

Population-scale immune multiome atlas reveals regulatory disease mechanisms

Corresponding Author: Dr Ramnik Xavier

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
General statement

In the manuscript entitled 'Population-scale multiome immune cell atlas reveals complex disease drivers', Kinai et al., generate paired single nucleus (sn) RNAseq and ATACseq from 1,108 Finnish individuals as part of the FinnGen Biobank. They then use this data to; i) systematically identify chromatin accessibility peak to gene expression links, ii) map contextually restricted expression and chromatin accessibility QTLs, by coupling this with paired germline genetic data, iii) devise a pipeline to delineate the specificity of these effects, called 'CASCADE', iv) categorise detectable regulatory mechanisms across molecular QTL layers and linkage information, v) nominate effects that are potentially causal of complex disease susceptibility by colocalisation, and vi) make informed hypotheses about how selective pressures shape the molecular impact of variants that underpin complex traits, while highlighting the value of conducting such a study in the FinnGen cohort.

I would like to congratulate the authors on what is a phenomenal body of work that has precipitated in a significant advance in our understanding of complex trait genetics. The data that has been generated, the associations that have been mapped, the methods that have been developed, and the insights that these provide are likely to be of great value to the scientific community, even beyond those interested in complex trait genetics specifically. I would consider these findings, especially the regulatory buffering of chromatin accessibility vs gene expression, to be breakthroughs, rather than incremental advances. I have some comments on some specific aspects of the manuscript text, the presentation of findings, methods and conclusions that I believe would strengthen the manuscript, but would like to highlight my belief that this work is of an extremely high standard, and after addressing these comments would be worthy for publication in Nature.

Major comments

1. At several points of the manuscript, some of the terminology and phrasing could be adjusted to improve accessibility for a general audience. For example, on L182, the authors first introduce '95% credible set', which will not be informative to a reader who does not have a background in statistical genetics. Similarly, on L185, 'colocalized' is first used in the results section but not defined. It would also be very useful if, when this is first defined, the authors highlight that this cannot determine causality between the two effects, as the claims based on this evidence could be misinterpreted by a reader who is unfamiliar with this approach and its limitations.

2. In its current form, the presentation of the peak-to-gene link information prior to the QTL associations is slightly misleading to the reader. On L1307-1311, the authors outline that they only test for links between genes with a cis-eQTL (eGenes) or with a cis-caQTL (caPeaks), however, in the results section this detail is not provided, and reads as though each gene (regardless of whether genetic effects are detectable), is tested for association with all peaks in cis. This would presumably have an impact on the multiple testing burden within each cis-window, and so I would recommend the author's either adjust their presentation so that this the link analysis is presented after the eQTL/caQTL is introduced, or this detail is made more clear in the results section. The current presentation is also inconsistent with the order in which regulatory cascades are

defined (Fig 4b), which is peak -> caQTL -> eQTL -> link.

3. In figure panels 2a-d, the authors make use of associations derived from the entire sample, which they call 'PBM'. Although I believe this is reasonable, its presentation alongside the 'L1' annotations is somewhat misleading, as it is an aggregation of all L1 data together in a single analysis. I would recommend referring to this data as something else, perhaps 'LO', and faceting it out of the plot with the L1 results. Relatedly, are peak-gene links from L2 cell types also folded into the results of Fig1e? If so, this could be misleading too.

4. There are a few details of the QTL mapping I find unusual and would appreciate further clarity and justification. For cis-eQTL mapping the authors use SAIGE-QTL, including PEER factors into their model. However, the number of PEER factors being utilised is optimised for the number of eGenes being detected by tensorQTL (L1371). I imagine this is to make the comparison between these two methods more equivalent (Extended Figure 5a), but these methods model gene expression in fundamentally different ways, and so I find the use of the tensor-QTL-optimised PEER factor number in SAIGE-QTL tests to be unusual, and potentially harmful to the number of associations that are actually being used later on. Can the authors confirm that if PEER factors were re-optimised for SAIGE-QTL tests, then this would lead to a similar number of PEER factors, and/or similar eGenes? Additionally, in the mapping of caQTLs, the authors use PCs instead of PEER factors as the use of PEER factors to model caQTLs has 'not been fully explored' L1391. This makes the covariates being tested for caQTLs and eQTLs not comparable – why not use PCA to detect covariates for both? The authors also pseudobulk and use tensorQTL for ca-QTL mapping because of the 'sparsity of snATACseq fragments' L195. Does this mean SAIGE-QTL is overly sensitive to sparsity, and therefore not useful for sc-caQTL mapping? If the authors have tested this, it would be a useful detail to include. Finally, the authors then use tensorQTL for trans-eQTL mapping, which results in further discrepancy of how different methods are being employed to test for different types of association.

5. The OneK1k study (Yazar et al., 2022, Science) and subsequent next phase (Cuomo et al., 2025, medRxiv) also map eQTLs in PBMCs of healthy individuals – in datasets of similar or larger sizes than this study. While I appreciate there is a methodological difference between these efforts with this one (single cell vs single nucleus), and biological difference in the ancestry of the study population, to better evaluate the reliability and novelty of the eQTLs being mapped in this study, it would be useful to compare the associations being detected here with these efforts.

6. For colocalization, the authors use a threshold of posterior probability for a shared causal variant (PPH4) of 0.5 (L420). Unfortunately, I believe this threshold is too relaxed and therefore likely to result in an excess of false positive associations being reported. This is because, in theory, a PPH4 of 0.5 could occur when the PPH3 (both traits under genetic control but with different causal variants) is still very close to 0.5, which would indicate near-equivalent evidence for traits to not share a causal variant. In line with this, comparable studies tend to use higher thresholds, such as; 0.8 (Cuomo et al., 2025 medRxiv; Vochteloo et al., 2025 medRxiv; Yazar et al., 2022 Science; The GTEx Consortium, 2020 Science; Hong et al., 2025 Nature Genetics; Kock et al., 2025 Cell); 0.75 (Alegbe and Harris et al., 2025 medRxiv), 0.7 (Bryois et al., 2022, Nature Neuroscience; Bian et al., 2025 Cell Genomics) or 0.6 (Natri et al., Nature Genetics). I would recommend the authors either acknowledge this possibility in the main text or raise the PPH4 threshold to that of similar studies and adjust the text/figures being presented based on this (e.g 5c). In addition, posterior probabilities for colocalization are never presented in the main text or figures, making it difficult for the reader to readily evaluate the statistical evidence in support of shared genetic effects. There is also no specification of what the prior probabilities are being used for colocalization (p_1 , p_2 , p_{12}), which is an important detail.

7. The trans-eQTL analysis provides a very surprising result that is currently under-explored by the authors and potentially an over-estimate. Identifying 20,826 trans-eGenes from 1,108 individuals is a very large number of effects given this sample size. Other studies have identified only a fraction of the number of these effects despite using a substantially larger sample size - e.g Vosa et al., (2021, Nature Genetics) identified 6,298 trans-eGenes from a meta-analysis that totalled 31,684 individuals. Can the authors provide additional evidence that these trans-eQTLs are not false positives? For example, false positive trans-eGenes can be detected due to cross-mapping of cis-eGenes with which they share homology (see Saha and Battle, 2018 F1000 Research), and removing genes with high probability of this would be useful. Additionally, although the authors employ what appears to be a strict p-value threshold to determine significant eQTL effects ($p < 5e-08$, L1365), this will account for the number of variants being tested per gene but not the number of genes being tested within each cell-type, which will be in the order of thousands. The expression of each gene is being tested as an independent phenotype, and so I would recommend the authors incorporate this in their multiple testing correction.

In addition, it would be useful to highlight whether any of these associations have been found in previous work, and whether there are any trans-eQTL effects relevant to disease. If, after performing the additional filters detailed above, the overlap with cis-effects really remains as low as the authors currently suggest (3-6% for cis-e/cis-caQTLs respectively, L275-276), this may potentially have a significant impact on how future trans-QTL analyses are performed, as these frequently rely on subsetting to cis-QTL or disease-causal variants only due to the limited statistical power.

8. While I appreciate the authors attempt to develop a method that better distinguishes cell-type specificity from a lack of statistical power, CASCADE, there are several caveats to using mash in the context of eQTL mapping, particularly at the single cell level. As the authors describe it (L1501-L1517), they are constructing the data-driven covariance matrices from strong variants ('CSs (PIP > 0.1) or, as an alternative when no CSs available, the top five variants with the strongest association ($Z > 4.89$) across any cell type'). However, the genes/variants that are tested in each condition will vary, often quite substantially, and the FLASH and PCA algorithms have limited robustness to missing data. Hence, it becomes necessary to either remove genes that are not tested in every condition, which leads to data-driven matrices being calculated on only a fraction of intersecting tests that may not accurately reflect the heterogeneity of genetic effects across

conditions, or input a value for missing tests which may strongly influence the data-driven matrix. What was the total proportion of tests that were either removed or imputed in the data-driven matrices, and have the authors tested how variation in this number influences their results? This is important if others are to utilise this method, and if the total number of non-imputed genetic effects is substantially smaller than the number of effects tested in each condition, I would doubt the reliability of the results from this analysis.

In addition, it is not clear to me how data from L2 annotations is used in this analysis (L1524 – ‘we mapped them to their corresponding L1 parent types’) – is mash applied at the L1 level, then separately within each L1 for L2 annotations? Or in one complete analysis? In general, I think the details of this approach could be clarified to aid interpretability.

9. Related to my previous comment on the novelty of eQTL associations, it would also be useful if the authors commented on which, and how many, of their colocalizations between QTLs and disease susceptibility were novel or supported by previous work. A good benchmark for this would be the number of loci/genes colocalised for each trait in this study compared with those detected by OpenTargets or Cuomo et al., (2025, MedRxiv). Even if eGenes have been colocalised with disease previously by eQTL-based approaches, the single-cell resolution, chromatin accessibility layer and CASCADE analysis provided in this study would still substantiate the residing therapeutic hypothesis, and this would be beneficial to highlight to the reader.

10. There are several examples of colocalisations being given with details on their detectable regulatory cascade (e.g. TICAM1, RNASET2, RNAH, GRAMD1B (later on) and others), and while there is comment about the cell-type where QTL effects are presented, there is not detail on the specificity of these effects, as determined by the author’s CASCADE approach. Including this detail, along with plots showing the gene expression and chromatin accessibility across genotypes would help the reader interpret the nature and strength of the potential biological mechanism being presented.

11. Although I agree with the statement on L591-595 (and L785-789) that the increased statistical power in this study is a contributing factor to the increased number of eQTLs detected for constrained genes, these statements come off a little strong given larger studies of bulk RNAseq-derived eQTLs exist (and so power cannot be the only factor driving this discovery), and the effect of power has not been directly tested by the authors. A good experiment to test this would be to repeat the eQTL mapping across multiple rounds of down sampling, where the number of cells/samples being included in the analysis is titrated, and then assessing whether the number of associations for constrained genes is lost as statistical power is reduced.

Additionally, previous studies have primarily relied on bulk or low-resolution capture of gene expression for QTL mapping. Even though these may have substantial statistical power, cell-type-specific effects are still likely to be lost due to the averaging of their effects across heterogeneous cell populations. However in this study, where associations are mapped at high resolutions, these effects are likely better captured, and this is also likely to contribute to the discovery the authors describe. It would therefore be beneficial if the authors either; i) investigated whether eQTLs that act on constrained genes are enriched for those identified at higher cellular resolutions (e.g. L2 vs L0) using either their own or existing datasets that map eQTLs at multiple resolutions, or ii) adjust the text at these positions to soften the claim that this effect is ‘primarily power based’ (L592).

Minor comments

1. What does ‘with genotypes of interest’ refer to in Extended Data Figure 1? This reads as if the samples were selected a priori based on specific germline variants.

2. I find the subset of the snATACseq data to the same cells that passed the snRNAseq QC to be an important point that strengthens the reliability of the comparisons between the two modalities (detailed on L1254). It would be good if the author’s highlight this detail in the results section so that it is readily available to the reader.

3. L146 – it would be useful to outline in the text the threshold being used to determine ‘significant’ peak–gene associations.

4. I believe there is no reference to Figure 1f. If this is missed, this should be rectified, or the panel should be omitted.

5. On L212, the term ‘propagation patterns’ is vague and should be clarified.

6. On L1477-1481 the authors detail how they use graph-based clustering to merge colocalised credible sets. Although I believe this could be a reasonable approach, for readers to properly understand and replicate this analysis, it would be useful if the authors provide examples.

7. On L222, the authors describe an ‘intronic variant’ in LILRB4, however this variant appears to not be in the gene (Extended Data Fig 5b).

8. On L280, when the authors introduce CASCADE, it would be useful if the authors introduce the problem they are seeking to address using this method in the results section. It is not yet obvious to a general reader the reason why CASCADE has been developed or what this might offer existing approaches.

9. In Figure 5c, the use of ‘NA’ to GWAS describe credible sets with unknown molecular impact is not very descriptive. I would recommend replacing this with ‘Unknown’.

10. On L456, the authors indicate that the relatively low colocalization rates in skin cancer and type 2 diabetes suggest 'their genetic architecture involves regulatory mechanisms in other cell types beyond PBMCs'. This is one possibility, but it is also possible that they do exist in PBMCs, but; i) are not active in healthy individuals, ii) this study does not have the statistical power to detect the molecular effects, or iii) are only active under specific stimulatory conditions. This statement should be softened or adjusted to incorporate these additional explanations.

11. In their presentation of associations across selective constraint bins (as defined by the LOEUF score), the authors present numerous especially interesting findings, but do not convey the uncertainty surrounding the scoring of genes in this manner. These statistics are useful, but imperfect measures of selection - which is difficult to reliably assess. It would be beneficial to convey this uncertainty to the reader.

12. In Figures 5e and 5g, the points are coloured by the percentage of genes with eQTLs or linked caPeaks respectively, with the range for these values being the minimum observed to maximum observed values (e.g 60-80% in 5e). I believe this may exaggerate the difference in these proportions that are interpreted by the reader, and would recommend this is updated to span the range of potentially observed values (0% to 100%). In addition, it would be useful if the direction of 'more constrained' was shown on the plots, with an arrow for example, to aid interpretability. These comments also affect Extended Data Figure 9h and i.

13. I do not understand the term 'expression-matched highly-expressed' (L584). How can genes be matched for expression but also subset to the top 20% most expressed?

Bradley T. Harris

(Remarks on code availability)

<https://github.com/mkanai/cascade>, <https://github.com/mkanai/finngen-multiome-flagship> and <https://github.com/FINNGEN/singlecell-pipeline> are not accessible, so I have been unable to test this code.

For <https://github.com/mkanai/fasthurdle>, which is available, I was unable to install the package due to GCC incompatibility on my cluster, and a lack of fortran package and sudo privileges on my local machine. I wonder if the authors could to provide a conda environment recipe, or singularity container that could be used to download the necessary dependencies instead?

Referee #2

(Remarks to the Author)

Ramnik J. Xavier and colleagues present a population-scale single-nucleus multi-omics atlas, profiling paired chromatin accessibility and gene expression in over 10 million immune cells from 1,108 Finnish individuals. The work provides a large-scale map of immune gene regulation, identifying numerous molecular quantitative trait loci and enhancer-gene links. A central conceptual contribution is the "CASCADE" framework, which categorizes genetic variants into mechanistic cascades linking chromatin alterations to gene expression. The authors report that variants involved in multi-layered regulatory pathways show stronger colocalization with disease, suggesting a biologically informed approach to prioritizing non-coding variants. The strategic focus on the Finnish population facilitates the study of variants which are not readily observed in other cohorts. This resource provides a substantial reference for functional genomics and may aid future studies of the genetic architecture of immune-mediated diseases. However, several issues require further improvement and justification for the work to meet the publication standards of a journal such as Nature.

1. The study's entire molQTL map is derived from a single cohort which, despite its size, lacks independent replication, thereby weakening confidence in the robustness of these foundational results. While the focus on the Finnish population provides a unique opportunity to discover population-enriched variants, it simultaneously limits the assessment of how generalizable the identified regulatory mechanisms are. To address the concern regarding independent and cross-ancestry validation for core QTLs, the authors should split their cohort into discovery and validation subsets to provide internal replication metrics for a defined set of high-confidence QTLs. Furthermore, they should evaluate the portability of their major findings by leveraging publicly available cross-ancestry molQTL resources, including the eQTL catalogue, and report what proportion of their lead QTLs replicate in both direction and significance in at least one other ancestral group. Ideally, validation should also be attempted in an independent cohort from a different population to assess cross-ancestry consistency, a point underscored by comparison to other studies (Yazar et al. Science, 2022, DOI: 10.1126/science.abf3041; Xue et al. MedRxiv, 2025 <https://www.medrxiv.org/content/10.1101/2025.08.27.25334533v1>).

2. A significant limitation of this otherwise comprehensive atlas is its focus on steady-state PBMCs from healthy donors, resulting in a static snapshot that fails to capture the dynamic gene regulation fundamental to immune cell function and disease pathology. As highlighted by comparison to prior work (Perez et al. Science, 2022) and (Yazar et al. Science, 2022), genetic effects can be state-dependent, emerging only upon cellular activation or along differentiation trajectories. To address this lack of dynamic and stimulus-responsive regulatory analysis, the authors should perform in silico trajectory inference on key lineages such as CD4+ T cells and monocytes from their existing data to identify dynamic QTLs. Subsequently, the authors should design a pilot study that involves exposing PBMCs from a subset of genotyped donors, comprising 50 to 100 individuals, to canonical immune stimuli like IFN- α or anti-CD3/CD28. Paired single-cell Multiome

profiling conducted pre- and post-stimulation would directly identify stimulus-responsive molQTLs and quantify the fraction of disease-associated variants whose regulatory effects are context-dependent.

3. While the CASCADE framework powerfully nominates candidate causal variants and cascades, functional validation is limited primarily to TNRC18, leaving the vast majority of the over 100,000 fine-mapped variants and proposed mechanisms without experimental support. This constitutes a notable shortfall in functional validation and causal mechanistic exploration, especially when compared to related studies that integrated deeper functional follow-up. To strengthen the mechanistic claims, we request targeted experimental validation for a selected set of high-priority findings. Specifically, the authors should perform high-throughput reporter assays such as MPRA on at least six to eight Finnish-enriched variants from distinct mechanistic categories, which could include full cascade, positional cascade, or opposing-effect caQTL examples. These assays should be conducted in a minimum of two relevant immortalized immune cell lines and, where feasible, one primary cell type to confirm cell-type-specific regulatory activity.

4. To more robustly establish a causal link from identified molecular QTLs to disease risk, we recommend applying formal causal inference frameworks. A Mendelian Randomization analysis, particularly a two-sample MR approach leveraging the cohort's rich GWAS data, would be highly suitable. For instance, using genetic variants associated with a specific chromatin accessibility peak or gene expression level as instrumental variables could test whether these molecular traits directly influence disease endpoints. This would significantly strengthen the mechanistic interpretation of the reported associations.

5. The study provides an exceptional map of molecular mechanisms, yet it does not fully explore the translational potential inherent in its rich single-cell data, a dimension that has been effectively demonstrated in related work such as (Perez et al., Science, 2022). Bridging this gap between foundational discovery and clinical insight represents a key opportunity. We recommend that the authors integrate the basic clinical phenotypes already available for this cohort, which could include common laboratory values like CRP or lymphocyte counts, or medication records. This integration would enable a proof-of-concept analysis to test whether single-cell-derived features, for example immune cell subset proportions or pathway activity scores for key pathways like the interferon-stimulated gene response, can predict clinical parameters or define molecular subtypes among donors. Such an analysis would significantly strengthen the manuscript because it would allow a discussion on the tangible utility of these molecular maps for patient stratification in autoimmune diseases, thereby enhancing the translational relevance of the work.

Minor issues :

6. The description of the two sample processing methods, multiplexed sample channels versus single-donor channels, in Extended Data Fig. 1a and the Methods (lines 1181–1183) is overly brief and raises methodological concerns. Typically, cryopreservation is required only when samples are to be pooled. It is unclear whether the multiplexed channels used cryopreserved samples while the single-donor channels used fresh samples. If so, the authors should explicitly state this and, more importantly, assess whether these different processing protocols introduced batch effects or technical bias. Any such bias must be accounted for in downstream analyses to ensure the robustness and reproducibility of the reported molecular findings.

7. The data availability statement must be detailed and actionable. Beyond a query portal for variant lookup, we strongly encourage the public deposition of the complete single-cell data matrices, including the cell-by-gene expression and cell-by-peak accessibility tables. A clear release timeline and the designated public repository should be specified.

8. We have tested the provided links from multiple locations. Our findings are consistent: while the main GitHub repository is accessible, the specific code links cited in the manuscript are currently inaccessible. This issue directly hinders the assessment of reproducibility for the analytical workflows. As a minimum requirement for the revision, all relevant code must be made functional and its access links verified to be publicly available.

9. The analysis in lines 176 to 184 uses two different methods SAIGE QTL for eQTL mapping and tensorQTL for caQTL mapping. The authors need to justify the use of different statistical frameworks for different molecular layers to ensure methodological consistency is clear.

10. The definition of the likely shared but underpowered category in lines 279 to 285 is conceptually important but its distinction from truly shared effects is vague. A clearer explanation of the classification threshold and its implications is needed.

11. The classification of variants into 25 patterns in lines 356 to 361 is highly complex. The authors should provide a more accessible summary or a conceptual diagram in the main text or somewhere to aid reader comprehension.

(Remarks on code availability)

Several key GitHub links provided by the authors are currently inaccessible, although their parent directories can be accessed:

<https://github.com/FINNGEN/singlecell-pipeline>

<https://github.com/mkanai/finngen-multiome-flagship>
<https://github.com/mkanai/cascade>

All the other GitHub links can be opened and display complete README files, in which the author thoroughly explains the installation and usage instructions for each software/pipeline, as well as listing the external software they depend on (with links provided to those dependencies).

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

(Remarks on code availability)

I co-reviewed the code with one of the reviewers who provided the listed reports. Here are the comments.

Several key GitHub links provided by the authors are currently inaccessible, although their parent directories can be accessed:

<https://github.com/FINNGEN/singlecell-pipeline>
<https://github.com/mkanai/finngen-multiome-flagship>
<https://github.com/mkanai/cascade>

All the other GitHub links can be opened and display complete README files, in which the author thoroughly explains the installation and usage instructions for each software/pipeline, as well as listing the external software they depend on (with links provided to those dependencies).

Referee #4

(Remarks to the Author)

Review: Population-scale multiome immune cell atlas reveals complex disease drivers

In this work, Kanai et al. present an admirably large scale multiomic single cell analysis of Finnish peripheral blood samples. The scale and scope of the work with over ~1000 individuals and 7-10M analyzed cells is truly commendable and I can't begin to fathom the time and resources required to conduct this effort. Congratulations to all the authors. This work falls into the pantheon of large published (1K1K) and ongoing (10K10K) single cell studies but adds great depth with multiomic analysis. I have no doubt that the data generated, if appropriately shared in raw data format (if possible, given the Finngen restrictions) will be widely used and highly cited. However, the biological implications are harder to determine and some major, but addressable, criticisms are outlined below related to the reporting and analysis. The major finding, pertaining, to the large heritability of caQTLs and the large discrepancies between eQTLs is certainly of great interest and has been reported, in less detail, by others.

Major comments:

1. There is a large and unexplained discrepancy between the total number cells in the single cell RNA (~10M) and ATAC(~7M) analysis that needs to be clearly addressed. The authors do a great job of identifying eQTLs, caQTLs, and linking peaks to genes. However, it is unclear in the main figures and manuscripts what final dataset was used for this analysis. Extended data figure 1 suggests that these were truly separate analyses. Given that this analysis was performed on multiome samples and single cell barcodes from RNA used as an additional measure for identifying quality cells in ATAC data, it is unclear why there is such a large difference (30%) in the total number of cells. Furthermore, a joint ATAC, RNA embedding is never shown. If eQTLs were called on just the snRNAseq data rather than the joint embedding, why was this done and do the authors think this is appropriate considering that the relationships between eQTLs and caQTLs and peak-gene is further characterized. Some explanation and clarification is warranted. I imagine for instance, that certain cell types in some individuals are not well represented in the caQTL data but are well represented in the eQTL data creating discrepancies. Extended data figure 2c shows total nuclei but not by individual which is more important for genetic associations.
2. The effects of batch effects on analysis are unclear. It is unclear how batch was handled in both the generation and analysis of the single cell data. I am particularly interested to know why multiomic run or processing date was not included as a co-variate in the identification of pseudobulk or single cell eQTLs. Is the assumption that PEER factors explain this batch effect without the need for further correction and if so, can a citation/justification be provided.
3. CASCADE as a method is not properly discussed. Given the importance of CASCADE for annotation of molQTLs, it warrants additional space in the manuscript. My understanding of this algorithm is that it classifies molQTLs into various categories based on intersection of genetic effects and cell type. It is unclear how much validation is required for this classification. Can the authors comment further?
4. No functional validation is performed. In multiple places in the manuscript and abstract(!) base editing validation is mentioned. This data is not shown in the manuscript and all mentions should be removed. The data cited is from a pre-print that this reviewer did not receive and has not examined. I want to say clearly, that I do not believe functional validation is necessarily required for this work as primary a large-resource and genetics study. I full appreciate the value of functional validation, and methods like STAGseq and other published single cell approaches bridging the V2F gap. Mention of these in the discussion is certainly warranted with the caveat below:
 - a. Caveat: The authors discuss at length and show ample evidence in support of the finding that caQTL are more valuable

(heritable with increased enrichment) than eQTLs. However, all proposed validation (which again, I think should be removed) is focused on discovery of eQTLs. Recent published methods are available that look at chromatin accessibility with CRISPR editing if this is the focus and one of the major findings of the manuscript.

5. No code was shared to examine the analysis. I believe this is a requirement. No code was shared. Statements in the manuscript and reporting summary point to github links that are non-functional.

A non-criticism for the authors. This work is massive and I truly appreciate that. The study was performed in PBMCs, which some may view as a criticism. I do not. I appreciate the effort that went in to highlighting examples that may apply to tissues (microglia) but I think this is unnecessary and takes away from the arguments in the paper. There is no clarity on why caQTLs account for much higher heritability and why no links to genes can be found. One discussed reason is that the wrong tissue or cell state (stimulation condition) is being assayed. This makes some biological sense to me. If a T cell is primed to respond to multiple factors given its genetics, then only regulates expression in distinct context that may explain the discrepancy. The discussion on buffering alludes to this.

Minor comments:

- the title on the submitted manuscript does not match the title on the online submission (an additional 1 has been randomly placed)
- Having read the manuscript, it's unclear to me what the title is trying to convey. What are these complex disease drivers, I would recommend changing it.
- There is a very large imbalance in the single cell depth of the first 339 donors versus the following ~800. Was any analysis performed between these two groups to show that no differences in power were noted. I.e, the eQTLs identified (particularly the single cell eQTLs) in the first batch are all replicated in the second (perhaps down-sampled). This is a minor comment.
- There is a statement on line 93-94 that "Importantly, the functional effects and regulatory mechanisms we characterize represent fundamental biology shared across human populations". What is the evidence for this statement.
- Is the large degree (4% and 2%) of opposite effect eQTL and caQTL (line 309) driven by differences in cell type. If you control for this, how many opposite effects are found in the same cell type. Are these then erroneous or biological?
- Overall the effect sizes for eQTL seem much smaller than the effects sizes for caQTLs (at least from the few examples discussed (example lines 489 – 500). Is there a reason this is the case? Does it reflect how the data was transformed prior to modeling or sparsity? Can the authors comment.
- Line 524-526 requires a citation. Talking about the shared origin and similarity of monocytes and microglia.
- LTBR effects, as the authors know, are multifaceted. The conclusions on lines 536-539 seem a bit of a biological stretch to me. It's unclear if a minor change in expression ($B = -0.05$) could by itself have profound changes on antimicrobial responses or tonsil hyperplasia.
- The findings around "constrained genes" are fascinating. There is statement on line 588-592 that minor changes in highly expressing genes can have large effects on disease pathology. Are there any evidence or citations that functionally confirm this with experimental evidence? Yes, they colocalize with GWAS variants but intuitively I would have presumed the opposite that highly expressed genes are less sensitive to minor changes in gene expression.
- Given that the findings with constrained genes seems to run opposite to studies from others, can the authors would analyze the 1K1K public data to show supporting evidence for GWAS localization in constrained genes. Is this not of comparable size if not cell number to discover these same effects?
- Lines 608-612 about FAS and B cells is missing a citation.
- Again, lines 703-710 should be removed from the paper. They are not presented as data and have not been peer-reviewed. Additionally, the identified cascade was in CD8 T cells, but functional validation performed in CD4 (Th1) suggesting a discrepancy.
- Line 734 -736 discusses this validation again and claims that chromatin changes were validated. Again, this should be removed.
- In the discussion regarding base-editing validation, it would be nice to cite work from others (not just STAGseq) as efforts from others are also paving the way in this V2F space.
- Figure 1d. The area under the curves looks different (ie number of individuals in snATAC vs snRNA but reporting in the supplement suggest that roughly equal number of individuals are present in both datasets. Please double check this graph.
- Figure 1c. CD8 naïve T cells are WAY out there. This is a little bizarre. They are very far from the other T and CD8 TEM populations. Please double check the labels. I understand that distance in UMAPs means nothing but it's still concerning.
- Figure 2e, does the number of caQTL and eQTL per cell type correlate directly to the number of cells per sample? If you examine just constrained highly expressed genes does that same pattern hold. Is there really more genetic regulation in a CD4 T cells than NK cells.
- Figure 2h suggests that most variants are found right in the middle of the annotated caQTL peak. Is this biological or a consequence of how peaks and caQTLs are called. And if it's the latter, is this graph informative?
- Figure 2k, it's a little unclear which blue dot pairs with which red dot. For clarity could these be connected.
- Figure 3d-j. Wonderful examples, for clarity can you facet by cell type in each graph. Overlapping box plots are perhaps not the best visual way to show this
- Extended data figure 1. Thank you for this figure, it should be required in all studies. Can you please include the number of cells passing each threshold. An integrated RNA-ATAC object is reported but never shown. Consider adding.
- The methods mention flow cytometry data (lines 1171). Why? Was this presented or used?
- Is filtering on MALAT1 expression appropriate? Line 1235. The authors cite a singular pre-print. I was not aware of this as an established strategy. Additionally, the threshold is set per run. Can this data be shared and further explanation of how many cells were lost explained.

(Remarks on code availability)

No code was shared and github links were not accessible. I cannot evaluate but feel confident in the expertise of the authors.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

I would like to thank the authors for the considerable amount of work that has been put in to respond to my previous comments. I believe the updates strengthen the manuscript and its conclusions. I have only a few remaining comments based on their revisions, which once addressed, would lead to me recommending the manuscript for acceptance.

1 (Previous comment 2). I believe the updated analysis is a large improvement on the previous presentation of peak-gene links. However, in the revised manuscript, the authors provide examples of what each peak category would look like using simulated data only (Supplementary Figure 3), or present odds-ratios (Supplementary Figure 4). I believe it would be useful to present the equivalent plot of Supplementary Figure 3 but using real data, so that the reader can visualise what the real associations look like.

In addition, I believe the authors may be somewhat overstating the categorisation of their peaks. Specifically, significance of hurdle model components is determined by 'Holm-corrected confirmation within each selected pair' – so if only one component is significant after this correction it is classified as either 'switch' or 'rheostat'. However, the ability to detect associations within each component will be somewhat determined by the number of expressing or non-expressing cells, and it is possible that a switch peak with only a few expressing cells in their study may in fact act as a dual peak (significant in both components), but would only be detectable in the count-component if they had collected more expressing cells from a larger number of donors. Hence, I think it is important to convey the potential uncertainty in their classifications to the reader.

2 (Previous comment 8). I would like to thank the authors for the analysis already performed to test the robustness of their CASCADE results. I believe the CASCADE method is potentially very useful for determining QTL cell type-specificity and regulatory mechanisms. However, I still feel that there are additional details that need to be presented and explored.

In their response, the authors state "Across L1 cell types, 18% and 32% of entries in the strong set used for data-driven covariance estimation were imputed for eQTLs and caQTLs, respectively. Missingness was concentrated in rarer cell types with fewer nuclei per donor and was negligible in well-powered cell types". I believe it is important to present to the reader what the missingness rates are per cell type and quantitatively show how this relates to cell-type abundances. I believe this is especially necessary for L2 annotations, where missingness rates may be higher but have not yet been quantified. In addition, although the authors present an analysis where they restricted to tests observed in only 50% of cell types, it would be important to test at what point the estimates become unreliable by titrating this across various rates of missingness, also showing how this relates to each cell type and its relative abundance (also at L2 level).

3 (Previous comment 9). I believe the additional benchmarking of colocalisations strengthens reliability in the authors conclusions. However, I still feel it would be of value to highlight in the results section of the manuscript which of the associations the authors describe were detectable in other studies. I agree that this study adds a lot of novelty, even for associations that are previously described, but believe this detail is important to make readily available to the reader.

4 (Previous comment 11). I appreciate the replication across independent datasets and down sampling experiments, I believe these tests strongly reinforce the conclusions. However, I believe the results presented in the new Supplementary Figure 18a are potentially misinterpreted. The authors now show that the slope for the relationship between gene constraint and the proportion of genes with colocalization is negative in each cell type level, but with quite dramatically different gradients – being much steeper for L1 and L2 annotations and with higher overall colocalization rates. If this is true, I believe this does indicate that the resolution at which eQTLs are mapped has an influence on: 1) overall colocalization rates, and 2) the likelihood that these effects are identified for a gene under higher constraint. Can the authors please test for a linear association between LOEUF decile and proportion of genes with a GWAS coloc with an interaction term for the cell-type level? If the interaction term is significant, this would indicate that the resolution at which eQTLs are mapped is a contributing factor, and their current presentation of the association being 'robust to cell-type resolution' is misleading.

5 (Previous minor comment 6). I appreciate the clarity on this, but not sure it makes sense. In the toy example, A–B, B–C, and D–E are colocalised, but the CS of C doesn't share a variant with B? How would it be colocalised in this instance?

Bradley T Harris

(Remarks on code availability)

I have installed the fasthurdle singularity container and the cascade package. Install of cascade was a slightly difficult as `cascade/src/variant\_qtl\_patterns.cpp` uses `IntegerVector::create()` with 25 arguments, which exceeds the 20-argument limit in older Rcpp versions. To overcome, I changed the default argument for `pattern\_to\_mechanism` from `IntegerVector::create(1, 2, ...)` to a C++11 initializer list `(={1, 2, ...})`. My exact commands were:

```
git clone https://github.com/mkanai/cascade.git
mkdir -p cascade_tmp
Rscript -e 'dir.create("cascade_tmp", showWarnings = FALSE); install.packages(c("R.utils", "qs2"), lib = "cascade_tmp",
repos = "https://cloud.r-project.org")'
R CMD INSTALL -l cascade_tmp cascade
```

I did not extensively test packages though.

Referee #2

(Remarks to the Author)

I would like to thank the authors for their thorough revision. They have addressed most of my previous concerns in a serious and data-driven manner. The new replication analyses, the integration of MPRA data, and the clinical proof-of-concept are valuable additions. I believe the manuscript is now closer to being acceptable for publication in Nature.

However, I still have two remaining issues that require attention. These are not fatal flaws, but they need to be clarified or further discussed. I recommend a minor revision.

First, the authors need to be clearer about the cell type specificity of their peak-to-gene links. The manuscript analyses eQTLs and caQTLs separately for each cell type. It is natural to ask whether the peak-gene links are also cell type specific. When the authors state that they identified 593,765 unique peak-gene links, does this mean each link is assigned to a specific cell type? Or are these links aggregated across all cell types? The current text does not make this clear. The same ambiguity affects the central claim that constrained genes have more but weaker enhancer-gene links. Is this statement based on the cell type level or the tissue level? For example, the FAS variant rs2147420 is reported to increase FAS expression in a cell type dependent manner. Does the corresponding peak-gene link also exist only in those specific cell types? I would like the authors to confirm that every peak-gene link is indeed cell type specific and that the buffering observation holds at that resolution. To help readers query this information, I suggest the authors provide a supplementary table with each significant peak-gene link as rows and each cell type as columns, indicating presence or absence. This would greatly improve the utility of the resource.

Second, the authors need to address alternative explanations for their buffering model. Their conclusion that constrained genes have more but weaker links, and that this buffers expression changes, is elegant. But it is not the only possible interpretation of the data. I would like the authors to discuss at least three alternatives in the main text or supplement. First, the observed gradient could partly reflect detection bias, because constrained genes have lower expression variance, which reduces standardized effect sizes even if the true biological effect is similar. Second, the high GWAS colocalisation rate could reflect reverse causation or confounding. The disease associated variant might act through other mechanisms or other tissues, and the eQTL signal in PBMCs might be an echo rather than the causal path. Third, the comparison of effect sizes between constrained and unconstrained genes might be affected by winner's curse, because only the most significant small effects survive in the constrained group. The authors have performed some sensitivity analyses, but they do not explicitly address these alternative interpretations. I encourage them to discuss these possibilities and, if possible, provide additional analyses such as bidirectional Mendelian randomisation to test the direction of causality. Even if they cannot completely rule out these alternatives, a balanced discussion would strengthen the paper.

Overall, I find the work impressive and the revision thorough. The two points above are important but manageable. I look forward to seeing the revised version with these clarifications and discussions added. I support publication after these minor changes.

(Remarks on code availability)

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

(Remarks on code availability)

The authors have made all code publicly available. All links listed in the 'Code Availability' section are currently accessible. Furthermore, the provided installation instructions are detailed and have been verified to be fully functional.

Referee #4

(Remarks to the Author)

The authors have addressed the majority of my concerns. I have no further comments.

Yuriy Baglaenko, PhD

(Remarks on code availability)

I examined the github files. They are accessible and well annotated. I did not run any analysis independently. I'm sorry I don't have time for that.

Version 2:

Reviewer comments:

Referee #1

(Remarks to the Author)

I would like to thank the authors for their detailed responses to my previous comments. I am now fully satisfied with the revised manuscript and would like to congratulate the authors on a brilliant study which I absolutely believe is worthy for publication in Nature.

Bradley T. Harris

(Remarks on code availability)

I did not run the code, but find the previous changes to satisfy my concerns

Referee #2

(Remarks to the Author)

All my previous questions have been well addressed. I am satisfied with the authors' responses and believe the manuscript now meets the criteria for acceptance.

(Remarks on code availability)

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

(Remarks on code availability)

Inspired by the comments of another reviewer, I conducted a more systematic evaluation of the software versions and dependencies. This testing revealed that the documentation for some of the software packages could be made clearer and more accessible. These issues do not require changes to the manuscript itself; instead, the authors could address them by improving the README files in the corresponding GitHub repositories.

1. Minimum R version requirements

The R packages cascade and fasthurdle failed to install or run under R 4.0.3, whereas they worked successfully under more recent versions of R, such as R 4.4 or R 4.5. I therefore recommend that the authors explicitly specify the minimum required R version in the README files.

2. More detailed installation instructions for cascade

To make the installation process more straightforward for new users, the authors could provide more detailed commands in the Installation section of the README. For example:

```
## Installation
# cascade depends on several Bioconductor packages.
# Please install these dependencies first.
# Install BiocManager if it is not already installed
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
# Install Bioconductor dependencies
BiocManager::install(c("BiocGenerics", "S4Vectors", "IRanges", "GenomicRanges"))
# Install cascade
install.packages("remotes")
remotes::install_github("mkanai/cascade")
```

3. Explicitly documenting dependencies for kanta_lab_preprocessing

For kanta_lab_preprocessing, the required software dependencies and version constraints are provided in the pyproject.toml file. I recommend that the authors also list these requirements explicitly in the README, particularly the requirement for Python ≥ 3.12 . This information is important because users may need to create a new Conda environment with an appropriate Python version before installing the software.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Revision Summary:

We thank the reviewers for their constructive feedback. We have addressed every comment in detail in the point-by-point responses below, and summarize the major revisions here.

1. **Revised title:** We changed the title to "**Population-scale immune multiome atlas reveals regulatory disease mechanisms**" to more clearly convey the study's focus.

2. **Improved peak–gene link analysis with regulatory mode classification:** We now test all peak–gene pairs in *cis*, using an improved hurdle model and a joint χ^2 score test with stage-wise confirmation. This identifies 593,765 unique links across 12,018 genes, each classified as switch (22%), rheostat (36%), or dual (42%). We show that constrained genes preferentially employ rheostat and dual regulation rather than binary switch (revised **Fig. 1e–h**, new **Supplementary Figs. 3,4, Extended Data Fig. 9h–j**).

3. **Massively parallel reporter assay (MPRA) functional validation at scale:** To orthogonally validate our fine-mapped molQTLs, we integrated recent MPRA results from Siraj et al. (2026), which tested 221,412 variants across five cell lines. We demonstrated that emVar rates and allelic effect concordance increase monotonically with PIP, confirming that our fine-mapping accurately identifies functional regulatory variants (new **Extended Data Fig. 6**). MPRA also showed that the intrinsic *cis*-regulatory activity of fine-mapped variants is constant across LOEUF, providing direct evidence that regulatory buffering acts downstream of the regulatory element rather than reflecting weaker upstream input (new **Extended Data Fig. 10c–f**).

4. **Refined CASCADE classification:** Motivated by MPRA emVar signatures, we partitioned the original "Full Cascade" into mechanistically distinct **Local**, **Distal**, and **Positional** cascades. All three remain enriched for GWAS colocalization, but only local and positional cascades show direct allele-specific MPRA activity (revised **Fig. 4,5**). Under the new framework and revised peak–gene analyses, we reclassified every vignette, including *TNRC18*, which is now a positional cascade in CD4⁺ T cells, consistent with the STAG-seq Th1 validation (revised **Fig. 6**).

5. **Replication of molQTLs and regulatory buffering:** We replicated *cis*-eQTLs, *cis*-caQTLs, and peak–gene links across five independent datasets spanning European and East Asian ancestries (new **Supplementary Figs. 6,8,9**). We further replicated regulatory buffering at constrained genes using 288 external datasets, recapitulating both the eQTL effect size gradient and the GWAS colocalization enrichment (new **Supplementary Figs. 19,20**).

6. **Stimulus-responsive regulation:** We demonstrated that our steady-state atlas captures a substantial fraction of fine-mapped stimulus-responsive eQTLs from the eQTL Catalogue (new **Supplementary Fig. 10**). Using paired naive/stimulated data from Alasoo et al. (2018), we further show that pathogen stimulation steepens the eQTL constraint gradient while leaving the caQTL gradient unchanged, indicating that enhancer priming acts independently of regulatory buffering (new **Extended Data Fig. 10a,b**).

Together, these revisions strengthen the analyses, mechanistic interpretation, and robustness of the study, and we hope the reviewers find the revised manuscript responsive to their comments.

Referee #1:

In the manuscript entitled 'Population-scale multiome immune cell atlas reveals complex disease drivers', Kinai et al., generate paired single nucleus (sn) RNAseq and ATACseq from 1,108 Finnish individuals as part of the FinnGen Biobank. They then use this data to; i) systematically identify chromatin accessibility peak to gene expression links, ii) map contextually restricted expression and chromatin accessibility QTLs, by coupling this with paired germline genetic data, iii) devise a pipeline to delineate the specificity of these effects, called 'CASCADE', iv) categorise detectable regulatory mechanisms across molecular QTL layers and linkage information, v) nominate effects that are potentially causal of complex disease susceptibility by colocalisation, and vi) make informed hypotheses about how selective pressures shape the molecular impact of variants that underpin complex traits, while highlighting the value of conducting such a study in the FinnGen cohort.

I would like to congratulate the authors on what is a phenomenal body of work that has precipitated in a significant advance in our understanding of complex trait genetics. The data that has been generated, the associations that have been mapped, the methods that have been developed, and the insights that these provide are likely to be of great value to the scientific community, even beyond those interested in complex trait genetics specifically. I would consider these findings, especially the regulatory buffering of chromatin accessibility vs gene expression, to be breakthroughs, rather than incremental advances. I have some comments on some specific aspects of the manuscript text, the presentation of findings, methods and conclusions that I believe would strengthen the manuscript, but would like to highlight my belief that this work is of an extremely high standard, and after addressing these comments would be worthy for publication in Nature.

We sincerely thank the reviewer for this generous assessment of our work and for recognizing the regulatory buffering findings as breakthroughs. We are grateful for the thoughtful and constructive comments, which have substantially improved the clarity of our manuscript. We have carefully addressed each point, as detailed in our responses below.

Major comments

1. At several points of the manuscript, some of the terminology and phrasing could be adjusted to improve accessibility for a general audience. For example, on L182, the authors first introduce '95% credible set', which will not be informative to a reader who does not have a background in statistical genetics. Similarly, on L185, 'colocalized' is first used in the results section but not defined. It would also be very useful if, when this is first defined, the authors highlight that this cannot determine causality between the two effects, as the claims based on this evidence could be misinterpreted by a reader who is unfamiliar with this approach and its limitations.

We thank the reviewer for this suggestion to improve accessibility. To clarify this point, we have revised the manuscript to define key statistical genetics terminology upon first use in the **Results**:

Statistical fine-mapping of *cis*-eQTL associations identified 86,536 and 89,350 independent 95% credible set (CS; a set of variants with 95% cumulative posterior probability of containing the causal variant)–gene–cell type trios for L1 and L2 cell types, respectively, ...

...

To assess sharing of *cis*-eQTLs across cell types, we performed colocalization analysis, which evaluates whether two genetic associations likely arise from the same causal variant.

2. In its current form, the presentation of the peak-to-gene link information prior to the QTL associations is slightly misleading to the reader. On L1307-1311, the authors outline that they only test for links between genes with a cis-eQTL (eGenes) or with a cis-caQTL (caPeaks), however, in the results section this detail is not provided, and reads as though each gene (regardless of whether genetic effects are detectable), is tested for association with all peaks in cis. This would presumably have an impact on the multiple testing burden within each cis-window, and so I would recommend the author's either adjust their presentation so that this the link analysis is presented after the eQTL/caQTL is introduced, or this detail is made more clear in the results section. The current presentation is also inconsistent with the order in which regulatory cascades are defined (Fig 4b), which is peak -> caQTL -> eQTL -> link.

We thank the reviewer for raising this point. We agree that our original presentation and description were misleading. Our implementation actually tested the union of eGenes/caPeaks and genes/peaks expressed in > 1% of nuclei, which effectively encompassed most genes and peaks. In the revised manuscript, we have corrected this by removing the eGene/caPeak restriction entirely and now test all peak–gene pairs in *cis* (< 1 Mb) regardless of QTL status, resolving both the inconsistency with the **Results** presentation and the ordering issue with **Fig. 4b**.

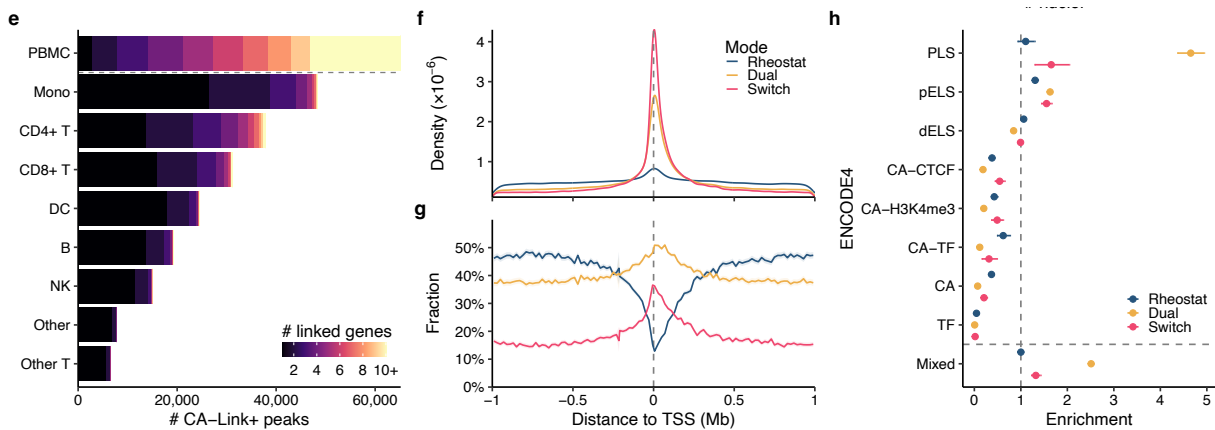
In addition, we took this opportunity to improve the peak–gene link analysis in two ways:

First, we improved the hurdle model specification and testing framework. We updated our *fasthurdle* package (<https://github.com/mkanai/fasthurdle>) to use score tests instead of Wald tests, which are better calibrated under the sparse and overdispersed count distributions typical of single-cell data and substantially faster by evaluating significance at the null model without fitting a full model for each peak–gene pair. We also refined the model specification in three ways: 1) Log-transformed total RNA counts are now used as an offset (rather than a covariate) in the count component to directly model expression rates; 2) Pseudobulk PEER factors (the same number optimized for *cis*-eQTL analysis) are included in both components to account for donor-level confounding; 3) ATAC spectral components are additionally included in the zero-inflation model to control for compositional confounding, where the zero component would otherwise conflate direct regulatory effects with indirect co-variation driven by sub-cell-type identity. Because the count component is fit on cells with non-zero expression only, we further restrict tests on peak–gene pairs with at least 5,000 expressing cells, providing sufficient sample size for coefficient estimation. Together, these changes yield more reliable and computationally efficient *P* values and effect size estimates.

Second, we replaced the Jiang-Doerge (JD) two-stage FDR procedure (Jiang & Doerge, 2006) with a joint χ^2 score test with stage-wise confirmation (Van den Berge et al., 2017). The original JD approach screened only on the zero-inflation component and confirmed on the count component, effectively assuming a single mode of regulation. Because the zero-inflation and count components of the hurdle model are independent by construction, the sum of their $\chi^2(1)$ score statistics follows a $\chi^2(2)$ distribution under the null, providing the principled omnibus screening statistic. After FDR control on the joint statistic, we apply Holm-corrected confirmation within each selected pair to classify the regulatory mode into three categories (new **Supplementary Fig. 3**): **switch** (significant only in the zero component, gating transcriptional activation), **rheostat** (significant only in the count component, tuning expression level among expressing cells), and **dual** (significant in both components). The joint test maintains power for peak–gene links driven by a single component, which the JD's sequential design structurally misses.

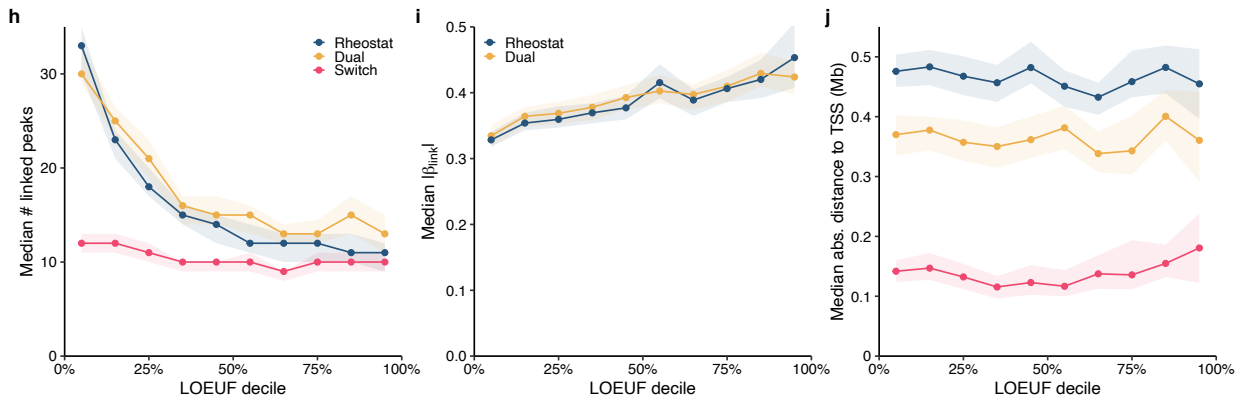
Together, these improvements yield 1,095,949 significant peak-gene-cell type trios (593,765 unique peak-gene pairs) across cell types (new **Fig. 1e**). The regulatory mode classification reveals that 22% of links act as a switch, 36% rheostat, and 42% dual, with biologically coherent variation across cell types and genomic context. Key new findings enabled by this classification include:

- **Spatial organization by regulatory mode:** Switch and dual links are enriched near the TSS with a downstream bias, while rheostat links are more uniformly distributed (median absolute distance to TSS = 167, 254, and 444 kb for switch, dual, and rheostat), consistent with promoter-proximal elements activating gene expression while distal enhancers fine-tune expression levels (new **Fig. 1f,g**).
- **Context-dependent regulatory architecture:** 4.1% of significant peaks alternate between switch and rheostat modes across genes and cell types, with a recurring pattern in which the same regulatory element gates expression in one cell type while tuning expression levels in another, with myeloid cells more frequently employing graded regulation. We illustrate this with *IRF4*, *TCF4*, and *CCDC50* (new **Supplementary Fig. 4**).
- **ENCODE4 validation by regulatory mode:** Dual peaks show the strongest enrichment for promoter-like signatures (RR = 4.65), followed by switch (1.65) and rheostat (1.10). Rheostat peaks show a slight enrichment for distal enhancer-like signatures (RR = 1.06), whereas switch and dual peaks do not (new **Fig. 1h**).
- **Constraint analysis by regulatory mode:** The excess of linked peaks at constrained genes is driven primarily by rheostat and dual links (median 33 and 30 vs. 11 and 13 between constrained and unconstrained genes for rheostat and dual, respectively, compared to 10 vs. 8 for switch). Both rheostat and dual links show smaller effect sizes at constrained genes, 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). Together, these patterns indicate that constrained genes preferentially employ fine-tuned expression control rather than binary gating (new **Extended Data Fig. 9h–j**).

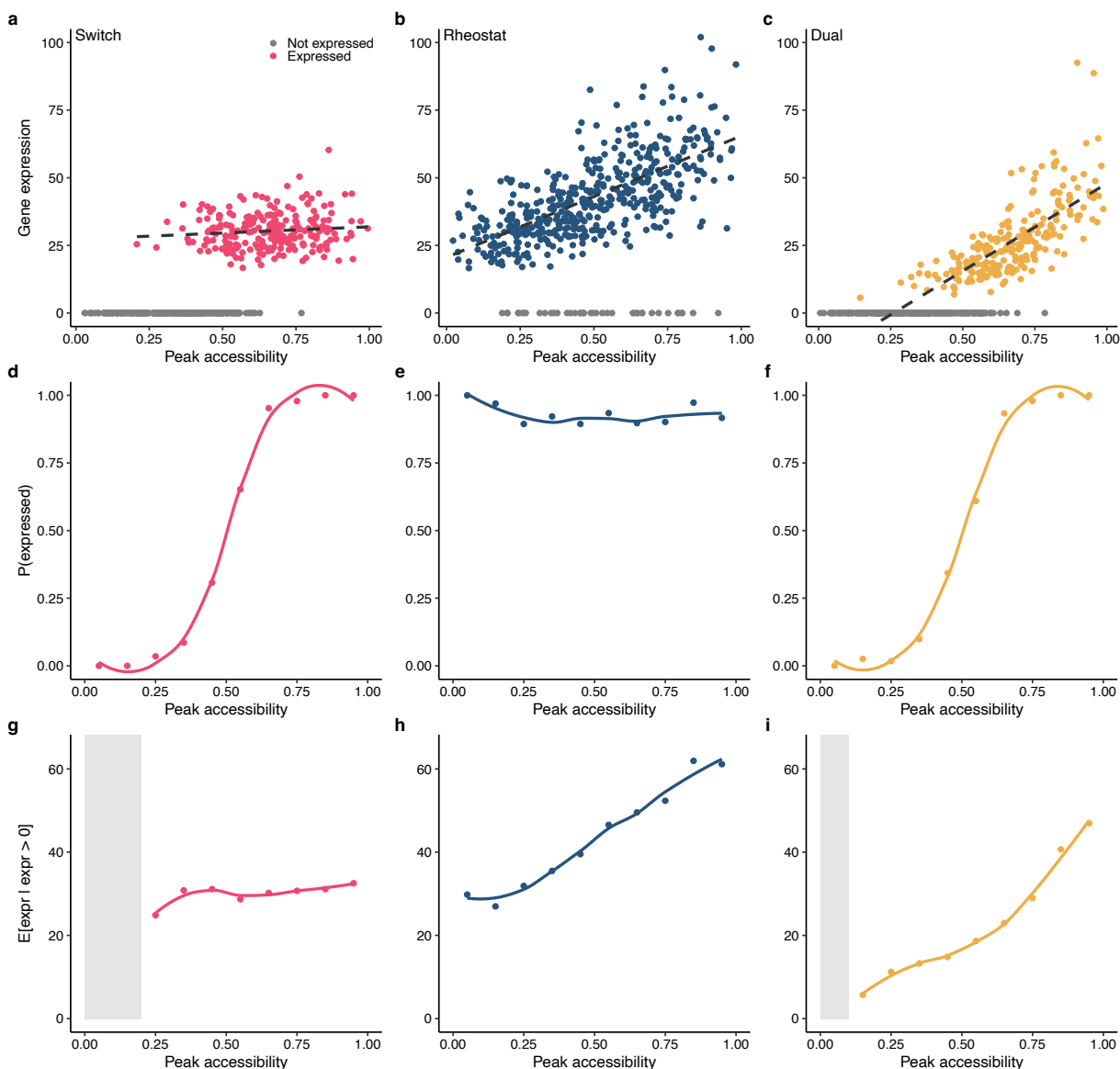


Revised Fig. 1e–h: Single-nucleus RNA-seq and ATAC-seq atlases of 1,108 FinnGen donors. **e.** The number of chromatin accessibility (CA) peaks with significant enhancer–gene links (CA–Link+; see **Methods**) across L0 (PBMC) and L1 cell types. **f.** Distribution of signed peak-to-TSS distances for significant peak–gene link, stratified by regulatory mode (switch, rheostat, or dual; see **Methods**). Negative distances are towards the 5' end and positive distances are

towards the 3' end of each gene, aligned by strand. The vertical dashed line at 0 indicates the TSS. **g**, Fraction of significant peak–gene links assigned to each regulatory mode, stratified by 20 kb distance bins. Shaded ribbons represent 95% confidence intervals within each bin. **h**, The enrichment of peak–gene links for each ENCODE4 cCRE classification, stratified by regulatory modes. Enrichment is defined as a risk ratio of having a significant link and being in an annotation (see **Methods**). Error bars represent 95% confidence intervals.



Revised Extended Data Fig. 9 | Multi-layered regulatory buffering at constrained genes. **h–j**, Median number of linked peaks per gene (**h**), median absolute peak–gene link effect sizes ($|\beta_{link}|$; **i**), and median absolute peak–gene distance to transcription start site (TSS; **j**) across LOEUF deciles, stratified by peak regulatory mode (dual, switch, rheostat).



New Supplementary Fig. 3: Schematic of the hurdle model decomposition 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, $E[\text{expr} \mid \text{expr} > 0]$, 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**).

We have updated the **Results** and **Methods** sections accordingly.

Results:

Leveraging our large-scale paired multiome measurements, we systematically mapped 593,765 enhancer-gene links with significant peak-gene associations in *cis* (i.e., < 1 Mb between TSS

and peak) across cell types using a hurdle model to handle both zero-inflation and overdispersion in single-cell gene expression³², with significance determined via a joint χ^2 score test with stage-wise confirmation controlling FDR at $\alpha = 0.05$ (**Fig. 1e, Supplementary Fig. 2, Supplementary Table 4, Methods**). For each significant link, we further classified the regulatory mode into three categories: “switch” links, where peak accessibility determines whether a gene is expressed; “rheostat” links, where it tunes expression levels among expressing cells; and “dual” links, where it controls both (**Supplementary Fig. 3, Methods**). These regulatory links encompass 104,780 consensus peaks (35%) that regulate 12,018 genes in *cis* (median 4 genes per peak, median 46 peaks per gene), totaling 1,095,949 significant peak–gene–cell type trios. At the trio level, 22% were classified as switch, 36% as rheostat, and 42% as dual (**Supplementary Fig. 4a,b**).

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 identity⁴², 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 specification vs. binary activation in monocytes. 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^{43,44}, 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 commitment⁴⁵. 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^{46,47}, 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**).

The three regulatory modes showed distinct relationships with gene promoters. Switch and dual links were enriched near the TSS with a downstream bias in the direction of transcription, while rheostat links were distributed more uniformly across the *cis* window (median absolute distance to TSS = 167, 254, and 444 kb for switch, dual, and rheostat, respectively; **Fig. 1f,g**). This is consistent with promoter-proximal elements activating gene expression while distal enhancers fine-tune expression levels among expressing cells.

Methods:

We filtered to peak–gene pairs where both the gene and peak had nonzero expression or accessibility of $\geq 1\%$ in nuclei from the tested cell type, ensuring sufficient statistical power. We

additionally required at least 5,000 expressing cells to ensure stable coefficient estimation in the hurdle count component, where the negative binomial log-link produces inflated point estimates when estimated from few expressing cells.

...

For each gene, we fitted a null hurdle model (without the peak variable) and then tested each candidate peak using score tests. Both components included pseudobulk PEER factors (using the same number optimized for the *cis*-eQTL analysis per cell type; **Methods**) to account for donor-level confounding, log-transformed total ATAC counts and the percentage of mitochondrial counts as covariates. Log-transformed total RNA counts were included as an offset in the count model to directly model expression rates, and as a covariate in the zero-inflation model. To control for highly expressed genes or peaks that could dominate normalization, genes or peaks exceeding 5% of total counts in any cell were excluded when computing size factors. In addition, the zero-inflation model included the top K ATAC spectral components to absorb the compositional confound whereby the zero component would otherwise conflate direct regulatory effects with passive co-variation driven by sub-cell-type identity (e.g., naive vs. memory sub-populations). These spectral components were computed separately per cell type using the same SnapATAC2 (ref. 96) spectral embedding procedure described above. The optimal K was determined per cell type by selecting the elbow of the mean delta-BIC curve from covariate-only hurdle models fitted on a stratified sample of 1,000 genes at $K = 0, 1, \dots, 10, 15, 20, 25, 30, 50$. Spectral components were not included in the count model, which conditions on expressing cells (RNA count > 0) and is therefore less susceptible to the compositional confound that drives zero-inflation differences across sub-populations.

...

To determine the significance of peak–gene links, we applied a joint χ^2 score test with stage-wise confirmation⁴⁴ to classify the regulatory mode. Because the zero-inflation and count components of the hurdle model are independent by construction, the sum of their $\chi^2(1)$ score statistics follows a $\chi^2(2)$ distribution under the null. In the screening stage, we computed this joint statistic for each peak–gene pair and applied FDR control using the Benjamini-Hochberg procedure. Peak–gene pairs with $\text{FDR} < 0.05$ were considered significant. In the confirmation stage, for each significant pair, we applied the Holm procedure to the two component P values to determine which component(s) drove the association. This stage-wise approach guarantees overall FDR control at the screening level while providing valid family-wise error rate control within each selected pair, and classifies significant links into mechanistic categories based on which component(s) are confirmed. We refer to these regulatory modes as “switch”, “rheostat”, and “dual” following the classical distinction between binary on/off control of transcriptional activation and graded modulation of transcriptional output among already-expressing cells^{122,123}. This terminology maps directly onto our two-component decomposition, and broadly aligns with the modern kinetic view in which *cis*-regulatory elements modulate transcription through changes in burst frequency, duration, or size^{124–126}.

- **Switch** (only the zero-inflation component significant): peaks where accessibility determines whether a gene is expressed at all but does not influence expression levels among expressing cells, consistent with regulatory elements that gate burst frequency or the fraction of time a gene spends in the transcriptionally active state, e.g., promoter-proximal elements or enhancers that recruit essential transcription factors for gene activation.
- **Rheostat** (only the count component significant): peaks where accessibility modulates expression magnitude without substantially affecting the probability of expression, consistent with regulatory elements that scale burst size or sustained transcriptional

output among already-expressing cells, *e.g.*, distal enhancers that boost transcription rate through additional transcription factor binding or chromatin looping.

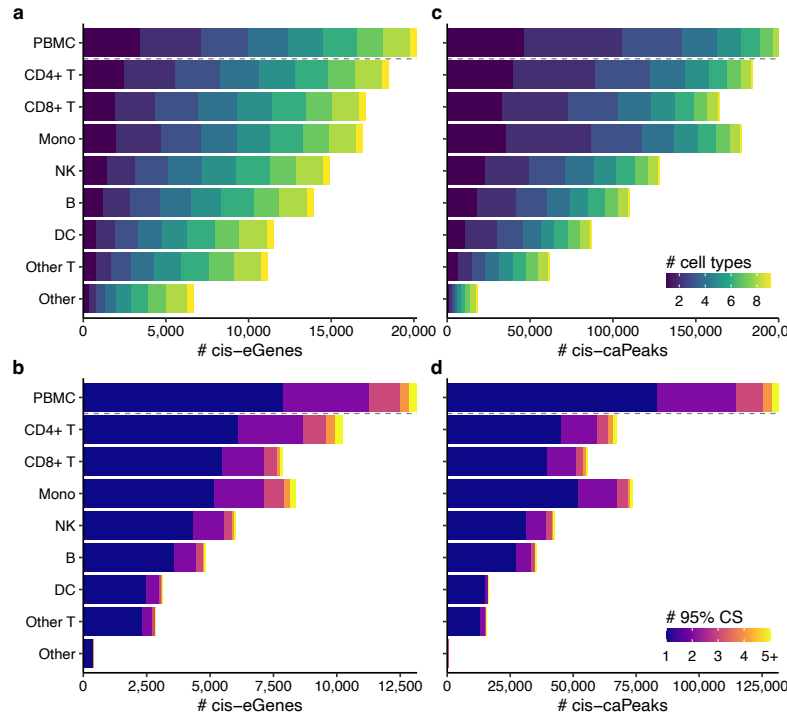
- **Dual** (both components significant): peaks where chromatin accessibility both activates gene expression and modulates its magnitude, consistent with regulatory elements that simultaneously gate burst frequency and scale burst size.

This classification assigns a regulatory mode to each peak–gene–cell type trio. We further aggregated these trio-level modes across all linked genes and cell types to summarize each peak’s overall regulatory behavior. A peak was classified exclusively switch or exclusively rheostat if every significant trio shared that mode, dual if at least one trio was classified as dual, and switch/rheostat (alternating switch and rheostat) if its trios spanned both switch and rheostat without any dual trio. Trios with joint FDR < 0.05 but lacking a confirmed component (omnibus-only) were tolerated within the exclusive categories. This per-peak summary captures whether a regulatory element operates in a fixed mode across contexts or alternates between gating and tuning depending on the target gene or cell type.

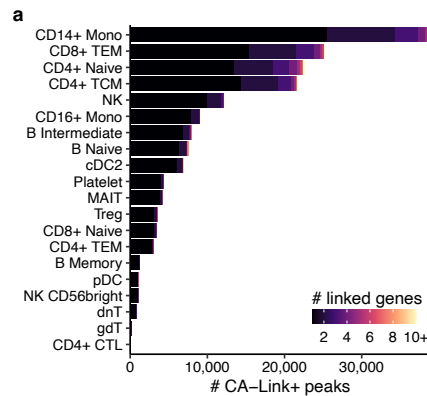
3. In figure panels 2a-d, the authors make use of associations derived from the entire sample, which they call ‘PBMC’. Although I believe this is reasonable, it’s presentation alongside the ‘L1’ annotations is somewhat misleading, as it is an aggregation of all L1 data together in a single analysis. I would recommend referring to this data as something else, perhaps ‘L0’, and faceting it out of the plot with the L1 results. Relatedly, are peak-gene links from L2 cell types also folded into the results of Fig1e? If so, this could be misleading too.

We thank the reviewer for this suggestion to improve clarity. We have revised our nomenclature to refer to this aggregated PBMC analysis as “L0” throughout the manuscript, and have updated **Fig. 2a–d** to visually separate L0 from L1 cell types.

Regarding **Fig. 1e**, this figure presents only the L0 + L1 peak–gene links as described in the legend. We have added a new **Supplementary Fig. 2a** showing L2 results separately.



Revised Fig. 2: Genetic architecture of molecular QTLs in FinnGen. a,c. The number of significant *cis*-eGenes (a) and *cis*-caPeaks (c) for each L1 cell type (FDR < 0.05; see **Methods**). Color represents the number of cell types that each eGene or caPeak is significantly detected for b,d. The number of *cis*-eGenes (b) and *cis*-caPeaks (d) with fine-mapped 95% CSs. Color represents the number of 95% CSs for each *cis*-eGene or *cis*-caPeak.



New 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.

4. There are a few details of the QTL mapping I find unusual and would appreciate further clarity and justification. For cis-eQTL mapping the authors use SAIGE-QTL, including PEER factors into their model. However, the number of PEER factors being utilised is optimised for the number of eGenes being detected by tensorQTL (L1371). I imagine this is to make the comparison between these two methods more equivalent (Extended Figure 5a), but these methods model gene expression in fundamentally different ways, and so I find the use of the tensor-QTL-optimised PEER factor number in SAIGE-QTL tests to be unusual, and potentially harmful to the number of associations that are actually being used later on. Can the authors confirm that if PEER factors were re-optimised for SAIGE-QTL tests, then this would lead to a similar number of PEER factors, and/or similar eGenes? Additionally, in the mapping of caQTLs, the authors use PCs instead of PEER factors as the use of PEER factors to model caQTLs has 'not been fully explored' L1391. This makes the covariates being tested for caQTLs and eQTLs not comparable – why not use PCA to detect covariates for both? The authors also pseudobulk and use tensorQTL for ca-QTL mapping because of the 'sparsity of snATACseq fragments' L195. Does this mean SAIGE-QTL is overly sensitive to sparsity, and therefore not useful for sc-caQTL mapping? If the authors have tested this, it would be a useful detail to include. Finally, the authors then use tensorQTL for trans-eQTL mapping, which results in further discrepancy of how different methods are being employed to test for different types of association.

We thank the reviewer for these detailed methodological questions. We are happy to provide clarification and additional analyses for each point:

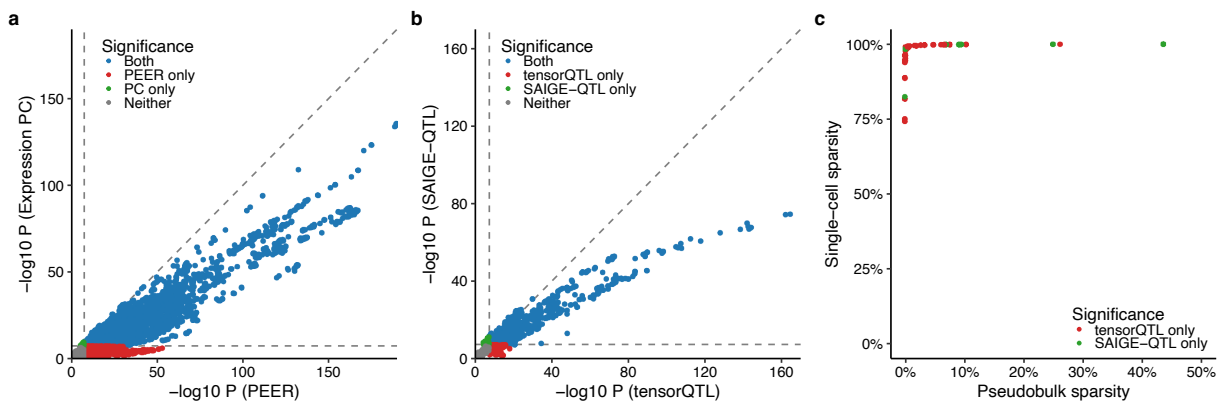
1. PEER factor optimization: We note that PEER factors are computed at the per-donor pseudobulk level, not at single-cell level, and are then used as covariates in both tensorQTL and SAIGE-QTL. This approach follows established practice in the original SAIGE-QTL publication (Zhou *et al.*, *medRxiv*) and the TenK10K study (Cuomo *et al.*, *medRxiv*). Because both methods utilized the same pseudobulk-derived PEER factors, the optimization procedure is appropriate for both frameworks.

2. PEER vs PC covariates: To address the reviewer's concern about consistency, we have performed an additional sensitivity analysis running SAIGE-QTL with expression PCs vs. PEER factors. We randomly selected 200 genes in CD4⁺ T cells stratified by expression sparsity (20 genes per sparsity decile) and ran SAIGE-QTL with the optimal number of expression PCs determined by the same approach used for PEER factors (**Response Fig. 1a**). Nearly all significant associations identified with expression PCs were also detected with PEER factors; 99% of genome-wide significant associations with expression PCs ($P < 5.0 \times 10^{-8}$) were recapitulated by PEER factors. However, PEER-factors identified approximately twice as many associations at genome-wide significance (37,107 vs. 16,618 variant–gene pairs), consistent with PEER capturing additional sources of latent variations. These results confirm that PEER factors provide greater statistical power and justify their use in our SAIGE-QTL analysis.

3. SAIGE-QTL for caQTL mapping: It is known that snATAC-seq data exhibits substantially greater sparsity than snRNA-seq. Pseudobulk aggregation addresses this sparsity by pooling fragments across cells within each donor, providing more robust quantification for caQTL mapping. This pseudobulk tensorQTL approach is consistent with other large-scale snATAC-seq QTL studies, including the TenK10K snATAC-seq study (Xue *et al.*, *medRxiv*). Additionally, the computational burden of testing ~300,000 peaks (compared to ~30,000 genes) made single-cell approaches impractical.

To address the reviewer's question directly, we ran SAIGE-QTL on 200 randomly selected peaks in CD4⁺ T cells stratified by chromatin accessibility sparsity (20 peaks per sparsity decile), and compared the results to our pseudobulk tensorQTL results. We found that tensorQTL identified more genome-wide significant associations overall (4,234 vs. 3,261 variant-peak pairs at $P < 5.0 \times 10^{-8}$), while 97.5% of SAIGE-QTL associations were recapitulated by tensorQTL (**Response Fig. 1b**). Because snATAC-seq peaks are uniformly far sparser at the single-cell level than snRNA-seq genes (**Response Fig. 1c**), SAIGE-QTL's Poisson mixed model is likely underpowered and pseudobulk sum aggregation gains power by pooling fragments across cells.

4. **tensorQTL for trans-eQTL:** We used tensorQTL for trans-eQTL mapping primarily due to computational efficiency.



Response Fig. 1 | Sensitivity analyses of eQTL covariate choice and caQTL mapping method in CD4⁺ T cells. **a**, Comparison of *cis*-eQTL $-\log_{10} P$ -values in CD4⁺ T cells when SAIGE-QTL is run with PEER factors (x-axis) vs. expression PCs (y-axis) as latent covariates, across 200 randomly selected genes stratified by expression non-zero proportion decile. Each point represents a variant-gene pair. Dashed grey lines represent $y = x$ and the genome-wide significance ($P < 5.0 \times 10^{-8}$). **b**, Comparison of *cis*-caQTL $-\log_{10} P$ -values in CD4⁺ T cells when using pseudobulk tensorQTL (x-axis) vs. single-cell SAIGE-QTL (y-axis), across 200 randomly selected chromatin accessibility peaks stratified by chromatin accessibility non-zero proportion decile. **c**, Relationship between pseudobulk sparsity (1 - fraction of donors with a nonzero pseudobulk count; x-axis) and single-cell sparsity (1 - fraction of cells with a nonzero fragment count; y-axis) for the variant-peak pairs from (**b**) that reach genome-wide significance in only one of the two methods.

5. The OneK1k study (Yazar et al., 2022, Science) and subsequent next phase (Cuomo et al., 2025, medRxiv) also map eQTLs in PBMCs of healthy individuals – in datasets of similar or larger sizes than this study. While I appreciate there is a methodological difference between these efforts with this one (single cell vs single nucleus), and biological difference in the ancestry of the study population, to better evaluate the reliability and novelty of the eQTLs being mapped in this study, it would be useful to compare the associations being detected here with these efforts.

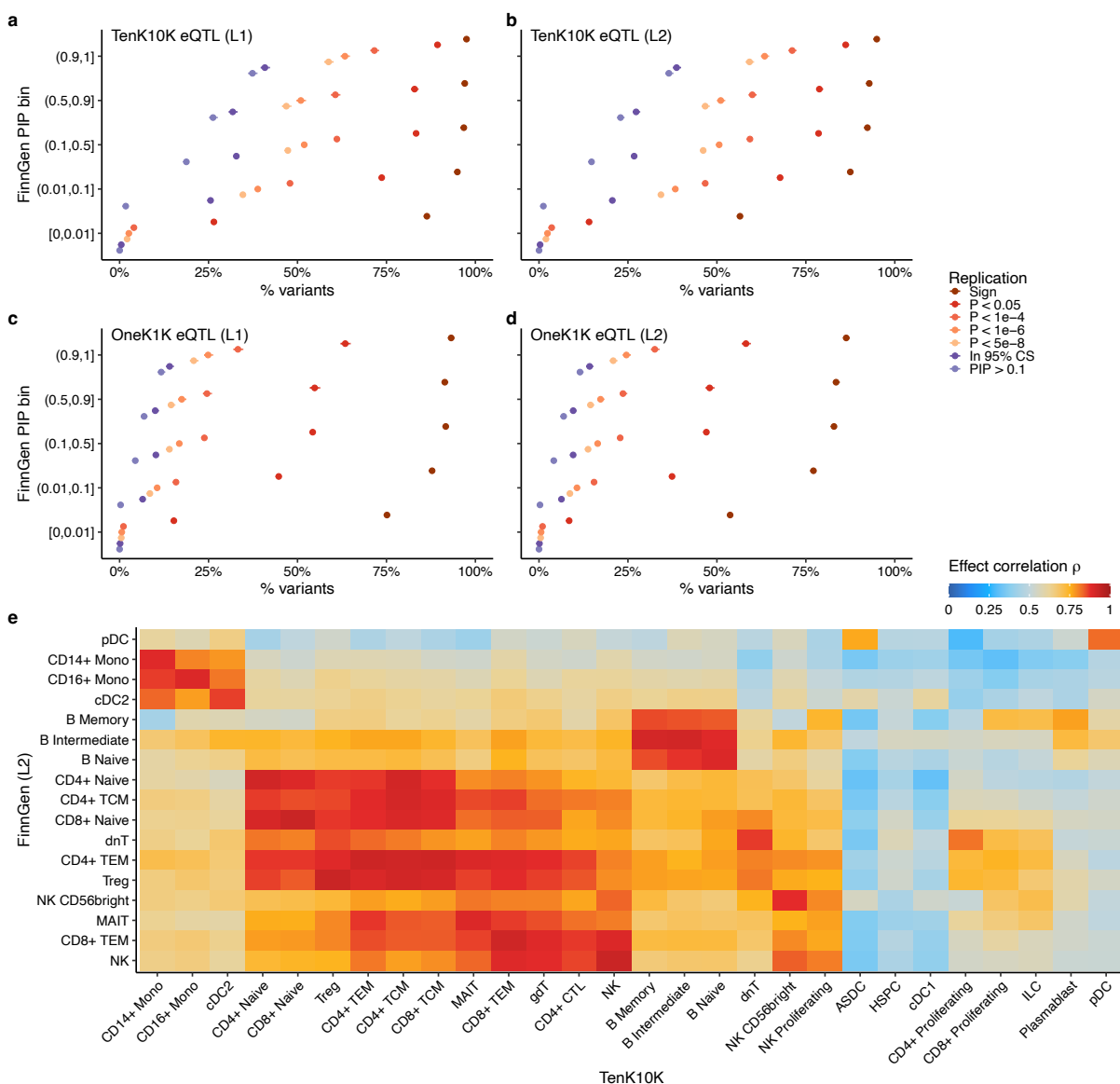
We thank the reviewer for this suggestion. We have now performed a comprehensive comparison of our *cis*-eQTL results with TenK10K (Cuomo et al., medRxiv) and OneK1K (Yazar et al., 2022), as well as three additional cross-ancestry datasets. We stratified FinnGen fine-mapped eQTLs by SuSiE PIP and evaluated replication across multiple metrics, pooled across L2 cell types.

For TenK10K (~5.4 M cells from 1,925 donors) and OneK1K (~1.2 M cells from 982 donors), which provide full summary statistics, replication rates increased monotonically with PIP across all metrics (new **Supplementary Fig. 8a–d**). Among fine-mapped variants with PIP > 0.9, 96% and 89% showed concordant effect direction in TenK10K and OneK1K, respectively; 59% and 21% reached genome-wide significance ($P < 5.0 \times 10^{-8}$), with fine-mapping replication (external

95% CS membership) at 40% and 14%, respectively. Effect size correlations (Spearman's ρ) at PIP > 0.9 were 0.92 and 0.80 for TenK10K and OneK1K among all matched variants, increasing to 0.96 and 0.86 among lead variants. Cross-cell-type correlation analysis confirmed cell-type-specific replication, with corresponding cell types between FinnGen and TenK10K clustering together (new **Supplementary Fig. 8e**).

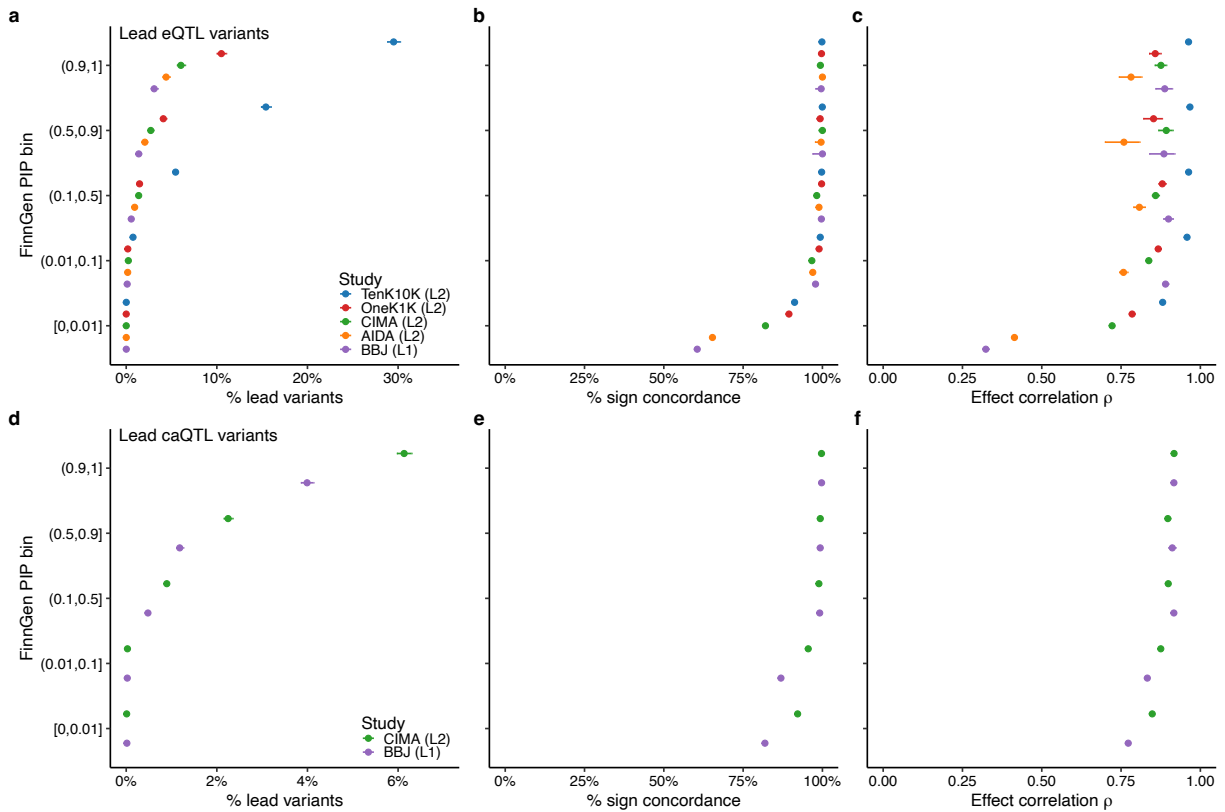
We additionally compared with three East Asian datasets (CIMA, AIDA, and BBJ), demonstrating that fine-mapped FinnGen eQTLs are replicated across ancestries (lead variant $\rho = 0.78$ –0.89 at PIP > 0.9; new **Supplementary Fig. 9**). Beyond eQTLs, we replicated caQTLs and peak–gene links in CIMA and/or BBJ, as detailed in our response to **Reviewer 2, Comment 1**.

To address these points, we have now added a new **Supplementary Fig. 8,9** and revised the **Results** accordingly.



New 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**).



New 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.

Results:

We replicated fine-mapped *cis*-eQTLs across five independent datasets spanning diverse ancestries (TenK10K, OneK1K, CIMA, AIDA, and BBJ; **Supplementary Figs. 8,9**), where replication rates increased monotonically with FinnGen PIP. Among fine-mapped *cis*-eQTLs with PIP > 0.9, 96% showed concordant effect direction and 59% reached genome-wide significance in TenK10K (Spearman's $\rho = 0.92$), with lead variant $\rho = 0.78$ – 0.96 across all five datasets.

6. For colocalization, the authors use a threshold of posterior probability for a shared causal variant (PPH4) of 0.5 (L420). Unfortunately, I believe this threshold is too relaxed and therefore likely to result in an excess of false positive associations being reported. This is because, in theory, a PPH4 of 0.5 could occur when the PPH3 (both traits under genetic control but with different causal variants) is still very close to 0.5, which would indicate near-equivalent evidence for traits to not share a causal variant. In line with this, comparable studies tend to use higher thresholds, such as; 0.8 (Cuomo et al., 2025 medRxiv; Vochteloo et al., 2025 medRxiv; Yazar et al., 2022 Science; The GTEx Consortium, 2020 Science; Hong et al., 2025 Nature Genetics; Kock et al., 2025 Cell); 0.75 (Alegbe and Harris et al., 2025 medRxiv), 0.7 (Bryois et al., 2022, Nature Neuroscience; Bian et al., 2025 Cell Genomics) or 0.6 (Natri et al., Nature Genetics). I would recommend the authors either acknowledge this possibility in the main text or raise the PPH4 threshold to that of similar studies and adjust the text/figures being presented based on this (e.g 5c). In addition, posterior probabilities for colocalization are never presented in the main text or figures, making it difficult for the reader to readily evaluate the statistical evidence in support of shared genetic effects. There is also no specification of what the prior probabilities are being used for colocalization (p_1 , p_2 , p_{12}), which is an important detail.

We thank the reviewer for raising this point. We have revised our colocalization analyses to use a PP.H4 threshold of 0.8, consistent with the recent studies. While the number of colocalizations are slightly reduced (from 17,317 95% CS–trait pairs to 15,947 pairs), our main conclusions remain the same. We have revised all relevant figures and text accordingly. We also added PP.H4 values for individual examples.

Regarding the prior probabilities, we used the default parameters specified in `coloc.bf_bf` function, i.e., $p_1 = p_2 = 1e-4$ and $p_{12} = 5e-6$. We have revised the **Methods** section to clarify this point.

We used the default prior probabilities specified in the `coloc.bf_bf` function: $p_1 = p_2 = 1.0 \times 10^{-4}$ (prior probability that a variant is associated with trait 1 or trait 2 only) and $p_{12} = 5.0 \times 10^{-6}$ (prior probability that a variant is associated with both traits).

7. The trans-eQTL analysis provides a very surprising result that is currently under-explored by the authors and potentially an over-estimate. Identifying 20,826 trans-eGenes from 1,108 individuals is a very large number of effects given this sample size. Other studies have identified only a fraction of the number of these effects despite using a substantially larger sample size - e.g Vosa et al., (2021, Nature Genetics) identified 6,298 trans-eGenes from a meta-analysis that totalled 31,684 individuals. Can the authors provide additional evidence that these trans-eQTLs are not false positives? For example, false positive trans-eGenes can be detected due to cross-mapping of cis-eGenes with which they share homology (see Saha and Battle, 2018 F1000 Research), and removing genes with high probability of this would be useful. Additionally, although the authors employ what appears to be a strict p-value threshold to determine significant eQTL effects ($p < 5e-08$, L1365), this will account for the number of variants being tested per gene but not the number of genes being tested within each cell-type, which will be in the order of thousands. The expression of each gene is being tested as an independent phenotype, and so I would recommend the authors incorporate this in their multiple testing correction.

In addition, it would be useful to highlight whether any of these associations have been found in previous work, and whether there are any trans-eQTL effects relevant to disease. If, after performing the additional filters detailed above, the overlap with cis-effects really remains as low as the authors currently suggest (3-6% for cis-e/cis-caQTLs respectively, L275-276), this may potentially have a significant impact on how future trans-QTL analyses are performed, as these

frequently rely on subsetting to cis-QTL or disease-casual variants only due to the limited statistical power.

We thank the reviewer for raising this important point. We have performed additional QC for the *trans*-eQTL analysis, including cross-mapping filtering and replication against the eQTLGen Phase 2 (Warmerdam et al., *medRxiv*). As the *trans*-eQTL results are not central to the main conclusions of the manuscript, we have moved them to **Supplementary Note** with these additional filters applied. Below we summarize the analyses performed:

1. **Cross-mapping filter:** Following Saha and Battle (2018), we removed *trans*-eQTL associations where the *trans*-eGene shares sequence homology with any gene within 1 Mb of the lead variant, using the precomputed genome-wide cross-mappability scores. This removed 36% of *trans*-eQTL associations and 4,271 *trans*-eGenes entirely, retaining 16,555 unique *trans*-eGenes.

2. **Multiple testing correction:** We additionally applied Bonferroni correction to account for the number of genes tested per cell type (2,648–27,236 genes; median 11,508), yielding per-cell-type thresholds of approximately 4.3×10^{-12} . At this stringent threshold, 17 unique *trans*-eGenes survived both Bonferroni and cross-mapping filters. We note that this correction is very conservative, and the eQTLGen Phase 2 ($n = 43,301$) also used $P < 5.0 \times 10^{-8}$ without gene-level Bonferroni correction to identify 9,406 *trans*-eGenes.

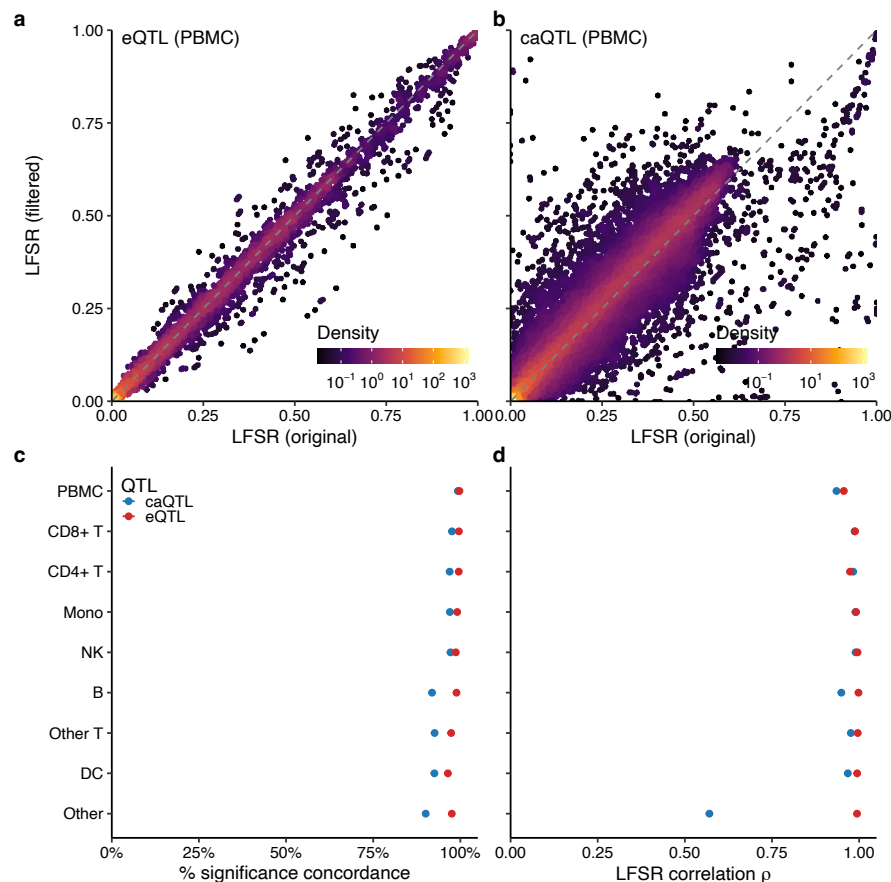
3. **Replication in eQTLGen Phase 2:** We looked up all 48,953 eQTLGen Phase 2 fine-mapped *trans*-eQTL variant-gene pairs in our summary statistics. Of these, 522 (1%) reached $P < 1.0 \times 10^{-5}$ in at least one FinnGen cell type, a modest fraction consistent with the ~40-fold difference in sample size and the small effect sizes of *trans*-eQTLs. Among these 522 pairs, effect sizes were highly concordant (98% sign concordance, Spearman's $\rho = 0.86$), providing evidence that the *trans*-eQTL signals present in our data represent genuine regulatory effects.

We applied the same Bonferroni correction to *trans*-caQTLs (215,949 *trans*-caPeaks at $P < 5.0 \times 10^{-8}$), yielding 36 *trans*-caPeaks at the corrected threshold. The revised manuscript presents all *trans*-molQTL results in **Supplementary Note** with these filters applied. A comprehensive characterization of cell-type-specific trans-regulatory effects will benefit from larger multiome cohorts and is an important direction for future work.

8. While I appreciate the authors attempt to develop a method that better distinguishes cell-type specificity from a lack of statistical power, CASCADE, there are several caveats to using mash in the context of eQTL mapping, particularly at the single cell level. As the authors describe it (L1501-L1517), they are constructing the data-driven covariance matrices from strong variants ('CSs (PIP > 0.1) or, as an alternative when no CSs available, the top five variants with the strongest association ($Z > 4.89$) across any cell type'). However, the genes/variants that are tested in each condition will vary, often quite substantially, and the FLASH and PCA algorithms have limited robustness to missing data. Hence, it becomes necessary to either remove genes that are not tested in every condition, which leads to data-driven matrices being calculated on only a fraction of intersecting tests that may not accurately reflect the heterogeneity of genetic effects across conditions, or input a value for missing tests which may strongly influence the data-driven matrix. What was the total proportion of tests that were either removed or imputed in the data-driven matrices, and have the authors tested how variation in this number influences their results? This is important if others are to utilise this method, and if the total number of non-imputed genetic effects is substantially smaller than the number of effects tested in each condition, I would doubt the reliability of the results from this analysis.

We thank the reviewer for these important methodological questions. As described in the **Methods**, missing effects for any variant–gene/peak–cell type combination were imputed as zero with large standard errors ($SE = 1000$). This is a standard approach in mash that effectively downweights missing observations. Because missingness occurs at the level of each variant–gene–cell type combination, rather than complete exclusion of genes/variants, the data-driven covariance matrices are estimated from the majority of available data.

To address the reviewer's concern, we have now quantified the proportion of missing values in our analyses. Across L1 cell types, 18% and 32% of entries in the strong set used for data-driven covariance estimation were imputed for eQTLs and caQTLs, respectively. Missingness was concentrated in rarer cell types with fewer nuclei per donor and was negligible in well-powered cell types. To assess robustness, we re-ran mash restricting the strong set to variant–gene/peak pairs observed in at least 50% of cell types. The resulting LFSR values were highly concordant with the original analysis, with median significance concordance of 99% and 97% across cell types and median LFSR correlation of 0.99 and 0.98 for eQTLs and caQTLs, respectively (new **Supplementary Fig. 25**). These results confirm that the imputation of missing values with large standard errors does not affect our findings.



New Supplementary Fig. 25 | 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.

In addition, it is not clear to me how data from L2 annotations is used in this analysis (L1524 – ‘we mapped them to their corresponding L1 parent types’) – is mash applied at the L1 level, then separately within each L1 for L2 annotations? Or in one complete analysis? In general, I think the details of this approach could be clarified to aid interpretability.

We have updated the **Methods** to make this explicit. Briefly, mash is applied separately to L1 and L2 cell types. For specificity classification, L2 results are then aggregated to their corresponding L1 parent cell types to maintain hierarchical consistency.

To clarify these points, we have revised the **Methods** as follows:

To distinguish true cell type specificity from limited statistical power, we applied mash to jointly model effect sizes across cell types, separately for L1 and L2 cell types.

...

Missing effect estimates in any cell types were set to zero with large standard errors (= 1000), effectively downweighting these observations in the mash model. Across L1 cell types, 18%

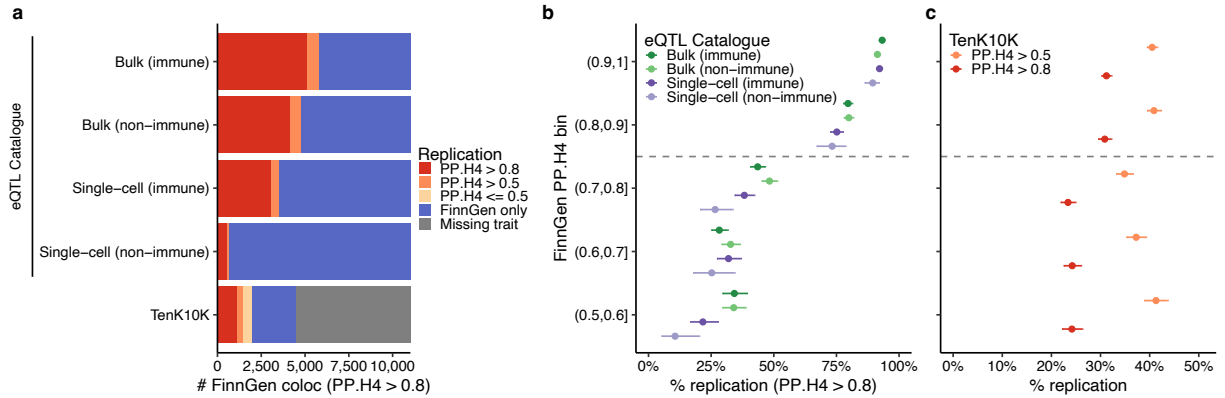
and 32% of entries in the strong set were imputed for eQTLs and caQTLs, respectively, with missingness concentrated in rarer cell types with fewer nuclei per donor. To assess robustness, we ran mash restricting the strong set to variant–gene/peak pairs observed in at least 50% of cell types and found highly concordant LFSR values (median significance concordance of 99% and 97%; median LFSR correlation of 0.99 and 0.98, for eQTLs and caQTLs, respectively; **Supplementary Fig. 25**).

9. Related to my previous comment on the novelty of eQTL associations, it would also be useful if the authors commented on which, and how many, of their colocalizations between QTLs and disease susceptibility were novel or supported by previous work. A good benchmark for this would be the number of loci/genes colocalised for each trait in this study compared with those detected by OpenTargets or Cuomo et al., (2025, MedRxiv). Even if eGenes have been colocalised with disease previously by eQTL-based approaches, the single-cell resolution, chromatin accessibility layer and CASCADE analysis provided in this study would still substantiate the residing therapeutic hypothesis, and this would be beneficial to highlight to the reader.

We thank the reviewer for this suggestion to better contextualize our colocalization findings. We have now systematically benchmarked the novelty and replication of our colocalizations (new **Supplementary Fig. 12**).

1. Comparison with the eQTL Catalogue: For a direct comparison, we colocalized 287 eQTL Catalogue datasets from 48 studies ($n = 65\text{--}4,729$; new **Supplementary Table 28**) against the same set of FinnGen GWAS summary statistics, using the identical SuSiE-based colocalization pipeline applied in our study (**Methods**). Of 11,015 unique gene-trait eQTL colocalizations in FinnGen ($PP.H4 > 0.8$), 6,808 (62%) were also detected in at least one external dataset, while 4,207 (38%) were novel (new **Supplementary Fig. 13a**). We classified eQTL Catalogue datasets along two axes: resolution (bulk vs. single-cell) and tissue relevance (blood/immune vs. non-immune). Bulk immune-cell datasets captured the most replicated colocalizations (5,122; 47%), primarily driven by INTERVAL ($n = 4,729$; Tokolyi, et al., 2025). Existing single-cell immune datasets (OneK1K, Perez, et al., and others) captured 3,054 (28%), reflecting their smaller sample sizes (new **Supplementary Fig. 13b**). Non-immune tissue datasets detected 4,265 (39%) of our colocalizations; however, only 958 (8.7%) were found exclusively in non-immune tissues and not in any immune dataset, consistent with ubiquitously expressed eQTL genes rather than tissue-specific regulatory mechanisms. Overall, 53% of our colocalizations were supported by at least one immune-cell eQTL dataset in the eQTL Catalogue.

2. Replication in TenK10K (Cuomo et al., medRxiv): As the reviewer suggested, we also compared our colocalizations with TenK10K (~5.4 M cells from 1,925 donors). We validated colocalizations between TenK10K eQTLs and GWAS for 57 traits with matching FinnGen traits. Of the 4,071 testable FinnGen gene-trait colocalizations, 1,618 (40%) replicated at $PP.H4 > 0.8$ and 2,084 (51%) at $PP.H4 > 0.5$ (new **Supplementary Fig. 13c**). The replication rates were relatively stable across FinnGen $PP.H4$ bins, rather than increasing monotonically with FinnGen $PP.H4$ as expected. This likely reflects methodological differences, *i.e.*, TenK10K applied single-causal-variant coloc.abf (Giambartolomei, et al., 2014) rather than multi-causal-variant coloc.susie (Wallace, 2021), and used external GWAS meta-analyses that can be miscalibrated for fine-mapping and colocalization (Kanai et al., 2022). For this reason, we consider the eQTL Catalogue comparison (which uses the same GWAS traits and the same colocalization pipeline) to be the more reliable benchmark.



New 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 ≤ 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**.

3. Added value of our study: Importantly, even for colocalizations previously identified by existing eQTL studies, our study provides three layers of additional resolution: cell-type assignment, chromatin accessibility, and CASCADE classification. Among the 4,207 novel gene-trait colocalizations not found in the eQTL Catalogue, 331 involve variants classified as full or positional cascade, spanning 131 genes including *IL21R* and *RHOH* in the manuscript vignettes (**Fig. 6** and **Extended Data Fig. 7**). We further highlight two additional examples below:

- **CD2 and RA** (rs798000, local cascade): *CD2* encodes a T cell surface glycoprotein essential for T cell activation and adhesion (Binder et al, 2020). A variant located 16 kb upstream of *CD2*, rs798000:C>T, colocalizes with RA ($\beta = -0.07$, $P = 1.8 \times 10^{-8}$; PP.H4 = 1.0) through a local cascade, which increases chromatin accessibility of an overlapping peak (chr1:116737244-116738320; $\beta = 0.18$, $P = 8.8 \times 10^{-8}$, PIP = 0.99) linked to *CD2* ($P_{\text{joint}} = 1.1 \times 10^{-14}$; dual) and thus increases *CD2* expression in CD4⁺ T cells ($\beta = 0.04$, $P = 4.1 \times 10^{-17}$, PIP = 1.0). The cell-type resolution pinpoints CD4⁺ T cells as the mediating population, consistent with the central role of CD4⁺ T cells in RA pathogenesis.
- **LCP1 and autoimmune disease** (rs568422359, distal cascade): *LCP1* encodes L-plastin, an actin-bundling protein that regulates immune synapse formation, T cell activation, and leukocyte migration (Joshi & Morley, 2022). An intronic variant of *LRRC63*, rs568422359:TA>T, located 3 kb upstream of *LCP1*, decreases *LCP1* expression ($\beta = -0.27$, $P = 5.3 \times 10^{-53}$, PIP = 0.89 in CD14⁺ Mono) but not *LRRC63*, and colocalizes with autoimmune disease ($\beta = -0.09$, $P = 1.5 \times 10^{-9}$; PP.H4 = 0.99) through a distal cascade across broad immune cell types, including CD4⁺ T, CD8⁺ T, NK, monocytes, and DC. The broad cell-type involvement is consistent with L-plastin's fundamental role in cytoskeletal remodeling across immune lineages.

Furthermore, our study uniquely enables the characterization of Finnish-enriched variants that are underpowered in other populations. The two vignettes in **Fig. 6** (*TNRC18* and *IL21R*) both involve Finnish-enriched full cascade variants regulating from chromatin to expression, with functional validation via STAG-seq for *TNRC18*. These discoveries, combined with the cell-type resolution, chromatin accessibility, and CASCADE classification for both novel and previously known colocalizations, represent a substantial advance over existing resources.

To address these points, we have now added a new **Supplementary Fig. 13** and revised the **Results** accordingly.

To benchmark our colocalizations against existing resources, we colocalized 287 eQTL Catalogue datasets⁷ against the same FinnGen GWAS using the identical colocalization pipeline (**Supplementary Fig. 13; 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 ($n = 4,729$)⁵⁸. The novel colocalizations were enriched for cell-type-specific effects and CASCADE mechanisms not detectable in bulk tissue studies (**Supplementary Table 29**). Our GWAS-eQTL colocalizations were further replicated in TenK10K¹⁹, despite the methodological differences in colocalization pipelines (**Supplementary Fig. 13c**).

10. There are several examples of colocalisations being given with details on their detectable regulatory cascade (e.g *TICAM1*, *RNASET2*, *RNAH*, *GRAMD1B* (later on) and others), and while there is comment about the cell-type where QTL effects are presented, there is not detail on the specificity of these effects, as determined by the author's CASCADE approach. Including this detail, along with plots showing the gene expression and chromatin accessibility across genotypes would help the reader interpret the nature and strength of the potential biological mechanism being presented.

We thank the reviewer for this suggestion. We have revised the **Results** to include CASCADE specificity classifications for each highlighted example:

- *TICAM1*, local cascade, lineage specific (myeloid)
- *RNASET2*, positional cascade, single cell type (monocytes)
- *RHOH*, local cascade, single cell type (CD4⁺ T cell)
- *PILRB*, local cascade, single cell type (monocytes)
- *APOBEC3A*, distal cascade, single cell type (monocytes)
- *CBX6*, distal cascade, cross-lineage shared
- *TNRC18*, positional cascade, single cell type (CD4⁺ T cells)
- *IL21R*, distal cascade, lineage specific (lymphoid)

To visualize the cell type specificity of these examples, we have added forest plots of eQTL, caQTL, and peak–gene link effect sizes across cell types for *TICAM1*, *RNASET2*, and *RHOH* (new **Supplementary Fig. 12**) as well as *PILRB*, *APOBEC3A*, and *CBX6* (new **Supplementary Fig. 14**). Forest plots for *TNRC18* and *IL21R* are already covered in **Fig. 6**. Full CASCADE classifications for all fine-mapped molQTL variants are available in **Supplementary Table 20**.

11. Although I agree with the statement on L591-595 (and L785-789) that the increased statistical power in this study is a contributing factor to the increased number of eQTLs detected for constrained genes, these statements come off a little strong given larger studies of bulk RNAseq-derived eQTLs exist (and so power cannot be the only factor driving this discovery), and the effect

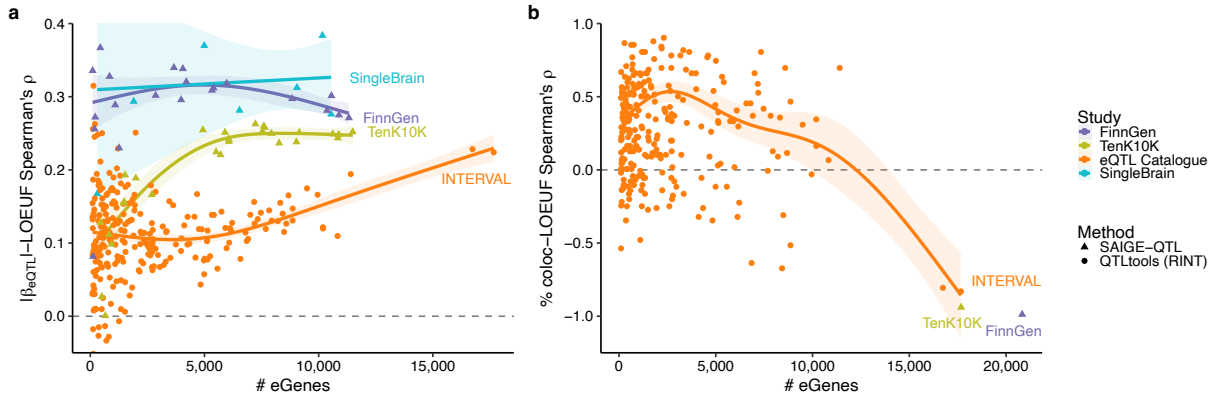
of power has not been directly tested by the authors. A good experiment to test this would be to repeat the eQTL mapping across multiple rounds of down sampling, where the number of cells/samples being included in the analysis is titrated, and then assessing whether the number of associations for constrained genes is lost as statistical power is reduced.

We thank the reviewer for this important suggestion. We have performed the downsampling experiment and a series of sensitivity analyses that provide a comprehensive account of the role of statistical power.

1. The eQTL effect size gradient is robust to all confounders tested. In our primary SAIGE-QTL analysis, eQTL effect sizes are systematically attenuated at constrained genes (Spearman's $\rho = 0.20$, $P = 1.7 \times 10^{-228}$). To confirm this is not driven by expression-level or variance differences, we performed additional sensitivity analyses. Matching genes for expression level preserved the gradient (Spearman's $\rho = 0.16$, $P = 3.3 \times 10^{-9}$); simultaneously matching for both mean expression and expression variance residuals still preserved it (Spearman's $\rho = 0.10$, $P = 3.7 \times 10^{-6}$; **Extended Data Fig. 9b–d**), attributing approximately two-thirds of the attenuation to regulatory redundancy. We additionally ran tensorQTL on pseudobulk PBMC expression to compare rank-based inverse normal transformation (RINT) normalization (which forces unit variance per gene) with $\log_2(\text{CP10K} + 1)$ normalization (which preserves variance). RINT attenuated the gradient ~2-fold (mean $\rho = 0.14$ vs. 0.27 ; new **Extended Data Fig. 9e**). This attenuation is itself biologically meaningful, as constrained genes have lower expression variance (**Extended Data Fig. 9c**), and RINT removes this gene-specific variance by design. The gradient remained significantly positive under RINT, confirming that regulatory redundancy attenuates eQTL effects independent of expression variance.

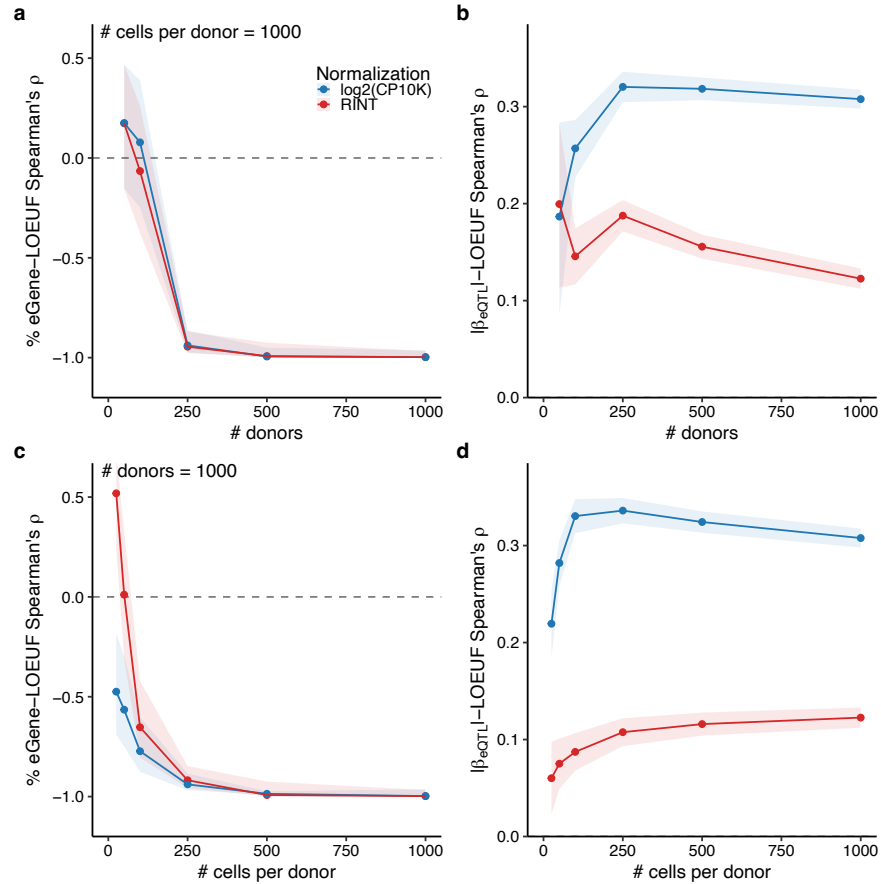
2. The effect size gradient and GWAS colocalization enrichment replicate across studies. To replicate our findings, we analyzed the constraint gradient across 287 eQTL datasets from the eQTL Catalogue ($n = 65\text{--}4,729$; Kerimov, et al., 2021), TenK10K ($N = 1,925$; Cuomo et al., *medRxiv*), and SingleBrain ($n = 983$; Jang et al., 2026). The effect size gradient was positive in 97% of datasets (median Spearman's $\rho = 0.12$, max = 0.38), holding across both analytical methods (single-cell SAIGE-QTL vs. bulk QTLtools with RINT) and tissue contexts (immune vs non-immune; new **Supplementary Fig. 20a**). The effect size gradient converged towards approximately $\rho = 0.2\text{--}0.3$ as the number of eGenes increased, consistent with power-dependent detection of regulatory buffering. Subsequent to our study, the eQTLGen Phase 2 ($n = 43,301$; Warmerdam et al., *medRxiv*) recapitulated 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. (2026) independently reported the same gradient across striatal and cortical cell types in 178 brain donors, citing our preprint.

The GWAS colocalization enrichment at constrained genes (**Fig. 5f**) also replicated in TenK10K (Spearman's $\rho = -0.94$, $P = 5.5 \times 10^{-5}$), closely recapitulating our result (Spearman's $\rho = -0.99$, $P = 9.3 \times 10^{-8}$). We further evaluated GWAS colocalization across 287 eQTL Catalogue datasets against FinnGen GWAS. At the typical colocalization rate of ~1%, individual datasets had too few colocalizations (~10 per LOEUF decile) to detect the gradient; the majority showed positive ρ_{coloc} (unconstrained genes favored), mirroring the eGene detection paradox (new **Supplementary Fig. 20b**). The colocalization enrichment emerged only in datasets with >2,000 colocalizations, including INTERVAL ($N = 4,729$, Spearman's $\rho = -0.83$, $P = 0.003$; Tokolyi, et al., 2025), confirming that detecting this signal requires the scale of colocalizations that only large-scale eQTL studies produce.



New 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.

3. Downsampling reveals a power-dependent transition. We systematically downsampled CD4⁺ T cells along two orthogonal axes: donor count (25–1,000 donors at 1,000 cells / donor) and per-donor cell depth (25–1,000 cells at 1,000 donors), with 5 replicates per condition (new **Supplementary Fig. 19**). Along both axes, Spearman's ρ between eGene fraction and LOEUF decile transitioned from positive to strongly negative as power increased. This transition reflects two opposing forces: constrained genes are more highly expressed (favoring detection) but have smaller eQTL effects (opposing detection), and at sufficient power the expression advantage overcomes the effect size disadvantage. In contrast, Spearman's ρ between eQTL effect size and LOEUF remained stable across all conditions, confirming it is a fixed property of the regulatory architecture rather than a power artifact.

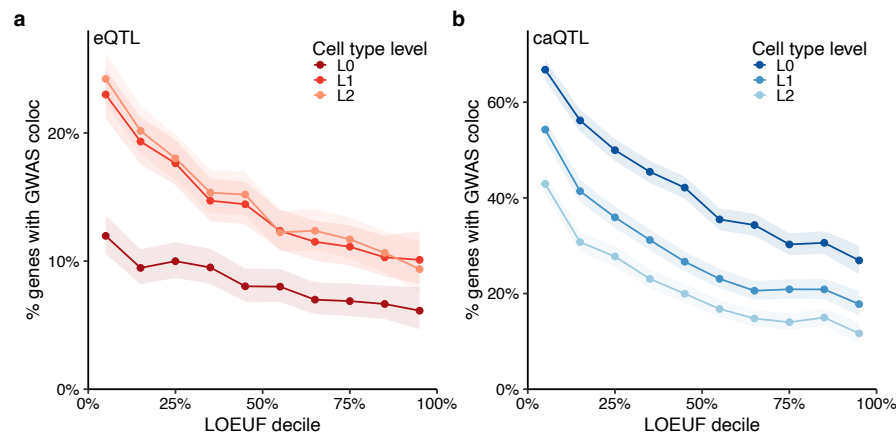


New Supplementary Fig. 19 | Downsampling reveals a power-dependent transition in eGene detection alongside a stable effect size gradient. Downsampling analysis of *cis*-eQTL mapping in CD4⁺ T cells. tensorQTL pseudobulk *cis*-eQTL mapping was performed under two normalization strategies: rank-based inverse 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 ($|\rho_{\text{eQTL}}|$) 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.

Additionally, previous studies have primarily relied on bulk or low-resolution capture of gene expression for QTL mapping. Even though these may have substantial statistical power, cell-type-specific effects are still likely to be lost due to the averaging of their effects across heterogeneous cell populations. However in this study, where associations are mapped at high resolutions, these effects are likely better captured, and this is also likely to contribute to the discovery the authors describe. It would therefore be beneficial if the authors either; i) investigated whether eQTLs that act on constrained genes are enriched for those identified at higher cellular resolutions (e.g L2 vs L0) using either their own or existing datasets that map eQTLs at multiple resolutions, or ii) adjust the text at these positions to soften the claim that this effect is 'primarily power based' (L592).

Because eGene detection rates saturate at high power and are confounded by expression level, we assessed this question using GWAS colocalization rates, which directly addresses the GWAS-

eQTL colocalization paradox. The eQTL colocalization gradient was already strongly present at L0 (PBM; 12% constrained vs. 6.1% unconstrained genes), before any cell-type deconvolution, and held at finer resolution with higher absolute rates (L1: 23% vs. 10%; L2: 24% vs. 9.4%). The caQTL colocalization gradient showed the opposite resolution dependence, with rates highest at L0 (67% vs. 27%) and declining with resolution (L1: 54% vs. 18%; L2: 43% vs. 12%). Total colocalizations therefore diverged in opposite directions across resolution levels (eQTL: L0 = 1,370, L2 = 2,469; caQTL: L0 = 6,912, L2 = 3,620; new **Supplementary Fig. 18**), while the constraint gradient was preserved at every level. This contrast demonstrates that resolution effects are primarily driven by statistical power dynamics rather than qualitatively distinct regulatory biology at constrained genes. We have revised the **Results** accordingly.



New Supplementary Fig. 18 | GWAS colocalization rates across LOEUF deciles and cell type resolution levels
a,b, The fraction of genes with *cis*-eQTL–GWAS colocalization (**a**) or *cis*-caQTL–GWAS colocalization (**b**) for each loss-of-function observed/expected upper bound fraction (LOEUF) decile, stratified by cell type resolution level. Color represents cell type level: L0 (PBM), L1 (major cell types), or L2 (fine-grained cell types; see **Methods**). Shaded ribbons represent 95% confidence intervals.

Consistent with previous findings, 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); 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 $\rho = 0.16$, $P = 3.7 \times 10^{-6}$), attributing approximately two-thirds of the residual gradient to regulatory redundancy rather than variance differences (**Extended Data Fig. 9d**). The gradient also recapitulated in pseudobulk tensorQTL across normalization choices and was robust to cell-type resolution (**Extended Data Fig. 9e**, **Supplementary Fig. 18**; **Methods**).

...
The apparent discrepancy with earlier findings is primarily power-based, confirmed by downsampling and cross-study replication (**Supplementary Fig. 19,20**).

Minor comments

1. What does 'with genotypes of interest' refer to in Extended Data Figure 1? This reads as if the samples were selected a priori based on specific germline variants.

This refers to an initial pilot study and is no longer applicable. We have revised **Extended Data Fig. 1** to remove this outdated language.

2. I find the subset of the snATACseq data to the same cells that passed the snRNAseq QC to be an important point that strengthens the reliability of the comparisons between the two modalities (detailed on L1254). It would be good if the author's highlight this detail in the results section so that it is readily available to the reader.

We have revised the **Results** to highlight this point as follows:

We next generated the snATAC-seq atlas containing 7,100,899 high-quality nuclei ($n = 1,093$), each with paired high-quality snRNA-Seq data, enabling direct comparison between chromatin accessibility and gene expression within the same nuclei (**Methods**).

3. L146 – it would be useful to outline in the text the threshold being used to determine 'significant' peak–gene associations.

As described in our response to **Comment 2**, we replaced the Jiang-Doerge two-stage FDR procedure (Jiang & Doerge, 2006) with a joint χ^2 score test with stage-wise confirmation (Van den Berge et al., 2017). We have revised the **Results** to clarify the significance threshold ($\text{FDR} < 0.05$ on the joint χ^2 statistic). Full details of the procedure are provided in the **Methods**.

Leveraging our large-scale paired multiome measurements, we systematically mapped 593,765 enhancer–gene links with significant peak–gene associations in *cis* (i.e., < 1 Mb between TSS and peak) using a hurdle model to handle both zero-inflation and overdispersion in single-cell gene expression³², with significance determined via a joint χ^2 score test with stage-wise confirmation⁴¹ controlling FDR at $\alpha = 0.05$ (**Fig. 1e, Supplementary Fig. 2, Supplementary Table 4, Methods**).

4. I believe there is no reference to Figure 1f. If this is missed, this should be rectified, or the panel should be omitted.

We thank the reviewer for catching this oversight. We replaced the old **Fig. 1f** with new **Fig. 1f,g** (see **Comment 2**) in the **Results** section and have ensured that all figure panels are referenced in the main text.

5. On L212, the term 'propagation patterns' is vague and should be clarified.

We have revised the **Results** to clarify this point:

Colocalization across molQTLs identified extensive sharing of genetic signals across regulatory layers, with notable differences in how genetic effects propagate across molecular layers.

6. On L1477-1481 the authors detail how they use graph-based clustering to merge colocalised credible sets. Although I believe this could be a reasonable approach, for readers to properly understand and replicate this analysis, it would be useful if the authors provide examples.

We thank the reviewer for this suggestion. We have added the following illustrative example to the **Methods** section.

For example, consider five CSs (A–E) with pairwise colocalization A–B, B–C, and D–E. Step 1 identifies two connected components: {A, B, C} and {D, E}. In step 2, suppose A and B share a common variant v_1 , while C doesn't; and D and E share another variant v_2 . The algorithm produces three merged CSs: {A, B} with shared variant v_1 , {C} as a singleton, and {D, E} with shared variant v_2 . This approach ensures that merged CSs contain only CSs sharing at least one variant, while preserving colocalizations within each component.

7. On L222, the authors describe an ‘intronic variant’ in *LILRB4*, however this variant appears to not be in the gene (Extended Data Fig 5b).

We thank the reviewer for identifying this inconsistency. The variant rs776736 is located in an intron of a non-canonical *LILRB4* transcript (ENST00000270452) but is upstream of the canonical transcript (ENST00000430952). We have revised the text to accurately reflect this:

For example, we identified a variant upstream of *LILRB4*, rs776736:A>G, that increases *LILRB4* gene expression in monocytes

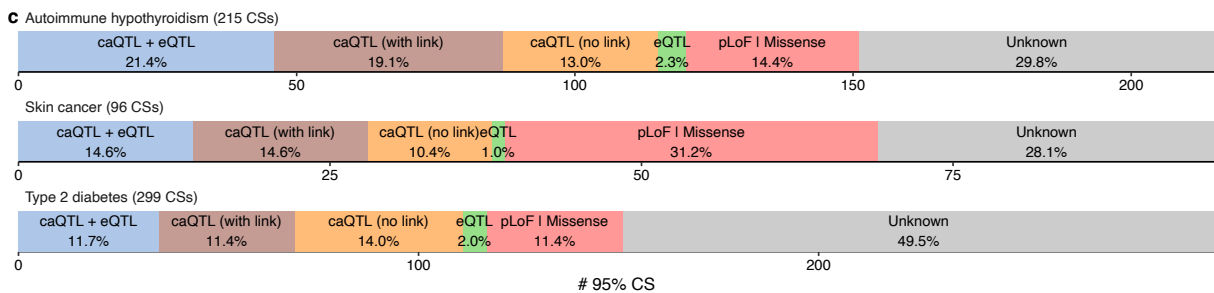
8. On L280, when the authors introduce CASCADE, it would be useful if the authors introduce the problem they are seeking to address using this method in the results section. It is not yet obvious to a general reader the reason why CASCADE has been developed or what this might offer existing approaches.

To clarify this point, we have added the following text in the **Results**:

Interpreting molQTLs presents two key challenges: distinguishing genuine cell type-specific effects from apparent specificity driven by power differences across cell types, and mapping how genetic variation connects to gene expression through chromatin accessibility.

9. In Figure 5c, the use of ‘NA’ to GWAS describe credible sets with unknown molecular impact is not very descriptive. I would recommend replacing this with ‘Unknown’.

We have replaced “NA” with “Unknown” in **Fig. 5c** to clarify this point.



Revised Fig. 5: Colocalization between molQTL and GWAS variants reveals multi-layered regulatory buffering at constrained genes. c. Proportion of 95% CSs classified by molQTL colocalization (caQTL & eQTL, only caQTL, and only eQTL) and presence of nonsynonymous coding variants (predicted loss-of-function [pLoF] or missense). Three

example traits, autoimmune hypothyroidism, skin cancer, and type 2 diabetes, are shown to demonstrate the difference in molQTL colocalization rates across trait categories.

10. On L456, the authors indicate that the relatively low colocalization rates in skin cancer and type 2 diabetes suggest ‘their genetic architecture involves regulatory mechanisms in other cell types beyond PBMCs’. This is one possibility, but it is also possible that they do exist in PBMCs, but; i) are not active in healthy individuals, ii) this study does not have the statistical power to detect the molecular effects, or iii) are only active under specific stimulatory conditions. This statement should be softened or adjusted to incorporate these additional explanations.

We thank the reviewer for raising this point. We have revised the text to acknowledge this point:

In contrast, skin cancer and type 2 diabetes show substantially lower molQTL colocalization rates (41% and 39%, respectively), particularly for caQTL and eQTL (15% and 12%, respectively; **Fig. 5c**), suggesting their genetic architecture involves regulatory mechanisms in other cell types beyond PBMCs, though other factors such as context-dependent effects or limited statistical power may also contribute to this observation.

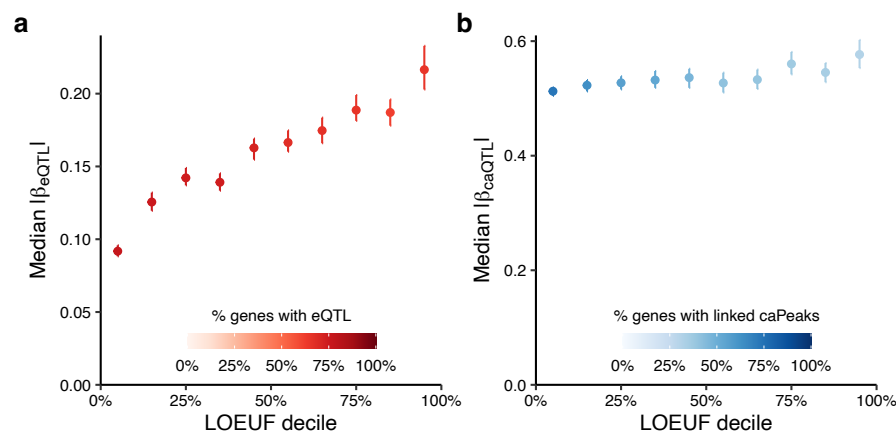
11. In their presentation of associations across selective constraint bins (as defined by the LOEUF score), the authors present numerous especially interesting findings, but do not convey the uncertainty surrounding the scoring of genes in this manner. These statistics are useful, but imperfect measures of selection - which is difficult to reliably assess. It would be beneficial to convey this uncertainty to the reader.

We appreciate this comment. We note that LOEUF was specifically designed to account for uncertainty by using the upper bound of the confidence interval, making it more robust than earlier constraint metrics. Furthermore, our analyses use LOEUF deciles rather than continuous values, which further reduces sensitivity to uncertainty at individual genes. We have added the following text acknowledging this property and the general limitations of constraint estimation.

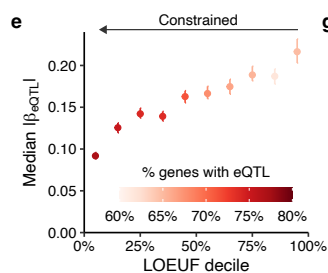
...using loss-of-function observed/expected upper bound fraction (LOEUF) deciles from gnomAD^{62,63} v4.1, which accounts for measurement uncertainty through confidence interval bounds (**Methods**).

12. In Figures 5e and 5g, the points are coloured by the percentage of genes with eQTLs or linked caPeaks respectively, with the range for these values being the minimum observed to maximum observed values (e.g 60-80% in 5e). I believe this may exaggerate the difference in these proportions that are interpreted by the reader, and would recommend this is updated to span the range of potentially observed values (0% to 100%). In addition, it would be useful if the direction of ‘more constrained’ was shown on the plots, with an arrow for example, to aid interpretability. These comments also affect Extended Data Figure 9h and i.

We thank the reviewer for this suggestion. However, we selected the observed data range to highlight biologically meaningful differences; a 0–100% scale would obscure the gradient central to our findings, as shown below (**Response Fig. 2**). We have ensured the color bar legends clearly indicate the actual range and added an arrow (revised **Fig. 5e**) indicating the direction of increasing constraint.



Response Fig. 2 | The constraint gradient for *cis*-eQTL and *cis*-caQTL effect sizes with 0–100% color scale. a,b, The relationship between loss-of-function observed/expected upper bound fraction (LOEUF) deciles and regulatory layers. Panels show median absolute effect sizes for *cis*-eQTL (a), *cis*-caQTL (b). Color represents the fraction of genes with either eQTL (blue) or linked peaks affected by caQTL (red) for each decile.



Revised Fig. 5e: Colocalization between molQTL and GWAS variants reveals multi-layered regulatory buffering at constrained genes. The relationship between loss-of-function observed/expected upper bound fraction (LOEUF) deciles and regulatory layers. Panels show median absolute effect sizes for *cis*-eQTL. Color represents the fraction of genes with linked peaks affected by caQTL (red) for each decile. Lower LOEUF decile corresponds to more constrained genes.

13. I do not understand the term ‘expression-matched highly-expressed’ (L584). How can genes be matched for expression but also subset to the top 20% most expressed?

We apologize for the confusion in terminology. This analysis first filters to highly-expressed genes (top 20%), then matches for expression levels across LOEUF deciles within this subset to control for residual expression differences (**Methods**).

We have revised the **Results** to clarify as follows:

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 $\rho = 0.23$, $P = 7.6 \times 10^{-18}$)

Referee #1 (Remarks on code availability):

<https://github.com/mkanai/cascade>, <https://github.com/mkanai/finngen-multiome-flagship> and <https://github.com/FINNGEN/singlecell-pipeline> are not accessible, so I have been unable to test this code.

For <https://github.com/mkanai/fasthurdle>, which is available, I was unable to install the package due to GCC incompatibility on my cluster, and a lack of fortran package and sudo privileges on my local machine. I wonder if the authors could provide a conda environment recipe, or singularity container that could be used to download the necessary dependencies instead?

We apologize for this oversight. All code repositories are now publicly available. For fasthurdle, we now provide pre-built binaries via R-universe (<https://mkanai.r-universe.dev/fasthurdle>), which eliminates the need for a C++ compiler or Fortran toolchain entirely. For environments where installing compiler toolchains is difficult, we also provide a pre-build docker image that can be converted to a Singularity container for use on HPC clusters. Full installation instructions, including source installation and Docker/Singularity options, are provided in the fasthurdle README.

Referee #2:

Ramnik J. Xavier and colleagues present a population-scale single-nucleus multi-omics atlas, profiling paired chromatin accessibility and gene expression in over 10 million immune cells from 1,108 Finnish individuals. The work provides a large-scale map of immune gene regulation, identifying numerous molecular quantitative trait loci and enhancer–gene links. A central conceptual contribution is the “CASCADE” framework, which categorizes genetic variants into mechanistic cascades linking chromatin alterations to gene expression. The authors report that variants involved in multi-layered regulatory pathways show stronger colocalization with disease, suggesting a biologically informed approach to prioritizing non-coding variants. The strategic focus on the Finnish population facilitates the study of variants which are not readily observed in other cohorts. This resource provides a substantial reference for functional genomics and may aid future studies of the genetic architecture of immune-mediated diseases. However, several issues require further improvement and justification for the work to meet the publication standards of a journal such as Nature.

We thank the reviewer for recognizing the value of our population-scale multiome atlas, the CASCADE framework, and the strategic advantage of the Finnish population for studying disease-causing regulatory variants. We appreciate the constructive feedback received and have addressed each comment in detail below.

1. The study's entire molQTL map is derived from a single cohort which, despite its size, lacks independent replication, thereby weakening confidence in the robustness of these foundational results. While the focus on the Finnish population provides a unique opportunity to discover population-enriched variants, it simultaneously limits the assessment of how generalizable the identified regulatory mechanisms are. To address the concern regarding independent and cross-ancestry validation for core QTLs, the authors should split their cohort into discovery and validation subsets to provide internal replication metrics for a defined set of high-confidence QTLs. Furthermore, they should evaluate the portability of their major findings by leveraging publicly available cross-ancestry molQTL resources, including the eQTL catalogue, and report what proportion of their lead QTLs replicate in both direction and significance in at least one other ancestral group. Ideally, validation should also be attempted in an independent cohort from a different population to assess cross-ancestry consistency, a point underscored by comparison to other studies (Yazar et al. Science, 2022, DOI: 10.1126/science.abf3041; Xue et al. MedRxiv, 2025 <https://www.medrxiv.org/content/10.1101/2025.08.27.25334533v1>).

We thank the reviewer for raising this point about validation and generalizability. We address this concern through comprehensive cross-ancestry comparison rather than internal cohort splitting, which would reduce power for both discovery and validation.

We have now performed a comprehensive comparison of our *cis*-eQTL and *cis*-caQTL results with multiple independent datasets spanning across diverse ancestries:

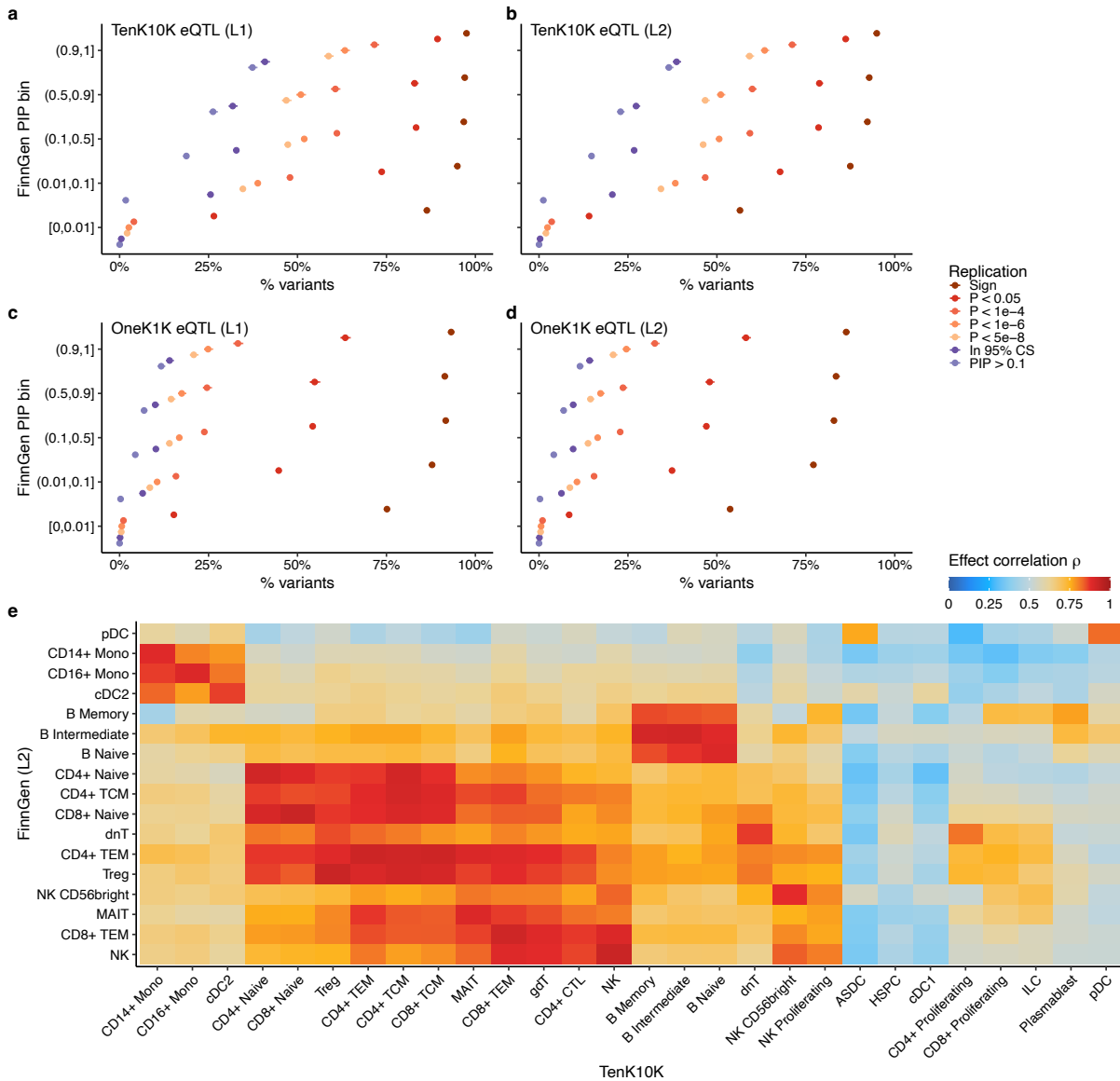
- **TenK10K**, ~5.4 M cells from 1,925 donors (Cuomo et al., medRxiv)
- **OneK1K**, ~1.2 M cells from 982 donors (Yazar et al., 2022)
- **Asian Immune Diversity Atlas (AIDA)**, ~1.2 M cells from 619 donors (Kock et al., 2025)
- **Chinese Immune Multi-Omics Atlas (CIMA)**, ~6.5 M cells from 428 donors (Yin et al., 2026)

In addition, to directly address the reviewer's request for validation in an independent cohort from a different population, we included an independent Japanese PBMC multiome dataset (**BioBank Japan [BBJ]**; ~0.7 M nuclei from 98 donors; *manuscript in preparation*), which enables replication of *cis*-eQTL, *cis*-caQTL, and peak–gene link results within a single independent cohort of East Asian ancestry.

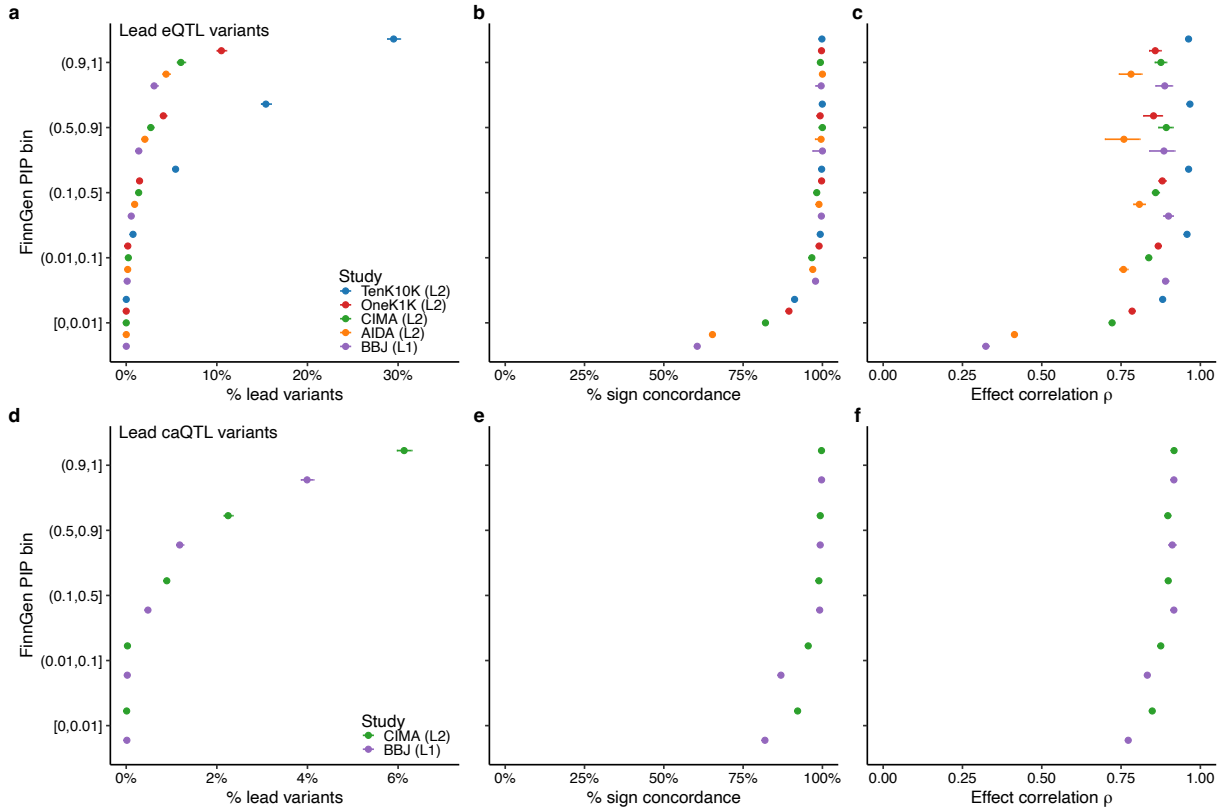
We stratified FinnGen fine-mapped eQTL variants by PIP and evaluated replication across multiple metrics, pooled across L2 cell types (or L1 for BBJ given the sample size). For TenK10K and OneK1K, which provide full summary statistics, replication rates increased monotonically with PIP across all metrics (new **Supplementary Fig. 8a–d**). Among fine-mapped variants with PIP > 0.9, 96% and 89% showed concordant effect direction in TenK10K and OneK1K, respectively; 59% and 21% reached genome-wide significance ($P < 5.0 \times 10^{-8}$), with fine-mapping replication (external 95% CS membership) at 40% and 14%, respectively. Effect size correlations (Spearman's ρ) at PIP > 0.9 were 0.92 and 0.80 for TenK10K and OneK1K among all matched variants, increasing to 0.96 and 0.86 among lead variants. Cross-cell-type correlation analysis confirmed cell-type-specific replication, with corresponding cell types between FinnGen and TenK10K clustering together (new **Supplementary Fig. 8e**). For cross-ancestry datasets (CIMA, AIDA, and BBJ), which provide lead variant statistics, overlapping lead eQTLs showed near-perfect sign concordance (99–100%) and strong effect correlations (Spearman's $\rho = 0.78$ –0.96 at PIP > 0.9), while lead variant overlap rates were modest (3–6% of PIP > 0.9 eQTLs), reflecting power difference (new **Supplementary Fig. 9a–c**).

We also replicated caQTL results using CIMA and BBJ (new **Supplementary Fig. 9d–f**, where overlapping lead caQTLs showed near-perfect sign concordance (99.7%) and strong effect correlations (Spearman's $\rho = 0.92$ for both). We further replicated peak–gene links in BBJ by running the same hurdle model across L1 cell types (new **Supplementary Fig. 6**). 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 correlations of $\rho = 0.69$ (count model) and 0.59 (zero model). These BBJ-replicating links also largely retained their FinnGen regulatory mode assignment (66% overall concordance: rheostat 87%, switch 63%, dual 57%), and switch/rheostat flips were rare (< 2%). Most discordant assignments reflected FinnGen dual links detected as a single mode in BBJ, consistent with reduced power at $n = 98$.

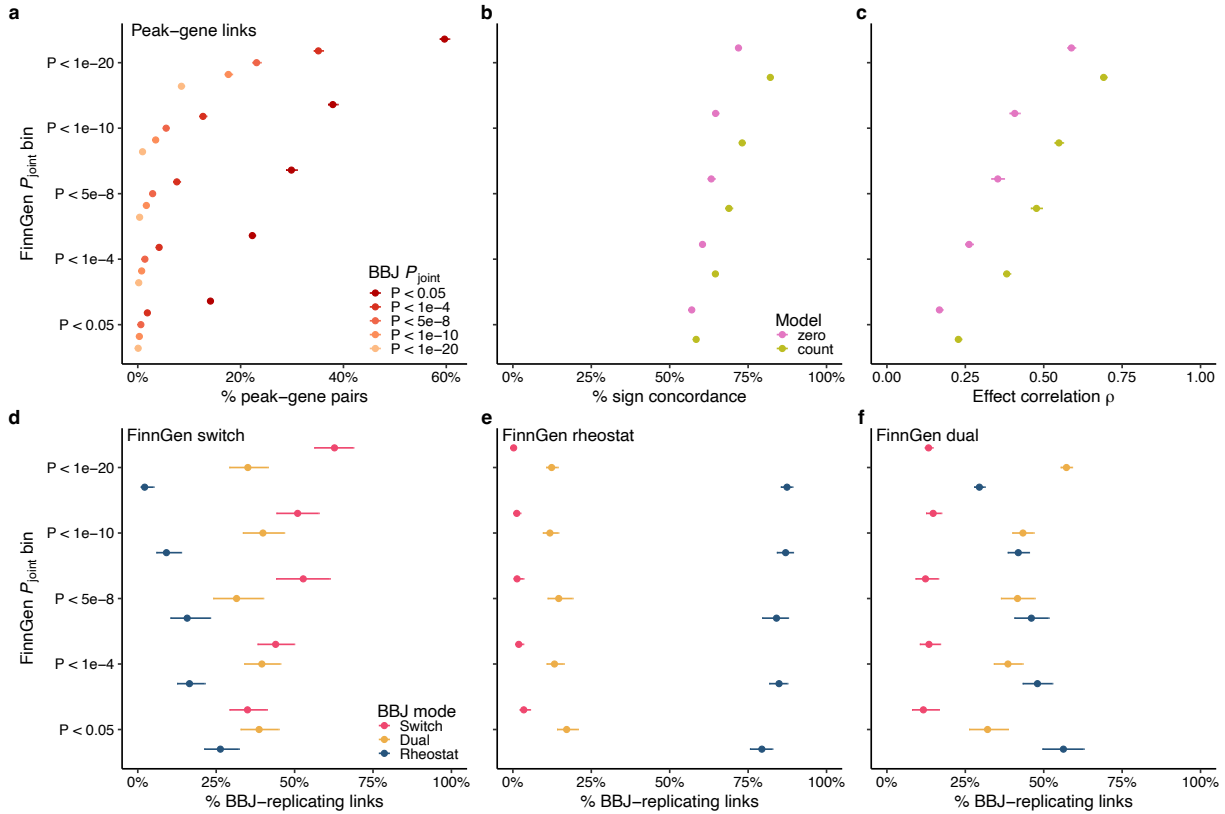
To address these points, we have now added new **Supplementary Figs. 6,8,9** and revised the **Results** accordingly.



New 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**).



New 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.



New 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 bar represents 95% confidence intervals.

We replicated peak-gene links in an independent Japanese multiome cohort (BioBank Japan, BBJ; $n = 98$), where replication rates increased monotonically with FinnGen significance and BBJ-replicating links largely retained their FinnGen regulatory-mode assignment (**Supplementary Fig. 6**).

...

We replicated fine-mapped *cis*-eQTLs across five independent datasets spanning diverse ancestries (TenK10K, OneK1K, CIMA, AIDA, and BBJ; **Supplementary Figs. 8,9**), where replication rates increased monotonically with FinnGen PIP. Among fine-mapped *cis*-eQTLs with $\text{PIP} > 0.9$, 96% showed concordant effect direction and 59% reached genome-wide significance in TenK10K (Spearman's $\rho = 0.92$), with lead variant $\rho = 0.78$ – 0.96 across all five datasets. We also replicated *cis*-caQTLs in CIMA and BBJ (Spearman's $\rho = 0.92$ for both).

2. A significant limitation of this otherwise comprehensive atlas is its focus on steady-state PBMCs from healthy donors, resulting in a static snapshot that fails to capture the dynamic gene regulation fundamental to immune cell function and disease pathology. As highlighted by comparison to prior work (Perez et al. Science, 2022) and (Yazar et al. Science, 2022), genetic effects can be state-dependent, emerging only upon cellular activation or along differentiation trajectories. To address this lack of dynamic and stimulus-responsive regulatory analysis, the authors should perform *in silico* trajectory inference on key lineages such as CD4⁺ T cells and monocytes from their existing data to identify dynamic QTLs. Subsequently, the authors should design a pilot study that involves exposing PBMCs from a subset of genotyped donors, comprising 50 to 100 individuals, to canonical immune stimuli like IFN- α or anti-CD3/CD28. Paired single-cell Multiome profiling conducted pre- and post-stimulation would directly identify stimulus-responsive molQTLs and quantify the fraction of disease-associated variants whose regulatory effects are context-dependent.

We thank the reviewer for highlighting the importance of context-dependent dynamic eQTLs. While we agree that stimulus-responsive molQTL mapping is a valuable direction, conducting a new study involving 50–100 donors with paired pre/post-stimulation multiome profiling is beyond the scope of this revision. Such experiments would require substantial additional time, resources, and represent a distinct study in their own right. Instead, we have taken two complementary approaches: (1) performing *in silico* dynamic eQTL analysis using T cell activation states, and (2) leveraging existing stimulus-response eQTL datasets to quantify what fraction of regulatory variation is context-dependent.

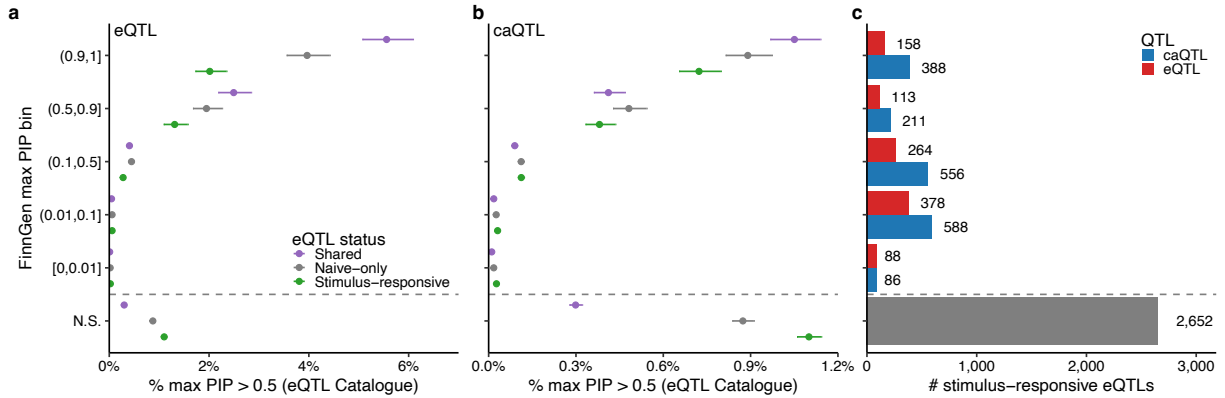
[REDACTED]

[REDACTED]

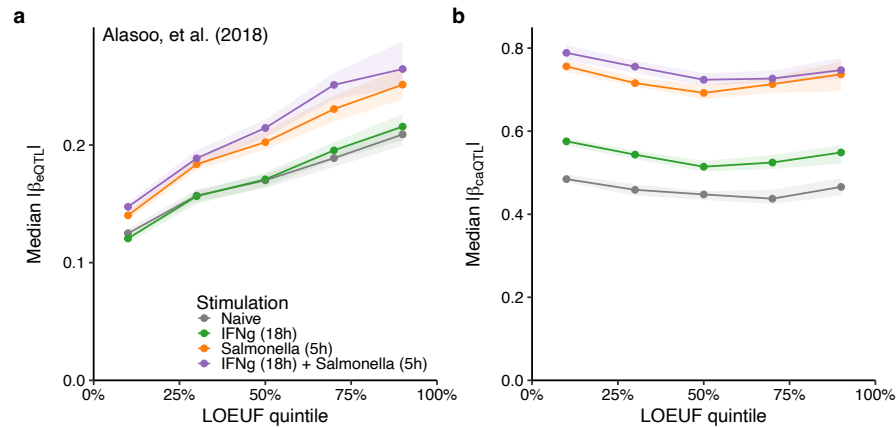
2. Quantifying stimulus-responsive disease variants: We intersected fine-mapped stimulus-responsive eQTLs from the eQTL Catalogue (eight studies with paired naive/stimulated immune conditions) with our FinnGen molQTLs. We defined 4,626 stimulus-responsive eQTLs that were fine-mapped under stimulation (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), while 48% were not in *cis* of any FinnGen eGenes or caPeaks (new **Supplementary Fig. 10**). Among our fine-mapped molQTLs, the fraction of stimulus-responsive eQTLs increased monotonically with FinnGen PIP, with caQTL variants capturing stimulus-responsive eQTLs at higher rates than eQTL variants, consistent with enhancer priming (Alasoo et al, 2018). We note that these fractions should represent a lower bound given the limited set of stimuli covered by the eight studies in the eQTL Catalogue.

Using paired naive/stimulated caQTLs and eQTLs from Alasoo et al. (2018), we further confirmed that the regulatory buffering we observe at constrained genes is not specific to steady-state conditions (new **Extended Data Fig. 10a,b**). Pathogen stimulation (Salmonella and IFN γ + Salmonella) increased caQTL effect sizes uniformly across the constraint spectrum (Spearman's ρ = -0.05 to -0.06; LOEUF x condition interaction P = 0.75), while the eQTL gradient steepened (Spearman's ρ = 0.23 in naive to 0.26 in IFN γ + Salmonella; interaction P = 2.4×10^{-3}), with eQTL effects increasing disproportionately at unconstrained genes (27% vs. 18% at constrained genes; P = 0.009). This indicates that enhancer priming and regulatory buffering operate at distinct regulatory layers. Stimulation increases caQTL effects regardless 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.

We have revised the **Results** to highlight these points:



New 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.



Extended Data Fig. 10 | The buffering gradient is independent of intrinsic *cis*-regulatory activity and cellular stimulation.

a, Median absolute *cis*-eQTL effect size ($|\beta_{eQTL}|$) across LOEUF quintiles in iPSC-derived macrophages from Alasoo et al. (2018), stratified by stimulation conditions (**Methods**). Shaded ribbons represent 95% confidence intervals. **b**, Median maximum absolute *cis*-caQTL effect size ($|\beta_{caQTL}|$) per gene across LOEUF quintiles, stratified by stimulation condition.

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 Catalogue⁷ 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 priming⁵². These fractions likely represent a lower bound, given the limited set of stimuli profiled across the eight eQTL Catalogue studies.

...

To assess whether this buffering architecture is specific to steady-state conditions, we analyzed paired naive/stimulated caQTLs and eQTLs from Alasoo et al. (macrophages, 84 donors, 4 conditions). 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 x condition interaction $P = 2.4 \times 10^{-3}$), 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-priming and regulatory buffering operate at distinct regulatory layers. Stimulation increases caQTL effects regardless 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.

3. While the CASCADE framework powerfully nominates candidate causal variants and cascades, functional validation is limited primarily to TNRC18, leaving the vast majority of the over 100,000 fine-mapped variants and proposed mechanisms without experimental support. This constitutes a notable shortfall in functional validation and causal mechanistic exploration, especially when compared to related studies that integrated deeper functional follow-up. To strengthen the mechanistic claims, we request targeted experimental validation for a selected set of high-priority findings. Specifically, the authors should perform high-throughput reporter assays such as MPRA on at least six to eight Finnish-enriched variants from distinct mechanistic categories, which could include full cascade, positional cascade, or opposing-effect caQTL examples. These assays should be conducted in a minimum of two relevant immortalized immune cell lines and, where feasible, one primary cell type to confirm cell-type-specific regulatory activity.

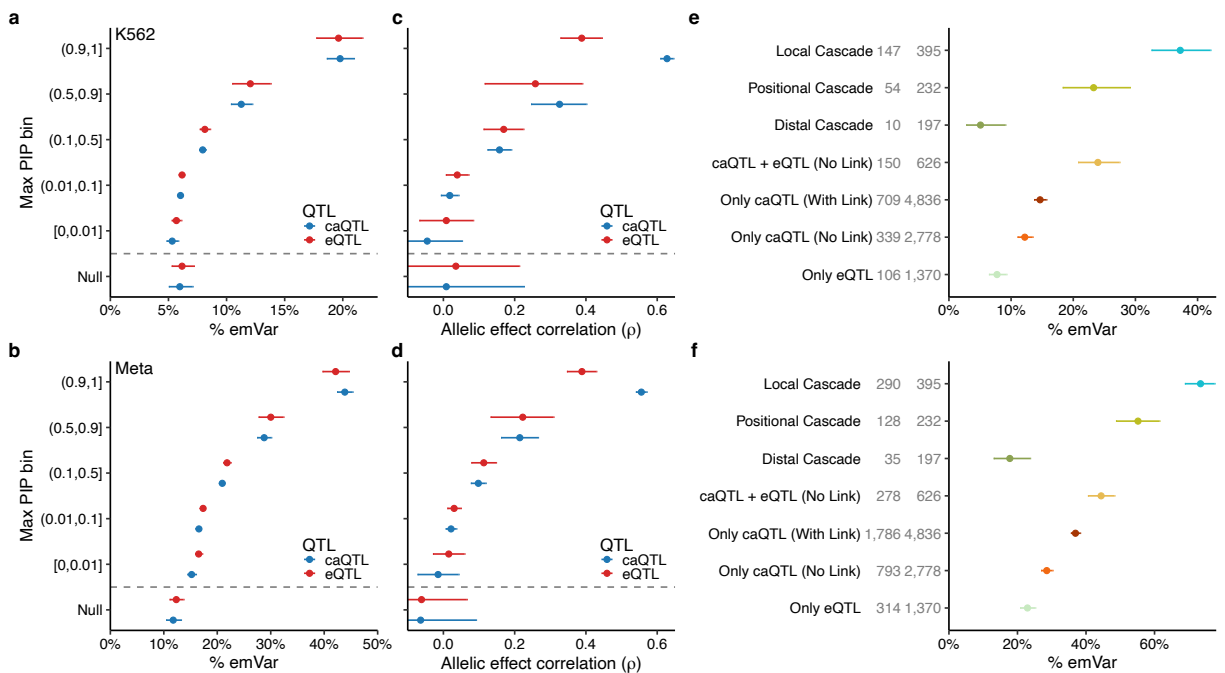
We thank the reviewer for emphasizing the importance of functional validation. Rather than conducting additional MPRA experiments for a small number of variants, we leverage systematic functional data from our recent study (Siraj et al., 2026), which tested the allelic effects of 221,412 fine-mapped variants (GTEx eQTL and UKBB/BBJ GWAS) using MPRA across 5 distinct cell lines, including K562 (blood/myeloid lineage relevant to PBMCs).

We integrated these MPRA results with our CASCADE classifications to provide orthogonal functional validation at scale (new **Extended Data Fig. 6**). Of the 221,412 fine-mapped and 86,064 control variants tested by MPRA, 172,042 were also tested in our molQTL analysis, of which 10,428 are fine-mapped molQTLs (PIP > 0.5). We found that expression-modulating variant (emVar) rates increase monotonically with our molQTL PIP for both eQTL and caQTL variants, from ~5–10% at null (Siraj's control) or low-PIP variants to ~20% (K562) and ~45% (meta-analyzed across cell types) at PIP > 0.9 (new **Extended Data Fig. 6a,b**), confirming that our fine-mapping accurately identifies functional regulatory variants.

Among emVars, the concordance between MPRA allelic effects ($\Delta \log_2$ FC) and endogenous molQTL effects (β_{eQTL} and β_{caQTL}) also scales with PIP (Spearman's $\rho = 0.63$ for caQTL and $\rho = 0.39$ for eQTL at PIP > 0.9 in K562; new **Extended Data Fig. 6c,d**), consistent with prior MPRA-eQTL concordance (Siraj et al., 2026). Notably, caQTL concordance is higher than eQTL concordance at PIP > 0.9, consistent with MPRA and caQTLs both measuring variant effects on *cis*-regulatory element activity, whereas eQTLs reflect the net endogenous impact on gene expression.

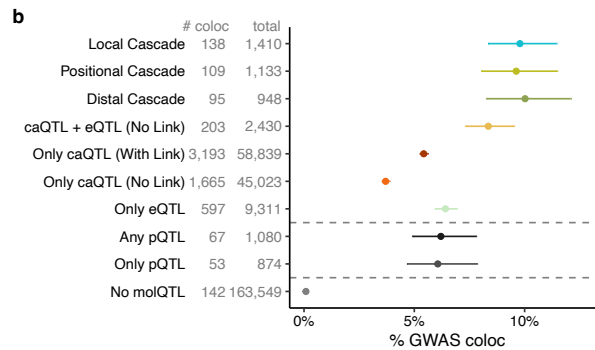
Probing the MPRA signal within cascade variants revealed that our original “Full Cascade” category encompassed two mechanistically distinct subclasses (revised **Fig. 4; Methods**). Variants whose caQTL and peak–gene link originated from the variant’s own overlapping peak ($n = 395$, now named **Local Cascade**) showed markedly elevated MPRA activity (37% emVar in K562, 73% in meta), consistent with direct allele-specific activity on the variant’s own regulatory sequence (new **Extended Data Fig. 6e,f**). In contrast, variants whose caQTL and peak–gene link traced through a non-overlapping distal peak ($n = 197$, now named **Distal Cascade**) showed emVar rates indistinguishable from only eQTL variants (5.1% in K562, 18% in meta vs. 7.7% and 23% for only eQTL), despite showing a complete caQTL–link–eQTL cascade. This depletion is mechanistically expected because our MPRA tests a 200-bp sequence centered on the variant, whereas the causal regulatory element for Distal Cascade variants resides elsewhere. Such variants likely reflect 3D genome contacts bridging the variant’s locus to the distal peak, or regulatory effects driven by distal non-SNV causal variants, such as SVs, not included in our current analysis. We therefore refined our classification to treat Local and Distal Cascade as distinct mechanistic categories, alongside a third Positional Cascade category (variant overlaps a linked peak but lacks a detectable caQTL at that peak; $n = 232$).

Importantly, this MPRA-driven refinement preserved the disease-relevance priority of cascade variants. GWAS colocalization rates were uniform across Local (9.8%), Positional (9.6%), and Distal Cascades (10%), all markedly above non-cascade categories (3.7–8.4%; revised **Fig. 5b**). This orthogonal comparison indicates that the Local–Distal split distinguishes how each variant acts on its own regulatory sequence while preserving which cascade variants are disease-relevant, resolving a mechanistic distinction that GWAS colocalization alone could not.



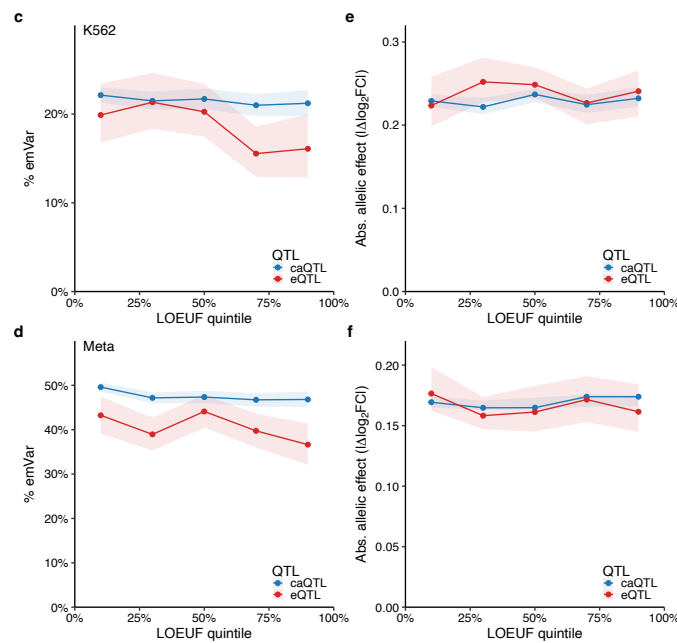
New Extended Data Fig. 6 | Massively parallel reporter assay (MPRA) validation of fine-mapped molQTL variants and CASCADE mechanism categories. **a,b**, Fraction of expression-modulating variants (emVars) among molQTL variants, stratified by maximum PIP bins across cell types and genes for *cis*-eQTL (blue) and *cis*-caQTL (red), in K562 (**a**) and meta-analysis across five cell lines (**b**; **Methods**). The “Null” bin comprises control variants from Siraj et al. (2026). **c,d**, Spearman rank correlation (ρ) between MPRA allelic effects ($\Delta\log_2$ FC) and molQTL effect sizes (β), computed among emVars and stratified by maximum PIP bin and QTL type, in K562 (**c**) and meta-analysis (**d**). **e,f**,

Fraction of emVars across CASCADE mechanism categories among fine-mapped molQTL variants (PIP > 0.5), in K562 (e) and meta-analysis (f). The number of emVars out of total MPRA-tested variants is annotated per category. Error bars represent 95% confidence intervals.



Revised Fig. 5 | Colocalization between molQTL and GWAS variants reveals multi-layered regulatory buffering at constrained genes. **b.** Colocalization rates of fine-mapped molQTL variants (PIP > 0.5), stratified by QTL mechanism categories. “Any pQTL” and “Only pQTL” represent any fine-mapped pQTL (PIP > 0.5) variants and those not fine-mapped as caQTL or eQTL (PIP ≤ 0.5), respectively. Numbers of colocated and total fine-mapped variants for each mechanism category are shown. Error bars represent binomial 95% confidence intervals.

Finally, we tested whether the intrinsic *cis*-regulatory activity measured by MPRA depends on target gene constraint, a key prediction of the regulatory buffering model. Because MPRA measures allelic effects on a reporter construct independently of endogenous chromatin context and gene regulation, it can distinguish whether the attenuation of eQTL effects at constrained genes reflects weaker *cis*-regulatory input or downstream buffering. We found that both the emVar rate and allelic effect magnitude ($|\Delta\log_2 FC|$) of fine-mapped molQTL variants were constant across LOEUF quintiles (new **Extended Data Fig. 10c–f**). This confirms that the intrinsic *cis*-regulatory activity of fine-mapped variants is independent of target gene constraint. The weaker endogenous eQTL effects at constrained genes are therefore attributable to downstream buffering at the chromatin-to-expression interface, not to an absence of regulatory function.



Extended Data Fig. 10 | The buffering gradient is independent of intrinsic *cis*-regulatory activity and cellular stimulation.

c,d, Fraction of emVars among fine-mapped molQTL variants (PIP > 0.5), stratified by the LOEUF quintile of the target gene, in K562 (**c**) and meta-analysis (**d**). **e,f**, Median absolute MPRA allelic effect ($|\Delta\log_2 FC|$) of fine-mapped molQTL variants across LOEUF quintiles of the target gene, K562 (**e**) and meta-analysis (**f**). Shaded ribbons represent 95% confidence intervals of the median.

We have revised the **Results** sections accordingly. For the CASCADE validation:

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 sits 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.

...

To provide orthogonal functional validation, we intersected molQTL variants with our recent massively parallel reporter assay (MPRA) data, which tested the allelic effects of 221,412 fine-mapped variants and 86,064 control variants across 5 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 concordance⁵². Notably, caQTL concordance exceeds eQTL concordance at every PIP level, 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 variant, 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.

For the buffering analysis:

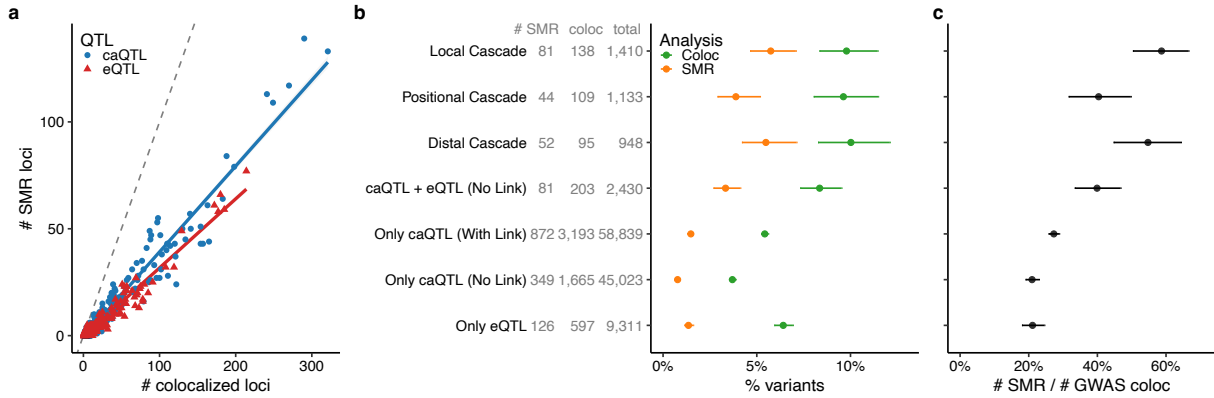
To provide orthogonal evidence for regulatory buffering, we leveraged our recent MPRA data⁵², which measures allelic effects on a reporter construct independently of endogenous chromatin context and gene regulation. If the attenuation of eQTL effects at constrained genes were driven by weaker *cis*-regulatory input, MPRA-measured regulatory activity should also decrease with constraint. Instead, both the emVar rate and allelic effect magnitude ($|\Delta\log_2 FC|$) of fine-mapped molQTL variants were constant across LOEUF quintiles (**Extended Data Fig. 10c-f**). This confirms that the intrinsic *cis*-regulatory activity of fine-mapped variants is independent of target gene constraint. The weaker endogenous eQTL effects at constrained genes are therefore attributable to downstream buffering at the chromatin-to-expression interface, not to an absence of regulatory function.

4. To more robustly establish a causal link from identified molecular QTLs to disease risk, we recommend applying formal causal inference frameworks. A Mendelian Randomization analysis, particularly a two-sample MR approach leveraging the cohort's rich GWAS data, would be highly suitable. For instance, using genetic variants associated with a specific chromatin accessibility peak or gene expression level as instrumental variables could test whether these molecular traits directly influence disease endpoints. This would significantly strengthen the mechanistic interpretation of the reported associations.

We thank the reviewer for this suggestion. We have now performed Summary-based Mendelian Randomization (SMR) analysis (Zhu et al., 2016) to formally test causal relationships between molecular QTLs and disease outcomes. 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. We applied SMR across 11,584 colocating GWAS loci, 3,868 colocating genes, and 23,025 colocating peaks, testing for 153,395 colocating trios across cell types (cell type–GWAS locus–molQTL locus; **Methods**).

Of 33,443 significant SMR associations (Bonferroni $P_{SMR} < 3.3 \times 10^{-7}$ and $P_{HEIDI} > 0.05$), 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 (new **Supplementary Fig. 11a**). 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 (new **Supplementary Fig. 11b,c**).

The higher caQTL colocalization rate is consistent with the regulatory buffering architecture we describe in the manuscript. caQTLs operate upstream of the chromatin-to-expression buffering checkpoint and capture disease-relevant genetic variation that is progressively attenuated before reaching gene expression. SMR confirms that these caQTLs are nonetheless causally linked to disease risk.



New 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 represent fitted slopes 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**). Error bars represent 95% confidence intervals. **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.

We have revised the **Results** section as follows:

To independently validate these colocalization results for causality, we performed summary-based Mendelian randomization (SMR)⁶⁰ across 11,584 colocating GWAS loci, 3,868 colocating genes, and 23,025 colocating peaks, testing for 153,395 colocating 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**). 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.

5. The study provides an exceptional map of molecular mechanisms, yet it does not fully explore the translational potential inherent in its rich single-cell data, a dimension that has been effectively demonstrated in related work such as (Perez et al., Science, 2022). Bridging this gap between foundational discovery and clinical insight represents a key opportunity. We recommend that the authors integrate the basic clinical phenotypes already available for this cohort, which could include common laboratory values like CRP or lymphocyte counts, or medication records. This integration would enable a proof-of-concept analysis to test whether single-cell-derived features, for example immune cell subset proportions or pathway activity scores for key pathways like the interferon-stimulated gene response, can predict clinical parameters or define molecular subtypes among donors. Such an analysis would significantly strengthen the manuscript because it would allow a discussion on the tangible utility of these molecular maps for patient stratification in autoimmune diseases, thereby enhancing the translational relevance of the work.

We thank the reviewer for highlighting the translational potential of the atlas. A comprehensive clinical integration, including longitudinal registry endpoints, laboratory measurements, and medical records, is beyond the scope of this revision and will be addressed in a dedicated translational study. As a proof of concept, we used 40 TCAT T-cell program scores (see our response to **Comment 2**; aggregated per donor), which summarize pathway-level activity across co-regulated genes, and tested their associations with clinical phenotypes in FinnGen (autoimmune, metabolic, and cardiovascular diseases, blood cell counts, laboratory tests) and 11 external polygenic scores (PGS). Among 35 significant associations ($FDR < 0.05$), we have found the following observations (**Response Fig. 4**):

[REDACTED]

These analyses demonstrate how our atlas can bridge molecular mechanisms to clinical phenotypes beyond QTL study in healthy blood donors, which by definition is restricted for rare autoimmune conditions at the time of enrollment. A systematic investigation on this matter is the subject of an upcoming study.

[REDACTED]

Minor issues :

6. The description of the two sample processing methods, multiplexed sample channels versus single-donor channels, in Extended Data Fig. 1a and the Methods (lines 1181–1183) is overly brief and raises methodological concerns. Typically, cryopreservation is required only when samples are to be pooled. It is unclear whether the multiplexed channels used cryopreserved samples while the single-donor channels used fresh samples. If so, the authors should explicitly state this and, more importantly, assess whether these different processing protocols introduced batch effects or technical bias. Any such bias must be accounted for in downstream analyses to ensure the robustness and reproducibility of the reported molecular findings.

We thank the reviewer for raising this point. To clarify, all samples in our study were cryopreserved, as they were shipped from Helsinki to Boston as frozen samples. Both single-donor and multiplexed channels used cryopreserved PBMCs. We have revised the **Methods** and **Extended Data Fig. 1a** to clarify this.

a

[REDACTED]

Because all samples underwent the same cryopreservation protocol, the processing differences between single-donor and multiplexed channels are limited to pooling strategy, minimizing potential for systematic bias between these approaches. We further assessed batch effects in our data using variance-decomposition and integration metrics, as detailed in our response to **Reviewer 4, Comment 2**. PEER factor adjustment for eQTLs and chromatin accessibility PC adjustment for caQTLs reduced batch-related pseudobulk variance to near zero across multiome channels, flow cells, and sequencers (new **Supplementary Fig. 23**), and integration Local Inverse Simpson's Index (iLISI) and cell-type LISI (cLISI) statistics (Korsunsky et al., 2019) confirmed appropriate batch mixing with near-perfect cell-type preservation (median cLISI = 1.01 for snRNA-seq and 1.02 for snATAC-seq, on a 1–8 scale across L1 cell types; new **Supplementary Fig. 24**).

To clarify these points, we have revised the **Methods** section:

Paired, single-nucleus gene expression and ATAC profiles were generated from cryopreserved PBMCs across both single-donor and multiplexed channels using the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression assay (10x Genomics, PN-1000283).

7. The data availability statement must be detailed and actionable. Beyond a query portal for variant lookup, we strongly encourage the public deposition of the complete single-cell data matrices, including the cell-by-gene expression and cell-by-peak accessibility tables. A clear release timeline and the designated public repository should be specified.

We appreciate the reviewer's emphasis on data availability. Under FinnGen's data governance policies, individual-level data cannot be publicly deposited and are instead available through controlled access procedures. Single-cell count matrices constitute individual-level data, and recent work has further demonstrated their vulnerability to linking attacks using eQTLs, enabling individual re-identification with accuracy exceeding 90% as exemplified using the publicly released OneK1K dataset (Walker et al., 2024). For these reasons, we have elected not to publicly deposit complete single-cell count matrices.

Instead, we provide comprehensive summary-level data, including complete caQTL and eQTL summary statistics, fine-mapping results, and peak–gene link statistics, which will be publicly available upon publication. We have revised the **Data Availability** statement to clarify access procedures for individual-level data.

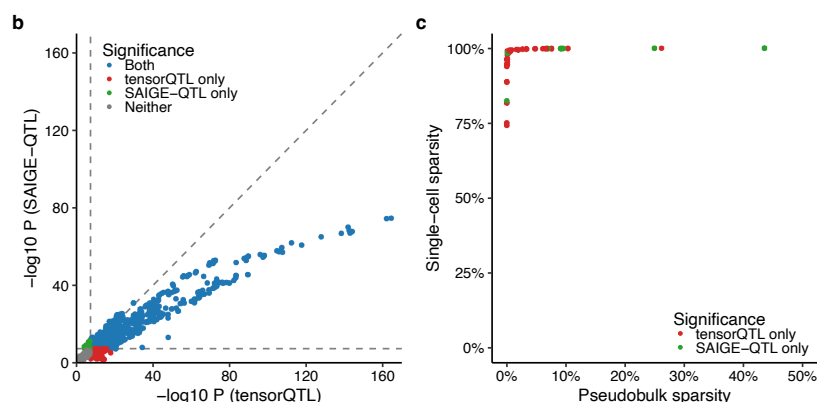
FinnGen as a research project is granted use of national healthcare data and biospecimens according to national and European regulations (GDPR), which preclude the research project from distributing individual level data. The FinnGen single-nucleus caQTL and eQTL summary statistics are currently subject to embargo according to the FinnGen consortium agreement. They are available upon request and being prepared for public release by Q4 2026 or at the time of publication. Individual-level data, including single-cell count matrices, are available to approved researchers through FinnGen's data access procedures (https://www.finnngen.fi/en/access_results). Academic users wishing to work with the FinnGen project resource directly can follow the procedures described at: <https://www.finnngen.fi/en/how-we-collaborate>.

8. We have tested the provided links from multiple locations. Our findings are consistent: while the main GitHub repository is accessible, the specific code links cited in the manuscript are currently inaccessible. This issue directly hinders the assessment of reproducibility for the analytical workflows. As a minimum requirement for the revision, all relevant code must be made functional and its access links verified to be publicly available.

We apologize for this oversight. All code repositories cited in the manuscript are now publicly available.

9. The analysis in lines 176 to 184 uses two different methods SAIGE QTL for eQTL mapping and tensorQTL for caQTL mapping. The authors need to justify the use of different statistical frameworks for different molecular layers to ensure methodological consistency is clear.

We address this concern in detail in our response to **Reviewer 1, Comment 4 (Response Fig. 1b,c)**. Briefly, snATAC-seq data exhibits substantially greater sparsity than snRNA-seq, and pseudobulk aggregation addresses this by pooling fragments across cells within each donor. This pseudobulk tensorQTL approach is consistent with other large-scale single-nucleus caQTL studies, including TenK10K (Xue et al., medRxiv). Additionally, the computational burden of testing ~300,000 peaks (compared to ~30,000 genes) made single-cell approaches impractical.



Response Fig. 1 | Sensitivity analyses of eQTL covariate choice and caQTL mapping method in CD4⁺ T cells.
b. Comparison of *cis*-caQTL $-\log_{10} P$ -values in CD4⁺ T cells when using pseudobulk tensorQTL (x-axis) vs. single-cell SAIGE-QTL (y-axis), across 200 randomly selected chromatin accessibility peaks stratified by chromatin accessibility non-zero proportion decile. **c.** Relationship between pseudobulk sparsity (1 – fraction of donors with a nonzero pseudobulk count; x-axis) and single-cell sparsity (1 – fraction of cells with a nonzero fragment count; y-axis) for the variant–peak pairs from (b) that reach genome-wide significance in only one of the two methods.

We have clarified this rationale in the revised **Methods** section.

We used the sum of accessibility counts to address the inherent sparsity of snATAC-seq data, where most nuclei show zero counts at a given peak. The substantially greater sparsity of snATAC-seq compared to snRNA-seq at the single-cell level, together with the computational cost of testing ~300,000 peaks, made pseudobulk tensorQTL the appropriate choice in this study over single-cell SAIGE-QTL.

10. The definition of the likely shared but underpowered category in lines 279 to 285 is conceptually important but its distinction from truly shared effects is vague. A clearer explanation of the classification threshold and its implications is needed.

We thank the reviewer for raising this point. The classification thresholds and their implications are fully specified in the **Methods**. Briefly, the “likely shared but underpowered” category captures effects that reach significance ($FDR < 0.05$) in one lineage or PBMC but show low local false sign rate ($LFSR < 0.05$ by mash) in other lineages where they do not reach nominal significance. Low LFSR indicates high posterior confidence that the effect has the same direction, suggesting that the effect is biologically present but statistically underpowered in those cell types. In contrast, “cross-lineage shared” category captures effects that reach FDR significance in both lineages, while “lineage-specific” category captures effects showing high $LFSR (\geq 0.5)$ in the non-significant lineage, indicating true absence rather than limited power.

To clarify the threshold, we have revised the **Results** section, with a pointer to the **Methods**.

We classified eGenes and caPeaks into five categories: 1) cross-lineage shared, with significant effects across myeloid and lymphoid lineages; 2) likely shared but underpowered, significant in one lineage (or only PBMC) but with low local false sign rate ($LFSR < 0.05$) across lineages; 3) lineage-specific, restricted to either myeloid or lymphoid cell types; 4) T-cell-specific; and 5) single cell-type specific (**Fig. 3a, Methods**).

11. The classification of variants into 25 patterns in lines 356 to 361 is highly complex. The authors should provide a more accessible summary or a conceptual diagram in the main text or somewhere to aid reader comprehension.

We appreciate this feedback. The 25 patterns are systematically derived from combinations of four features (peak overlap, caQTL status, peak–gene link, and eQTL status), which we collapse into seven interpretable mechanistic categories for presentation. We believe the existing figures already address the reviewer’s concern: **Fig. 4a** provides a conceptual schematic of the classification workflow, and **Extended Data Fig. 5a** presents the complete 25-pattern classification table with color-coded annotations. The **Methods** section provides detailed definitions for each category.

Referee #2 (Remarks on code availability):

Several key GitHub links provided by the authors are currently inaccessible, although their parent directories can be accessed:

<https://github.com/FINNGEN/singlecell-pipeline>

<https://github.com/mkanai/finngen-multiome-flagship>

<https://github.com/mkanai/cascade>

All the other GitHub links can be opened and display complete README files, in which the author thoroughly explains the installation and usage instructions for each software/pipeline, as well as listing the external software they depend on (with links provided to those dependencies).

We apologize for this oversight. All code repositories are now publicly available with detailed README files.

Referee #4:

Review: Population-scale multiome immune cell atlas reveals complex disease drivers

In this work, Kanai et al. present an admirably large scale multiomic single cell analysis of Finnish peripheral blood samples. The scale and scope of the work with over ~1000 individuals and 7-10M analyzed cells is truly commendable and I can’t begin to fathom the time and resources required to conduct this effort. Congratulations to all the authors. This work falls into the pantheon of large published (1K1K) and ongoing (10K10K) single cell studies but adds great depth with multiomic analysis. I have no doubt that the data generated, if appropriately shared in raw data format (if possible, given the Finngen restrictions) will be widely used and highly cited. However, the biological implications are harder to determine and some major, but addressable, criticisms are outlined below related to the reporting and analysis. The major finding, pertaining, to the large heritability of caQTLs and the large discrepancies between eQTLs is certainly of great interest and has been reported, in less detail, by others.

We thank the reviewer for this positive assessment of our work and for recognizing the scale, scope, and contribution of our multiome atlas to the field. We appreciate the acknowledgement that our findings on caQTL-mediated heritability and regulatory buffering add depth to previous studies. We have carefully reviewed and addressed each point, as detailed in our responses below.

Major comments:

1. There is a large and unexplained discrepancy between the total number cells in the single cell RNA (~10M) and ATAC(~7M) analysis that needs to be clearly addressed. The authors do a great job of identifying eQTLs, caQTLs, and linking peaks to genes. However, it is unclear in the main figures and manuscripts what final dataset was used for this analysis. Extended data figure 1 suggests that these were truly separate analyses. Given that this analysis was performed on multiome samples and single cell barcodes from RNA used as an additional measure for identifying quality cells in ATAC data, it is unclear why there is such a large difference (30%) in the total number of cells. Furthermore, a joint ATAC, RNA embedding is never shown. If eQTLs were called on just the snRNAseq data rather than the joint embedding, why was this done and do the authors think this is appropriate considering that the relationships between eQTLs and caQTLs and peak-gene is further characterized. Some explanation and clarification is warranted. I imagine for instance, that certain cell types in some individuals are not well represented in the caQTL data but are well represented in the eQTL data creating discrepancies. Extended data figure 2c shows total nuclei but not by individual which is more important for genetic associations.

We thank the reviewer for raising this point. The difference in nucleus numbers between snRNA-seq (~10.6M) and snATAC-seq (~7.1M) reflects both technical and QC factors. As described in the **Methods**, snATAC-seq data was subset to barcodes retained in the snRNA-seq atlas, ensuring all snATAC-seq nuclei have paired snRNA-seq data. First, ~1.0 M nuclei yielded no snATAC-seq output despite successful snRNA-seq capture, which is a known limitation of multiome assays where chromatin accessibility and transcriptome capture efficiency can differ. Second, additional ATAC-specific QC filters excluded ~2.5 M nuclei that passed snRNA-seq QC but had suboptimal snATAC-seq profiles.

We did not use joint embeddings. snRNA-seq and snATAC-seq were processed separately, with cell type annotations transferred from the snRNA-seq reference. Regarding the reviewer's concern about cell type representation discrepancies, eQTL and caQTL analyses are conducted independently within each modality, so differences in cell type representation do not create analytical discrepancies. Cross-modality analyses, such as peak-gene link analysis, were performed using the 7.1 M nuclei with matched snRNA-seq and snATAC-seq data.

Per-individual nuclei counts for both modalities are shown in **Fig. 1d**.

To clarify these points, we have revised the **Methods**:

Per-channel snATAC-seq peak-barcode matrices were first subset to barcodes retained in the snRNA-seq atlas, leveraging established empty droplet and doublet prediction approaches to ensure all snATAC-seq nuclei had paired snRNA-seq data. Of the snRNA-seq QC-passing barcodes, ~1.0M had no detectable ATAC signal (fewer than 200 peak fragments, or absent from the Cell Ranger ARC outputs), reflecting the differential capture efficiency of chromatin accessibility vs. transcriptome in the Multiome assay. For peak calling, we computed standard QC metrics for each sample using per-sample peaks called by Cell Ranger ARC and Signac¹⁰⁵ v1.14.0, and excluded a further ~2.5 M nuclei not meeting all of the following criteria: 2,000 < the number of fragments in peaks < 100,000, the ratio of reads in the ENCODE's genomic blacklist regions < 0.01, nucleosome signal < 2, and TSS enrichment score > 4.

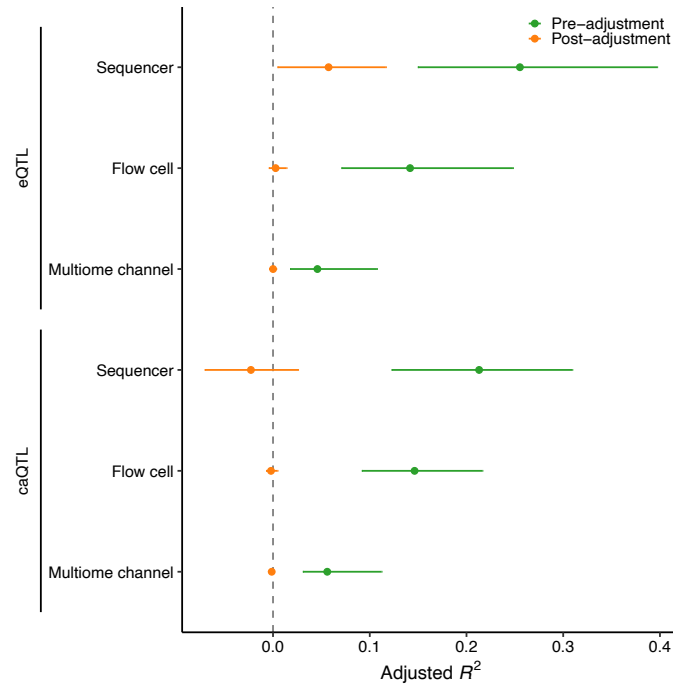
2. The effects of batch effects on analysis are unclear. It is unclear how batch was handled in both the generation and analysis of the single cell data. I am particularly interested to know why multiomic run or processing date was not included as a co-variate in the identification of pseudobulk or single cell eQTLs. Is the assumption that PEER factors explain this batch effect without the need for further correction and if so, can a citation/justification be provided.

We thank the reviewer for this question. After quality control, our study retains 718 multiome channels, sequenced across 31 flow cells on seven sequencers. The large number of channels precludes their inclusion as categorical covariates in the regression models. Because SAIGE-QTL requires raw count data as input, we also cannot apply batch-corrected embeddings (e.g., Harmony) directly for QTL mapping.

We instead rely on PEER factors (for eQTLs) and expression PCs (for caQTLs) computed on pseudobulk data to capture latent technical variation, including batch effects. This follows established practice in large-scale eQTL or caQTL studies (The GTEx Consortium, 2020; Yazar, et al., 2022; Zhou, et al., *medRxiv*; Cuomo, et al., *medRxiv*; Xue, et al., *medRxiv*).

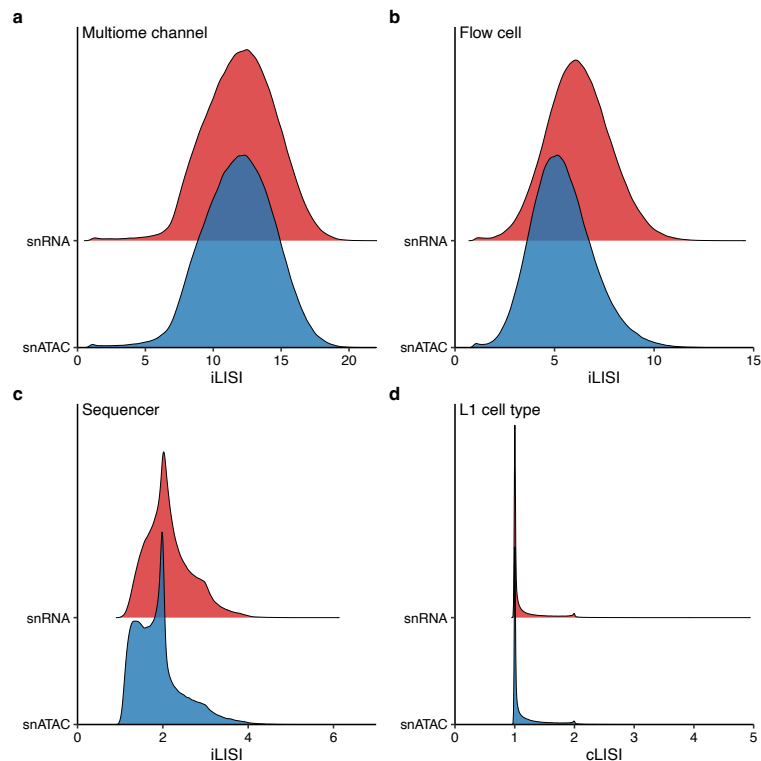
To verify that PEER factors capture batch variation in our own data, we quantified the fraction of pseudobulk expression variance (adjusted R^2) explained by batch identifiers (multiome channel, flow cell, and sequencer) before and after PEER adjustment at the PBMC level, using donor-level composition over each identifier as covariates to account for donors spanning multiple batches. Adjusted R^2 dropped substantially at every granularity after PEER adjustment, from 26% to 5.7% for multiome channel, from 14% to 0.3% for flow cell, and from 4.6% to 0.0% for sequencer (new **Supplementary Fig. 23**), demonstrating that PEER factors absorb the bulk of batch-related variation.

To further assess batch mixing at the single-cell level, we computed integration Local Inverse Simpson's Index (iLISI) and cell-type LISI (cLISI) statistics (Korsunsky, et al., 2019) for multiple batch identifiers on our integrated snRNA-seq and snATAC-seq atlases. We observed near-perfect cell-type preservation (median cLISI = 1.01 for snRNA-seq and 1.02 for snATAC-seq, on a 1–8 scale across L1 cell types) and iLISI values concordant between the two atlases across all batch identifiers, consistent with expected mixing under the observed batch-size distribution (new **Supplementary Fig. 24**). For example, median sequencer-level iLISI was 2.1 (snRNA-seq) and 1.8 (snATAC-seq) across seven sequencers; residual fine-grained structure at the channel level (median iLISI = 12.1 and 11.9 across 718 channels) is absorbed by PEER factors in the QTL models, as shown above.



New Supplementary Fig. 23 | 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.



New Supplementary Fig. 24 | 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.

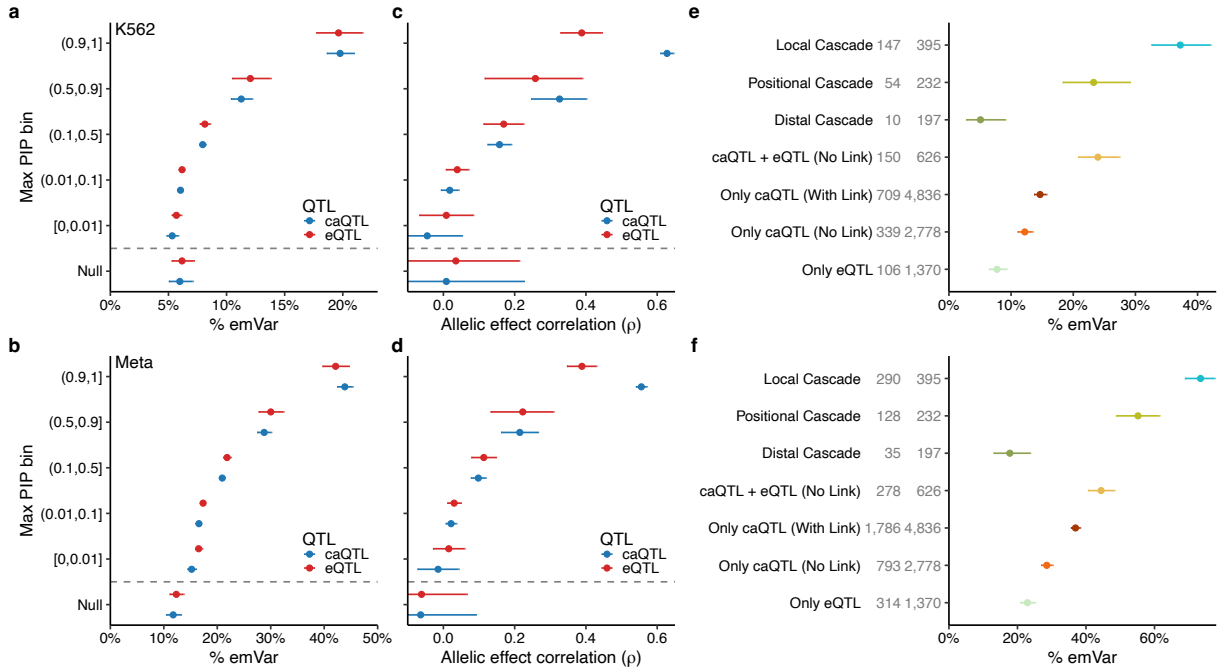
To clarify these points, we have revised the **Methods** section:

Sequencing batch identifiers (multiome channel, flow cell, sequencer) were not included as categorical covariates given their large number (718 channels across 31 flow cells and 7 sequencers after QC) and the SAIGE-QTL requirement for raw count input, which precludes batch-corrected embeddings. Latent technical variation, including batch effects, was instead captured by PEER factors, consistent with previous large-scale studies^{6,16,114}. To further confirm that batch effects were adequately controlled, we quantified the fraction of pseudobulk variance attributable to each batch identifier before and after adjustment, and computed integration LISI (iLISI) and cell-type LISI (cLISI) statistics¹¹⁵ on the integrated snRNA-seq and snATAC-seq atlases (**Supplementary Figs. 23,24**).

3. CASCADE as a method is not properly discussed. Given the importance of CASCADE for annotation of molQTLs, it warrants additional space in the manuscript. My understanding of this algorithm is that it classifies molQTLs into various categories based on intersection of genetic effects and cell type. It is unclear how much validation is required for this classification. Can the authors comment further?

We thank the reviewer for this comment. CASCADE is a classification framework that categorizes fine-mapped molQTL variants into mechanistic classes based on observed features: chromatin accessibility peak overlap, caQTL status, peak–gene links, and eQTL status.

The validity of CASCADE classifications rests on the robustness of the underlying data layers (caQTLs, eQTLs, and peak–gene links), which we have validated through cross-ancestry replication, comparison with external datasets, and MPRA functional validation (see our responses to **Reviewer 2, Comments 1 and 3**).



New Extended Data Fig. 6 | Massively parallel reporter assay (MPRA) validation of fine-mapped molQTL variants and CASCADE mechanism categories.

a,b, Fraction of expression-modulating variants (emVars) among molQTL variants, stratified by maximum PIP bins across cell types and genes for *cis*-eQTL (blue) and *cis*-caQTL (red), in K562 (**a**) and meta-analysis across five cell lines (**b**; **Methods**). The “Null” bin comprises control variants from Siraj et al. (2026). **c,d**, Spearman rank correlation (ρ) between MPRA allelic effects ($\Delta\log_2$ FC) and molQTL effect sizes (β), computed among emVars and stratified by maximum PIP bin and QTL type, in K562 (**c**) and meta-analysis (**d**). **e,f**, Fraction of emVars across CASCADE mechanism categories among fine-mapped molQTL variants (PIP > 0.5), in K562 (**e**) and meta-analysis (**f**). The number of emVars out of total MPRA-tested variants is annotated per category. Error bars represent 95% confidence intervals.

We have expanded the description of CASCADE in the **Results** section (see our response to **Reviewer 1, Minor comment 8**) to better introduce the motivation and framework. Additionally, the Methods section provides detailed definitions of all classification categories, and **Extended Data Fig. 5a** presents the complete 25-pattern classification scheme.

a	Pattern	CA peak overlap	caQTL	Peak-gene link	eQTL	QTL mechanism
1	Peak overlap	For overlapping peak	For overlapping peak	For caQTL peak	For linked gene	Local cascade
2	Peak overlap	For non-overlapping peak	For non-overlapping peak	For overlapping peak	For linked gene	Positional cascade
3	Peak overlap	No detected caQTL	No detected caQTL	For overlapping peak	For linked gene	Positional cascade
4	Peak overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	For linked gene	Positional cascade
5	No overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	For linked gene	Distal cascade
6	Peak overlap	For overlapping peak	For overlapping peak	For caQTL peak	For non-linked gene	caQTL + eQTL (No link)
7	Peak overlap	For non-overlapping peak	For non-overlapping peak	For overlapping peak	For non-linked gene	caQTL + eQTL (No link)
8	Peak overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	For non-linked gene	caQTL + eQTL (No link)
9	No overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	For non-linked gene	caQTL + eQTL (No link)
10	Peak overlap	For overlapping peak	For overlapping peak	No link	For non-linked gene	caQTL + eQTL (No link)
11	Peak overlap	For non-overlapping peak	For non-overlapping peak	No link	For non-linked gene	caQTL + eQTL (No link)
12	No overlap	For non-overlapping peak	For non-overlapping peak	No link	For non-linked gene	caQTL + eQTL (No link)
13	Peak overlap	For overlapping peak	For overlapping peak	For caQTL peak	No detected eQTL	Only caQTL (With link)
14	Peak overlap	For non-overlapping peak	For non-overlapping peak	For overlapping peak	No detected eQTL	Only caQTL (With link)
15	Peak overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	No detected eQTL	Only caQTL (With link)
16	No overlap	For non-overlapping peak	For non-overlapping peak	For caQTL peak	No detected eQTL	Only caQTL (With link)
17	Peak overlap	For overlapping peak	For overlapping peak	No link	No detected eQTL	Only caQTL (No link)
18	Peak overlap	For non-overlapping peak	For non-overlapping peak	No link	No detected eQTL	Only caQTL (No link)
19	No overlap	For non-overlapping peak	For non-overlapping peak	No link	No detected eQTL	Only caQTL (No link)
20	Peak overlap	No detected caQTL	No detected caQTL	For overlapping peak	For non-linked gene	Only eQTL
21	Peak overlap	No detected caQTL	No detected caQTL	No link	For non-linked gene	Only eQTL
22	No overlap	No detected caQTL	No detected caQTL	No link	For non-linked gene	Only eQTL
23	Peak overlap	No detected caQTL	No detected caQTL	For overlapping peak	No detected eQTL	No molQTL
24	Peak overlap	No detected caQTL	No detected caQTL	No link	No detected eQTL	No molQTL
25	No overlap	No detected caQTL	No detected caQTL	No link	No detected eQTL	No molQTL

Extended Data Fig. 5 | CASCADE classification framework.

a, Classification scheme showing 25 mutually exclusive patterns based on i) chromatin accessibility (CA) peak overlap, ii) caQTL, iii) peak-gene links, and iv) eQTL, grouped into seven mechanistic categories (see **Methods**).

4. No functional validation is performed. In multiple places in the manuscript and abstract(!) base editing validation is mentioned. This data is not shown in the manuscript and all mentions should be removed. The data cited is from a pre-print that this reviewer did not receive and has not examined. I want to say clearly, that I do not believe functional validation is necessarily required for this work as primary a large-resource and genetics study. I full appreciate the value of functional validation, and methods like STAGseq and other published single cell approaches bridging the V2F gap. Mention of these in the discussion is certainly warranted with the caveat below:

a. Caveat: The authors discuss at length and show ample evidence in support of the finding that caQTL are more valuable (heritable with increased enrichment) than eQTLs. However, all proposed validation (which again, I think should be removed) is focused on discovery of eQTLs. Recent published methods are available that look at chromatin accessibility with CRISPR editing if this is the focus and one of the major findings of the manuscript.

We thank the reviewer for this comment and for acknowledging that functional validation is not necessarily required for this study. We have revised the framing throughout the manuscript and abstract to make clear that STAG-seq (Liu, et al., *bioRxiv*) is a separate recent study from our group, cited to contextualize our QTL findings and to corroborate the functional relevance of a CASCADE-nominated variant, such as *TNRC18*.

We acknowledge that STAG-seq validates transcriptomic consequences of base-editing rather than chromatin accessibility changes directly. While recent methods (e.g., TDAC-seq, Roh et al, 2025; UNIChro-seq, Kono et al., *medRxiv*) enable direct chromatin accessibility validation after base-editing, we agree with the reviewer that such validation is beyond the scope of this paper.

We note, however, that precise transcriptomic readout remains important even for validating caQTL-only variants. Our study underscores the value of having both caQTL and eQTL evidence for GWAS colocalization, but many caQTL effects may not manifest as detectable eQTLs in steady-state PBMCs. Base-editing in a specific cell type under specific conditions can reveal

eQTL effects that are otherwise undetectable, making transcriptomic validation relevant for caQTL variants as well.

We have revised the manuscript to clarify the relationship between the two papers and ensure that references to STAG-seq validation are appropriately framed as results from our recent study rather than data presented in this manuscript.

In our recent paper⁹⁰, we developed a novel base-editing and phenotyping platform called Sensitive Transcriptomics And Genotyping by sequencing (STAG-seq),

5. No code was shared to examine the analysis. I believe this is a requirement. No code was shared. Statements in the manuscript and reporting summary point to github links that are non-functional.

We apologize for this oversight. All code repositories cited in the manuscript are now publicly available.

A non-criticism for the authors. This work is massive and I truly appreciate that. The study was performed in PBMCs, which some may view as a criticism. I do not. I appreciate the effort that went in to highlighting examples that may apply to tissues (microglia) but I think this is unnecessary and takes away from the arguments in the paper. There is no clarity on why caQTLs account for much higher heritability and why no links to genes can be found. One discussed reason is that the wrong tissue or cell state (stimulation condition) is being assayed. This makes some biological sense to me. If a T cell is primed to respond to multiple factors given its genetics, then only regulates expression in distinct context that may explain the discrepancy. The discussion on buffering alludes to this.

We thank the reviewer for this thoughtful perspective and for appreciating the scope of our study. We agree that the PBMC focus is strength rather than a limitation, providing a tractable system for population-scale multiome profiling while capturing immune cells relevant to many diseases.

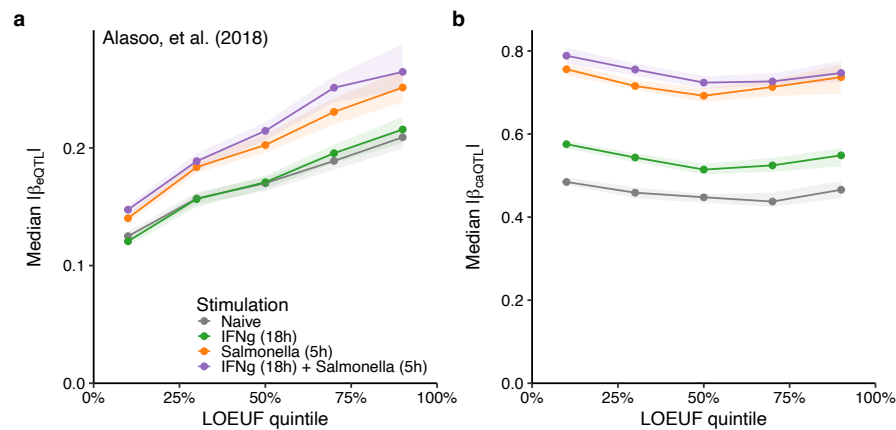
Regarding the microglia example, we have retained this locus because the integrated multiome surfaces a convergence that our framework is specifically designed to capture. A common local cascade variant at *PILRB* and the Finnish-enriched *TYROBP* deletion from the same cohort both attenuate the DAP12-coupled microglial sensome (Haure-Mirande et al., 2022) in monocytes. We have revised the text to make this convergence clear and removed the previous appeal to shared developmental origin, which we agree was distracting and not the strongest framing of the example.

PILRB and *TYROBP*/DAP12 are paired components of the DAP12-coupled microglial sensome implicated in AD pathology⁷², and the convergence of two distinct AD risk variants at this locus, a common local cascade variant and a Finnish-enriched structural deletion, on this shared signaling axis points to DAP12 attenuation as a likely mechanism for impaired amyloid- β clearance and neuroinflammation.

We appreciate the reviewer's biological intuition regarding the "priming" model, *i.e.*, genetic variants may alter chromatin accessibility in steady-state PBMCs, preparing cells for context-dependent gene expression responses that only manifest under specific stimulation conditions.

This interpretation aligns with our regulatory buffering model, where caQTLs capture regulatory potential that is not always transmitted to detectable eQTLs in baseline conditions.

To empirically assess the alignment between the “priming” model and our regulatory buffering framework, we leveraged publicly available paired naive/stimulated molQTL datasets as a proof-of-concept (see our response to **Reviewer 2, Comment 2** for full details). Using caQTLs and eQTLs from Alasoo et al. (2018), pathogen stimulation uniformly increased *cis*-caQTL effect sizes regardless of target gene constraint (Spearman’s $\rho = -0.05$ to -0.06 ; LOEUF x condition interaction $P = 0.75$), whereas *cis*-eQTL effect sizes increased disproportionately at unconstrained genes (Spearman’s $\rho = 0.23$ in naive to 0.26 in IFN γ + Salmonella; interaction $P = 2.4 \times 10^{-3}$). This indicates that enhancer priming and regulatory buffering operate at distinct regulatory layers. Stimulation increases caQTL effects regardless 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.



Extended Data Fig. 10 | The buffering gradient is independent of intrinsic *cis*-regulatory activity and cellular stimulation.

a, Median absolute *cis*-eQTL effect size ($|\beta_{eQTL}|$) across LOEUF quintiles in iPSC-derived macrophages from Alasoo et al. (2018), stratified by stimulation conditions (**Methods**). Shaded ribbons represent 95% confidence intervals. **b**, Median maximum absolute *cis*-caQTL effect size ($|\beta_{caQTL}|$) per gene across LOEUF quintiles, stratified by stimulation condition.

Complementary analyses in our response to **Reviewer 2, Comment 2** (eQTL Catalogue overlap with stimulus-responsive eQTL variants; dynamic *cis*-eQTL mapping with SAIGE-QTL-Dynamic) further quantify the context-dependent regulatory architecture captured within our steady-state atlas.

To clarify these points, we have revised the **Results**:

To assess whether this buffering architecture is specific to steady-state conditions, we analyzed paired naive/stimulated caQTLs and eQTLs from Alasoo et al. (macrophages, 84 donors, 4 conditions). 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 x condition interaction $P = 2.4 \times 10^{-3}$), 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-priming and

regulatory buffering operate at distinct regulatory layers. Stimulation increases caQTL effects regardless 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.

Minor comments:

- the title on the submitted manuscript does not match the title on the online submission (an additional 1 has been randomly placed)

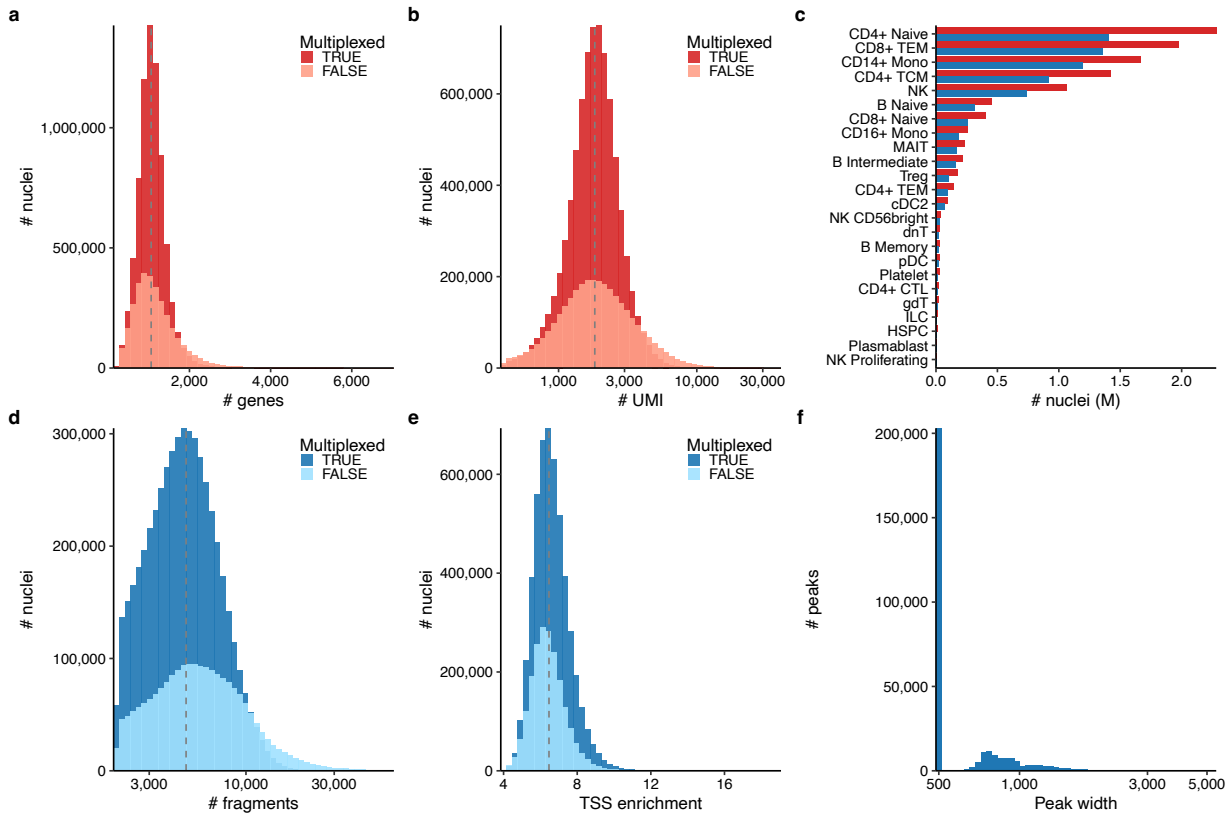
We thank the reviewer for catching this typo on the submission system. We have ensured the title matches between the manuscript and the online system.

- Having read the manuscript, it's unclear to me what the title is trying to convey. What are these complex disease drivers, I would recommend changing it.

We have revised the title to the following to more clearly convey the study's focus: **Population-scale immune multiome atlas reveals regulatory disease mechanisms**

- There is a very large imbalance in the single cell depth of the first 339 donors versus the following ~800. Was any analysis performed between these two groups to show that no differences in power were noted. I.e., the eQTLs identified (particularly the single cell eQTLs) in the first batch are all replicated in the second (perhaps down-sampled). This is a minor comment.

We acknowledge the potential imbalance between the first 339 donors profiled in single channels vs. the following 832 multiplexed donors. Empirically, per-nucleus depth is comparable between the two groups (medians for single-donor vs. multiplexed: snRNA counts 1,846 vs. 1,824; genes detected 1,070 vs. 1,053; snATAC counts 5,455 vs. 4,543; TSS enrichment 6.34 vs. 6.52), and the median number of nuclei per donor is slightly higher in the multiplexed group (snRNA: 8,933 vs. 9,571; snATAC: 6,432 vs. 6,731; **Supplementary Fig. 1, Supplementary Table 1**). Batch-related pseudobulk variance across multiome channels, flow cells, and sequencers was also reduced to near zero after PEER factor or chromatin accessibility PC adjustment (see our response to **Comment 2**). Moreover, because SAIGE-QTL directly models cell-level RNA counts with log-transformed total RNA counts as covariates, our eQTL mapping is robust to this imbalance, if any.

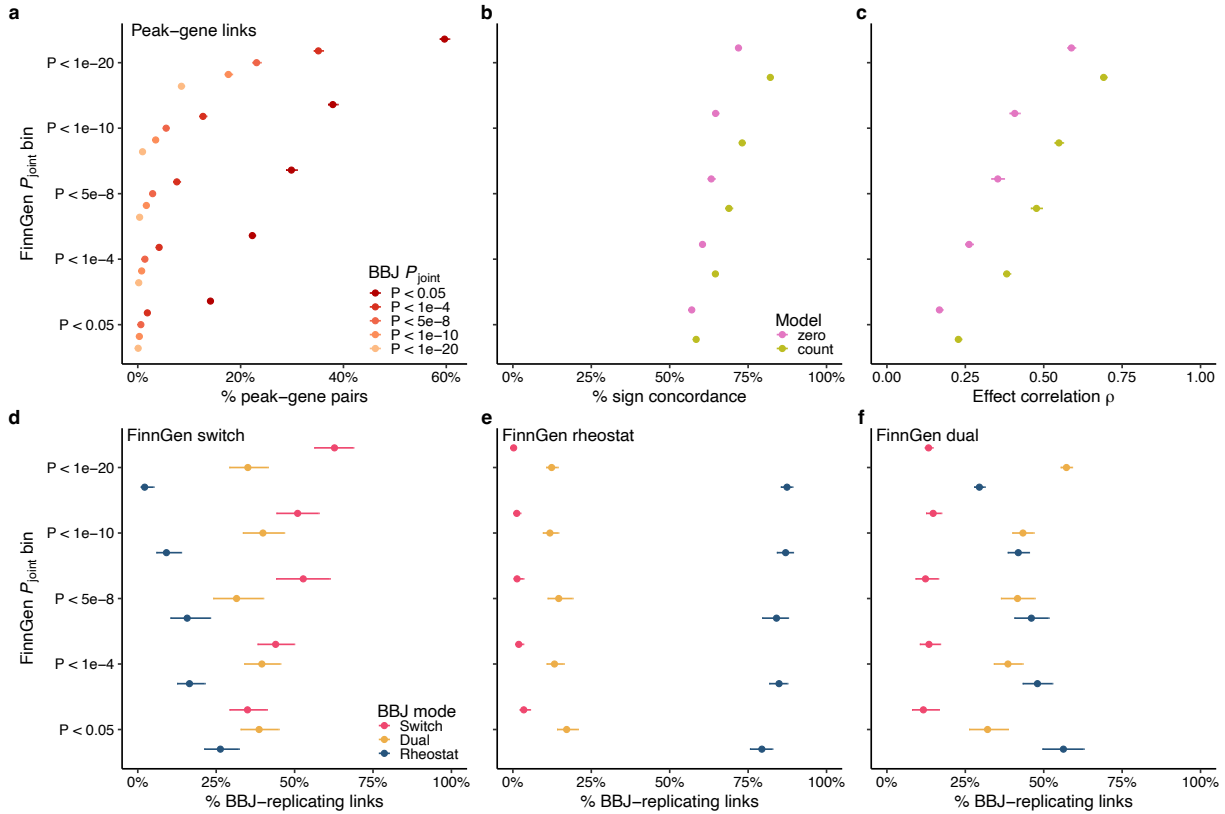


Supplementary Fig. 1 | Quality control metrics and cell type composition of snRNA-seq and snATAC-seq atlases. **a,b**, The number of genes (**a**) and the number of UMIs (**b**) per nucleus in the snRNA-seq atlas (10,612,211 nuclei from 1,108 donors), stratified by multiplexing status (see **Methods**). **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**).

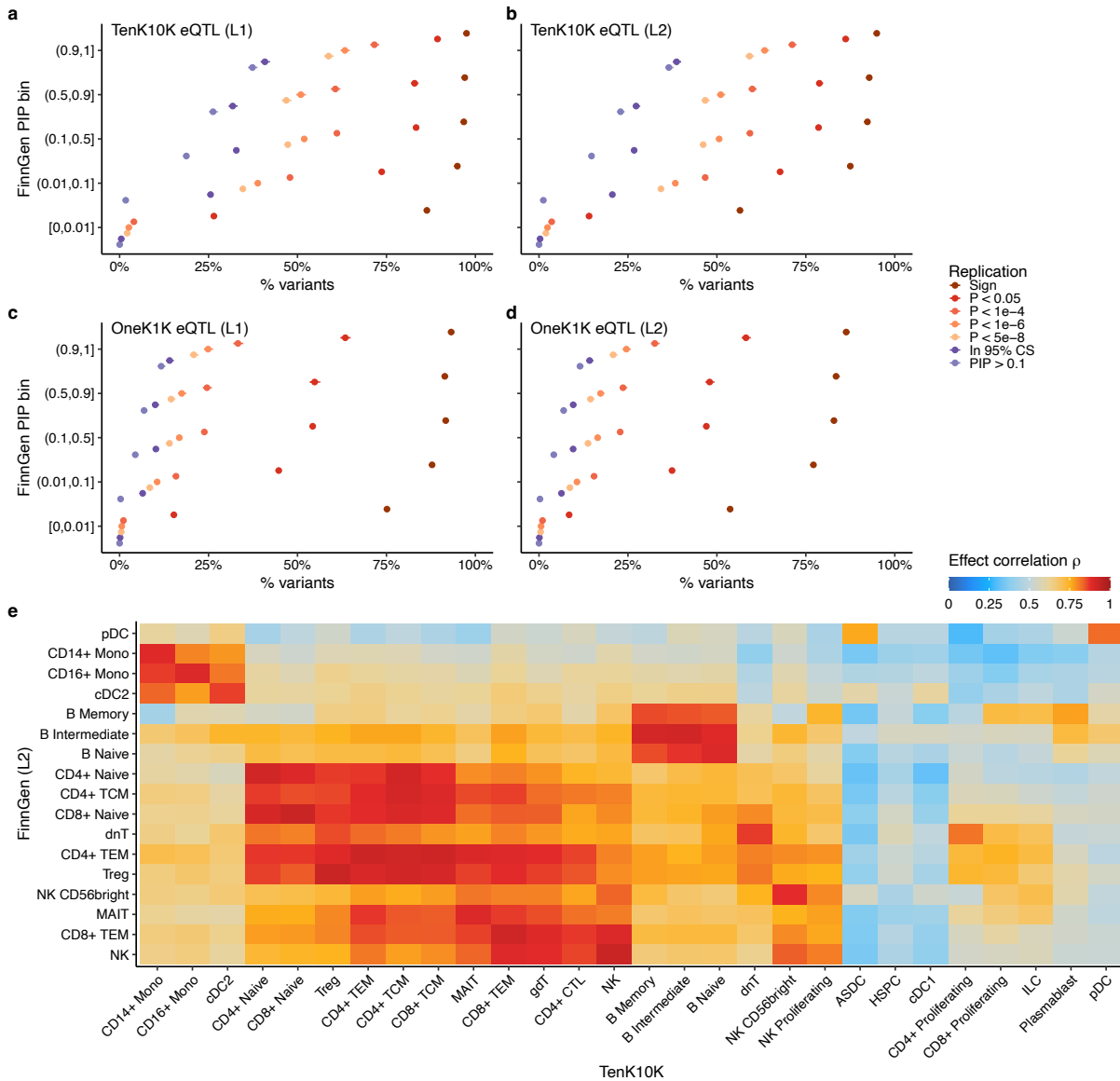
• There is a statement on line 93-94 that “Importantly, the functional effects and regulatory mechanisms we characterize represent fundamental biology shared across human populations”. What is the evidence for this statement.

Previous studies have demonstrated that causal genetic architecture on complex traits (GWAS) and molecular phenotypes (molQTLs) are largely shared across human populations (Brown et al., 2016; Shi et al., 2021; Taylor et al., 2024). Additionally, this statement is now supported by our comprehensive cross-ancestry replication analysis (see our responses to **Reviewer 1, Comment 5** and **Reviewer 2, Comment 1**), which demonstrates high concordance of eQTL effects across European and East Asian ancestries.

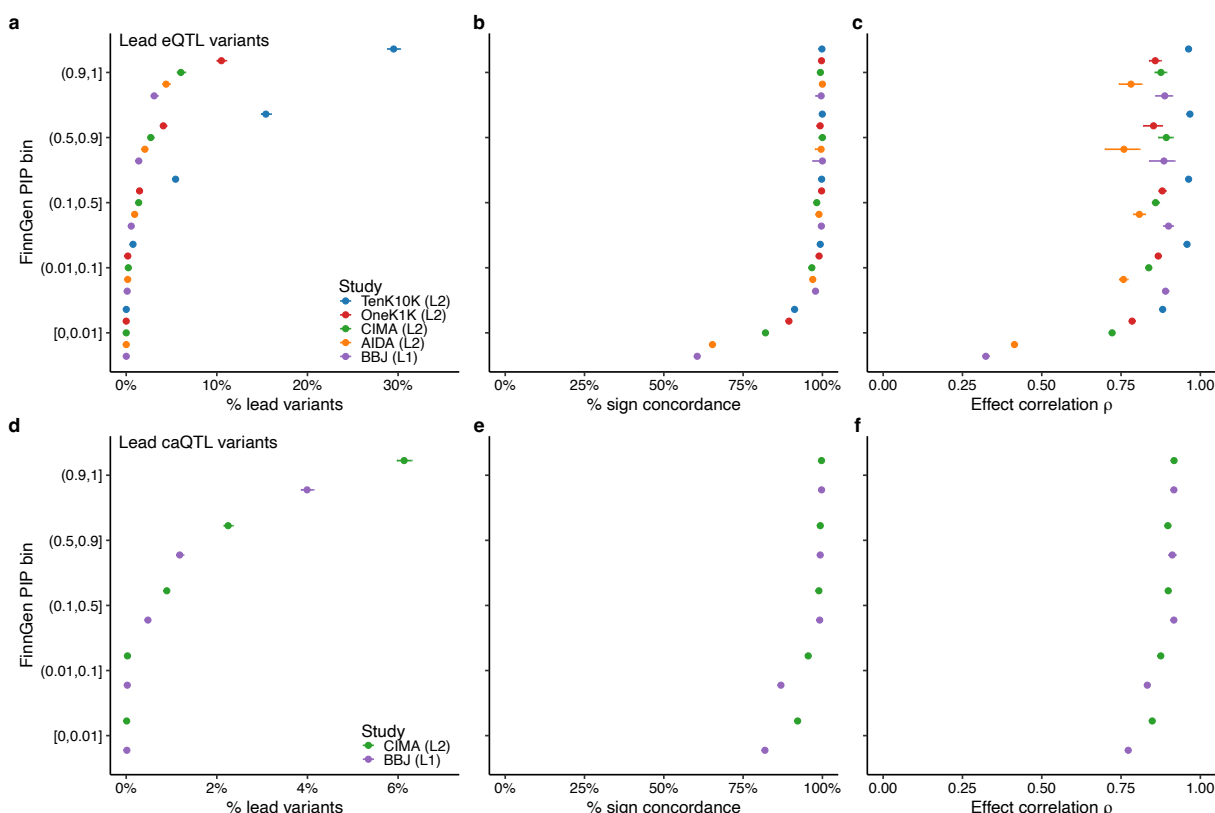
To clarify these points, we have added **Supplementary Figs. 6,8,9** and revised the **Results**:



New 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 bar represents 95% confidence intervals.



New 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**).



New 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.

Results:

We replicated peak–gene links in an independent Japanese multiome cohort (BBJ; $n = 98$), where replication rates increased monotonically with FinnGen significance and BBJ-replicating links largely retained their FinnGen regulatory-mode assignment (**Supplementary Fig. 6**).

...

We replicated fine-mapped *cis*-eQTLs across five independent datasets spanning diverse ancestries (TenK10K, OneK1K, CIMA, AIDA, and BBJ; **Supplementary Figs. 8,9**), where replication rates increased monotonically with FinnGen PIP. Among fine-mapped *cis*-eQTLs with $PIP > 0.9$, 96% showed concordant effect direction and 59% reached genome-wide significance in TenK10K (Spearman's $\rho = 0.92$), with lead variant $\rho = 0.78$ – 0.96 across all five datasets. We also replicated *cis*-caQTLs in CIMA and BBJ (Spearman's $\rho = 0.92$ for both).

- Is the large degree (4% and 2%) of opposite effect eQTL and caQTL (line 309) driven by differences in cell type. If you control for this, how many opposite effects are found in the same cell type. Are these then erroneous or biological?

We appreciate the reviewer's question. To clarify, the opposite-effect eQTLs and caQTLs refer to cases where the same causal variant has opposite effect directions across different cell types,

not within the same cell type. We identified shared causal variants across cell types using fine-mapping and then characterized their effect consistency using Cochran's Q heterogeneity test. The opposite-effect cases thus likely represent genuine biological complexity where a single causal variant can increase expression in one cell type while decreasing it in another, reflecting cell-type specific regulatory contexts.

To clarify this point, we have revised the **Results**:

Among eGenes and caPeaks with shared causal variants, we further characterized their effect consistency **across cell types** using Cochran's Q heterogeneity test ($P < 5.0 \times 10^{-8}$). Specifically, 50% of eGenes and 43% of caPeaks showed consistent effects **across cell types**, while 13% and 13% showed heterogeneous effects (with the same direction), and 4% and 2% exhibited opposite effects despite sharing the same causal variants (**Fig. 3b,c, Supplementary Table 18,19**).

- Overall the effect sizes for eQTL seem much smaller than the effects sizes for caQTLs (at least from the few examples discussed (example lines 489 – 500). Is there a reason this is the case? Does it reflect how the data was transformed prior to modeling or sparsity? Can the authors comment.

The apparent difference in effect sizes between eQTLs and caQTLs reflects differences in units rather than biological magnitude. As noted in our **Methods**, these effect sizes are derived from different statistical models: β_{eQTL} represents change in log gene expression counts per allele (SAIGE-QTL Poisson mixed model), whereas β_{caQTL} represents change in pseudobulk chromatin accessibility in standard deviation units per allele (tensorQTL Gaussian model). Because of these different units, absolute effect sizes cannot be directly compared across modalities.

- Line 524-526 requires a citation. Taking about the shared origin and similarity of monocytes and microglia.

We thank the reviewer for flagging this. We have revised the **Results** to remove the appeal to shared monocyte–microglia developmental origin, which we agree was an imprecise generalization. The revised text instead grounds the cross-cell-type relevance directly in the DAP12-coupled microglial sensome (Haure-Mirande et al., 2022), the receptor signaling network through which *PILRB* and *TYROBP* are causally implicated in Alzheimer's pathology and onto which both AD risk variants at this locus converge in our monocyte data.

- LTBR effects, as the authors know, are multifaceted. The conclusions on lines 536-539 seem a bit of a biological stretch to me. It's unclear if a minor change in expression ($B = -0.05$) could by itself have profound changes on antimicrobial responses or tonsil hyperplasia.

We appreciate this critique. Following an updated peak–gene analysis (see our response to **Reviewer 1, Comment 2**), the *LTBR* example became less significant and less central to our CASCADE demonstration. We have therefore replaced the *LTBR* vignette in the **Results** with a stronger *APOBEC3A/CBX6* example, where CASCADE analysis nominates *CBX6* as an additional candidate effector at the well-characterized *APOBEC3A/B* breast cancer susceptibility locus (revised **Extended Data Fig. 8c** and new **Supplementary Figs. 14,15**).

Similarly, 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*^{73–75}. 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^{76,77}. *CBX6* encodes a Polycomb Repressive Complex 1 chromobox component reported as a potential tumor suppressor in breast cancer⁷⁸, consistent with its broad down-regulation by the risk allele across PBMCs.

- The findings around “constrained genes” are fascinating. There is statement on line 588-592 that minor changes in highly expressing genes can have large effects on disease pathology. Are there any evidence or citations that functionally confirm this with experimental evidence? Yes, they colocalize with GWAS variants but intuitively I would have presumed the opposite that highly expressed genes are less sensitive to minor changes in gene expression.

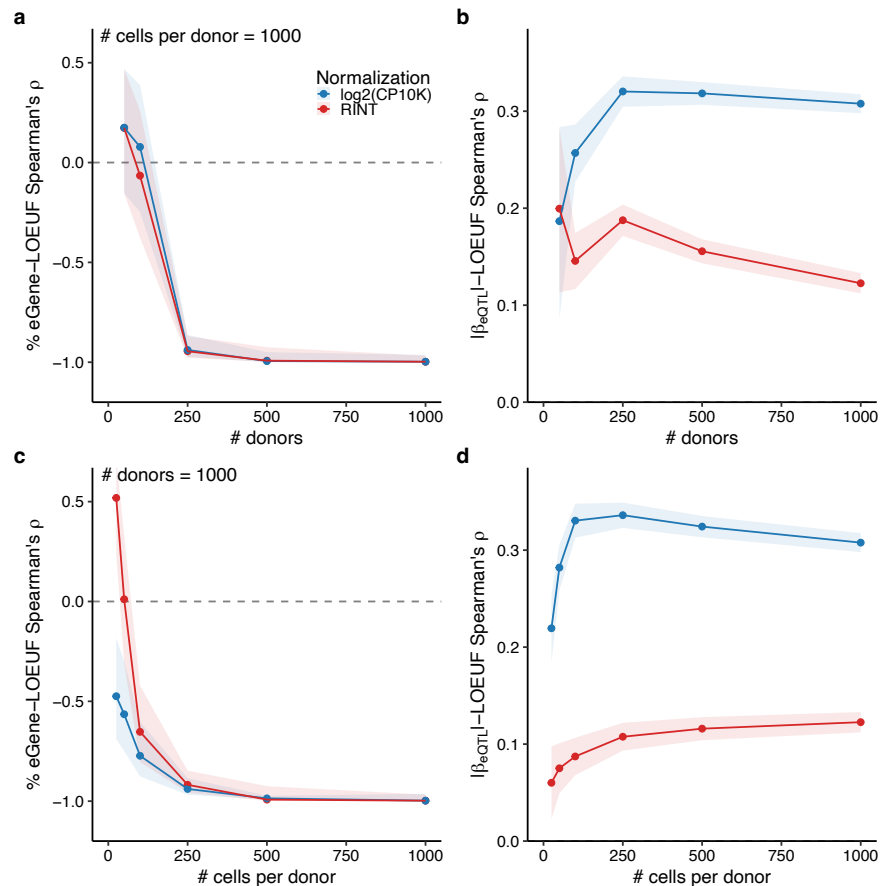
We thank the reviewer for this question. We would like to clarify that high expression does not buffer constrained genes against dosage perturbation. Rather, these genes are highly expressed because their function requires specific quantities, making them dosage-sensitive. Consistent with previous studies (Karczewski et al., 2020; Zeng et al., 2024), our data show that constrained genes have higher mean expression with lower variance (**Extended Data Fig. 9c**), reflecting evolutionary pressure to maintain stable expression levels. Furthermore, constrained genes exhibit both haploinsufficiency and triplosensitivity, indicating bidirectional dosage constraints that tightly control their expression range (Morrill & Amon, 2019; Collins et al., 2022).

Experimental evidence from Naqvi et al. (2023) directly demonstrated this sensitivity using precise titration of SOX9 TF levels. While most SOX9-dependent targets are buffered against small dosage changes, a “core” subset of direct, primary targets exhibit heightened sensitivity to SOX9 dosage. This confirms that high expression does not confer robustness to minor expression changes but rather reflects a requirement for precise dosage.

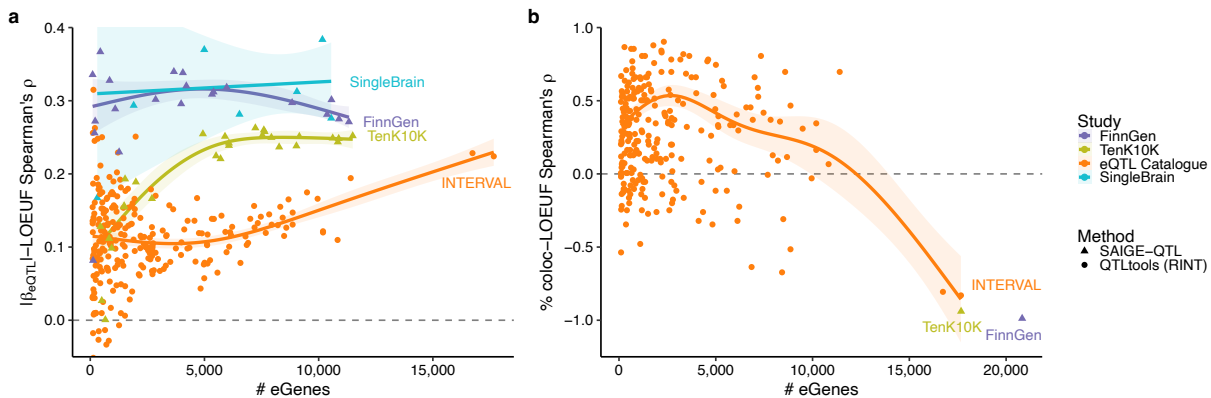
- Given that the findings with constrained genes seems to run opposite to studies from others, can the authors would analyze the 1K1K public data to show supporting evidence for GWAS localization in constrained genes. Is this not of comparable size if not cell number to discover these same effects?

We thank the reviewer for this suggestion. We analyzed the constraint gradient in OneK1K ($n = 982$; Yazar et al., 2022). The effect size gradient was positive in all cell types (Spearman's $\rho = 0.04$), confirming that constrained genes have smaller eQTL effect sizes. While OneK1K has a

comparable donor count to FinnGen (982 vs. 1,108), its per-donor cell depth is ~40-fold lower (~60 vs. ~2,300 cells/donor for CD4 T cells), limiting statistical power for small-effect eQTLs. Our downsampling analysis (new **Supplementary Fig. 19**) shows that this low-cell-depth regime precedes the power threshold at which the GWAS colocalization gradient at constrained genes fully emerges, explaining why OneK1K alone cannot adjudicate the coloc enrichment. Our cross-study replication (new **Supplementary Fig. 20**) resolves this directly: both the effect size gradient and the GWAS colocalization enrichment replicate in TenK10K ($n = 1,925$) and in INTERVAL ($n = 4,729$), validating our core finding (see also our response to **Reviewer 1, Comment 11**).



New 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 ($|\beta_{\text{eQTL}}|$) 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.



New 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.

- Lines 608-612 about FAS and B cells is missing a citation.

We have added two primary citations. Chu *et al.* (2002) reports that CD40-activated CLL B cells express high levels of CD95/FAS yet remain resistant to FAS-mediated apoptosis, with the resistance attributable to compensatory anti-apoptotic mechanisms. Rathmell *et al.* (1996) provides the primary demonstration that Fas/FasL signaling has dual roles in B cells, with BCR-modulated FasL engagement driving clonal proliferation rather than deletion, supporting the non-canonical FAS pathways implicated in B cell proliferation.

- Again, lines 703-710 should be removed from the paper. They are not presented as data and have not been peer-reviewed. Additionally, the identified cascade was in CD8 T cells, but functional validation performed in CD4 (Th1) suggesting a discrepancy.

As noted in our response to the **Comment 4**, the STAG-seq study is our recent paper, which is currently under review and available as a preprint (Liu, et al., *bioRxiv*).

Regarding the apparent cell type discrepancy, our revised peak-gene link analysis (see our response to **Reviewer 1**, **Comment 2**) reclassified the cascade. rs748670681 is now identified as a positional cascade variant in CD4⁺ T cells, where the *cis*-eQTL for *TNRC18* ($\beta = -0.47$, $P = 1.4 \times 10^{-127}$, PIP = 1.0) co-occurs with a significant peak-gene link from the overlapping peak (chr7:5397119-5397618; rheostat link, $\beta_{\text{count}} = 0.04$, $P_{\text{count}} = 1.6 \times 10^{-3}$) within the same cell type (**Fig. 6c-f**). The cascade is therefore now anchored in CD4⁺ T cells rather than CD8⁺ T cells.

This reclassification is also consistent with the STAG-seq results (Liu, et al., *bioRxiv*), which comprehensively evaluated rs748670681 across multiple immune cell types, including monocyte-derived macrophages and dendritic cells, granulocytes, and activated CD4⁺ and CD8⁺ T cells, as well as ex vivo differentiated Th1 cells. *TNRC18* downregulation was observed in both CD4⁺ and CD8⁺ T cells, with the strongest effect in Th1 cells. Together, our endogenous QTL identification

in CD4⁺ T cells and the STAG-seq validation in Th1 cells consistently anchor the regulatory mechanism within the CD4⁺ lineage.

We have clarified this in the **Results**:

In our recent paper⁸⁴, 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.

- Line 734 -736 discusses this validation again and claims that chromatin changes were validated. Again, this should be removed.

We have revised the text to clarify that STAG-seq validates transcriptomic consequences rather than chromatin accessibility changes directly.

We thank the reviewer for catching this. rs748670681 was, and still is, classified as a positional cascade variant, meaning the variant overlaps a regulatory peak linked to *TNRC18* but is not itself a *cis*-caQTL for that peak. The original sentence inaccurately described this as a mechanism “where chromatin changes transmit to gene expression,” which conflated positional with full cascade. We have corrected the summary to reflect the positional cascade classification and to remove the implication that chromatin accessibility changes were validated as follows. This phrasing restricts the STAG-seq validation to its transcriptomic consequences.

The *TNRC18* variant exemplifies a positional cascade in which a non-coding variant within a regulatory peak linked to the target gene modulates its expression without itself altering chromatin accessibility, with functional validation revealing downstream transcriptomic consequences such as retroviral reactivation and inflammatory memory states⁸³.

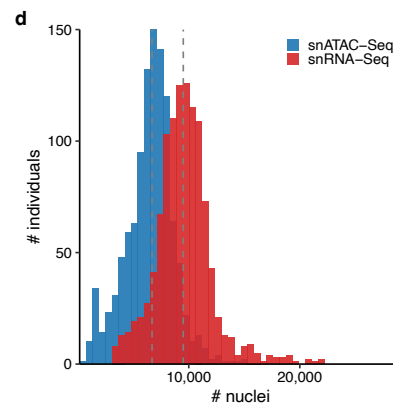
- In the discussion regarding base-editing validation, it would be nice to cite work from others (not just STAGseq) as efforts from others are also paving the way in this V2F space.

We have expanded our **Discussion** to cite additional variant-to-function methods, including Findlay et al (2018), Nasser et al (2021), Schmidt et al (2022), Martin-Rufino et al (2025), and Zhu et al (*bioRxiv*).

Our population-scale molQTLs provide a foundation for variant-to-function (V2F) research that complements, rather than substitutes for, direct functional perturbation. CASCADE nominates putative causal variants, cell types, and chromatin-to-expression mechanisms, while a growing suite of base-editing and CRISPR perturbation platforms now bridges this gap in primary, disease-relevant cells^{91,94–98}.

- Figure 1d. The area under the curves looks different (ie number of individuals in snATAC vs snRNA but reporting in the supplement suggest that roughly equal number of individuals are present in both datasets. Please double check this graph.

We thank the reviewers for catching this visualization error. We have revised **Fig. 1d** to correctly display the overlapping distributions, confirming comparable sample sizes between snRNA-seq and snATAC-seq.

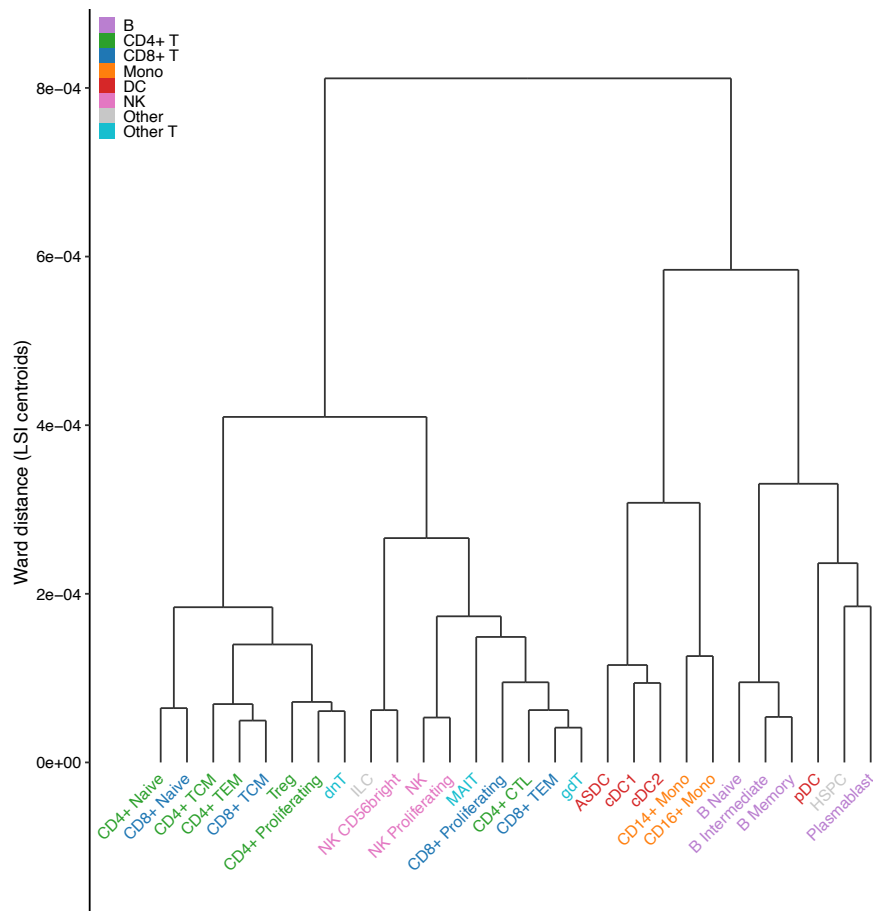


Revised Fig. 1 | Single-nucleus RNA-seq and ATAC-seq atlases of 1,108 FinnGen donors. d. The distribution of the number of nuclei per individual for snRNA-Seq and snATAC-Seq. Vertical lines represent the median of the number of nuclei per individual.

- Figure 1c. CD8 naïve T cells are WAY out there. This is a little bizzare. They are very far from the other T and CD8 TEM populations. Please double check the labels. I understand that distance in UMAPs means nothing but it's still concerning.

We thank the reviewers for raising this point. We have confirmed the cell type labels are correct. The apparent separation of CD8⁺ Naïve T cells from other T cell populations is likely exacerbated by a UMAP visualization artifact rather than a reflection of true biological distance.

To empirically validate this, we examined the snATAC-seq spectral (LSI) embedding directly. In hierarchical clustering of L2 cell type centroids in LSI space (**Response Fig. 5**), CD8⁺ Naïve T cells form a mutual nearest-neighbor pair with CD4⁺ Naïve T cells at the base of the T cell branch, grouped with memory T cells (CD4⁺ TCM, CD4⁺ TEM, CD8⁺ TCM, Treg) within a broader lymphoid block. CD8⁺ TEM cells instead group with the cytotoxic/effector populations (CD4⁺ CTL, CD8⁺ Proliferating, MAIT, gdT), reflecting the T cell activation state over CD4/CD8 lineage in shaping chromatin accessibility. The peripheral placement of CD8⁺ Naïve T cells in **Fig. 1c** therefore reflects a projection of transcriptionally and epigenetically intermediate populations rather than mislabeling or true outlier biology.



Response Fig. 5 | Ward hierarchical clustering of L2 cell type centroids in snATAC-seq LSI space. Centroids were computed from the integrated snATAC-seq LSI embedding. Pairwise Euclidean distances between centroids were used as input to Ward.D2 hierarchical clustering. Leaves are colored by L1 cell type.

- Figure 2e, does the number of caQTL and eQTL per cell type correlate directly to the number of cells per sample? If you examine just constrained highly expressed genes does that same pattern hold. Is there really more genetic regulation in a CD4 T cells than NK cells.

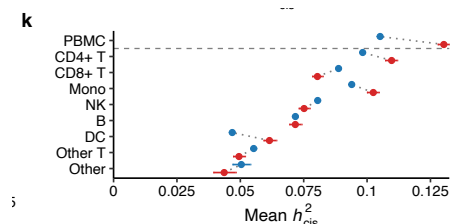
We examined the correlation between the number of eGenes/caPeaks and mean number of nuclei per sample and found strong correlations for both eQTLs and caQTLs (Spearman's $\rho = 0.96$ and 0.95 , respectively). This pattern held when restricted to constrained, highly-expressed genes (LOEUF decile 0 and top 20% by mean expression) and caPeaks linked to them (Spearman's $\rho = 0.97$ and 0.95 , respectively), consistent with cell-type differences in QTL counts reflecting per-sample power rather than biological differences in genetic regulation.

- Figure 2h suggests that most variants are found right in the middle of the annotated caQTL peak. Is this biological or a consequence of how peaks and caQTLs are called. And if it's the latter, is this graph informative?

The enrichment of caQTL variants at peak centers is largely biological. Peak summits represent regions of maximal chromatin accessibility where TF binding sites and other regulatory motifs are concentrated, so functional variants are expected to cluster there. This pattern demonstrates that caQTL variants preferentially affect the core of regulatory elements rather than at peak edges, consistent with their regulatory roles and previous studies.

- Figure 2k, it's a little unclear which blue dot pairs with which red dot. For clarity could these be connected.

We have revised **Fig. 2k** to connect paired blue and red dots with lines as shown below.



Revised Fig. 2 | Genetic architecture of molecular QTLs in FinnGen. k. The distribution of mean estimated *cis*-heritabilities for caQTL and eQTL across L1 cell types. The y-axis is ordered by the number of eGenes for each cell type using the same order as in (a). Dotted line connects the eQTL and caQTL mean *cis*-heritabilities within each cell type.

- Figure 3d-j. Wonderful examples, for clarity can you facet by cell type in each graph. Overlapping box plots are perhaps not the best visual way to show this

We thank the reviewer for this suggestion. We have revised **Fig. 3d-j** to facet by cell type as shown below.

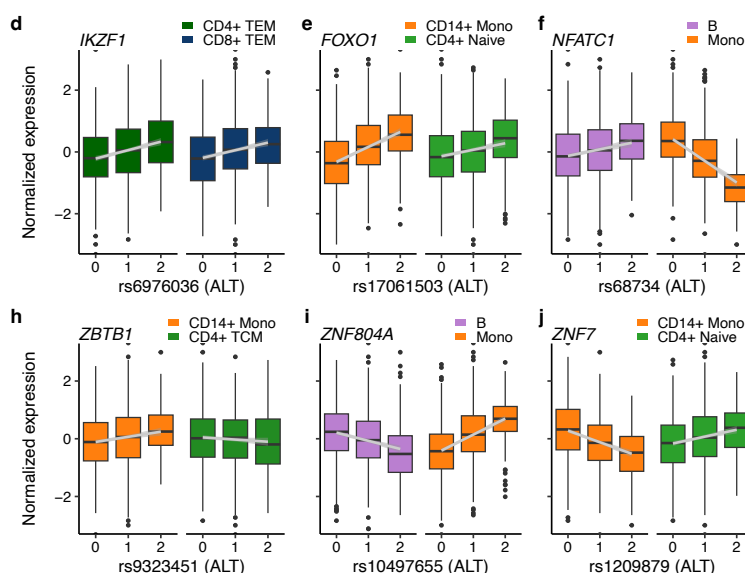
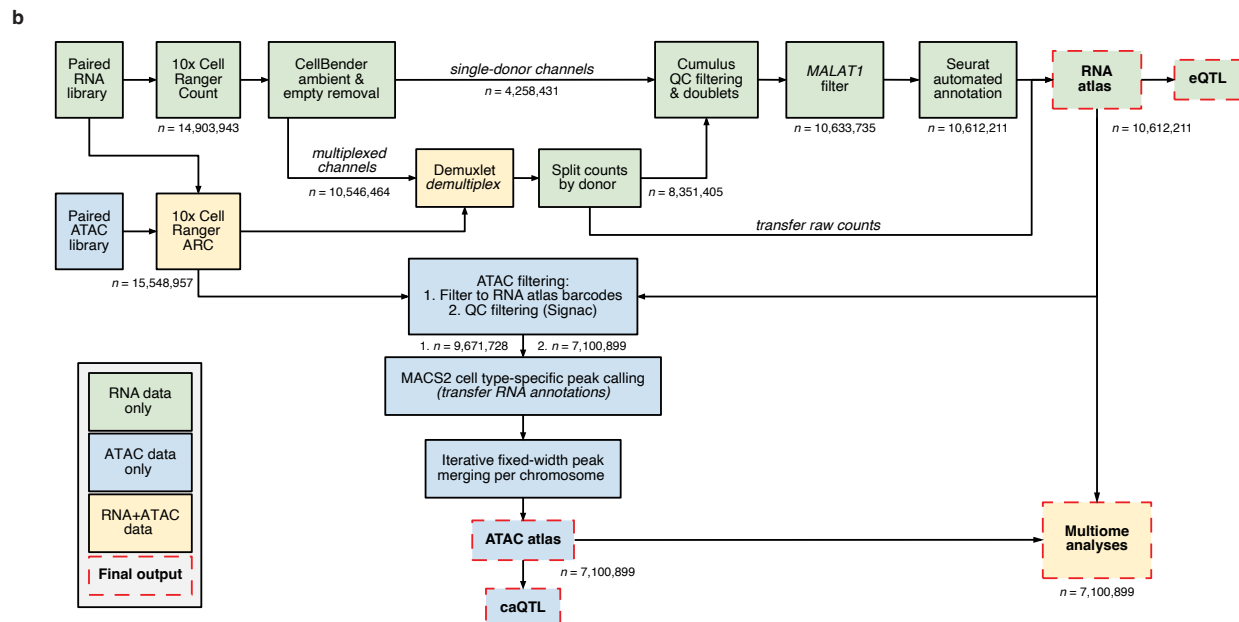


Fig. 3: Cell type specificity of molecular QTLs. d-f. Examples of eGenes that are driven by the same variant with consistent (d, *IKZF1*), heterogeneous (e, *FOXO1*), and opposite (f, *NFATC1*) effects across cell types. Effect sizes are with respect to an alternative allele (ALT).

- Extended data figure 1. Thank you for this figure, it should be required in all studies. Can you please include the number of cells passing each threshold. An integrated RNA-ATAC object is reported but never shown. Consider adding.

We have revised **Extended Data Fig. 1b** to annotate the number of nuclei entering each pipeline step. We have also removed the previously labeled “integrated RNA-ATAC object” from the figure. We use the two atlases directly via shared barcodes for multiome analyses.



Extended Data Fig. 1 | Schematic overview of experimental and computational pipelines.

b, Multiome analysis pipeline. Counts under each box indicate the number of nuclei flowing along that path. The snRNA-seq atlas contains 10,612,211 nuclei from 1,108 donors, and the snATAC-seq atlas contains 7,100,899 nuclei from 1,093 donors with paired snRNA-seq data. QC, quality control. eQTL, expression quantitative trait locus.

- The methods mention flow cytometry data (lines 1171). Why? Was this presented or used?

We apologize for the confusion. Flow cytometry was performed as a pilot study for a subset of individuals but never used in the present study. We have removed the corresponding texts from the **Methods** section.

- Is filtering on MALAT1 expression appropriate? Line 1235. The authors cite a singular pre-print. I was not aware of this as an established strategy. Additionally, the threshold is set per run. Can this data be shared and further explanation of how many cells were lost explained.

We appreciate the reviewer raising this point. While the initially cited preprint (Clarke & Bader, *bioRxiv*) was the primary reference for our QC procedure, *MALAT1* expression is well established as a marker of cell quality and nuclear integrity (Muskovic & Powell, 2021; Montserrat-Ayuso & Esteve-Codina, 2024). Indeed, the official 10x Genomics support page specifically discusses *MALAT1* as a cell quality indicator, noting that high *MALAT1* expression reflects dead or dying cells while low *MALAT1* expression corresponds to empty droplets or cytosolic debris (<https://kb.10xgenomics.com/s/article/360004729092-What-does-the-expression-level-of-Mal1-in-my-gene-expression-data-indicate>).

Because *MALAT1* abundance varies with nucleus quality and sequencing depth, the threshold was set per channel from its *MALAT1* distribution (Clarke & Bader, *bioRxiv*). Across all channels, *MALAT1* filtering removed 21,524 of 10,633,735 nuclei (0.2%) entering this stage, leaving the 10,612,211 QC-passing nuclei (**Extended Data Fig. 1b**).

Referee #4 (Remarks on code availability):

No code was shared and github links were not accessible. I cannot evaluate but feel confident in the expertise of the authors.

We apologize for this oversight. All code repositories cited in the manuscript are now publicly available.

References:

Alasoo, K. *et al.* (2018) 'Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response', *Nature genetics*, 50(3), pp. 424–431.

Binder, C. *et al.* (2020) 'CD2 immunobiology', *Frontiers in Immunology*, 11, p. 1090.

Bønnelykke, K. *et al.* (2014) 'A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations', *Nature Genetics*, 46(1), pp. 51–55.

Brown, B.C. *et al.* (2016) 'Transethnic Genetic-Correlation Estimates from Summary Statistics', *American journal of human genetics*, 99(1), pp. 76–88.

Burger, S. *et al.* (2026) 'Inter-individual variation of cellular and gene-expression properties of the human striatum', *bioRxiv*. bioRxiv. Available at: <https://doi.org/10.64898/2026.03.20.713160>.

Chu, P. *et al.* (2002) 'Latent sensitivity to Fas-mediated apoptosis after CD40 ligation may explain activity of CD154 gene therapy in chronic lymphocytic leukemia', *Proceedings of the National Academy of Sciences of the United States of America*, 99(6), pp. 3854–3859.

Clarke, Z.A. and Bader, G.D. (2024) 'MALAT1 expression indicates cell quality in single-cell RNA sequencing data', *bioRxiv*. Available at: <https://doi.org/10.1101/2024.07.14.603469>.

Collins, R.L. *et al.* (2022) 'A cross-disorder dosage sensitivity map of the human genome', *Cell*, 185(16), pp. 3041–3055.e25.

Cuomo, A.S.E. *et al.* (2025) 'Impact of rare and common genetic variation on cell type-specific gene expression', *medRxiv*. Available at: <https://doi.org/10.1101/2025.03.20.25324352>.

Findlay, G.M. *et al.* (2018) 'Accurate classification of BRCA1 variants with saturation genome editing', *Nature*, 562(7726), pp. 217–222.

Giambartolomei, C. *et al.* (2014) 'Bayesian test for colocalisation between pairs of genetic association studies using summary statistics', *PLoS genetics*, 10(5), p. e1004383.

Haure-Mirande, J.-V. *et al.* (2022) 'Microglial TYROBP/DAP12 in Alzheimer's disease: Transduction of physiological and pathological signals across TREM2', *Molecular Neurodegeneration*, 17(1), p. 55.

Jang, B. *et al.* (2026) 'A meta-analysis of single-nucleus expression quantitative trait loci linking genetic risk to brain disorders', *Nature Genetics*, pp. 1–11.

Jiang, H. and Doerge, R.W. (2006) 'A two-step multiple comparison procedure for a large number of tests and multiple treatments', *Statistical applications in genetics and molecular biology*, 5(1), p. Article28.

- Joshi, H. and Morley, S.C. (2022) 'Efficient T cell migration and activation require L-plastin', *Frontiers in Immunology*, 13, p. 916137.
- Kanai, M. *et al.* (2022) 'Meta-analysis fine-mapping is often miscalibrated at single-variant resolution', *Cell Genomics*, (100210), p. 100210.
- Karczewski, K.J. *et al.* (2020) 'The mutational constraint spectrum quantified from variation in 141,456 humans', *Nature*, 581(7809), pp. 434–443.
- Kerimov, N. *et al.* (2021) 'A compendium of uniformly processed human gene expression and splicing quantitative trait loci', *Nature genetics*, 53(9), pp. 1290–1299.
- Kock, K.H. *et al.* (2025) 'Asian diversity in human immune cells', *Cell*, 188(8), pp. 2288–2306.e24.
- Kono, M. *et al.* (2025) 'Accurate, sensitive, and efficient chromatin accessibility quantification at target loci using UNICHro-seq', *medRxiv*, p. 2025.07.29.25332340.
- Korsunsky, I. *et al.* (2019) 'Fast, sensitive and accurate integration of single-cell data with Harmony', *Nature Methods*, 16(12), pp. 1289–1296.
- Kotliar, D. *et al.* (2025) 'Reproducible single-cell annotation of programs underlying T cell subsets, activation states and functions', *Nature Methods*, 22(9), pp. 1964–1980.
- Kumasaka, N. *et al.* (2023) 'Mapping interindividual dynamics of innate immune response at single-cell resolution', *Nature genetics*, 55(6), pp. 1066–1075.
- Liu, K. *et al.* (2025) 'Sensitive Transcriptomics and Genotyping reveals function of genetic variants in immunity', *bioRxiv*. Available at: <https://doi.org/10.1101/2025.03.12.642879>.
- Martin-Rufino, J.D. *et al.* (2025) 'Transcription factor networks disproportionately enrich for heritability of blood cell phenotypes', *Science (New York, N.Y.)*, 388(6742), pp. 52–59.
- Montserrat-Ayuso, T. and Esteve-Codina, A. (2024) 'High content of nuclei-free low-quality cells in reference single-cell atlases: a call for more stringent quality control using nuclear fraction', *BMC Genomics*, 25(1), p. 1124.
- Morrill, S.A. and Amon, A. (2019) 'Why haploinsufficiency persists', *Proceedings of the National Academy of Sciences of the United States of America*, 116(24), pp. 11866–11871.
- Muskovic, W. and Powell, J.E. (2021) 'DropletQC: improved identification of empty droplets and damaged cells in single-cell RNA-seq data', *Genome Biology*, 22(1), p. 329.
- Naqvi, S. *et al.* (2023) 'Precise modulation of transcription factor levels identifies features underlying dosage sensitivity', *Nature Genetics*, 55(5), pp. 841–851.
- Nasser, J. *et al.* (2021) 'Genome-wide enhancer maps link risk variants to disease genes', *Nature*, 593(7858), pp. 238–243.
- Perez, R.K. *et al.* (2022) 'Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus', *Science (New York, N.Y.)*, 376(6589), p. eabf1970.
- Rathmell, J.C. *et al.* (1996) '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(2), pp. 319–329.
- Roh, H. *et al.* (2025) 'Coupling CRISPR scanning with targeted chromatin accessibility profiling using a double-stranded DNA deaminase', *Nature Methods*, 22(10), pp. 2083–2093.
- Saha, A. and Battle, A. (2018) 'False positives in trans-eQTL and co-expression analyses arising from RNA-sequencing alignment errors', *F1000Research*, 7(1860), p. 1860.
- Schmidt, R. *et al.* (2022) 'CRISPR activation and interference screens decode stimulation responses in primary human T cells', *Science (New York, N.Y.)*, 375(6580), p. eabj4008.
- Shi, H. *et al.* (2021) 'Population-specific causal disease effect sizes in functionally important regions impacted by selection', *Nature communications*, 12(1), p. 1098.
- Siraj, L. *et al.* (2026) 'Functional dissection of complex trait variants at single-nucleotide resolution', *Nature*, pp. 1–11.
- Soskic, B. *et al.* (2022) 'Immune disease risk variants regulate gene expression dynamics during CD4+ T cell activation', *Nature Genetics*, 54(6), pp. 817–826.
- Taylor, D.J. *et al.* (2024) 'Sources of gene expression variation in a globally diverse human cohort', *Nature*, 632(8023), pp. 122–130.
- The GTEx Consortium (2020) 'The GTEx Consortium atlas of genetic regulatory effects across human tissues', *Science*, 369(6509), pp. 1318–1330.
- Tokolyi, A. *et al.* (2025) 'The contribution of genetic determinants of blood gene expression and splicing to molecular phenotypes and health outcomes', *Nature genetics*, pp. 1–10.
- Van den Berge, K. *et al.* (2017) 'stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage', *Genome Biology*, 18(1), p. 151.
- Võsa, U. *et al.* (2021) 'Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression', *Nature Genetics*, 53(9), pp. 1300–1310.
- Walker, C.R. *et al.* (2024) 'Private information leakage from single-cell count matrices', *Cell*, 187(23), pp. 6537–6549.e10.
- Wallace, C. (2021) 'A more accurate method for colocalisation analysis allowing for multiple causal variants', *PLoS genetics*, 17(9), p. e1009440.
- Warmerdam, C.A.R. *et al.* (2026) 'Trans-eQTLs reveal the architecture of human gene regulatory networks', *medRxiv*, p. 2026.02.04.26343575.
- Xue, A. *et al.* (2025) 'Genetic regulation of cell type-specific chromatin accessibility shapes immune function and disease risk', *medRxiv*. Available at: <https://doi.org/10.1101/2025.08.27.25334533>.
- Yazar, S. *et al.* (2022) 'Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease', *Science (New York, N.Y.)*, 376(6589), p. eabf3041.

Yin, J. *et al.* (2026) 'Chinese Immune Multi-Omics Atlas', *Science (New York, N.Y.)*, 391(6781), p. eadt3130.

Zeng, T. *et al.* (2024) 'Bayesian estimation of gene constraint from an evolutionary model with gene features', *Nature genetics*, 56(8), pp. 1632–1643.

Zhou, W. *et al.* (2024) 'Efficient and accurate mixed model association tool for single-cell eQTL analysis', *medRxiv: the preprint server for health sciences*, p. 2024.05.15.24307317.

Zhu, R. *et al.* (2025) 'Genome-scale perturb-seq in primary human CD4+ T cells maps context-specific regulators of T cell programs and human immune traits', *bioRxiv*. bioRxiv. Available at: <https://doi.org/10.64898/2025.12.23.696273>.

Zhu, Z. *et al.* (2016) 'Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets', *Nature genetics*, 48(5), pp. 481–487.

Referee #1:

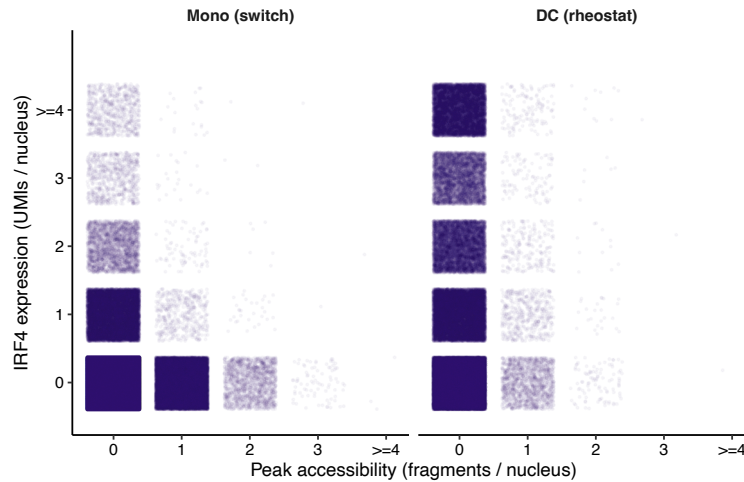
I would like to thank the authors for the considerable amount of work that has been put in to respond to my previous comments. I believe the updates strengthen the manuscript and its conclusions. I have only a few remaining comments based on their revisions, which once addressed, would lead to me recommending the manuscript for acceptance.

We thank the reviewer for the positive assessment of our revisions. We are pleased that the reviewer found the changes to have strengthened the manuscript, and we address each remaining comment below.

1 (Previous comment 2). I believe the updated analysis is a large improvement on the previous presentation of peak-gene links. However, in the revised manuscript, the authors provide examples of what each peak category would look like using simulated data only (Supplementary Figure 3), or present odds-ratios (Supplementary Figure 4). I believe it would be useful to present the equivalent plot of Supplementary Figure 3 but using real data, so that the reader can visualise what the real associations look like.

We thank the reviewer for raising this point. We agree a real-data view is desirable, but single-cell chromatin accessibility and expression are both sparse and zero-inflated, so the per-cell relationship is dominated by zeros and does not visually resolve the three modes even for strongly significant links. At the strongly significant *IRF4* link (switch in monocytes, rheostat in DCs; **Supplementary Fig. 4d**), only 0.04% and 1.04% of nuclei, respectively, have the peak accessible and the gene expressed, leaving the modes indistinguishable in the raw per-nucleus data (**Response Fig. 1**). This is why we use a hurdle model to separate the components statistically, and we retain the schematic in **Supplementary Fig. 3** as the clearest illustration of the regulatory modes. We add a clarifying sentence to the **Supplementary Fig. 3** legend to make this rationale explicit for the reader.

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.



Response Fig. 1 | Empirical single-cell data do not visually resolve the regulatory mode of a strong peak–gene link.

Per-nucleus chromatin accessibility vs. *IRF4* expression for the *IRF4* promoter-linked peak (chr6:246,344–246,843) in monocytes (Mono; switch) and dendritic cells (DC; rheostat). Each point is a nucleus (jittered; $n = 1,379,011$ Mono, 91,022 DC). Only 0.04% (Mono) and 1.04% (DC) of nuclei have the peak accessible and the gene expressed, so the modes are indistinguishable visually.

In addition, I believe the authors may be somewhat overstating the categorisation of their peaks. Specifically, significance of hurdle model components is determined by 'Holm-corrected confirmation within each selected pair' – so if only one component is significant after this correction it is classified as either 'switch' or 'rheostat'. However, the ability to detect associations within each component will be somewhat determined by the number of expressing or non-expressing cells, and it is possible that a switch peak with only a few expressing cells in their study may in fact act as a dual peak (significant in both components), but would only be detectable in the count-component if they had collected more expressing cells from a larger number of donors. Hence, I think it is important to convey the potential uncertainty in their classifications to the reader.

Since the count component is estimated only among expressing cells, a peak–gene link called "switch" in a low-abundance context may be an underpowered "dual". Indeed, our BBJ replication shows exactly this pattern (**Supplementary Fig. 6f**): among the most significant FinnGen dual links ($P_{\text{joint}} < 1 \times 10^{-20}$), 57% retained the dual assignment in BBJ and 43% were detected as a single mode under BBJ's reduced power ($n = 98$), predominantly rheostat (30%) over switch (13%). Switch/rheostat flips remained rare ($< 2\%$). Reduced power thus collapses dual links to a single mode rather than reshuffling switch and rheostat. We partially mitigated this by restricting count-component tests to pairs with $\geq 5,000$ expressing cells. We modified the **Methods** to add this caveat.

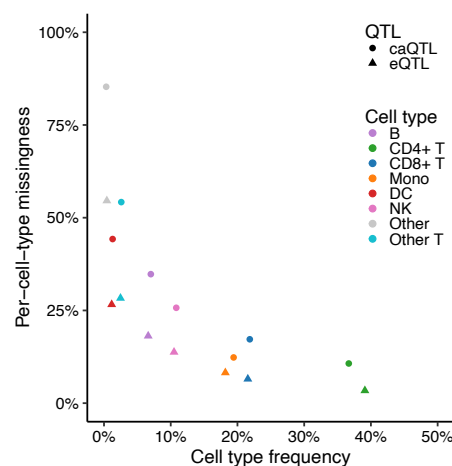
Regulatory-mode assignments reflect the components detectable at current statistical power. Because the count component is estimated only among expressing cells, a link with few expressing cells may be classified as switch (or, symmetrically, a link with few non-expressing cells as rheostat) when it would be dual with greater power. Single-mode assignments in low-abundance contexts should therefore be interpreted as a lower bound on the number of active components rather than as evidence that the other component is absent.

2 (Previous comment 8). I would like to thank the authors for the analysis already performed to test the robustness of their CASCADE results. I believe the CASCADE method is potentially very useful for determining QTL cell type-specificity and regulatory mechanisms. However, I still feel that there are additional details that need to be presented and explored.

In their response, the authors state “Across L1 cell types, 18% and 32% of entries in the strong set used for data-driven covariance estimation were imputed for eQTLs and caQTLs, respectively. Missingness was concentrated in rarer cell types with fewer nuclei per donor and was negligible in well-powered cell types”. I believe it is important to present to the reader what the missingness rates are per cell type and quantitatively show how this relates to cell-type abundances. I believe this is especially necessary for L2 annotations, where missingness rates may be higher but have not yet been quantified. In addition, although the authors present an analysis where they restricted to tests observed in only 50% of cell types, it would be important to test at what point the estimates become unreliable by titrating this across various rates of missingness, also showing how this relates to each cell type and it’s relative abundance (also at L2 level).

We thank the reviewer and apologize for any confusion in our earlier reply. To clarify, peak- and gene-level cell type specificity is determined jointly by per-cell-type significance (FDR) and mash LFSR. Significance is assessed at both L1 and L2, and significant L2 subtypes are mapped to their L1 parents before LFSR is applied, so the separately fit L2 mash results actually never enter categorization. At L1, LFSR only demotes a candidate specific feature to “likely shared but underpowered”, so imputation can never create specificity and its influence is bound to this single demotion regardless of the missingness rate. We now state this explicitly in the **Methods** and address the two requests below.

We now show per-cell-type missingness for eQTLs and caQTLs (new **Supplementary Fig. 26**). Missingness is near-perfectly anti-correlated with cell type frequency (Spearman’s $\rho = -0.98$ for eQTL and -0.95 for caQTL), negligible in well-powered cell types and concentrated in the rare aggregate categories that contribute little signal.

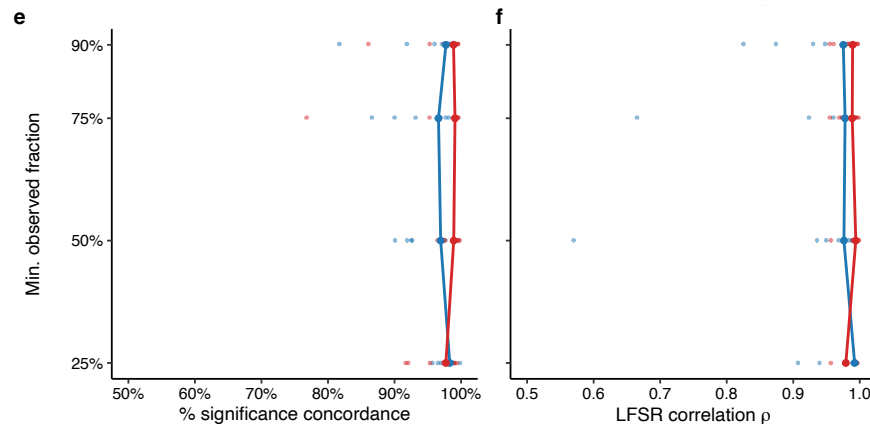


Supplementary Fig. 26 | Per-cell-type missingness in the mash strong set and its relationship to cell type frequency.

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.

We then titrated the observed-fraction threshold (variants observed in $\geq 25\%$, 50% , 75% , and 90% of L1 cell types) for both eQTLs and caQTLs, re-estimating the covariance and posteriors at

each threshold. Agreement with the full run stayed high across the entire range. Overall significance-call concordance (LFSR < 0.05) was $\geq 96\%$ for eQTLs (97%, 99%, 96%, and 97% at 25%, 50%, 75%, and 90% observed, respectively) and $\geq 95\%$ for caQTLs (98%, 95%, 95%, and 95%), and median per-cell-type LFSR correlations were ≥ 0.97 for both (new **Supplementary Fig. 27e,f**). The only appreciable degradation was confined to the rare aggregate “other” category, the same low-signal bin in which missingness concentrates. We did not find any threshold that makes the estimates unreliable.



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

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.

To clarify these points, we modified the **Methods** as follows:

Across L1 cell types, 18% and 32% of entries in the strong set were imputed for eQTLs and caQTLs, respectively, with missingness anti-correlated with cell-type abundance and concentrated in rarer cell types with fewer nuclei per donor (**Supplementary Fig. 26**). To assess robustness, we titrated the strong set to variant–gene/peak pairs observed in at least 25%, 50%, 75%, and 90% of L1 cell types, re-estimating the covariance and posteriors at each threshold. LFSR values remained highly concordant with the full run across the entire titration, with no threshold at which estimates became unreliable (**Supplementary Fig. 27**).

...

Feature significance is set by per-cell-type FDR, and mash LFSR is used only to distinguish genuinely specific features (categories 3–5) from those likely shared but underpowered (category 2). Consequently, imputation in the strong set can only shift a feature toward the shared, underpowered call and cannot create or sharpen specificity. This LFSR consolidation is performed at L1, with L2 subtypes aggregated to their L1 parents (see above).

3 (Previous comment 9). I believe the additional benchmarking of colocalisations strengthens reliability in the authors conclusions. However, I still feel it would be of value to highlight in the results section of the manuscript which of the associations the authors describe were detectable in other studies. I agree that this study adds a lot of novelty, even for associations that are previously described, but believe this detail is important to make readily available to the reader.

We thank the reviewer for this suggestion. Our colocalization benchmark against 287 eQTL Catalogue datasets already annotates every gene–trait colocalization for external replication, and

Supplementary Table 29 provides these per-association annotations, including the maximum PP.H4, whether the colocalization was replicated (PP.H4 > 0.8) in bulk immune, bulk non-immune, single-cell immune, and single-cell non-immune datasets, the best external study for each, and an overall novelty flag (6,808 of 11,015 [62%] detectable in eQTL Catalogue, 4,207 [38%] novel). To make this readily available to the reader, we now point to this per-association annotation in the **Results**, so the external replication status of any highlighted association can be looked up directly. We prefer this to labelling each locus as novel at its introduction, which would clutter the narrative, while still making the detail accessible.

That said, we summarize below the external replication status of every colocalization highlighted in the **Results**, benchmarked against the 287 eQTL Catalogue datasets (**Supplementary Table 29**). Five of the ten were detectable in at least one external eQTL dataset, and five are novel to this study. Novelty is concentrated among cell-type-specific cascades (*TICAM1* in CD16⁺ monocytes, *RHOH* in CD4⁺ T cells) and Finnish-enriched variants (*TNRC18*, *IL21R*) that bulk or non-Finnish datasets cannot resolve. Even the detectable colocalizations gain a cell-type assignment, a chromatin-accessibility layer, and a CASCADE mechanism absent from the external bulk studies.

Gene	Disease	Cell type	External replication
<i>GRAMD1B</i>	CLL	B	Detectable
<i>FAS</i>	CLL	CD16 ⁺ Mono, B	Detectable
<i>RNASET2</i>	Skin cancer	Mono	Detectable
<i>APOBEC3A</i>	Breast cancer	Mono	Detectable
<i>CBX6</i>	Breast cancer	Mono	Detectable
<i>TICAM1</i>	AIHT, skin cancer	CD16 ⁺ Mono	Novel
<i>RHOH</i>	AIHT	CD4 ⁺ T	Novel
<i>PILRB</i>	Alzheimer's disease	Mono	Novel
<i>TNRC18</i>	IBD	CD4 ⁺ T	Novel
<i>IL21R</i>	AIHT	NK, CD4 ⁺ T	Novel

We have revised the **Results** accordingly.

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).

4 (Previous comment 11). I appreciate the replication across independent datasets and down sampling experiments, I believe these tests strongly reinforce the conclusions. However, I believe the results presented in the new Supplementary Figure 18a are potentially misinterpreted. The authors now show that the slope for the relationship between gene constraint and the proportion of genes with colocalization is negative in each cell type level, but with quite dramatically different gradients – being much steeper for L1 and L2 annotations and with higher overall colocalization rates. If this is true, I believe this does indicate that the resolution at which eQTLs are mapped has an influence on: 1) overall colocalization rates, and 2) the likelihood that these effects are identified for a gene under higher constraint. Can the authors please test for a linear association between LOEUF decile and proportion of genes with a GWAS coloc with an interaction term for the cell-type level? If the interaction term is significant, this would indicate that the resolution at which eQTLs are mapped is a contributing factor, and their current presentation of the association being ‘robust to cell-type resolution’ is misleading.

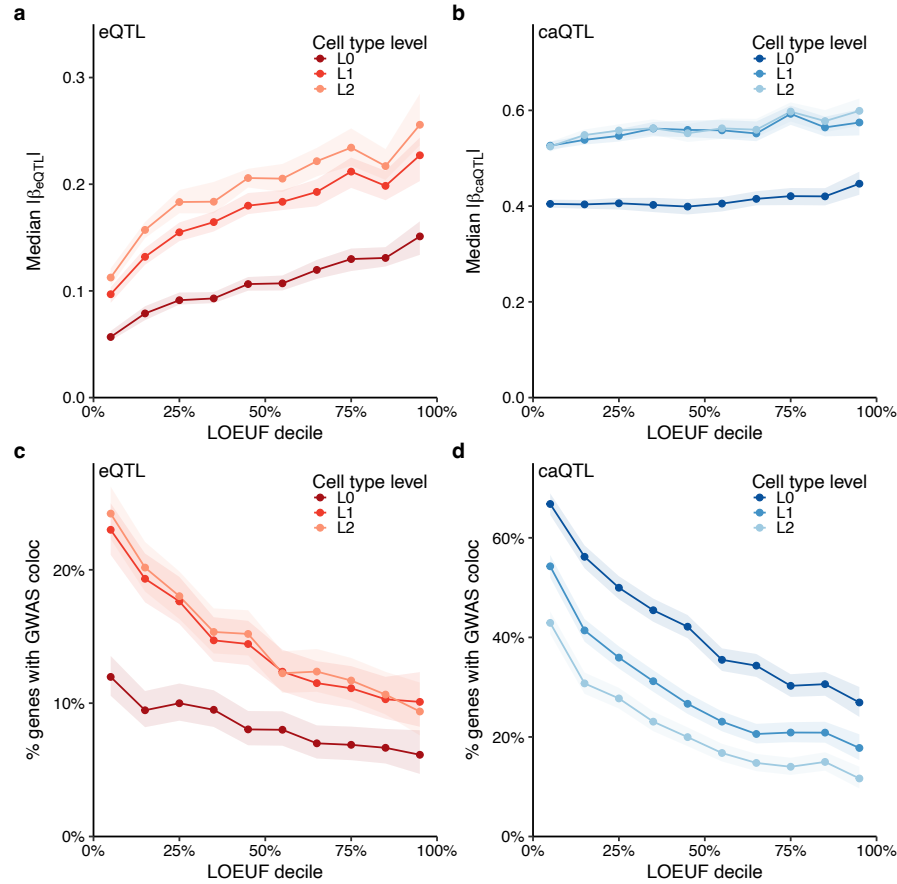
We would first clarify that "robust to cell-type resolutions" was intended to refer to the eQTL effect-size gradient (regulatory buffering), for which we unintentionally cited an incomplete figure (original **Supplementary Fig. 18**; cell-type-level colocalization gradient only). We have now added panels showing that this constraint gradient in $|\beta_{\text{eQTL}}|$ is near-constant across cell-type resolutions (Spearman's $\rho = 0.26, 0.23, \text{ and } 0.22$ for L0, L1, L2, respectively; new **Supplementary Fig. 18a,b**). The buffering architecture is thus present with essentially the same gradient at any resolution.

As requested, we also performed the LOEUF-decile $\times$ cell-type-level interaction test on the proportion of genes with a GWAS colocalization (gene-level logistic regression across L0, L1, and L2; 15,976 genes). The interaction is significant (likelihood-ratio $P = 3.1 \times 10^{-3}$), which we now report in **Results**, and we address the reviewer's two questions below:

1. **Overall colocalization rates.** Resolution does raise the overall rate for eQTLs (8.5%, 14.8%, and 15.4% for L0, L1, and L2, respectively), but through detection power, not the buffering architecture. The constraint gradient is negative and highly significant at every resolution (slope = $-0.08, -0.12, -0.12$; $P = 1.4 \times 10^{-13}$ to 3.1×10^{-48} ; Spearman's $\rho = -0.96$ to -1.0).

2. **Identifying effects at constrained genes.** Higher resolution reveals rather than creates these effects. The effect-size gradient is invariant across resolutions ($\rho = 0.26, 0.23, 0.22$; new **Supplementary Fig. 18a,b**), so the constrained-gene attenuation is not resolution-dependent. The steeper colocalization slope at finer resolution reflects recovery of more of this existing buffered signal as PBMC (L0) mapping stops masking cell-type-specific eQTLs, with saturation at cell-type resolution (L1 to L2 negligible), consistent with our downsampling and cross-cohort replication (**Supplementary Figs. 18c,d, 19, 20**).

We therefore retain "robust to cell-type resolutions", now backed by the effect-size gradient panels (new **Supplementary Fig. 18a,b**), and add one clarifying sentence after the colocalization result (**Fig. 5f**) reporting the interaction (**Supplementary Fig. 18c,d**).



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.

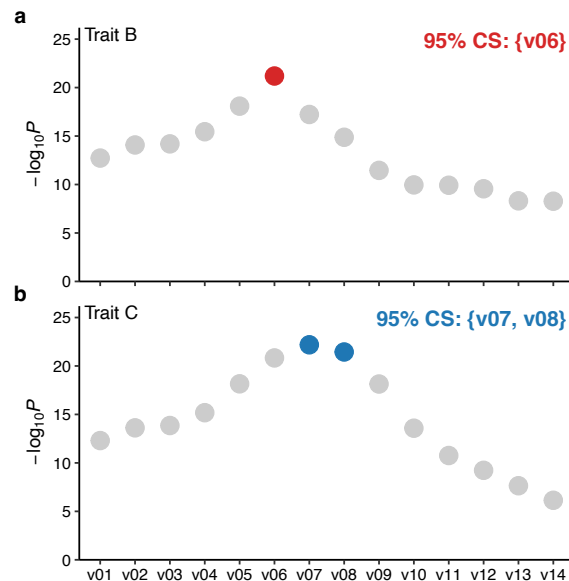
To address these points, we modified the **Results** accordingly.

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 $|\beta_{eQTL}|$ 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.

5 (Previous minor comment 6). I appreciate the clarity on this, but not sure it makes sense. In the toy example, A–B, B–C, and D–E are colocalised, but the CS of C doesn't share a variant with B? How would it be colocalised in this instance?

We thank the reviewer for asking this question. We would like to clarify that colocalization by PP.H4 is a probabilistic statement that two signals share a causal variant given the local LD structure. It does not by itself require the two 95% CSs to contain a common variant. In high LD, two colocalizing signals can be fine-mapped to distinct but correlated variants, so their 95% CSs need not intersect.

To make this concrete, we simulated an example locus and applied coloc.susie (Wallace et al., 2021): two signals colocalize at PP.H4 = 0.95, yet their 95% CSs, {v06} and {v07, v08}, share no variant (**Response Fig. 2**). This is exactly why we enforce variant sharing as a separate, more stringent step, so that each merged CS is anchored by a genuinely shared variant and chains of pairwise colocalizations do not inflate merged CSs into overly large clusters. In the toy example, C therefore colocalizes with B (PP.H4 > 0.8) without sharing a variant with it, and is retained as a singleton because no variant is common to A, B, and C.



Response Fig. 2 | Colocalization does not require shared credible-set variants.

Two simulated signals in a single LD block, fine-mapped and colocalized with coloc.susie (Wallace et al., 2021). Both traits show peaks in the same region and colocalize at PP.H4 = 0.95, yet their 95% credible sets (B, {v06}, red; C, {v07, v08}, blue) contain no variant in common. The two causal variants are in high LD ($r = 0.88$), so the shared-signal (H4) hypothesis is strongly favored even though fine-mapping assigns each trait a distinct credible set. Points show $-\log_{10}P$ per variant.

We have modified the **Methods** to make this explicit.

We note that colocalization by PP.H4 is a probabilistic statement that two signals share a causal variant given the local LD structure and does not by itself require the two 95% credible sets to contain a common variant, since in high LD two colocalizing signals can be fine-mapped to distinct but correlated variants.

Referee #1 (Remarks on code availability):

I have installed the fasthurdle singularity container and the cascade package. Install of cascade was a slightly difficult as ``cascade/src/variant_qtl_patterns.cpp`` uses ``IntegerVector::create()`` with 25 arguments, which exceeds the 20-argument limit in older Rcpp versions. To overcome, I changed the default argument for ``pattern_to_mechanism`` from ``IntegerVector::create(1, 2, ...)`` to a C++11 initializer list (``={1, 2, ...}``). My exact commands were:

```
git clone https://github.com/mkanai/cascade.git
mkdir -p cascade_tmp
Rscript -e 'dir.create("cascade_tmp", showWarnings = FALSE); install.packages(c("R.utils",
"qs2"), lib = "cascade_tmp", repos = "https://cloud.r-project.org")'
R CMD INSTALL -l cascade_tmp cascade
```

I did not extensively test packages though.

We thank the reviewer for testing both packages and suggesting a solution. We have modified the CASCADE package in the latest release (<https://github.com/mkanai/cascade>) to fix this Rcpp version incompatibility issue.

Referee #2:

I would like to thank the authors for their thorough revision. They have addressed most of my previous concerns in a serious and data-driven manner. The new replication analyses, the integration of MPRA data, and the clinical proof-of-concept are valuable additions. I believe the manuscript is now closer to being acceptable for publication in Nature.

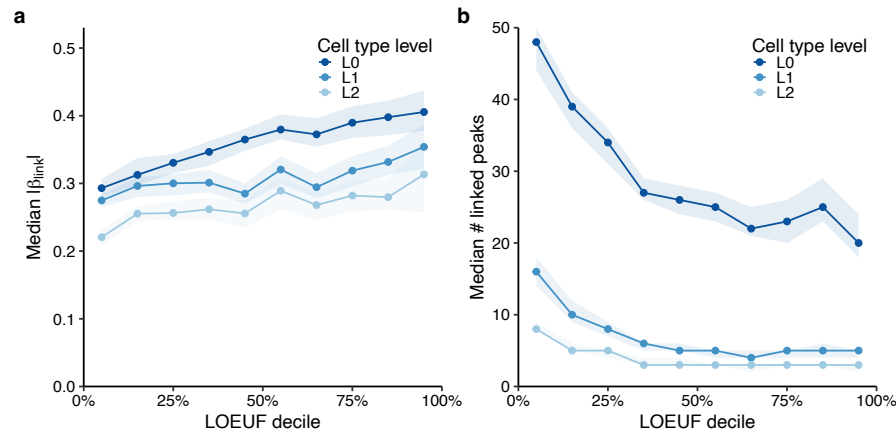
We thank the reviewer for the positive assessment and for recognizing the value of the new replication, MPRA, and clinical analyses. We address the two remaining issues below.

However, I still have two remaining issues that require attention. These are not fatal flaws, but they need to be clarified or further discussed. I recommend a minor revision.

First, the authors need to be clearer about the cell type specificity of their peak-to-gene links. The manuscript analyses eQTLs and caQTLs separately for each cell type. It is natural to ask whether the peak-gene links are also cell type specific. When the authors state that they identified 593,765 unique peak-gene links, does this mean each link is assigned to a specific cell type? Or are these links aggregated across all cell types? The current text does not make this clear. The same ambiguity affects the central claim that constrained genes have more but weaker enhancer-gene links. Is this statement based on the cell type level or the tissue level? For example, the FAS variant rs2147420 is reported to increase FAS expression in a cell type dependent manner. Does the corresponding peak-gene link also exist only in those specific cell types? I would like the authors to confirm that every peak-gene link is indeed cell type specific and that the buffering observation holds at that resolution. To help readers query this information, I suggest the authors provide a supplementary table with each significant peak-gene link as rows and each cell type as columns, indicating presence or absence. This would greatly improve the utility of the resource.

We thank the reviewer and confirm that the peak-gene links are cell-type-specific. Each link is determined within an individual cell type using that per-cell-type hurdle model. The 593,765 unique peak-gene pairs are the union of 1,095,949 significant peak-gene-cell-type trios, so every link is assigned to one or more specific cell types.

The buffering observation ("more but weaker links at constrained genes") likewise holds across cell-type resolutions and is not an artifact of cross-cell-type aggregation. Since **Fig. 5h,i** display aggregated per-gene summaries, we recomputed the link-count and $|\beta_{\text{link}}|$ gradients separately at each resolution and recapitulated constrained genes have more but weaker links at PBMC (L0), L1, and L2 alike (new **Supplementary Fig. 22**). The gradient therefore reflects genuine cell-type-level regulatory architecture, not the union of links across cell types. We also verified that the *FAS* peak–gene link is significant in the same cell types in which rs2147420 exerts its eQTL effect, including CD16⁺ monocytes and B cells, matching the eQTL's cell-type dependence.



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 ($|\beta_{\text{link}}|$; **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.

On the requested table, a static presence/absence matrix of all 593,765 links $\times$ 29 cell types would be too large and too sparse to be useful. We instead make the per-cell-type link information available in two ways. First, we release the complete peak–gene link summary statistics (all significant peak–gene–cell-type trios with their regulatory mode and zero/count-component effect sizes and *P* values) upon publication, distributed alongside our eQTL and caQTL summary statistics. Second, our CASCADE browser (for reviewer access: <http://35.238.169.158:8080>; to be published at <https://cascade.finngen.fi/>) now displays, for any gene or peak, its peak–gene links across all cell types interactively, with per-cell-type chromatin-accessibility tracks, zero/count effect sizes and *P* values, regulatory modes, and CASCADE categories, directly delivering the query utility the reviewer requested.

We modified the **Results** accordingly.

Leveraging our large-scale paired multiome measurements, we systematically mapped 593,765 peak–gene links with significant peak–gene associations in *cis* (i.e., < 1 Mb between TSS and peak) within each cell type using a hurdle model to handle both zero-inflation and overdispersion in single-cell gene expression³⁵, with significance determined via a joint χ^2 score test with stage-wise confirmation⁴⁴ controlling FDR at $\alpha = 0.05$ (**Fig. 1e**, **Supplementary Fig. 2**, **Supplementary Table 4**, **Methods**).

On buffering:

This more-but-weaker-links pattern was robust to cell-type resolutions, holding at PBMC (L0), L1, and L2 alike (**Supplementary Fig. 22**).

Data availability:

The FinnGen single-nucleus **caQTL**, **eQTL**, **peak-gene link** summary statistics are currently subject to embargo according to the FinnGen consortium agreement. They are available upon request and being prepared for public release by Q4 2026 or at the time of publication. ... The interactive CASCADE browser is available for reviewer access at <http://35.238.169.158:8080/> and will be publicly available upon public release.

Second, the authors need to address alternative explanations for their buffering model. Their conclusion that constrained genes have more but weaker links, and that this buffers expression changes, is elegant. But it is not the only possible interpretation of the data. I would like the authors to discuss at least three alternatives in the main text or supplement. First, the observed gradient could partly reflect detection bias, because constrained genes have lower expression variance, which reduces standardized effect sizes even if the true biological effect is similar. Second, the high GWAS colocalisation rate could reflect reverse causation or confounding. The disease associated variant might act through other mechanisms or other tissues, and the eQTL signal in PBMCs might be an echo rather than the causal path. Third, the comparison of effect sizes between constrained and unconstrained genes might be affected by winner's curse, because only the most significant small effects survive in the constrained group. The authors have performed some sensitivity analyses, but they do not explicitly address these alternative interpretations. I encourage them to discuss these possibilities and, if possible, provide additional analyses such as bidirectional Mendelian randomisation to test the direction of causality. Even if they cannot completely rule out these alternatives, a balanced discussion would strengthen the paper.

Overall, I find the work impressive and the revision thorough. The two points above are important but manageable. I look forward to seeing the revised version with these clarifications and discussions added. I support publication after these minor changes.

We thank the reviewer for raising this point and the generous overall assessment. We address each alternative explicitly below, drawing on analyses that were in part designed to discriminate against them. We also add a forward-looking discussion of candidate mechanisms to the **Discussion**.

1. Detection bias from lower expression variance. The gradient survives simultaneous matching on mean expression and expression-variance residuals (attributing approximately two-thirds of the residual to regulatory redundancy; **Extended Data Fig. 9d**), and MPRA-measured intrinsic regulatory activity is constant across LOEUF (**Extended Data Fig. 10c–f**), so equal *cis*-regulatory input is transmitted into progressively smaller expression changes at constrained genes. The pattern also holds for MAF-normalized effect sizes (**Extended Data Fig. 9f,g**).

2. Reverse causation and confounding. Because the instrument is a germline variant fixed at conception, the direction from variant to expression is anchored. Disease can alter expression, but not the germline genotype used as the instrument, so the eQTL is not an artifact of reverse causation, and a reverse-direction MR would test a different question. The analytical concern is

horizontal pleiotropy or tissue mismatch, which we address with SMR-HEIDI (distinguishing mediation from linkage or pleiotropy; **Methods**) and cell-type-matched colocalization, with mediation strongest for complete cascades (SMR-to-colocalization ratio rising from 21% for chromatin-only effects to 58% for local cascades; **Supplementary Fig. 11c**). The buffering gradient is, moreover, measured independently of disease.

3. Winner's curse. Our estimand compares effect-size distributions between constrained and unconstrained genes, and constrained-gene effects are smaller. Winner's curse inflates just-significant small effects upward and would therefore attenuate, not create, the gradient. Empirically it replicates out of sample and under downsampling (**Supplementary Figs. 19, 20**).

We also add a brief, forward-looking discussion of mechanism. Our data localize the buffer to the chromatin-to-expression transmission interface (normal caQTL, attenuated eQTL, constant MPRA activity) but do not identify the carrier. Consistent with our finding that constrained genes carry more but weaker enhancer–gene links, regulatory redundancy (multiple enhancers averaging out any single perturbed element) may contribute (Osterwalder et al. 2018; Wang & Goldstein 2020; Floc'hlay et al. 2021); additional, non-exclusive layers include intrinsic promoter architecture (Boukas et al. 2020; Einarsson et al. 2022), transcriptional burst kinetics (Larsson et al. 2019; Zuin et al. 2022), reduced expression noise and canalization (Siegal & Bergman 2002; Hermisson & Wagner 2004; Lehner 2008), and post-transcriptional regulation (Brümmer & Bergmann 2024). We present these as open hypotheses in the **Discussion**, with which layer predominates left to future work.

The molecular carrier of this buffering remains to be resolved. Consistent with our finding that constrained genes carry more but individually weaker peak–gene links, regulatory redundancy, in which multiple enhancers average out the effect of any single perturbed element, may contribute^{93,98,99}, additional, non-exclusive layers include intrinsic promoter architecture^{100,101}, transcriptional burst kinetics^{102,103}, reduced expression noise and canalization^{104–106}, and post-transcriptional regulation¹⁰⁷. Which of these predominates at constrained genes remains an open question for future work.

Referee #3 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

Referee #3 (Remarks on code availability):

The authors have made all code publicly available. All links listed in the 'Code Availability' section are currently accessible. Furthermore, the provided installation instructions are detailed and have been verified to be fully functional.

We thank the reviewer for co-reviewing our manuscript and assessing our code repositories.

Referee #4:

The authors have addressed the majority of my concerns. I have no further comments.

Referee #4 (Remarks on code availability):

I examined the github files. They are accessible and well annotated. I did not run any analysis independently. I'm sorry I don't have time for that.

We thank the reviewer for the positive assessment and all the review comments that have strengthened our manuscript.

References:

(Siegal and Bergman, 2002; Hermisson and Wagner, 2004; Lehner, 2008; Osterwalder *et al.*, 2018; Larsson *et al.*, 2019; Boukas, Bjornsson and Hansen, 2020; Wang and Goldstein, 2020; Floc'hlay *et al.*, 2021; Wallace, 2021; Einarsson *et al.*, 2022; Zuin *et al.*, 2022; Brümmer and Bergmann, 2024)

Boukas, L., Bjornsson, H.T. and Hansen, K.D. (2020) 'Promoter CpG density predicts downstream gene loss-of-function intolerance', *American Journal of Human Genetics*, 107(3), pp. 487–498.

Brümmer, A. and Bergmann, S. (2024) 'Disentangling genetic effects on transcriptional and post-transcriptional gene regulation through integrating exon and intron expression QTLs', *Nature Communications*, 15(1), p. 3786.

Einarsson, H. *et al.* (2022) 'Promoter sequence and architecture determine expression variability and confer robustness to genetic variants', *eLife*, 11(e80943). Available at: <https://doi.org/10.7554/eLife.80943>.

Floc'hlay, S. *et al.* (2021) 'Cis-acting variation is common across regulatory layers but is often buffered during embryonic development', *Genome Research*, 31(2), pp. 211–224.

Hermisson, J. and Wagner, G.P. (2004) 'The population genetic theory of hidden variation and genetic robustness', *Genetics*, 168(4), pp. 2271–2284.

Larsson, A.J.M. *et al.* (2019) 'Genomic encoding of transcriptional burst kinetics', *Nature*, 565(7738), pp. 251–254.

Lehner, B. (2008) 'Selection to minimise noise in living systems and its implications for the evolution of gene expression', *Molecular Systems Biology*, 4(1), p. 170.

Osterwalder, M. *et al.* (2018) 'Enhancer redundancy provides phenotypic robustness in mammalian development', *Nature*, 554(7691), pp. 239–243.

Siegal, M.L. and Bergman, A. (2002) 'Waddington's canalization revisited: developmental stability and evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 99(16), pp. 10528–10532.

Wallace, C. (2021) 'A more accurate method for colocalisation analysis allowing for multiple causal variants', *PLoS genetics*, 17(9), p. e1009440.

Wang, X. and Goldstein, D.B. (2020) 'Enhancer Domains Predict Gene Pathogenicity and Inform Gene Discovery in Complex Disease', *American journal of human genetics*, 106(2), pp. 215–233.

Zuin, J. *et al.* (2022) 'Nonlinear control of transcription through enhancer-promoter interactions', *Nature*, 604(7906), pp. 571–577.