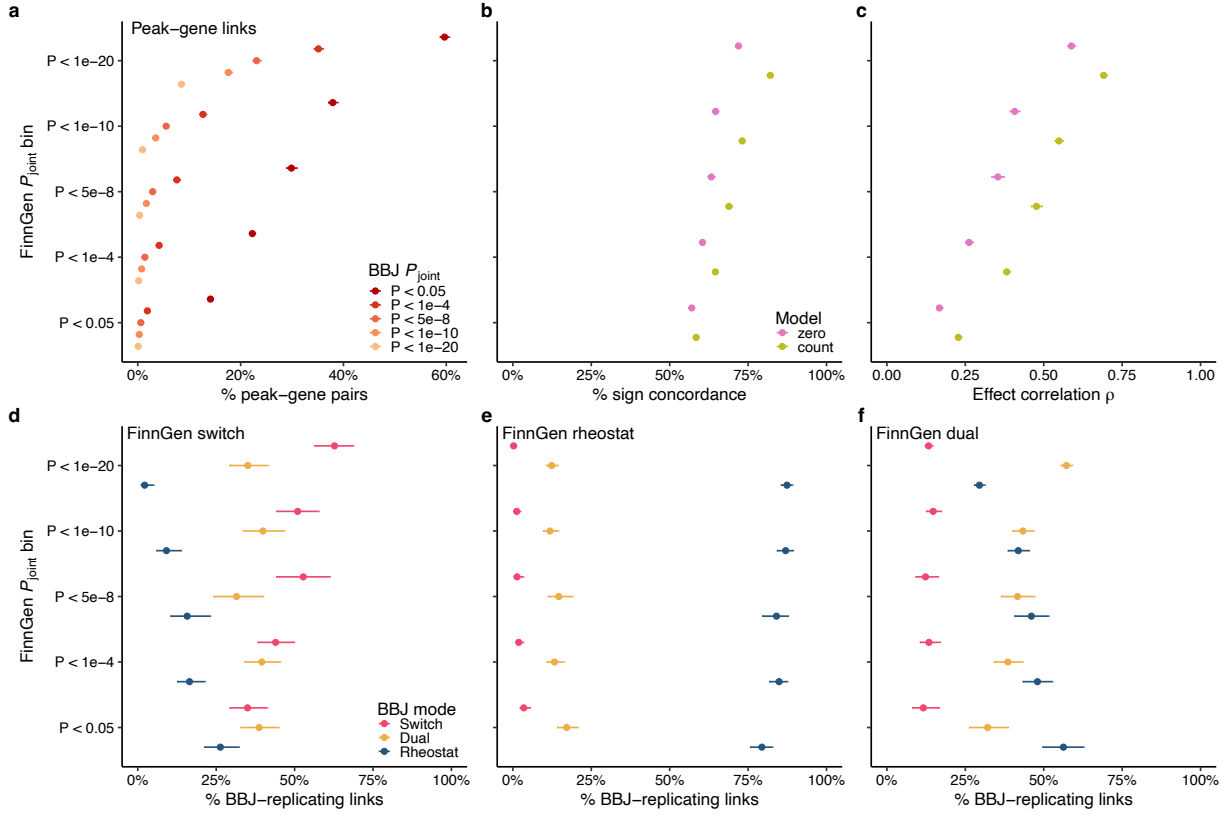
Supplementary Fig. 4 | Regulatory mode classification of peak-gene links.

a,b, The number of significant peak-gene links across level 1 (L1; **a**) and level 2 (L2; **b**) cell types, stratified by regulatory mode (see **Methods**). Horizontal dotted line in (**a**) separates PBMC and L1 cell types. **c**, The number of chromatin accessibility peaks classified by their regulatory mode across all peak-gene links across genes and cell types. “Dual” denotes peaks with at least one link classified as dual. “Switch/Rheostat” denotes peaks whose links alternate between switch and rheostat exclusively. “Omnibus only” denotes peaks reaching significance only via the joint χ^2 score test without passing stage-wise confirmation. **d–f**, Forest plots of hurdle model effect sizes (β) for the zero-inflation and count components for three illustrative peaks whose regulatory mode alternates between switch and rheostat across cell types: *IRF4* linked to chr6:246344-246843 (**d**), *TCF4* linked to chr18:55235635-55236760 (**e**), and *CCDC50* linked to chr3:192063473-192063972 (**f**). Error bars represent 95% confidence intervals.



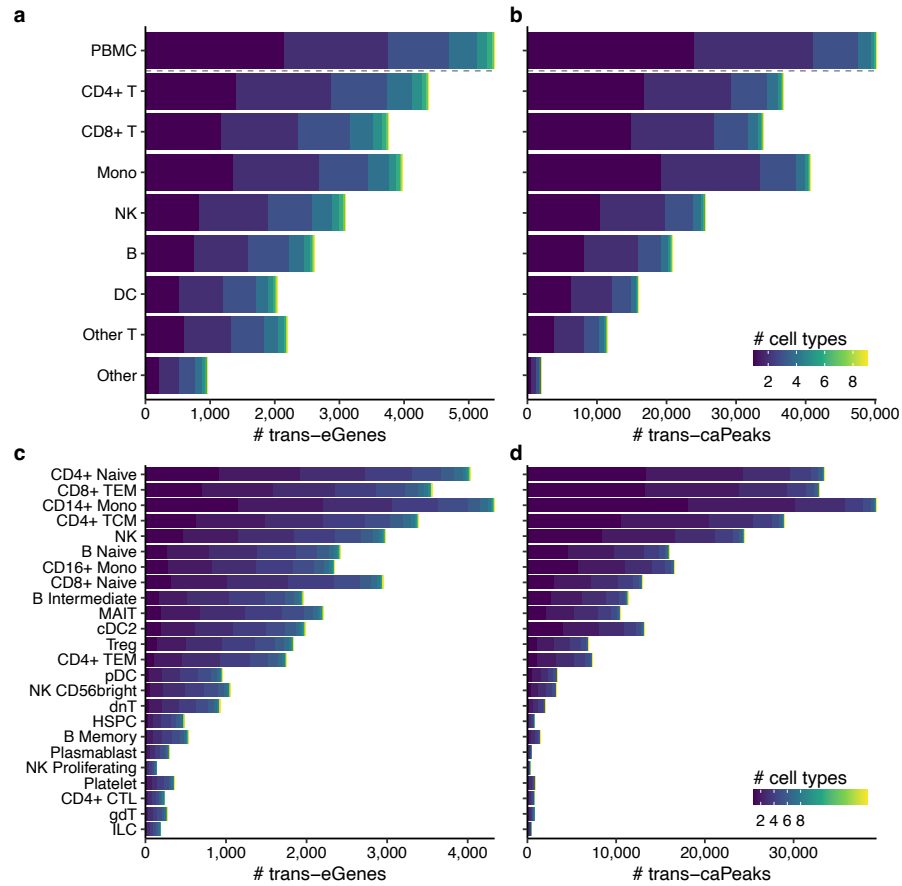
Supplementary Fig. 5 | Enrichment of chromatin accessibility peaks for ENCODE4 cCRE classifications.

a. The number of FinnGen chromatin accessibility (CA) peaks for each ENCODE4 cCRE classifications, split by peak–gene link status (CA–Link– and CA–Link+). The mixed category represents the CA peak that overlaps with multiple cCRE classifications. **b.** The enrichment of CA–Link+ for each ENCODE4 cCRE classification. Enrichment is defined as a risk ratio of having a significant link (CA–Link+) and being in an annotation (see **Methods**). Error bars represent 95% confidence intervals.



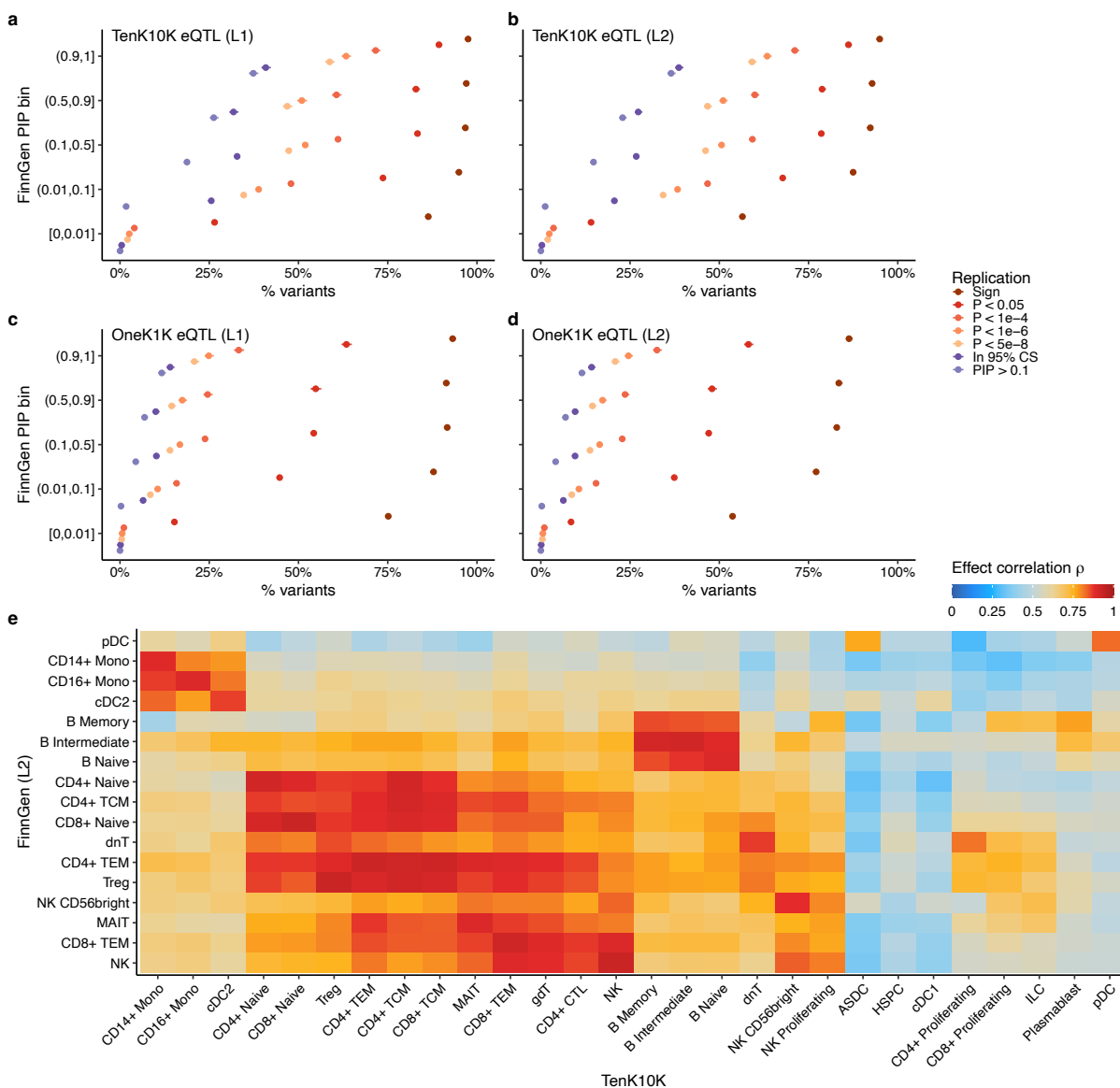
Supplementary Fig. 6 | Replication of peak-gene links in an independent Japanese multiome cohort.

a–c, Replication of FinnGen peak-gene links in the BBJ multiome cohort ($n = 98$), stratified by FinnGen P_{joint} bin and pooled across L1 cell types. The fraction of matched peak-gene pairs reaching significance in BBJ at multiple P_{joint} thresholds (**a**), sign concordance (**b**), and Spearman rank correlation (ρ) of effect sizes (**c**) are shown. Sign concordance and Spearman's ρ are computed separately for the zero-inflation and count model components of the hurdle model. **d–f**, Concordance of regulatory modes between FinnGen and BBJ for BBJ-replicating peak-gene links (BBJ $P_{\text{joint}} < 0.05$), stratified by FinnGen P_{joint} bin. For each FinnGen mode (switch in **d**, rheostat in **e**, dual in **f**), the fraction of BBJ-replicating links assigned to each BBJ mode is shown. Error bars represent 95% confidence intervals. All P values are two-sided.



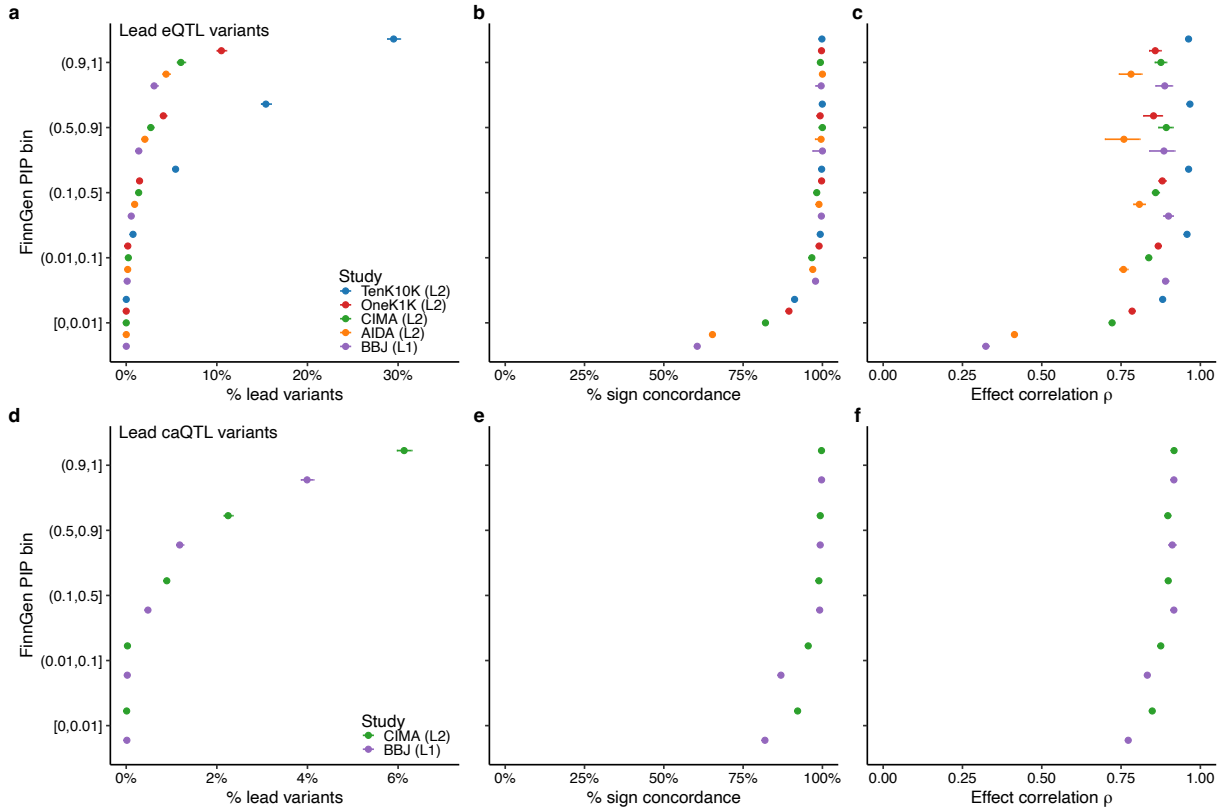
Supplementary Fig. 7 | Trans-molQTL analyses

a,b, The number of significant *trans*-eGenes (**a**) and *trans*-caPeaks (**b**) for each L1 cell type ($P < 5.0 \times 10^{-8}$; see **Methods**). Color represents the number of cell types that each eGene or caPeak is significantly detected for. **c,d**, Same as **a,b** for each L2 cell type.



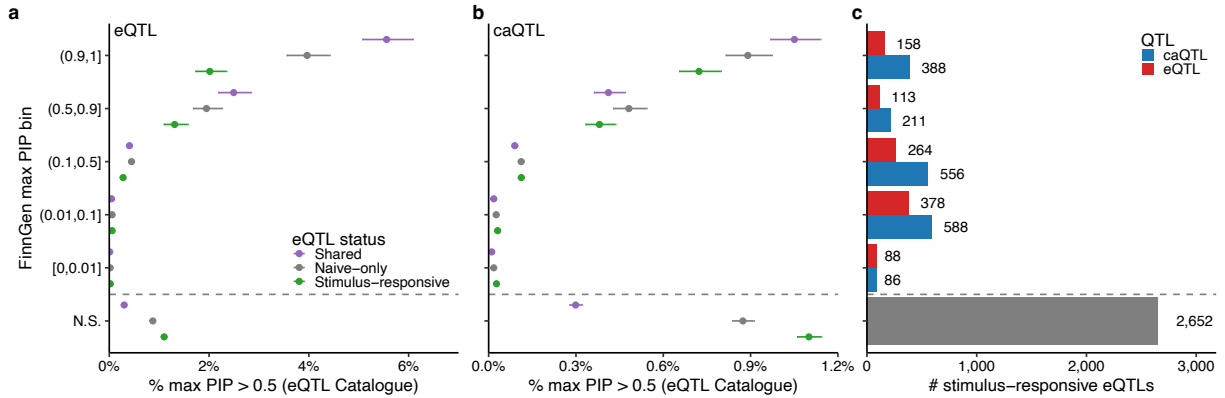
Supplementary Fig. 8 | Replication of cis-eQTL results in independent cohorts.

a,b, Posterior inclusion probability (PIP)-stratified replication of FinnGen fine-mapped *cis*-eQTL variants in TenK10K at (a) level 1 (L1) and (b) level 2 (L2) cell type resolution. FinnGen variants are binned by PIP and the fraction meeting each replication criterion is shown, pooled across cell types (**Methods**). Replication metrics include sign concordance, nominal significance at multiple *P*-value thresholds in TenK10K, membership in a TenK10K 95% credible set (CS), and TenK10K PIP > 0.1. Error bars represent binomial 95% confidence intervals. **c,d**, Same as **a,b** for OneK1K. **e**, Cross-cell-type consistency of eQTL effects between FinnGen and TenK10K. Each cell shows the Spearman rank correlation (ρ) of eQTL effect sizes for PIP > 0.5 variants between a TenK10K cell type (x-axis) and a FinnGen L2 cell type (y-axis). Rows and columns are independently ordered by hierarchical clustering (**Methods**).



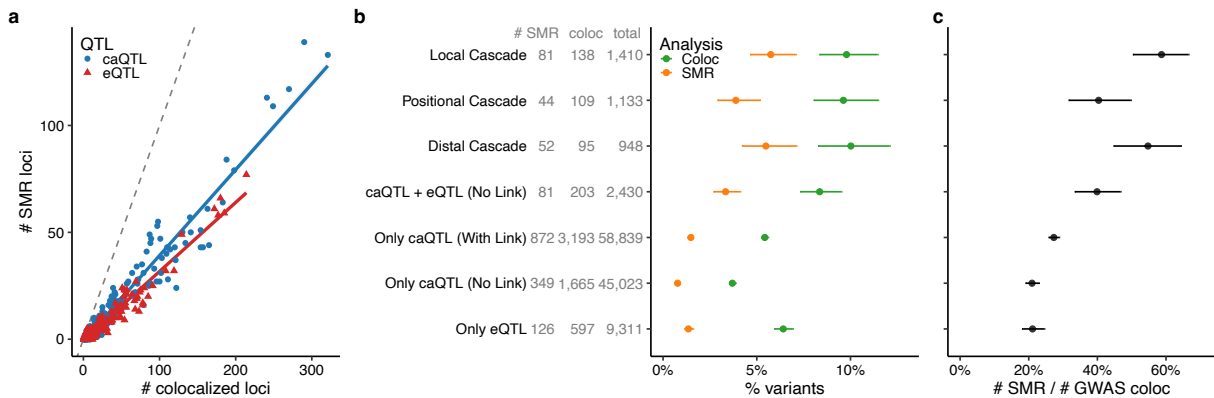
Supplementary Fig. 9 | Cross-ancestry replication of lead *cis*-eQTL and *cis*-caQTL results.

a–c, Replication of FinnGen fine-mapped *cis*-eQTL variants across five independent datasets (TenK10K, OneK1K, CIMA, AIDA, and BBJ), stratified by FinnGen PIP bin and pooled across cell types (**Methods**). The fraction of FinnGen variants that are lead eQTLs in each external dataset (**a**), sign concordance among overlapping lead variants (**b**), and Spearman rank correlation (ρ) of effect sizes among lead variants (**c**) are shown. Error bars represent 95% confidence intervals. **d–f**, Same as **a–c** for *cis*-caQTL replication in CIMA and BBJ.



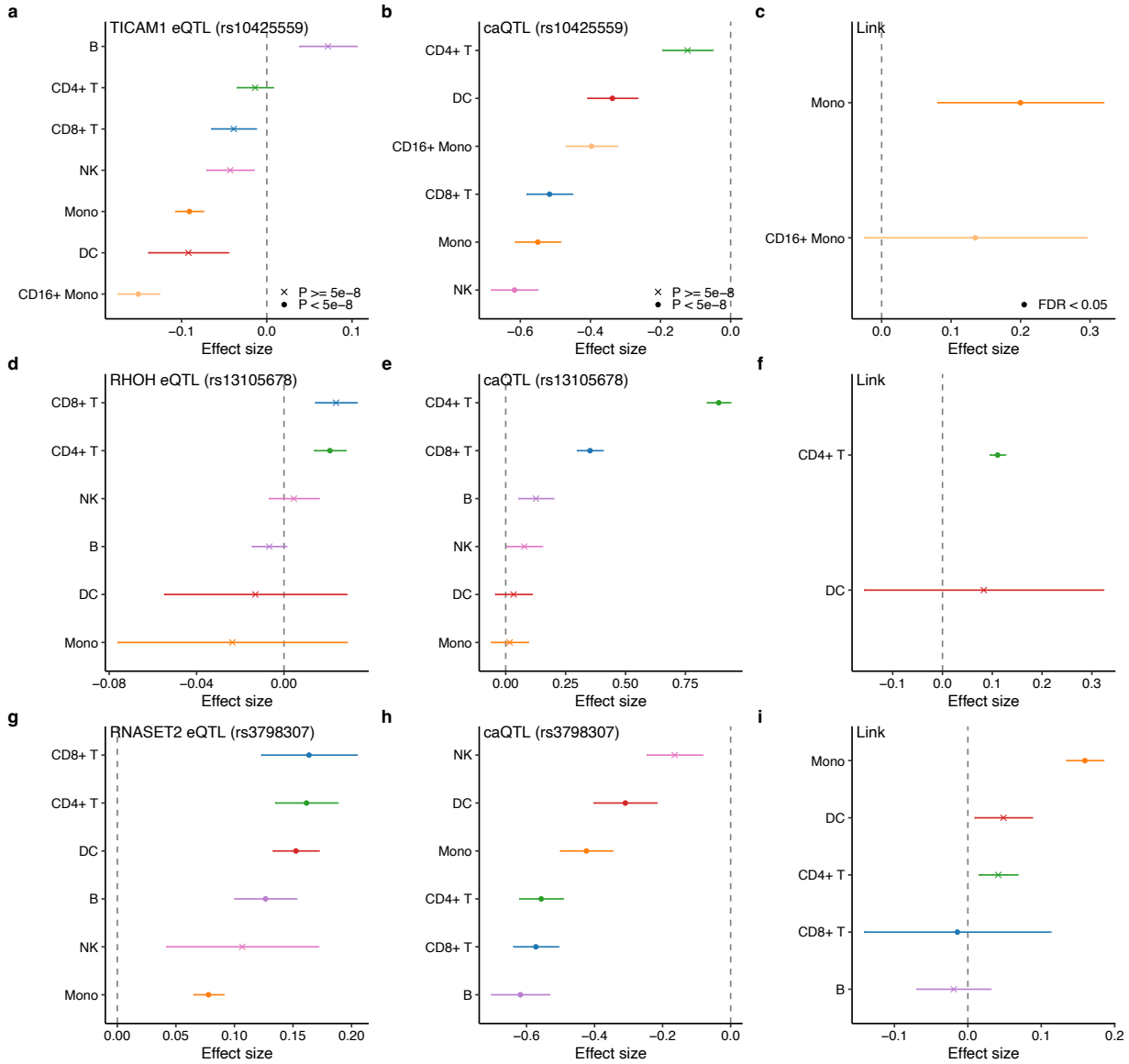
Supplementary Fig. 10 | Overlap between fine-mapped FinnGen molQTL variants and stimulus-responsive eQTLs from the eQTL Catalogue.

a,b, The fraction of FinnGen *cis*-eQTL (**a**) or *cis*-caQTL (**b**) fine-mapped variants that overlap stimulus-responsive eQTL fine-mapped variants from the eQTL Catalogue. Stimulus-responsive variant categories are defined across eQTL Catalogue studies with paired naive and stimulated immune cell conditions (see **Methods**): Shared (max PIP > 0.5 in both naive and stimulated conditions), stimulus-responsive (max PIP > 0.5 under stimulation and < 0.01 or untested in naive), and naive-only (max PIP > 0.5 in naive and < 0.01 or untested under stimulation). Error bars represent 95% confidence intervals. N.S., non-significant variants in FinnGen. **c,** The number of stimulus-responsive eQTL variants ($n = 4,626$) that overlap FinnGen molQTL variants for each FinnGen maximum PIP bin.



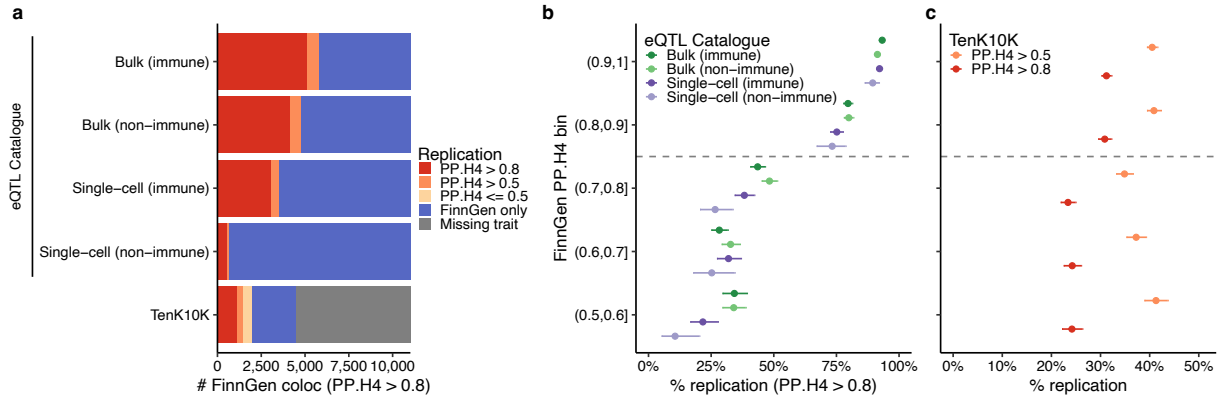
Supplementary Fig. 11 | Summary-based Mendelian Randomization (SMR) between molQTL and GWAS associations.

a, The number of SMR-significant loci (Bonferroni $P_{\text{SMR}} < 3.3 \times 10^{-7}$ and $P_{\text{HEIDI}} > 0.05$) vs. the number of colocated 95% credible sets (PP.H4 > 0.8) between molQTL and 439 complex traits in FinnGen (**Methods**). Colored lines are linear fits with 95% confidence bands for caQTL and eQTL, whereas the dotted slope represents $Y = X$. **b,** The fraction of fine-mapped molQTL variants (PIP > 0.5) supported by either colocalization (PP.H4 > 0.8) or SMR (Bonferroni $P_{\text{SMR}} < 3.3 \times 10^{-7}$ and $P_{\text{HEIDI}} > 0.05$), stratified by CASCADE mechanism category (as in **Fig. 5**). **c,** Ratio of SMR-significant to colocated variants (#SMR / # coloc) for each CASCADE category from (**b**), representing the fraction of colocated signals that are additionally supported by SMR. Error bars represent 95% confidence intervals.



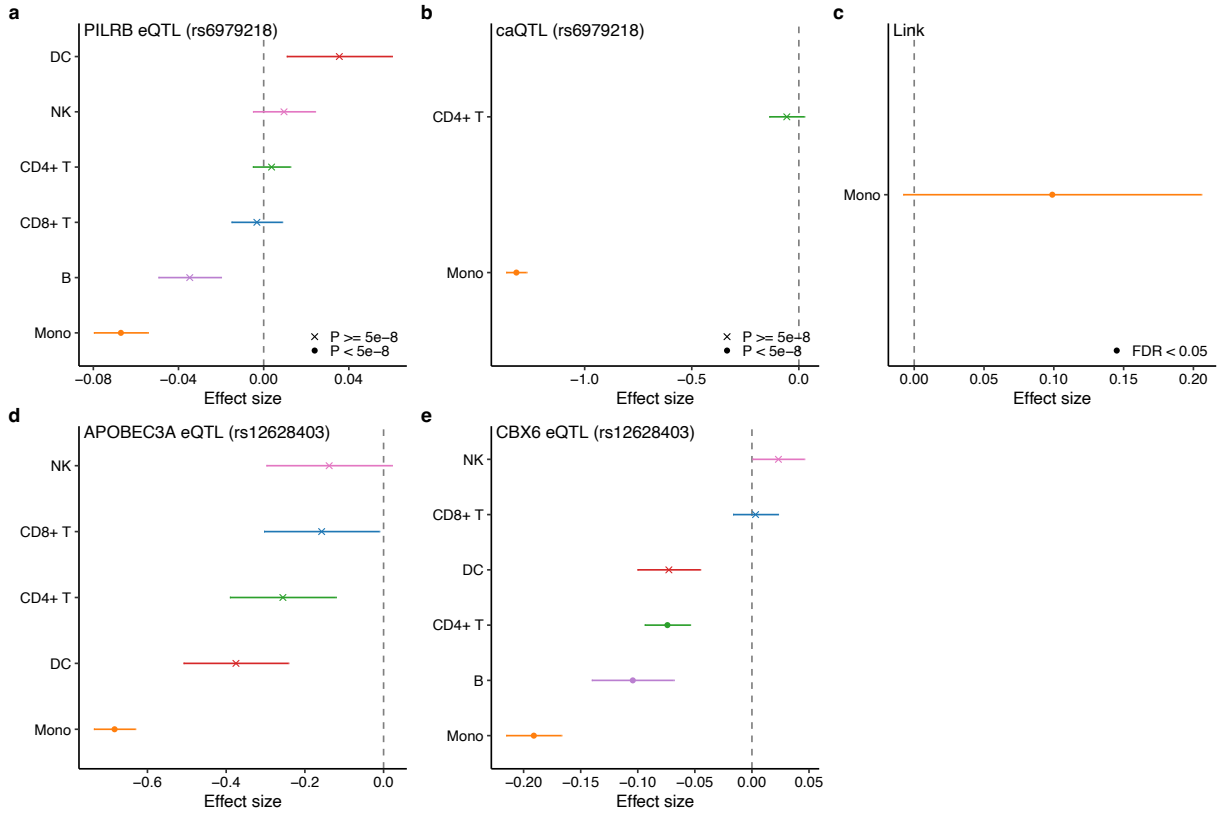
Supplementary Fig. 12 | Forest plots of fine-mapped molQTL variants colocalizing with autoimmune hypothyroidism at *TICAM1*, *RHOH*, and *RNASET2*.

a–c, Forest plots of rs10425559:A>G effect sizes on *TICAM1* expression (**a**), chromatin accessibility at chr19:4837478-4837977 (**b**), and the peak–gene link between this peak and *TICAM1* (**c**) across cell types. Effect sizes for **c** correspond to the count component of the hurdle model. **d–f**, Same as **a–c** for rs13105678:C>A at *RHOH* and chr4:40258141-40259058. **g–i**, Same as **a–c** for rs3798307:G>A at *RNASET2* and chr6:166950262-166950954. Significance is shown at the genome-wide significance threshold, $P < 5.0 \times 10^{-8}$ for eQTL and caQTL and at FDR < 0.05 for peak–gene link. Error bars represent 95% confidence intervals. All P values are two-sided.



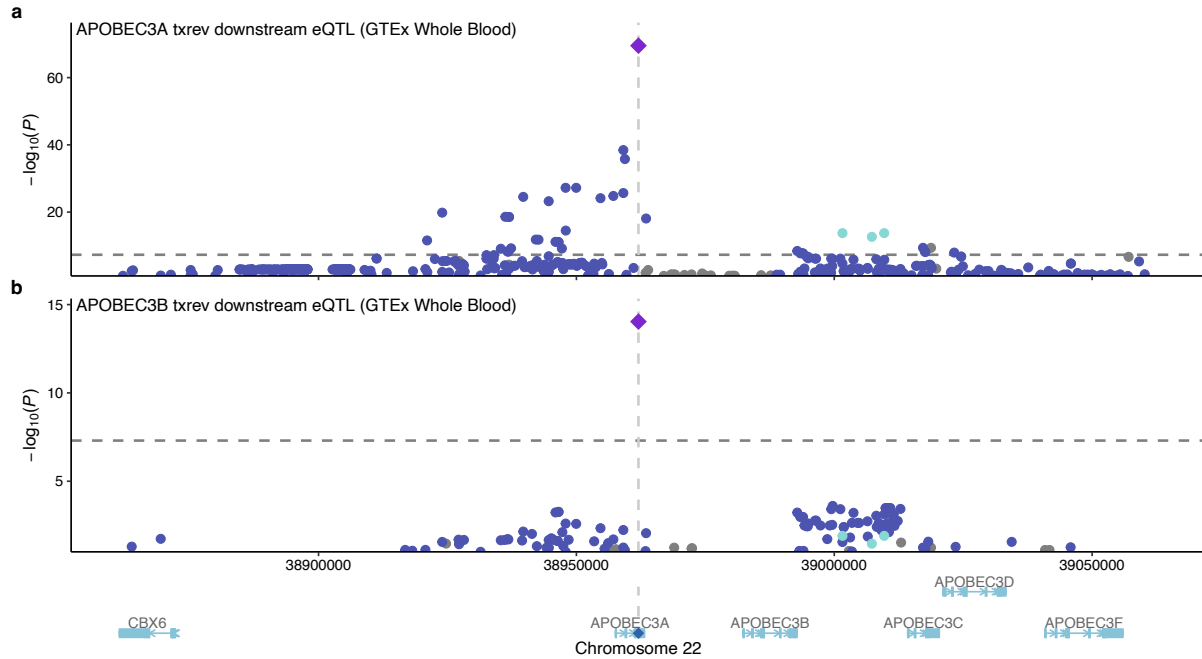
Supplementary Fig. 13 | Replication GWAS-eQTL colocalizations.

a, Number of FinnGen GWAS-eQTL colocalizations ($PP.H4 > 0.8$; $n = 11,015$ unique gene-trait pairs; **Methods**) classified by their replication status in external eQTL datasets (eQTL Catalogue, top four rows) or TenK10K (bottom row). For each eQTL Catalogue category (bulk or single-cell, immune or non-immune tissue), colocalizations are classified as replicated ($PP.H4 > 0.8$ or > 0.5 in at least one dataset within that category), not replicated ($PP.H4 \leq 0.5$), or FinnGen only (gene-trait pair not tested in that category). For TenK10K, “missing trait” indicates FinnGen traits without an equivalent trait tested in TenK10K. **b**, Replication rate of FinnGen GWAS-eQTL colocalizations in eQTL Catalogue datasets, stratified by FinnGen PP.H4 bin (x-axis) and eQTL Catalogue category (color). Replication is defined as $PP.H4 > 0.8$ in at least one dataset within each category. The denominator is restricted to FinnGen colocalizations with a testable gene-trait pair in that category. Error bars represent 95% confidence intervals. The dashed line separates FinnGen colocalizations above and below the $PP.H4 > 0.8$ threshold. **c**, Same as **b** for TenK10K, showing replication at $PP.H4 > 0.8$ and $PP.H4 > 0.5$ thresholds. TenK10K applied coloc.abf with external GWAS meta-analyses, in contrast to the coloc.susie based colocalization with FinnGen GWAS used in **b**.



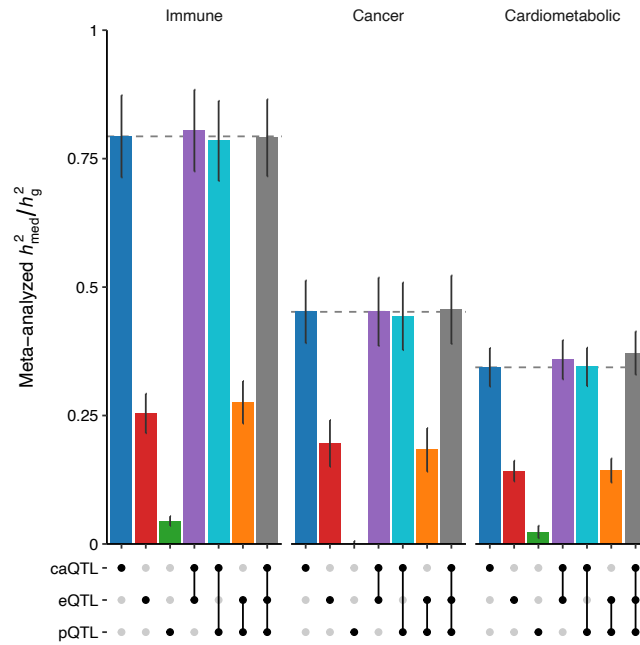
Supplementary Fig. 14 | Forest plots of fine-mapped molQTL variants at *PILRB*, *APOBEC3A*, and *CBX6*.

a–c, Forest plots of rs6979218:C>G effect sizes on *PILRB* expression (**a**), chromatin accessibility at chr7:100295229-100295728 (**b**), and the peak–gene link between this peak and *PILRB* (**c**) across cell types. Effect sizes for **c** correspond to the count component of the hurdle link model. **d,e**, Forest plots of rs12628403:A>C effect sizes on *APOBEC3A* (**d**) and *CBX6* (**e**) expression across cell types. Because there are 12 caQTL peaks that are linked to *APOBEC3A*/*CBX6*, their effect sizes are not shown. Significance is shown at the genome-wide significance threshold, $P < 5.0 \times 10^{-8}$ for eQTL and caQTL and at $FDR < 0.05$ for peak–gene link. Error bars represent 95% confidence intervals. All P values are two-sided.



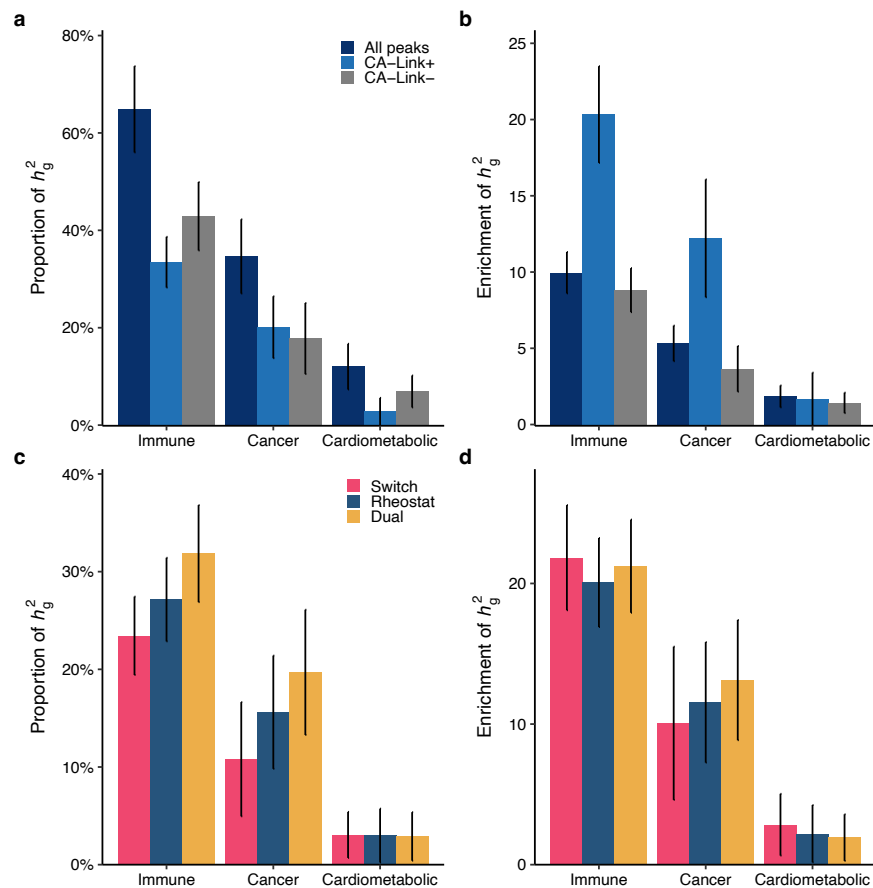
Supplementary Fig. 15 | Alternative 3' end usage (txrevise downstream) eQTL signals for *APOBEC3A* and *APOBEC3B* in GTEx whole blood.

a,b, LocusZoom plots of *cis*-QTL signals for alternative 3' end usage events at *APOBEC3A* (ENST00000495988; **a**) and *APOBEC3B* (ENST00000333467; **b**) at the *APOBEC3A/APOBEC3B* locus. Events were defined by txrevise, which decomposes Ensembl transcript annotations into independent alternative-promoter (upstream), internal-splicing (contained), and 3' end usage (downstream) events using shared exons as a scaffold (eQTL Catalogue, **Methods**). Color represents r^2 to the lead variant, rs12628403:A>C. Horizontal dotted line represents the genome-wide significance threshold ($P < 5.0 \times 10^{-8}$). All P values are two-sided.



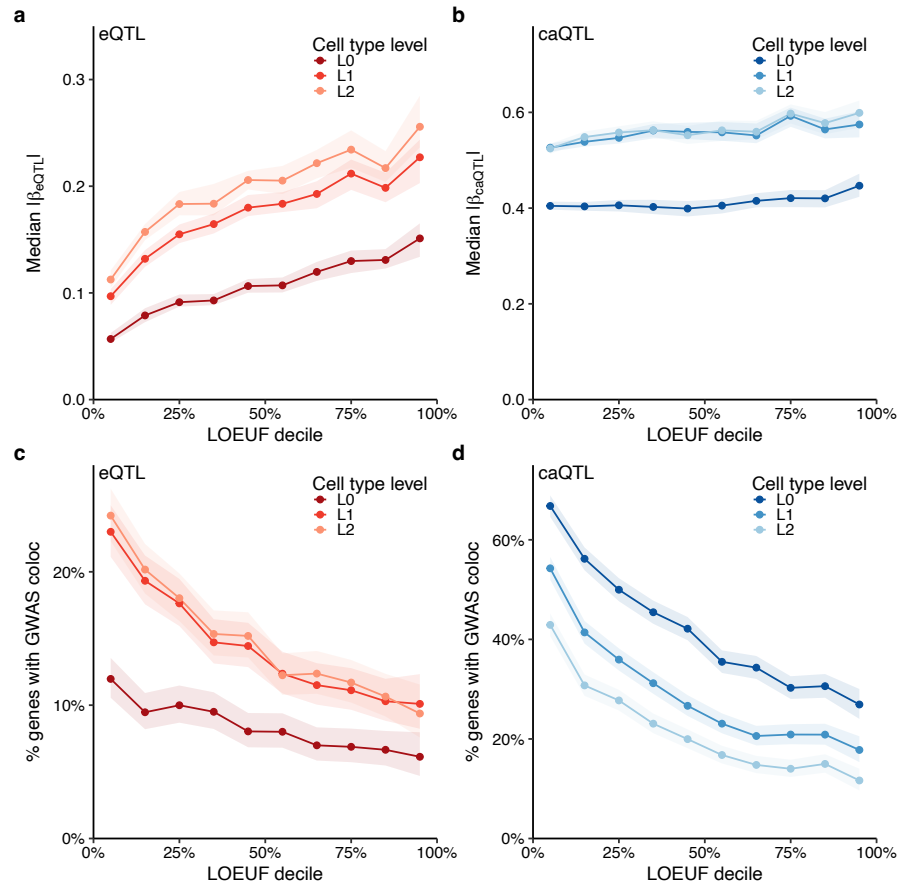
Supplementary Fig. 16 | QTL-mediated disease heritability.

Estimated proportion of disease heritability mediated by molecular traits. For immune, cancer, and cardiometabolic traits, MESC-estimated h^2_{med} / h^2_g within each trait category are meta-analyzed using random-effect model (**Methods**). The x-axis represents the combination of molQTLs for each estimate. Error bars represent 95% confidence intervals. Horizontal dotted lines represent the point estimate for caQTL in each trait category.



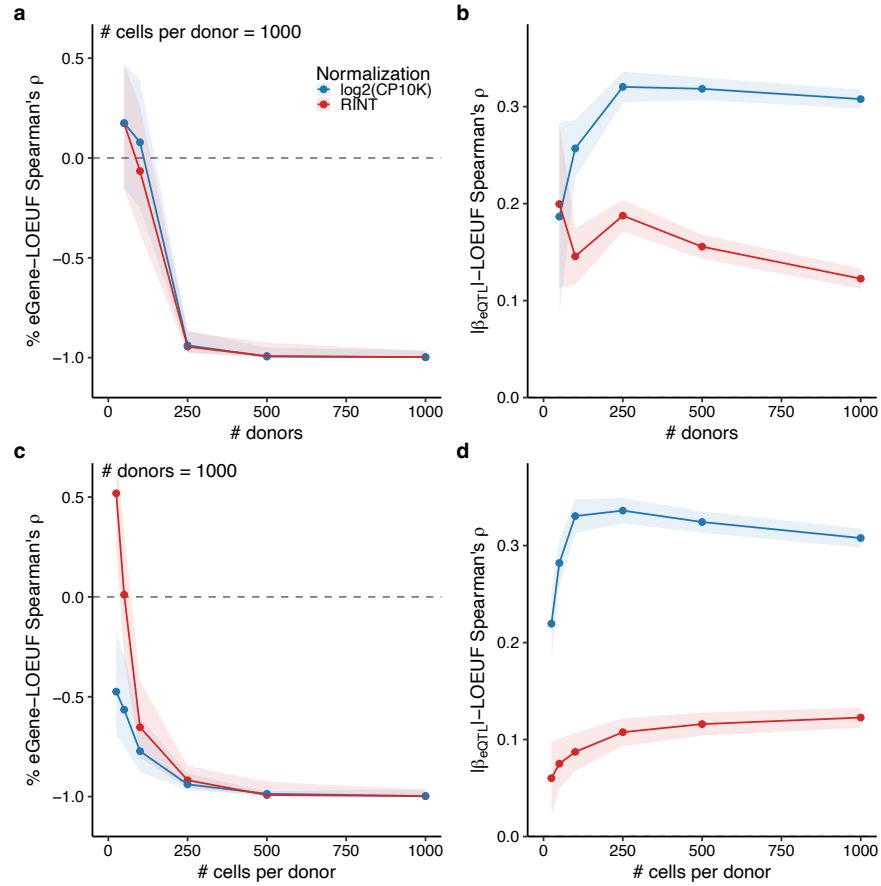
Supplementary Fig. 17 | Heritability enrichment of chromatin accessibility annotations.

a,b, Estimated proportion of disease heritability (**a**) and heritability enrichment (**b**), stratified by chromatin accessibility (CA) peak with links (CA-Link+), without links (CA-Link-), and all peaks combined. **c,d**, Same as **a,b** but stratified by regulatory modes (switch, rheostat, and dual). For immune, cancer, and cardiometabolic traits, S-LDSC estimates within each category are meta-analyzed using a random-effects model (REML; see **Methods**). Error bars represent 95% confidence intervals.



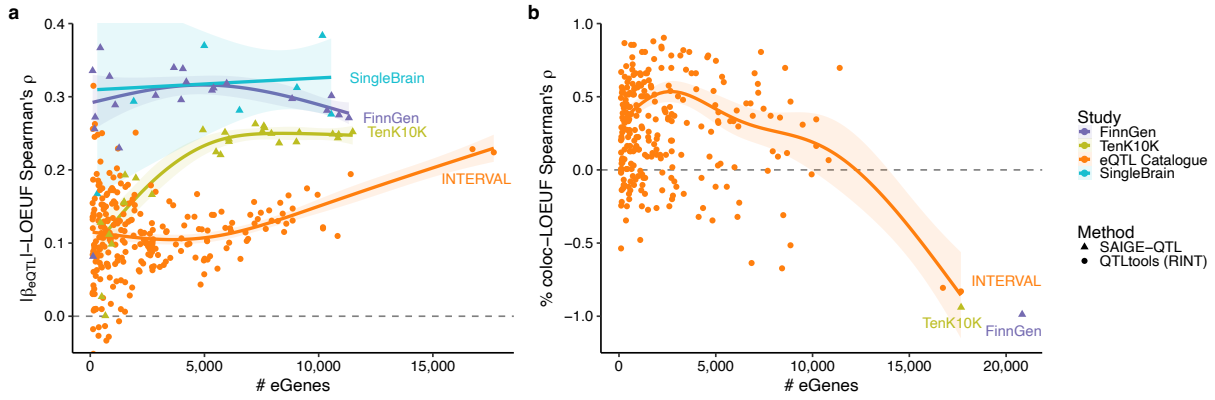
Supplementary Fig. 18 | Constraint gradients in molQTL effect size and GWAS colocalization across LOEUF deciles and cell type resolution levels.

a,b, Median absolute effect size of fine-mapped *cis*-eQTLs (**a**) and *cis*-caQTLs (**b**) for each loss-of-function observed/expected upper bound fraction (LOEUF) decile, stratified by cell type resolution level. **c,d**, The fraction of genes with *cis*-eQTL–GWAS colocalization (**c**) or *cis*-caQTL–GWAS colocalization (**d**) for each LOEUF decile, stratified by cell type resolution level. Color represents cell type level: L0 (PBMC), L1 (major cell types), or L2 (fine-grained cell types; see **Methods**). Shaded ribbons represent 95% confidence intervals.



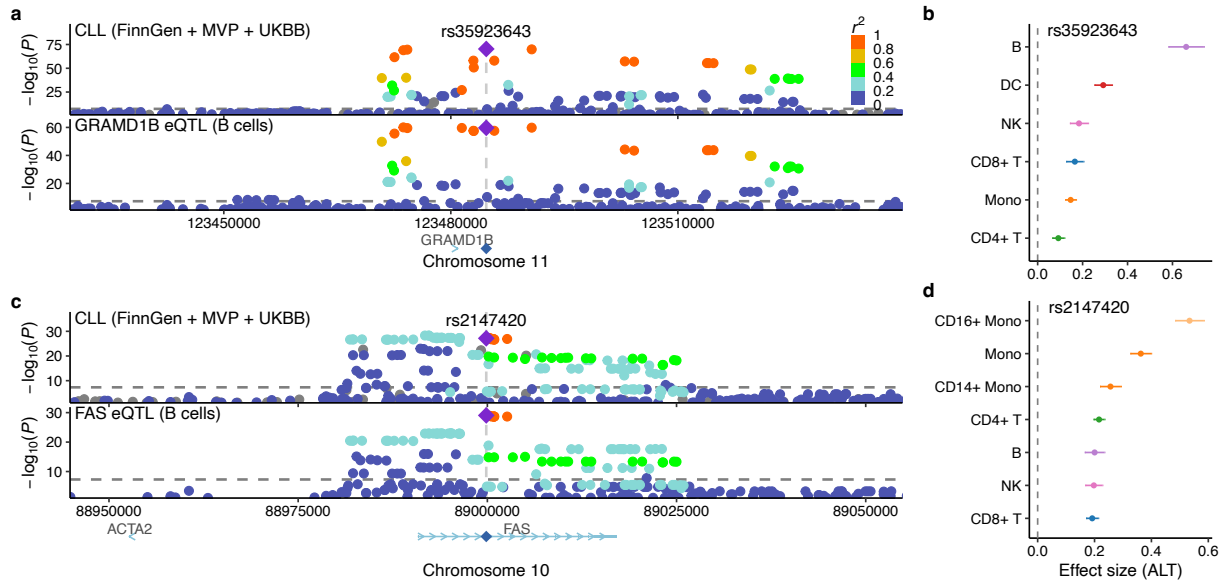
Supplementary Fig. 19 | Downsampling reveals a power-dependent transition in eGene detection alongside a stable effect size gradient.

Downsampling analysis of *cis*-eQTL mapping in CD4⁺ T cells. tensorQTL pseudobulk *cis*-eQTL mapping was performed under two normalization strategies: rank-based inverse normal transformation (RINT) and log2(CP10K + 1). For each condition, five independent downsampled replicates were combined by random-effects meta-analysis on Fisher z-transformed Spearman's ρ (REML; **Methods**). Points represent meta-analyzed estimates and shaded ribbons represent 95% confidence intervals. **a,b**, Donor-axis titration at fixed 1,000 cells per donor, with the number of donors varied from 25 to 1,000. **a**, Spearman's ρ between eGene detection fraction and LOEUF decile (ρ_{frac}) vs. the number of donors. **b**, Spearman's ρ between absolute lead *cis*-eQTL effect sizes ($|\beta_{eQTL}|$) and LOEUF (ρ_{eQTL}) across significant eGenes (FDR < 0.05) vs. the number of donors. **c,d**, Cell-axis titration at fixed 1,000 donors, with per-donor cell depth varied from 25 to 1,000 cells. **c**, ρ_{frac} vs. the number of cells per donor. **d**, ρ_{eQTL} vs. the number of cells per donor. Across both titration axes, ρ_{frac} (**a**, **c**; eGene detection rate) transitions from positive at low power to strongly negative at high power, while ρ_{eQTL} (**b**, **d**) remains stable.



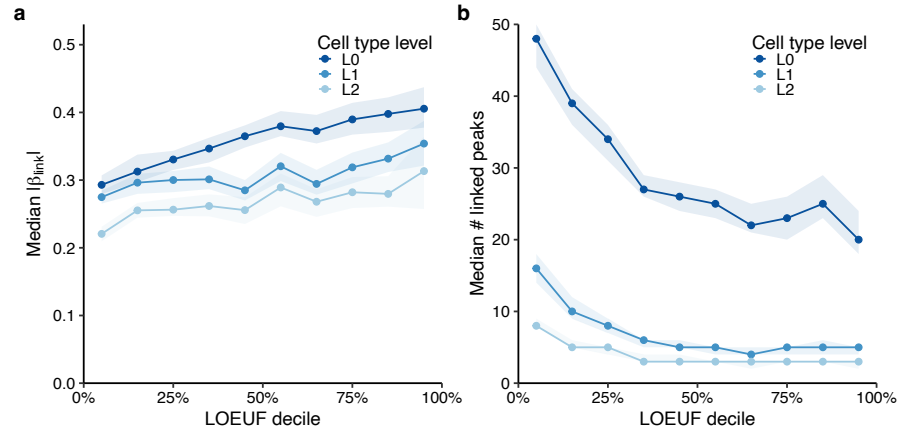
Supplementary Fig. 20 | Replication of the regulatory buffering gradient across independent *cis*-eQTL studies.

a, Spearman rank correlation (ρ) between absolute lead *cis*-eQTL effect sizes ($|\beta_{\text{eQTL}}|$) and LOEUF across significant eGenes (FDR < 0.05), plotted against the number of eGenes in each dataset (**Methods**). Smooth curves are GAM fits weighted by the inverse variance of each ρ estimate; shaded ribbons represent 95% confidence intervals. Horizontal dashed line indicates $\rho = 0$; positive values denote smaller eQTL effects at more constrained genes. **b**, Same as **a**, with the y-axis showing the Spearman's ρ between GWAS colocalization rate and LOEUF decile. FinnGen and eQTL Catalogue colocalizations are computed against 439 FinnGen GWAS traits (**Methods**); TenK10K colocalizations are reported by the original study using coloc.abf with external GWAS summary statistics. SingleBrain colocalizations were not available and are not shown. Negative values denote higher colocalization rates at more constrained genes.



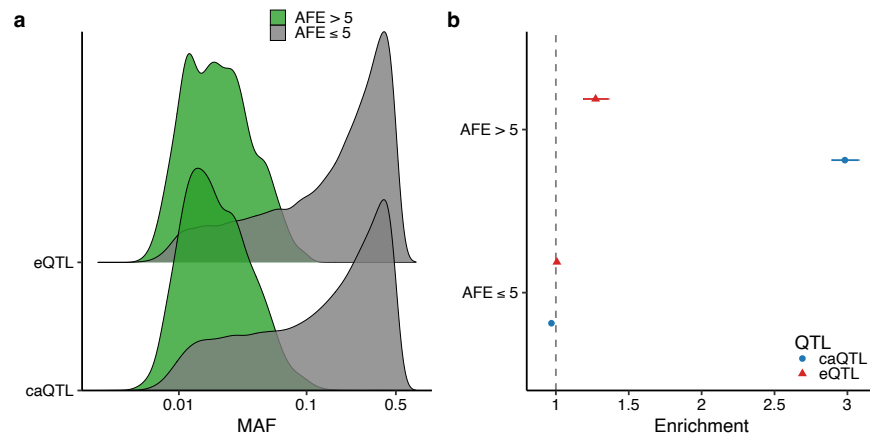
Supplementary Fig. 21 | Constrained genes with substantial *cis*-eQTL effects.

a,c, LocusZoom plots of *GRAMD1B* (**a**) and *FAS* (**c**) loci, colocalizing with chronic lymphocytic leukemia (CLL). For each locus, the top panel represents a LocusZoom plot of CLL in the FinnGen + MVP + UKBB meta-analysis. The bottom panel represents *cis*-eQTL effects of *GRAMD1B* or *FAS* expression in B cells. A lead variant for each locus is shown as a diamond. Color represents r^2 to the lead variant. Horizontal dotted line represents the genome-wide significance threshold ($P < 5.0 \times 10^{-8}$). **b,d**, Forest plots of rs35923643 (**b**) and rs2147420 (**d**) for *GRAMD1B* or *FAS* expression, respectively, across cell types. Error bars represent 95% confidence intervals. All P values are two-sided.



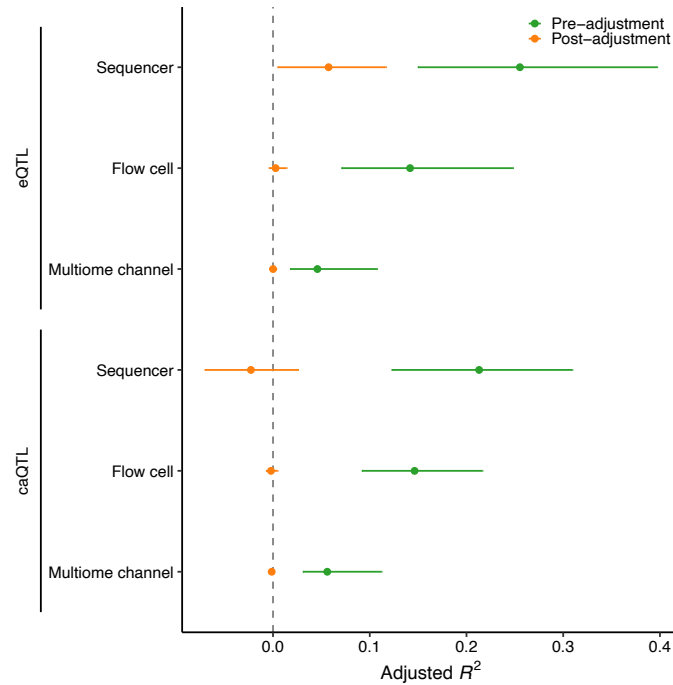
Supplementary Fig. 22 | Peak-gene link effect-size and count constraint gradients across cell type resolution levels.

a,b, Median absolute peak-gene link effect size ($|\beta_{link}|$; **a**) and median number of linked peaks per gene (**b**) for each loss-of-function observed/expected upper bound fraction (LOEUF) decile, stratified by cell type resolution level. Effect sizes are from the negative binomial count component of the hurdle model, restricted to rheostat and dual links. Color represents cell type level: L0 (PBMC), L1 (major cell types), or L2 (fine-grained cell types; see **Methods**). Shaded ribbons represent 95% confidence intervals.



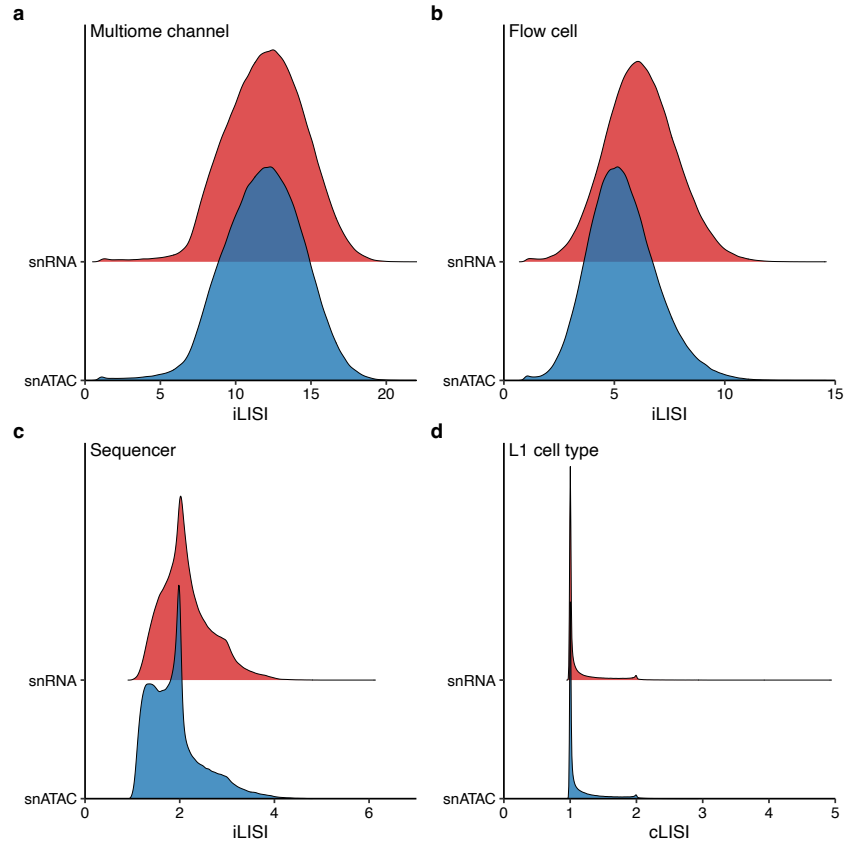
Supplementary Fig. 23 | Finnish allele frequency enrichment (AFE) and MAF distribution of fine-mapped molQTL variants.

a. The distribution of MAF for fine-mapped molQTL variants (PIP > 0.5) for Finnish-enriched (AFE > 5) and non-Finnish-enriched (AFE ≤ 5) variants. **b.** Enrichment of fine-mapped caQTL or eQTL variants (PIP > 0.5) for Finnish allele frequency enrichment (AFE > 5). Enrichment is defined as a risk ratio of being fine-mapped and being AFE > 5 (**Methods**). Error bars represent 95% confidence intervals.



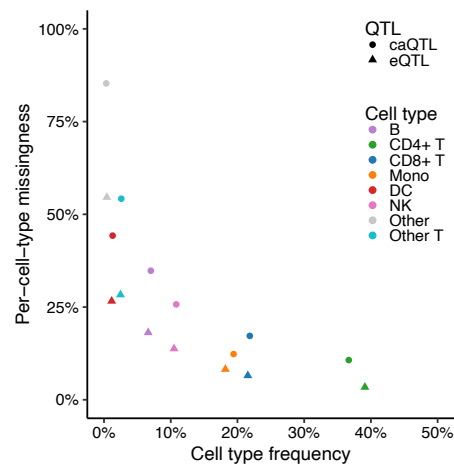
Supplementary Fig. 24 | Residual technical-batch variance in pseudobulk *cis*-molQTL features before and after latent-factor adjustment.

Adjusted R^2 of pseudobulk PBMC *cis*-eQTL or *cis*-caQTL features regressed on each donor's fractional cell composition across multiome channel ($n = 718$), flow cell ($n = 31$), or sequencer ($n = 7$) levels (**Methods**). Pre-adjustment uses the inverse-rank-normalized pseudobulk values; post-adjustment uses the same values after residualizing against the latent factors used in QTL mapping (PEER factors for eQTL, chromatin accessibility PCs for caQTL) and base covariates (age, sex, and top four genetic PCs). Points represent the median across features and horizontal bars span the interquartile range.



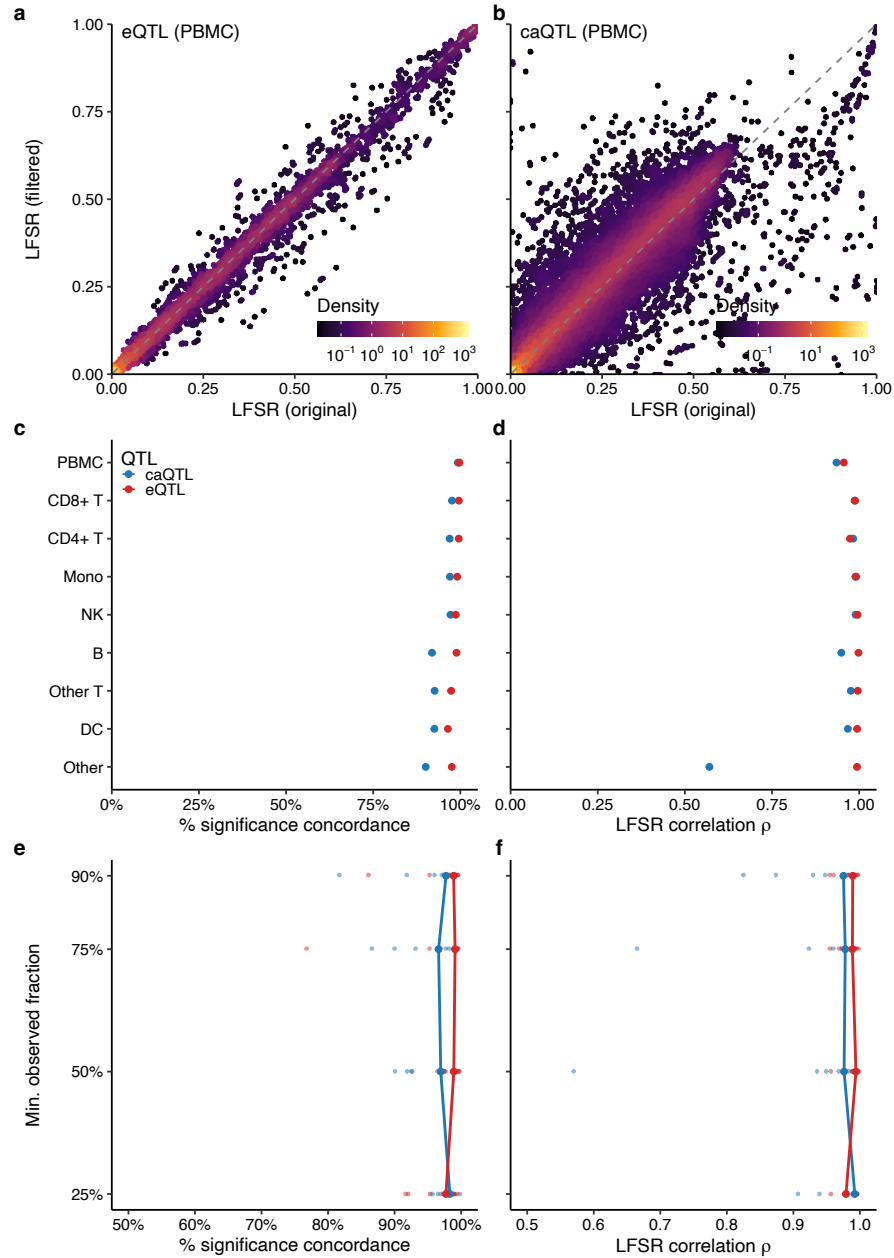
Supplementary Fig. 25 | Batch and cell-type mixing of the integrated snRNA-seq and snATAC-seq atlases assessed by LISI.

Local Inverse Simpson's Index (LISI) computed per nucleus at perplexity = 30 on a random subsample of 500,000 nuclei from the integrated snRNA-seq (PCA embedding) and snATAC-seq (snapATAC2 spectral embedding) atlases. iLISI approximates the effective number of batch labels in each nucleus's local neighborhood (higher = better mixing); cLISI approximates the effective number of cell-type labels (lower = cleaner separation). **a–c**, iLISI across multiome channels (a), flow cells (b), and sequencers (c). **d**, cLISI across L1 cell types.



Supplementary Fig. 26 | Per-cell-type missingness in the mash strong set and its relationship to cell type frequency.

Per-cell-type missingness, defined as the fraction of imputed entries in the strong set used for data-driven covariance estimation, vs. cell type frequency, for *cis*-eQTLs and *cis*-caQTLs. Each point represents one cell type; color represents cell type and shape represents QTL type.



Supplementary Fig. 27 | Robustness of mashr LFSR estimates to input missingness.

a,b, Density scatter plots of local false sign rates (LFSR) from the original mashr analysis vs. a sensitivity analysis using a stricter missingness threshold, for eQTL (**a**) and caQTL (**b**) in PBMC. Each point represents a variant–feature–cell type trio. Color represents point density (log10 scale). Dashed line indicates $y = x$. **c**, Percentage of variant–feature–cell type trios with concordant significance calls (LFSR < 0.05) between the original and sensitivity analyses, shown for each L1 cell type and stratified by QTL type. **d**, Spearman rank correlation (ρ) of LFSR values between the original and sensitivity analyses for each L1 cell type, stratified by QTL type. Error bars represent 95% confidence intervals. **e,f**, Titration of the observed-fraction threshold (variants observed in $\geq 25\%$, 50%, 75%, and 90% of L1 cell types), showing significance-call concordance (**e**) and LFSR Spearman correlation (**f**) vs. the full run at each threshold, stratified by QTL type. Points represent individual L1 cell types and lines connect the per-threshold medians.

Supplementary Note

Cohort and atlas composition

The first 339 donors were profiled in individual microfluidic Multiome assay channels (10x Genomics). To facilitate cohort scale-up, the remaining 832 samples were multiplexed in batches of 16 across eight channels, reducing both time and resource expenditure while maintaining high-quality data (**Extended Data Fig. 1a, Supplementary Table 1, Methods**).

Per-nucleus quality metrics for the two modalities were as follows. The snRNA-seq atlas had a median of 9,506 nuclei/donor, 1,056 genes/nucleus, and 1,828 unique molecular identifier (UMI)/nucleus (**Supplementary Fig. 1a,b**). The snATAC-seq atlas had a median of 6,692 nuclei/donor, 4,745 fragments/nucleus, and a transcription start site (TSS) enrichment score of 6.47 per nucleus (**Supplementary Fig. 1d,e**).

Cell type composition matched expectations for peripheral blood. The most frequent cell types recovered were CD14⁺ monocytes, CD4⁺ subsets (naive, T central memory), CD8⁺ subsets (naive, T effector memory), naive B cells, and natural killer (NK) cells (**Supplementary Fig. 1c, Supplementary Table 2**). The final atlas comprised 1,108 QC-passing donors, 29,602 genes, and 297,024 consensus peaks.

Peak–gene links and their regulatory modes

Annotation of linked peaks

The majority of peaks were annotated as distal enhancer-like signatures (dELS, 63.7%), while proximal enhancer-like signatures (pELS, 9.0%) and promoter-like signatures (PLS, 1.7%) showed the strongest enrichment for peak–gene links (risk ratio [RR] = 1.59 and 3.86, respectively; **Supplementary Fig. 5, Supplementary Table 6, Methods**).

Peak–gene links showed a median of 4 genes per peak and 46 peaks per gene. Among ENCODE4 cCRE classifications, rheostat peaks showed a slight enrichment for distal enhancer-like signatures (RR = 1.06), whereas switch and dual peaks did not, consistent with the more distal placement of rheostat links relative to promoters (**Fig. 1h**).

Regulatory modes vary across genes and cell types

A regulatory mode is a property of a peak–gene–cell type trio rather than of a peak alone, so a single peak can act in different modes in different contexts. At the peak level, regulatory modes often vary across multiple genes and cell types in which a peak is active, and only 20% of the 104,780 linked peaks acted in a single mode throughout (15% exclusively as rheostat and 5.2% exclusively as switch). Notably, 4.1% alternated between these two modes across genes and cell types, illustrating that the same peak can gate expression in one context while tuning expression levels in another (**Supplementary Fig. 4c, Supplementary Table 5, Methods**).

For example, a distal peak (chr6:246344-246843) located 145 kb upstream of *IRF4*, a transcription factor whose concentration specifies dendritic cell (DC) subset identity¹, acts as a rheostat in DCs, quantitatively modulating *IRF4* expression ($\beta_{\text{count}} = -0.40$, $P_{\text{count}} = 2.8 \times 10^{-16}$), but as a switch in monocytes, where it determines whether *IRF4* is expressed at all ($\beta_{\text{zero}} = 0.15$, $P_{\text{zero}} = 3.9 \times 10^{-4}$; **Supplementary Fig. 4d**), consistent with the dose-dependent role of *IRF4* in DC specification¹ vs. binary activation in monocytes². The same pattern extends across the

lymphoid-myeloid boundary. A distal peak (chr18:55235635-55236760) located 429 kb downstream of *TCF4*, which encodes E2-2, a transcription factor whose continuous expression is required to maintain plasmacytoid DC identity^{3,4}, acts as a switch in B cells ($\beta_{\text{zero}} = 0.15$, $P_{\text{zero}} = 4.3 \times 10^{-7}$) but as a rheostat in DCs ($\beta_{\text{count}} = 0.28$, $P_{\text{count}} = 2.0 \times 10^{-8}$; **Supplementary Fig. 4e**), consistent with dose-sensitive E2-2 tuning of DC subset fate vs. more binary E-protein activity during B-cell lineage commitment⁵. This principle extends beyond transcription factors. A distal peak (chr3:192063473-192063972) located 735 kb downstream of *CCDC50*, a selective autophagy cargo receptor that restrains type-I interferon signaling by delivering K63-polyubiquitinated RIG-I, MDA5, and STING for degradation^{6,7}, acts as a rheostat in B cells ($\beta_{\text{count}} = 0.23$, $P_{\text{count}} = 1.1 \times 10^{-23}$) but as a switch in DCs ($\beta_{\text{zero}} = 0.13$, $P_{\text{zero}} = 1.3 \times 10^{-3}$; **Supplementary Fig. 4f**), consistent with *CCDC50* providing graded dampening of antiviral signaling in B cells vs. switch-like engagement of the same brake in DCs. These examples illustrate a recurring theme in which the same regulatory element can gate expression in one cell type while tuning expression levels in another, with myeloid cells more frequently employing graded regulation (**Supplementary Table 5**).

Regulatory hotspots at gene clusters

Peak–gene links are unevenly distributed, concentrating at a few gene clusters. Notably, we identified genes with extensive regulatory links at 19p13.11, including *JUND*, *UBA52*, and *IFI30* (314, 307, and 268 linked peaks, respectively; **Extended Data Fig. 2a**), with the majority of links at this locus classified as rheostat (47%), suggesting these densely regulated genes are primarily fine-tuned across cell types rather than subject to binary gating. This region contains zinc finger clusters that evolve rapidly through copy number variation to generate transcriptional repressors against endogenous retroviruses⁸. Similarly, TNF receptor superfamily (TNFRSF) members demonstrated remarkable regulatory complexity, particularly at the 8p21.3 cluster containing TRAIL receptors (*TNFRSF10A–D*), collectively regulated by 133 peaks, of which 64% were shared across family members (**Extended Data Fig. 2b**). Among the four receptors, *TNFRSF10A* was predominantly switch-regulated, particularly in T cells, while *TNFRSF10B* was predominantly rheostat-regulated in myeloid cells, suggesting that pro-apoptotic and signaling functions within the same receptor family are controlled through different regulatory architectures.

Replication in an independent Japanese cohort

We replicated peak–gene links in an independent Japanese multiome cohort (BioBank Japan, BBJ; ~0.7 M nuclei, $n = 98$) by running the same hurdle model across L1 cell types (**Supplementary Fig. 6, Methods**). Replication rates increased monotonically with FinnGen significance. Among the most significant FinnGen links ($P_{\text{joint}} < 1.0 \times 10^{-20}$), 60% replicated at $P_{\text{joint}} < 0.05$ in BBJ, with effect size correlations of Spearman's $\rho = 0.69$ for the count component and 0.59 for the zero-inflation component (**Supplementary Fig. 6a–c**).

BBJ-replicating links also largely retained their FinnGen regulatory-mode assignment, with 66% overall concordance (rheostat 87%, switch 63%, and dual 57%; **Supplementary Fig. 6d–f**). The residual discordance was directional rather than random. Among the most significant FinnGen dual links, 57% remained dual in BBJ while 43% were detected as a single mode under reduced power, predominantly rheostat (30%) rather than switch (13%), and direct switch-to-rheostat flips remained rare (< 2%). Reduced power therefore collapses dual links to a single mode rather than reshuffling switch and rheostat, which bounds the direction of the mode-classification uncertainty described in the **Methods**.

Genetic architecture of molQTLs

Trans-molQTL analysis

We tested associations between molecular traits and genetic variants located in *trans* (*trans*-eQTLs, > 5 Mb away or on different chromosomes) using tensorQTL to explore distal genetic regulation (**Supplementary Fig. 7**). At $P < 5.0 \times 10^{-8}$, we identified 20,826 unique *trans*-eGenes and 215,949 unique *trans*-caPeaks in at least one cell type, of which 85% and 76% are also *cis*-eGenes and *cis*-caPeaks, respectively, suggesting the majority of genes and peaks are both *cis*- and *trans*-regulated. Among 86,536 *cis*-eQTL and 636,566 *cis*-caQTL 95% CS-gene-cell type trios, only 2,704 (3.1%) and 37,916 (6.0%) trios showed at least one significant *trans*-eQTL or *trans*-caQTL association with another gene or peak for the same cell type, respectively, suggesting distinct genetic regulation of *cis* and *trans* components. To additionally account for the number of features tested per cell type, we applied Bonferroni correction (dividing 5.0×10^{-8} by the number of genes or peaks tested per cell type), yielding 17 *trans*-eGenes and 36 *trans*-caPeaks at this conservative threshold (**Supplementary Tables 10,11**).

To identify potential false positives arising from read cross-mapping between homologous genes⁹, we filtered *trans*-eQTL associations where any gene within 1 Mb of the lead variant had cross-mappability > 0 with the *trans*-eGene, using precomputed genome-wide cross-mappability scores (GRCh38, GENCODE v26; <https://doi.org/10.6084/m9.figshare.7315136>). This removed 36% of *trans*-eQTL associations and 4,271 *trans*-eGenes, retaining 16,555 unique *trans*-eGenes.

Coordinated and layer-specific effects across molecular layers

Comparing caQTLs, eQTLs, and plasma pQTLs at the level of fine-mapped CSs shows both variants that act across layers and variants confined to a single layer (**Fig. 2f**). These patterns reveal that molQTLs operate through both coordinated effects across molecular layers and layer-specific mechanisms, highlighting the complex relationship between chromatin state, gene expression, and protein abundance.

For example, we identified a variant upstream of *LILRB4*, rs776736:A>G, that increases *LILRB4* gene expression in monocytes ($\beta = 0.16$, $P = 1.7 \times 10^{-25}$, PIP = 1.0) and protein expression ($\beta = 0.34$, $P = 6.5 \times 10^{-16}$, PIP = 1.0) in plasma (**Extended Data Fig. 4b**). *LILRB4* is an inhibitory immune checkpoint receptor expressed primarily on monocytes and dendritic cells that suppresses T cell activation^{10,11}.

Functional enrichment of fine-mapped molQTL variants

Fine-mapped molQTL variants were annotated against ENCODE4 candidate cis-regulatory elements (cCREs) and variant consequence categories, with enrichment estimated as a relative risk (**Methods**).

Fine-mapped molQTL variants were dramatically enriched for chromatin accessibility peaks with peak–gene links (CA-Link+), particularly for caQTLs (RR = 20) compared to eQTLs (RR = 8.9), while peaks without such links (CA-Link–) showed modest enrichment (**Fig. 2i, Supplementary Table 14**). Stratifying CA-Link+ peaks by their regulatory mode revealed a monotonic enrichment gradient: dual peaks were most enriched (RR = 11 and 23 for eQTL and caQTL, respectively), followed by peaks that alternated between switch and rheostat across genes or cell types (RR = 6.7 and 17), exclusively switch (RR = 5.9 and 14), and exclusively rheostat peaks (RR = 3.0 and 12; **Extended Data Fig. 4f, Supplementary Table 16**). Among ENCODE4 cCRE classifications, PLS exhibited the highest enrichment (RR = 16 and 19 for eQTL and caQTL, respectively), followed by pELS (RR = 4.7 and 5.6, respectively) and dELS (RR = 2.0 and 3.3, respectively),

while CA-only elements showed little or no enrichment (RR = 0.49–0.69 and 0.51–1.1, respectively). These functional enrichment patterns underscore the regulatory nature of most molQTL variants and highlight the utility of our peak–gene links for nominating functional variants.

Replication across independent datasets

Among FinnGen fine-mapped *cis*-eQTLs with PIP > 0.9, effect direction was concordant for 96% and 89% of variants in TenK10K¹² and OneK1K¹³ respectively, 59% and 21% reached genome-wide significance ($P < 5.0 \times 10^{-8}$), and 40% and 14% fell within an external 95% CS (**Supplementary Fig. 8a–d**). Effect size correlations at PIP > 0.9 were Spearman's $\rho = 0.92$ and 0.80 across all matched variants, rising to 0.96 and 0.86 among lead variants. Replication was also cell-type-resolved. Comparing effect sizes for fine-mapped variants (PIP > 0.5) across every pair of FinnGen and TenK10K cell types, corresponding cell types clustered together (**Supplementary Fig. 8e**).

For the lead-variant-only datasets (CIMA¹⁴, AIDA¹⁵, and BBJ¹⁶), overlapping lead *cis*-eQTLs showed near-perfect sign concordance (99–100%; **Supplementary Fig. 9a–c**). Lead-variant overlap itself was modest, at 3–6% of FinnGen *cis*-eQTLs with PIP > 0.9, reflecting the power difference between the studies rather than discordance among the variants that do overlap. Overlapping lead *cis*-caQTLs in CIMA and BBJ showed 99.7% sign concordance (**Supplementary Fig. 9d–f**).

Context-dependent regulation

To assess the extent to which our steady-state PBMC atlas captures context-dependent regulation, we intersected fine-mapped FinnGen molQTLs with stimulus-responsive eQTLs from eight eQTL Catalogue¹⁷ studies with paired naive and stimulated immune conditions (**Methods**).

We defined 4,626 stimulus-responsive eQTLs that were fine-mapped under stimulation (PIP > 0.5) but absent at baseline (PIP < 0.01 in naive). Of these, 16% were fine-mapped in our FinnGen molQTLs (PIP > 0.5), whereas 48% fell outside *cis* of any FinnGen eGenes or caPeaks (**Supplementary Fig. 10**). The fraction of stimulus-responsive eQTLs captured by FinnGen molQTLs increased monotonically with FinnGen PIP, with *cis*-caQTL variants capturing stimulus-responsive eQTLs at higher rates than *cis*-eQTL variants, consistent with enhancer priming¹⁸. These fractions likely represent a lower bound, given the limited set of stimuli profiled across the eight eQTL Catalogue studies.

Cell type specificity of molecular QTLs

Full cell-type specificity breakdown

Among 20,829 eGenes, 73% showed shared regulation (51% cross-lineage and 23% likely shared but underpowered), while 27% exhibited cell type specific patterns (9% lineage-specific, 4% T-cell-specific, and 14% single cell-type; **Fig. 3b, Supplementary Table 18**). Chromatin accessibility showed comparable sharing: of 210,584 caPeaks, 77% showed sharing (33% cross-lineage and 44% likely shared), with 23% displaying cell type specific patterns (7% lineage-specific, 2% T-cell-specific, and 14% single cell-type; **Fig. 3c, Supplementary Table 19**). The large proportion of likely shared but underpowered caPeaks (44%) suggests that the genetic regulation of chromatin accessibility is broadly shared across immune cell types, with observed specificity solely based on significance often reflecting limited statistical power rather than true biological differences.

Zinc finger genes

Among 375 ZNF genes analyzed (including *ZNF*, *ZBTB*, *ZSCAN*, *ZKSCAN*, and *ZFP* families; **Methods**), 345 (92%) were identified as eGenes. Of these 271 showed shared regulation (220 cross-lineage and 51 likely shared but underpowered), while 74 exhibited cell-type specific patterns (38 lineage-specific, nine T-cell-specific, and 27 single cell-type).

The main text provides the variant, cell types, effect sizes, and *P* values for *IKZF1*, *FOXO1*, *NFATC1*, *NFATC2*, and *ZBTB1* (**Fig. 3d–h**). Similarly, the *cis*-eQTL for *ZNF804A* (rs10497655:T>C) decreased *ZNF804A* expression in B cells but increased expression in monocytes ($\beta = -0.18$ and 0.13 , $P = 1.0 \times 10^{-17}$ and 4.6×10^{-87} , respectively; **Fig. 3i**), while the *cis*-eQTL for *ZNF7* (rs1209879:C>T) decreased *ZNF7* expression in CD14⁺ monocytes and increased in CD4⁺ naive T cells ($\beta = -0.16$ and 0.06 , $P = 2.0 \times 10^{-45}$ and 3.3×10^{-11} , respectively; **Fig. 3j**), demonstrating opposing effects between lymphoid and myeloid lineages.

CASCADE mechanism classification

Each fine-mapped molQTL variant was classified into one of 25 mutually exclusive patterns and grouped into seven mechanism categories (**Fig. 4a**, **Extended Data Fig. 5a**, **Methods**).

Full classification breakdowns

Among all 119,094 fine-mapped molQTL variants, 39% overlapped chromatin accessibility peaks, of which 71% (28% of all variants) showed caQTL effects for those overlapping peaks (**Fig. 4b**, **Supplementary Table 20**). Conversely, among the 61% of variants not overlapping peaks, 89% (54% of all variants) affected non-overlapping peaks as caQTLs. An additional 10% of all variants overlapped peaks but affected different, non-overlapping peaks, demonstrating complex accessibility regulation. Strikingly, 54% of molQTL variants have peak–gene links (18% via overlapping peak caQTL, 35% via non-overlapping peak caQTL, and 1% via peak overlap alone), highlighting functional connectivity between chromatin and gene regulation.

The 25 patterns collapsed into seven mechanistic categories (**Fig. 4a**). We defined a chromatin-to-expression cascade as a variant connected to its eGene through a peak–gene link, and resolved three cascade tiers by where the causal peak locates relative to the variant (**Methods**). Local cascade variants ($n = 1,410$; 1.2% of fine-mapped molQTL variants) complete a full caQTL–link–eQTL path through the variant’s own overlapping peak, the canonical case of direct chromatin perturbation at the variant’s regulatory element. Distal cascade variants ($n = 948$; 0.8%) complete the same caQTL–link–eQTL path through a non-overlapping distal peak, with the variant itself typically outside any peak. Positional cascade variants ($n = 1,133$; 1.0%) overlap a linked peak and show an eQTL effect but no detectable caQTL at that peak, likely reflecting TF binding or other mechanisms invisible to caQTL analysis. Another 2% ($n = 2,430$) showed both caQTL and eQTL effects without peak–gene links, suggesting either incomplete regulatory mapping or genuine pleiotropy. The vast majority exhibited partial effects: 87% showed only caQTL (49% with links and 38% without), while 8% showed only eQTL (**Fig. 4b**, **Supplementary Table 20**). These fractions change substantially when analyzing only fine-mapped eQTL variants ($n = 16,477$), where 23% demonstrate complete cascades (local, positional, and distal), 16% showed both caQTL and eQTL effects without links, and 61% remain as only eQTL variants, reflecting the excess of only caQTL variants ($n = 103,862$) in our fine-mapped molQTLs (**Extended Data Fig. 5b**). Beyond the 3,491 variants completing a full cascade, a further 70,580 fine-mapped variants connect to a target gene through a peak–gene link or an eQTL effect without completing one.

Variant-level cell type specificity varied across mechanisms (**Fig. 4b, Supplementary Table 20**). Overall, 27% of variants showed shared effects across cell types (12% cross-lineage and 15% likely shared), whereas 73% displayed cell type specific patterns (10% lineage-specific, 7% T-cell-specific, and 56% single cell-type). Cascade variants (local, positional, distal cascades) were slightly more cell type-specific (**Extended Data Fig. 5e**): 23% were shared (7% cross-lineage and 16% likely shared) and 77% were cell type-specific (9% lineage-specific, 8% T-cell-specific, 61% single cell-type), suggesting that individual regulatory variants tend to operate through cell-type specific mechanisms even though eGenes and caPeaks themselves are largely shared across cell types.

A local cascade at BACH2

rs72928038:G>A was independently shown to modulate expression in a massively parallel reporter assay (MPRA) of autoimmune-associated regulatory variants in T cells¹⁹. Because the enhancer and promoter peaks are co-accessible in individuals carrying neither variant, co-accessible peaks should be fine-mapped jointly rather than independently.

MPRA validation of fine-mapping and cascade tiers

To provide orthogonal functional validation, we intersected molQTL variants with recent systematic MPRA data²⁰, which tested the allelic effects of 221,412 fine-mapped variants and 86,064 control variants across five cell lines including K562. Of the MPRA-tested variants, 172,042 were also tested in our molQTL analysis, of which 10,428 are fine-mapped (PIP > 0.5). Expression-modulating variant (emVar) rates increased monotonically with our molQTL PIP for both eQTL and caQTL variants, from ~5–10% at null or low-PIP variants to ~20% (K562) and ~45% (meta-analyzed across cell types) at PIP > 0.9 (**Extended Data Fig. 6a,b**), confirming that our fine-mapping accurately identifies functional regulatory variants. Among emVars, the concordance between MPRA allelic effects and endogenous QTL effects also scales with PIP (Spearman's $\rho = 0.63$ for caQTL and $\rho = 0.39$ for eQTL at PIP > 0.9 in K562; **Extended Data Fig. 6c,d**), consistent with prior MPRA-eQTL concordance²⁰. Notably, caQTL concordance is higher than eQTL concordance at PIP > 0.9, consistent with MPRA and caQTLs both measuring variant effects on *cis*-regulatory elements, whereas eQTLs reflect the net endogenous impact on gene expression.

Across CASCADE categories, emVar rates differed sharply among the three cascade tiers (**Extended Data Fig. 6e,f**). Local cascade variants showed the highest MPRA activity (37% in K562 and 73% in meta), consistent with direct allele-specific activity on the variant's own regulatory sequence. Distal cascade variants, despite completing the caQTL–link–eQTL path, showed emVar rates indistinguishable from only eQTL variants (5.1% and 18% vs. 7.7% and 23%). This depletion is expected because MPRA tests a 200-bp sequence centered on the variant²⁰, whereas the regulatory element in distal cascades likely lies elsewhere; such variants may reflect 3D genome contacts bridging the variant's locus to the distal peak, or regulatory effects driven by distal tagging non-SNV causal variants, such as SVs, not included in our current analysis. Positional cascade variants showed intermediate activity (23% and 55%), consistent with direct regulatory activity through mechanisms such as TF binding that produces no detectable caQTL. These MPRA signatures independently validate our three-tier cascade categories, confirming that local and positional cascade variants act directly on their own sequence whereas distal cascade variants do not.

Colocalization with GWAS

Colocalization statistics and controls

From the perspective of molecular traits, of 208,659 peaks and 20,598 genes tested for colocalization, 11% and 19% respectively had at least one GWAS colocalization, indicating that a substantial fraction of genetically regulated molecular traits are directly relevant to diseases.

When incorporating plasma *cis*-pQTLs from FinnGen (PIP > 0.5), pQTLs showed comparable colocalization rates to other molQTL-only categories (6.2% for any pQTL variants regardless of CASCADE status, 6.1% for pQTL-only variants without caQTL/eQTL effects; **Methods**), though with substantially fewer variants available. As a negative control, we asked how often variants with no detectable molQTL effect colocalize at all. Variants with no detected molQTL effects (linkage disequilibrium [LD]-independent variants with max. PIP < 0.01 in both eQTL and caQTL, **Methods**) had minimal colocalization (0.09%).

Summary-based Mendelian randomization (SMR)

To independently validate these colocalization results for causality, we performed summary-based Mendelian randomization (SMR)²¹ across 11,584 colocalizing GWAS loci, 3,868 colocalizing genes, and 23,025 colocalizing peaks, testing for 153,395 colocalizing trios across cell types (cell type–GWAS locus–molQTL locus; **Methods**). SMR tests whether a genetic variant mediates a molecular QTL's effect on disease risk, with the HEIDI (heterogeneity in dependent instruments) test distinguishing pleiotropy from linkage. Of 33,443 significant SMR associations (Bonferroni $P_{\text{SMR}} < 3.3 \times 10^{-7}$ and $P_{\text{HEIDI}} > 0.05$), 4,048 of 11,104 caQTL-colocalizing GWAS loci (36%) and 1,698 of 5,974 eQTL-colocalizing loci (28%) had at least one SMR-significant association in the same cell type (**Supplementary Fig. 11a, Supplementary Table 24**). SMR rates mirrored the CASCADE colocalization hierarchy, with the SMR-to-colocalization ratio increasing from 21% for only eQTL or caQTL without peak–gene links to 58% for local cascade variants (**Supplementary Fig. 11b,c, Supplementary Table 22**). The higher caQTL colocalization rate likely reflects genetic effects on chromatin accessibility that do not fully propagate to detectable expression changes, and SMR confirms that these caQTLs are nonetheless causally linked to disease risk.

Benchmarking against the eQTL Catalogue

To benchmark our colocalizations against existing resources, we colocalized 287 eQTL Catalogue datasets¹⁷ against the same FinnGen GWAS using the identical colocalization pipeline (**Supplementary Fig. 13a,b, Supplementary Table 28; Methods**). Of 11,015 gene-trait eQTL colocalizations, 6,808 (62%) were also detected in at least one external dataset, while 4,207 (38%) were novel, with the shared colocalizations primarily captured by large bulk immune-cell studies such as INTERVAL ($n = 4,729$)²². The novel colocalizations were enriched for cell-type-specific effects and CASCADE mechanisms not detectable in bulk tissue studies (**Supplementary Table 29**). To make this readily available, the external replication status of each colocalization, including those highlighted throughout this study, is annotated per dataset category (bulk and single-cell, immune and non-immune, with the best external study for each). Our GWAS-eQTL colocalizations were further replicated in TenK10K¹², despite the methodological differences in colocalization pipelines (**Supplementary Fig. 13c**).

Regulatory cascades for autoimmune hypothyroidism

Of 215 fine-mapped AIHT 95% CSs in FinnGen, only two CSs harbor pLoF variants (0.9%) and 29 contain missense variants (14%), while the majority (56%) are colocalized with molQTLs: 46 colocalized with both caQTL and eQTL (21%), 41 (19%) and 28 (13%) colocalized with caQTL only with and without link, respectively, and 5 colocalized with eQTL only (2.3%; **Fig. 5c**, **Supplementary Tables 25,26**).

The two completely linked variants upstream of *TICAM1* act through a proximal peak at chr19:4837478-4837977, which overlaps rs10424978 and lies 3 bp from rs10425559. Both the peak and *TICAM1* expression colocalize with AIHT and with skin cancer (PP.H4 = 1.0 for all four pairs; **Extended Data Fig. 7a**, **Supplementary Fig. 12a–c**).

The regulatory architecture of AIHT risk is further illustrated by rs13105678:C>A near *RHOH*, which demonstrates a local cascade mechanism in CD4⁺ T cells. This variant increases AIHT risk (OR = 1.04, $P = 1.3 \times 10^{-8}$, PIP = 0.70) by increasing *RHOH* expression ($\beta = 0.02$, $P = 1.3 \times 10^{-8}$, PIP = 0.99; PP.H4 = 1.0) and chromatin accessibility of an overlapping peak (chr4:40258141-40259058; $\beta = 0.89$, $P = 1.8 \times 10^{-176}$, PIP = 1.0; PP.H4 = 1.0) that is linked to *RHOH* (dual link; $\beta_{\text{zero}} = 0.14$, $P_{\text{zero}} = 2.4 \times 10^{-30}$; $\beta_{\text{count}} = 0.11$, $P_{\text{count}} = 3.2 \times 10^{-43}$; **Extended Data Fig. 7b**, **Supplementary Fig. 12d–f**). *RHOH* encodes a GTPase-deficient Rho family protein that recruits ZAP70 kinase to the T cell receptor complex²³. Intriguingly, we identified a Finnish-enriched missense variant in *ZAP70* itself (rs145955907:C>T, p.Thr155Met, MAF = 2%) that independently increases AIHT risk (OR = 1.26, $P = 2.3 \times 10^{-23}$, PIP = 0.25)²⁴. While the interaction between these variants was not significant (OR = 1.04, $P = 0.1$), the convergence of regulatory and coding variation on T cell receptor signaling components highlights how multiple genetic mechanisms can converge on shared pathways.

Similarly, an intronic variant of *RNASET2*, rs3798307:G>A, demonstrates a positional cascade mechanism for *RNASET2* expression in monocytes with opposite effects on AIHT (OR = 0.94, $P = 1.8 \times 10^{-16}$) and skin cancer (OR = 1.06, $P = 5.6 \times 10^{-9}$; **Extended Data Fig. 7c**). In monocytes, rs3798307 increases *RNASET2* expression ($\beta = 0.08$, $P = 1.9 \times 10^{-32}$, PIP = 0.16; PP.H4 = 0.99 for skin cancer) and overlaps with a chromatin accessibility peak (chr6:166950262-166950954) that is linked to *RNASET2* (dual link; $\beta_{\text{zero}} = 0.03$, $P_{\text{zero}} = 0.04$; $\beta_{\text{count}} = 0.16$, $P_{\text{count}} = 3.7 \times 10^{-35}$; **Supplementary Fig. 12g–i**). *RNASET2* encodes a secreted ribonuclease that degrades extracellular RNA, a damage associated molecular pattern (DAMP) and potent immune activation signal, to limit inflammation²⁵. Its increased expression could shift immune responses toward anti-inflammatory phenotypes, reducing thyroid inflammation while creating a tumor-permissive microenvironment.

molQTLs beyond immune diseases

The *APOBEC3A/CBX6* locus illustrates how our cascade mechanisms can nominate co-regulated candidate genes beyond a known cancer susceptibility deletion polymorphism. An intronic variant of *APOBEC3A*, rs12628403:A>C, increases risk of breast cancer (OR = 1.1, $P = 2.6 \times 10^{-12}$, PIP = 1.0; **Extended Data Fig. 8c**) and tags the *APOBEC3A/B* deletion polymorphism, which elevates *APOBEC3A* enzymatic activity through 3' UTR fusion with *APOBEC3B*^{26–28}. Through a distal cascade mechanism across immune cells, rs12628403 decreases chromatin accessibility at 12 peaks (PIP = 1.0 for all) that are linked to both *APOBEC3A* and *CBX6* (six peaks shared, four in *cis* of *APOBEC3A* linked to *APOBEC3A* alone, and two in *cis* of *CBX6* linked to *CBX6* alone, located ~90 kb upstream of rs12628403). The risk allele decreases *APOBEC3A* expression in monocytes ($\beta = -0.68$, $P = 2.0 \times 10^{-137}$, PIP = 1.0), and decreases *CBX6* expression across lymphoid and myeloid lineages, with a stronger effect in

monocytes than in CD4⁺ T cells ($\beta = -0.19$ vs. -0.07 , $P = 8.8 \times 10^{-54}$ vs. 3.1×10^{-13} , PIP = 1.0 for both; **Supplementary Fig. 14d,e**). Although the risk allele decreases total *APOBEC3A* expression in monocytes, it shifts *APOBEC3A* transcription away from the canonical *APOBEC3A* 3' end ($\beta = -1.1$, $P = 3.1 \times 10^{-70}$, PIP = 1.0) and toward use of the *APOBEC3B* 3' UTR ($\beta = 0.70$, $P = 9.2 \times 10^{-15}$, PIP = 1.0; GTEx whole blood; **Supplementary Fig. 15**), directly capturing the chimeric *APOBEC3A-3B* transcript that drives the APOBEC mutational signature pervasive in breast cancer^{29,30}. *CBX6* encodes a Polycomb Repressive Complex 1 chromobox component reported as a potential tumor suppressor in breast cancer³¹, consistent with its broad down-regulation by the risk allele across PBMCs.

Heritability mediated by molQTLs

To quantify the proportion of disease heritability mediated by molecular traits, we applied mediated expression score regression (MESC)³² using *cis*-eQTLs and *cis*-caQTLs from PBMC cell types as well as plasma *cis*-pQTLs from FinnGen (**Methods**). For immune diseases, *cis*-caQTLs mediated $79 \pm 4\%$ of disease heritability, with *cis*-eQTLs mediating $25 \pm 2\%$ and *cis*-pQTLs $4 \pm 0.4\%$ (**Supplementary Fig. 16, Supplementary Table 30**). Notably, the combined contribution of all three QTL types did not exceed that of caQTLs alone, indicating that caQTLs are the primary mediator through which regulatory variation manifests as gene expression and protein abundance differences. For cancer and cardiometabolic diseases, caQTLs mediated $45 \pm 3\%$ and $34 \pm 2\%$ of disease heritability, respectively, highlighting the importance of quantifying disease-relevant cell types.

To partition heritability by chromatin accessibility peaks, we applied stratified LD score regression (S-LDSC)³³, which captures heritability enrichment of genomic annotations regardless of mediation (**Methods**). All chromatin accessibility peaks captured $65 \pm 5\%$ of immune disease heritability, $35 \pm 4\%$ for cancer, $12 \pm 2\%$ for cardiometabolic diseases (**Supplementary Fig. 17a, Supplementary Table 31**), compared to 79%, 45%, and 34% from MESC estimates for caQTL-mediated heritability, respectively. When stratified by peak–gene link status, peaks linked to genes (CA-Link+) showed substantially higher enrichment (21-fold) compared to peaks without links (CA-Link–; 9-fold) for immune diseases (**Supplementary Fig. 17b**). This enrichment difference diminished for non-immune diseases, consistent with the cell-type specificity of immune regulatory effects. Regulatory modes (switch, rheostat, and dual) showed comparable heritability enrichment for immune diseases (**Supplementary Fig. 17c,d**). Together, these analyses establish chromatin accessibility as a primary mediator of disease heritability, particularly for immune diseases where disease-relevant cell types were directly profiled. Nevertheless, eQTL-only variants retain non-negligible disease colocalization (6.4%; **Fig. 5b**), indicating regulatory mechanisms independent of detectable chromatin accessibility.

Regulatory buffering at constrained genes

Robustness of the constraint gradient in cis-eQTLs

Consistent with previous findings^{34–36}, constrained genes showed higher mean expression (Spearman's $\rho = -0.33$, $P = 2.0 \times 10^{-386}$) with lower variance (Spearman's $\rho = 0.07$, $P = 2.7 \times 10^{-13}$ for correlation between LOEUF and expression variance residuals corrected by mean expression; **Extended Data Fig. 9b,c**); however, the eQTL effect size gradient remained robust even among highly-expressed genes (top 20%) subsequently matched for expression levels across constraint bins (Spearman's $\rho = 0.16$, $P = 3.3 \times 10^{-9}$). Simultaneously matching for both mean expression and expression variance residuals further attenuated but did not eliminate the gradient (Spearman's $\rho = 0.10$, $P = 3.7 \times 10^{-6}$), attributing approximately two-thirds of the residual

gradient to regulatory redundancy rather than variance differences (**Extended Data Fig. 9d**). The gradient was also recapitulated in pseudobulk tensorQTL across normalization choices (**Extended Data Fig. 9e**) and was robust to cell-type resolutions, with a near-constant constraint gradient (Spearman's $\rho = 0.26, 0.23$, and 0.22 between LOEUF and $|\beta_{\text{eQTL}}|$ for PBMC [L0], L1, L2, respectively; **Supplementary Fig. 18a,b; Methods**).

Despite these weaker individual effects, *cis*-eQTLs at highly constrained genes were more likely to colocalize with GWAS traits (28% of constrained genes showed colocalization compared to only 12% of unconstrained genes; **Fig. 5f**). This colocalization gradient was negative and significant at every mapping resolution (L0, L1, and L2; Spearman's $\rho = -0.96$ to -1.0), with its detectable magnitude increasing from L0 to L1 and then saturating at L2 (LOEUF decile $\times$ cell-type level interaction $P = 3.1 \times 10^{-3}$; **Supplementary Fig. 18c,d**), consistent with greater mapping power resolving cell-type-restricted effects.

This gradient was not apparent in several earlier studies. The largest eQTLs detectable in small, bulk RNA studies indeed appear enriched at unconstrained and non-disease relevant genes^{34,37–39}, whereas increased power in our study reveals that constrained genes are more likely to harbor eQTLs, albeit with much smaller average effects. Downsampling resolves this directly (**Supplementary Fig. 19**), and the transition it reveals reflects two opposing forces. The higher mean expression of constrained genes noted above favors their detection, whereas their smaller eQTL effects oppose it, and only at sufficient power does the expression advantage overcome the effect size disadvantage.

The effect size gradient replicated broadly across independent studies. Across 287 eQTL Catalogue datasets¹⁷, TenK10K¹², and SingleBrain⁴⁰, it was positive in 97% of datasets (median Spearman's $\rho = 0.12$, maximum 0.38), holding across both analytical strategies (single-cell SAIGE-QTL and bulk QTLtools with rank-based inverse normal transformation) and both tissue contexts (immune and non-immune; **Supplementary Fig. 20a, Methods**). The gradient converged toward $\rho = 0.2$ – 0.3 as the number of eGenes per dataset increased, consistent with power-dependent detection of regulatory buffering.

The GWAS colocalization enrichment at constrained genes (**Fig. 5f**) replicated in TenK10K¹² (Spearman's $\rho = -0.94$, $P = 5.5 \times 10^{-5}$), closely recapitulating our own estimate ($\rho = -0.99$, $P = 9.3 \times 10^{-8}$). Across the 287 eQTL Catalogue datasets, by contrast, most individual datasets showed a positive ρ favoring unconstrained genes, mirroring the eGene detection paradox, because, at a typical colocalization rate of $\sim 1\%$, each dataset yields only ~ 10 colocalizations per LOEUF decile (**Supplementary Fig. 20b**). The enrichment emerged only in datasets contributing more than 2,000 colocalizations, including INTERVAL ($n = 4,729$; Spearman's $\rho = -0.83$, $P = 0.003$)²², indicating that resolving this signal requires a scale of colocalization that only large eQTL studies reach.

Two studies that appeared after our preprint independently support the same architecture. eQTLGen Phase 2 ($n = 43,301$)⁴¹ reported that both *cis*- and *trans*-eQTL effect sizes are negatively correlated with loss-of-function constraint, and that constrained genes tolerate fewer independent regulatory variants. Burger et al. reported the same gradient across striatal and cortical cell types in 178 brain donors⁴².

The contrast with chromatin accessibility

In contrast to eQTLs, *cis*-caQTL effect sizes showed only a weak relationship with constraint (**Fig. 5g**). This difference between eQTL and caQTL patterns persists when examining MAF-

normalized effect sizes, indicating that this pattern is not driven by differences in statistical power (**Extended Data Fig. 9f,g**).

Constrained genes carry more peak–gene links, each with smaller effects (**Fig. 5h,i**). This more-but-weaker-links pattern was robust to cell-type resolutions, holding at PBMC (L0), L1, and L2 alike (**Supplementary Fig. 22**). Stratifying by regulatory mode, the excess of linked peaks at constrained genes was primarily driven by rheostat and dual links (median = 33 and 30 vs. 11 and 13, respectively), while switch links showed a much weaker gradient (median 12 vs. 10; **Extended Data Fig. 9h**). Decomposing the effect size gradient by mode, both rheostat and dual links showed smaller effects at constrained genes (median $|\beta_{\text{link}}| = 0.33$ and 0.34 vs. 0.45 and 0.42 , respectively), with similar gradients for both modes (Spearman's $\rho = 0.12$ for both, $P = 1.1 \times 10^{-30}$ vs 1.9×10^{-31} , respectively; **Extended Data Fig. 9i**). Within each mode, peak–gene distances were largely invariant across the constraint spectrum (**Extended Data Fig. 9j**). Together, these patterns indicate that constrained genes preferentially employ fine-tuned expression control rather than binary gating.

cis-eQTL examples at constrained genes

For example, we found strong *cis*-eQTLs at constrained genes *GRAMD1B* and *FAS* that colocalize with chronic lymphocytic leukemia (CLL) risk. An intron variant of *GRAMD1B*, rs35923643:A>G, substantially increases *GRAMD1B* expression in B cells ($\beta = 0.66$, $P = 3.7 \times 10^{-59}$), while increasing CLL risk (OR = 1.58, $P = 5.0 \times 10^{-71}$ in the FinnGen+MVP+UKB meta-analysis; **Supplementary Fig. 21a,b**). *GRAMD1B* encodes a cholesterol transfer protein that regulates lipid homeostasis at ER-plasma membrane contact sites⁴³, a pathway implicated in CLL biology, where LDL-derived cholesterol amplifies pro-survival STAT3 signaling⁴⁴. Similarly, an intron variant of *FAS*, rs2147420:A>G increases *FAS* expression across immune cells, with the strongest effect in CD16⁺ monocytes ($\beta = 0.53$, $P = 4.9 \times 10^{-89}$) and a moderate effect in B cells ($\beta = 0.20$, $P = 1.2 \times 10^{-29}$), while increasing CLL risk (OR = 1.27, $P = 6.9 \times 10^{-28}$; **Supplementary Fig. 21c,d**). *FAS* encodes a death receptor that triggers apoptosis upon engagement by FAS ligand, and paradoxically, its increased expression in CLL may reflect compensatory upregulation in response to apoptosis resistance mechanisms⁴⁵ or could promote survival signals through non-canonical FAS pathways that have been implicated in B cell proliferation⁴⁶.

Peak-level evolutionary constraint

We next asked whether the regulatory peaks themselves are under evolutionary constraint. Constrained peaks (as measured by genomic non-coding constraint of haploinsufficient variation [Gnocchi]⁴⁷) preferentially linked to constrained genes, with the most constrained peak per gene having higher median Gnocchi at constrained than unconstrained genes (5.1 vs 4.6), yet the strongest peak per gene showed no correlation with gene constraint (**Extended Data Fig. 9k**). Stratifying all peak–gene links by both peak and gene constraint revealed that unconstrained peaks ($-4 < \text{Gnocchi} < 4$) showed stronger regulatory effects than constrained peaks ($\text{Gnocchi} > 4$), with both peak constraint categories showing similar gene constraint (LOEUF)-dependent increases (**Extended Data Fig. 9l**), suggesting that constrained peaks are systematically weakened to maintain essential gene expression.

Stimulation-dependent buffering

To assess whether this buffering architecture is specific to steady-state conditions, we analyzed paired naive/stimulated eQTLs and caQTLs from Alasoo et al. (macrophages, 84 donors, 4 conditions)¹⁸. The *cis*-eQTL effect size gradient steepened under pathogen stimulation (Spearman's $\rho = 0.23$ in naive to 0.26 in IFN γ + Salmonella; LOEUF $\times$ condition interaction $P =$

2.4×10^{-3}), with eQTL effects increasing disproportionately at unconstrained genes (27% vs. 18% at constrained genes; $P = 0.009$), while the *cis*-caQTL effect sizes increased uniformly under stimulation regardless of target gene constraints (Spearman's $\rho = -0.05$ to -0.06 ; interaction $P = 0.75$; **Extended Data Fig. 10a,b**). This indicates that enhancer-priming¹⁸ and regulatory buffering operate at distinct regulatory layers. Stimulation increases caQTL effects independently of gene constraint, but the activation of peak–gene links through priming is itself subject to the buffering architecture, limiting the downstream transmission to gene expression at constrained genes.

Finnish-enriched disease variants

Fine-mapped caQTL variants are enriched for Finnish-enriched ($AFE > 5\times$) variants far more than fine-mapped eQTL variants ($RR = 2.98$ vs. 1.27), while MAF distributions are similar between caQTLs and eQTLs across both Finnish-enriched ($AFE > 5\times$) and non-enriched ($AFE \leq 5\times$) strata, suggesting the difference is not a power artifact (**Supplementary Fig. 23, Supplementary Table 33**). Only 78 of these Finnish-enriched variants (1%) were fine-mapped ($PIP > 0.5$) in the eQTL Catalogue¹⁷.

In our recent paper⁴⁸, we developed a novel base-editing and phenotyping platform called Sensitive Transcriptomics And Genotyping by sequencing (STAG-seq), which validated that the minor allele for rs748670681 is associated with *TNRC18* down-regulation and provided deeper mechanistic insight showing the strongest effect in Th1, a $CD4^+$ T helper subset, among primary immune cells, converging with our endogenous $CD4^+$ T cell cascade. In Th1, we demonstrated that rs748670681 reactivated endogenous retroviral elements, up-regulated interferon-stimulated genes, and suppressed cell-cycle transcripts, shifting the population toward a memory-like inflammatory state.

References

1. Kim, S. *et al.* High amount of transcription factor IRF8 engages AP1-IRF composite elements in enhancers to direct type 1 conventional dendritic cell identity. *Immunity* **53**, 759–774.e9 (2020).
2. Briseño, C. G. *et al.* Distinct transcriptional programs control cross-priming in classical and monocyte-derived dendritic cells. *Cell Rep.* **15**, 2462–2474 (2016).
3. Ghosh, H. S., Cisse, B., Bunin, A., Lewis, K. L. & Reizis, B. Continuous expression of the transcription factor e2-2 maintains the cell fate of mature plasmacytoid dendritic cells. *Immunity* **33**, 905–916 (2010).
4. Cisse, B. *et al.* Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135**, 37–48 (2008).
5. Kwon, K. *et al.* Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity* **28**, 751–762 (2008).
6. Hou, P. *et al.* A novel selective autophagy receptor, CCDC50, delivers K63 polyubiquitination-activated RIG-I/MDA5 for degradation during viral infection. *Cell Res.* **31**, 62–79 (2021).
7. Hou, P. *et al.* Autophagy receptor CCDC50 tunes the STING-mediated interferon response in viral infections and autoimmune diseases. *Cell. Mol. Immunol.* **18**, 2358–2371 (2021).
8. Lukic, S., Nicolas, J.-C. & Levine, A. J. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell Death Differ.* **21**, 381–387 (2014).
9. Saha, A. & Battle, A. False positives in trans-eQTL and co-expression analyses arising from RNA-sequencing alignment errors. *F1000Res.* **7**, 1860 (2018).
10. Cella, M. *et al.* A novel inhibitory receptor (ILT3) expressed on monocytes, macrophages, and dendritic cells involved in antigen processing. *J. Exp. Med.* **185**, 1743–1751 (1997).
11. Deng, M. *et al.* LILRB4 signalling in leukaemia cells mediates T cell suppression and tumour infiltration. *Nature* **562**, 605–609 (2018).
12. Cuomo, A. S. E. *et al.* Impact of rare and common genetic variation on cell type-specific gene expression in human blood. *medRxiv* 2025.03.20.25324352 (2025) doi:10.1101/2025.03.20.25324352.
13. Yazar, S. *et al.* Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* **376**, eabf3041 (2022).
14. Yin, J. *et al.* Chinese Immune Multi-Omics Atlas. *Science* **391**, eadt3130 (2026).
15. Kock, K. H. *et al.* Asian diversity in human immune cells. *Cell* **188**, 2288–2306.e24 (2025).
16. Nagai, A. *et al.* Overview of the BioBank Japan Project: Study design and profile. *J. Epidemiol.* **27**, S2–S8 (2017).
17. Kerimov, N. *et al.* A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat. Genet.* **53**, 1290–1299 (2021).
18. Alasoo, K. *et al.* Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat. Genet.* **50**, 424–431 (2018).
19. Mouri, K. *et al.* Prioritization of autoimmune disease-associated genetic variants that perturb regulatory element activity in T cells. *Nat. Genet.* **54**, 603–612 (2022).
20. Siraj, L. *et al.* Functional dissection of complex trait variants at single-nucleotide resolution. *Nature* 1–11 (2026).
21. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
22. Tokolyi, A. *et al.* The contribution of genetic determinants of blood gene expression and splicing to molecular phenotypes and health outcomes. *Nat. Genet.* 1–10 (2025).
23. Gu, Y. *et al.* RhoH GTPase recruits and activates Zap70 required for T cell receptor signaling and thymocyte development. *Nat. Immunol.* **7**, 1182–1190 (2006).

24. Reeve, M. P. *et al.* Genome-wide association analyses of autoimmune hypothyroidism reveal autoimmune and thyroid-specific contributions and an inverse relationship with cancer risk. *Nat. Genet.* **58**, 550–559 (2026).
25. Gomez-Diaz, C. *et al.* RNase T2 restricts TLR13-mediated autoinflammation in vivo. *J. Exp. Med.* **222**, e20241424 (2025).
26. Kidd, J. M., Newman, T. L., Tuzun, E., Kaul, R. & Eichler, E. E. Population stratification of a common APOBEC gene deletion polymorphism. *PLoS Genet.* **3**, e63 (2007).
27. Caval, V., Suspène, R., Shapira, M., Vartanian, J.-P. & Wain-Hobson, S. A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3'UTR enhances chromosomal DNA damage. *Nat Commun* **5**, 5129 (2014).
28. Middlebrooks, C. D. *et al.* Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. *Nat. Genet.* **48**, 1330–1338 (2016).
29. Petljak, M. *et al.* Mechanisms of APOBEC3 mutagenesis in human cancer cells. *Nature* **607**, 799–807 (2022).
30. Roberts, S. A. *et al.* An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* **45**, 970–976 (2013).
31. Deng, H. *et al.* CBX6 is negatively regulated by EZH2 and plays a potential tumor suppressor role in breast cancer. *Sci. Rep.* **9**, 197 (2019).
32. Yao, D. W., O'Connor, L. J., Price, A. L. & Gusev, A. Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nat. Genet.* **52**, 626–633 (2020).
33. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
34. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
35. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
36. Zeng, T., Spence, J. P., Mostafavi, H. & Pritchard, J. K. Bayesian estimation of gene constraint from an evolutionary model with gene features. *Nat. Genet.* **56**, 1632–1643 (2024).
37. Mostafavi, H., Spence, J. P., Naqvi, S. & Pritchard, J. K. Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nat. Genet.* **55**, 1866–1875 (2023).
38. Connally, N. J. *et al.* The missing link between genetic association and regulatory function. *Elife* **11**, e74970 (2022).
39. Vösa, U. *et al.* Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
40. Jang, B. *et al.* A meta-analysis of single-nucleus expression quantitative trait loci linking genetic risk to brain disorders. *Nat. Genet.* **58**, 737–747 (2026).
41. Warmerdam, C. A. R. *et al.* Trans-eQTLs reveal the architecture of human gene regulatory networks. *medRxiv* 2026.02.04.26343575 (2026).
42. Burger, S. *et al.* Inter-individual variation of cellular and gene-expression properties of the human striatum. *bioRxiv* 2026.03.20.713160 (2026) doi:10.64898/2026.03.20.713160.
43. Sandhu, J. *et al.* Aster proteins facilitate nonvesicular plasma membrane to ER cholesterol transport in mammalian cells. *Cell* **175**, 514–529.e20 (2018).
44. McCaw, L., Shi, Y., Wang, G., Li, Y.-J. & Spaner, D. E. Low density lipoproteins amplify cytokine-signaling in Chronic Lymphocytic Leukemia cells. *EBioMedicine* **15**, 24–35 (2017).
45. Chu, P. *et al.* Latent sensitivity to Fas-mediated apoptosis after CD40 ligation may explain activity of CD154 gene therapy in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3854–3859 (2002).

46. Rathmell, J. C., Townsend, S. E., Xu, J. C., Flavell, R. A. & Goodnow, C. C. Expansion or elimination of B cells in vivo: dual roles for CD40- and Fas (CD95)-ligands modulated by the B cell antigen receptor. *Cell* **87**, 319–329 (1996).
47. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).
48. Liu, K. *et al.* Sensitive Transcriptomics and Genotyping reveals function of genetic variants in immunity. *bioRxiv* 2025.03.12.642879 (2025) doi:10.1101/2025.03.12.642879.