Atlas of cell type-specific post-transcriptional genetic impacts in immune-related diseases

Ting Zhang[1]†, Hui Chen[1,2]†, Shuxin Chen[1]†, Ruofan Ding[1], Dinglin Luo[1], Xudong Zou[1], Kewei Xiong[1], Yinuo Wang[1], Shibin Hu[3], Jian Yang[4], Gao Wang[5], Stephen Montgomery[6], Qin Li[7]*, Lei Li[1]*

[1]Institute of Systems and Physical Biology, Shenzhen Bay Laboratory; Shenzhen, 518055, China.
[2]Department of Physiological Science, School of Medicine, Universitat de Barcelona (UB); Barcelona, 08907, Spain.
[3]Institute of Bio-Architecture and Bio-Interactions (IBABI), Shenzhen Medical Academy of Research and Translation (SMART); Shenzhen, 518107, China.
[4]School of Life Sciences, Westlake University; Hangzhou, 310024, China.
[5]Department of Neurology, The Gertrude H. Sergievsky Center, Columbia University; New York, NY 10032, USA.
[6]Departments of Pathology and Genetics, Stanford School of Medicine; Stanford, CA 94305, USA.
[7]Department of Genetics, Perelman School of Medicine, University of Pennsylvania; Philadelphia, PA 19104, USA.

*Corresponding author. Email: lei.li@szbl.ac.cn or qinli@pennmedicine.upenn.edu

†These authors contributed equally to this work

Abstract

Genetic variants contribute to immune disease risk through complex, cell type-specific regulatory mechanisms, yet post-transcriptional processes including alternative polyadenylation (APA) and RNA editing remain poorly characterized. Using single-cell RNA-sequencing of >7.27 million peripheral blood mononuclear cells from 2,022 individuals, we mapped 32,763 single-cell quantitative trait loci (sc-xQTLs) across 31 immune cell types and diverse stimulation conditions. Unexpectedly, APA and RNA editing QTLs are more strongly enriched for immune disease heritability than expression QTLs. Immune stimulation, rather than sex or ancestry, drove 73.65% context-dependent sc-xQTLs. We identified 382 putative causal genes, 85.08% acting through expression-independent mechanisms. Integration with protein QTLs further identified 32 disease risk genes, including 19 druggable targets. Functional assays confirmed the allelic effect of sc-xQTL in shortening *DLD* 3′UTR, leading to reduced protein levels and increased ulcerative colitis risk. This study reveals a previously underappreciated role for post-transcriptional regulation in immune disease susceptibility.

Main Text

Many immune-related diseases, such as autoimmune diseases, are highly heritable, with a median heritability of approximately 60% (*1*). While genome-wide association studies (GWASs) have identified thousands of genetic loci associated with various immune-related diseases, such as autoimmune thyroid disease (AITD) (*2*), type 1 diabetes (T1D) (*3*), and rheumatoid arthritis (RA) (*4*), deciphering the underlying mechanisms remains a significant challenge. Notably, while some disease-associated variants affect coding regions of immune genes, over 90% reside in non-coding regions (*5*), especially active regulatory regions, such as enhancers and untranslated regions, underscoring their predominant roles in transcriptional and post-transcriptional regulation.

Expression quantitative trait loci (eQTL) analysis has been successful in linking genetic variations to their molecular mechanisms (*6*). However, most eQTL studies have been performed using bulk tissues, limiting the identification of cell-specific regulatory effects and their functional interpretation (*7*). Recent advances in single-cell eQTL (sc-eQTL) analyses across multiple cell types and stimulation conditions (*7-10*) have significantly enhanced our understanding of the genetic regulation of immune-related diseases. For instance, the systemic lupus erythematosus (SLE)-associated eQTL rs2736336 modulates *BLK* gene expression specifically in memory B cells, promoting B-cell hyperactivity and enhancing T-cell costimulatory capacities (*11*). In addition, many disease-associated variants only manifest their effects in response to external stimuli (*5*). These context-specific eQTL studies represent critical gene-by-environment interactions where genetic effects are significantly altered by the cellular environment. Additionally, recent efforts have also investigated the role of cell-type-specific regulation of alternative splicing in autoimmune disease (*12*). Despite these advances, current cell-type-specific eQTL studies are primarily limited to gene expression alone, overlooking the potentially critical role of post-transcriptional processes in mediating disease risk.

Beyond transcriptional control, post-transcriptional processes, particularly APA and Adenosine-to-Inosine (A-to-I) RNA editing, have emerged as critical regulators of immune function (*13-15*). APA, which affects ~70% of human genes, generates mRNA isoforms of varying 3′UTR lengths through alternative poly(A) sites, thereby significantly affecting the mRNA stability, translation, and cellular localization (*16*). A-to-I RNA editing, catalyzed by adenosine deaminases acting on RNA (ADARs) that bind to double-stranded RNA (dsRNA) substrates and convert adenosines to inosines, plays a crucial role in immune regulation (*17*). The ubiquitously expressed ADAR1 suppresses endogenous dsRNA sensing mediated by MDA5, a cytosolic dsRNA sensor (*18*). Our and other previous works demonstrated that genetic variants affecting APA and RNA editing levels are enriched in immune-related disease-associated loci (*13, 14, 19*) as exemplified by APA-mediated regulation of *IRF5* in SLE disease (*20*) and the association between reduced RNA editing levels and increased susceptibility to multiple autoimmune diseases (*14*). However, their cell type-specific contributions to immune-related disease pathogenesis remain unclear.

Here, we conducted a large-scale and comprehensive analysis of cell type-specific genetic regulation through generating sc-xQTLs, which include single-cell expression quantitative trait loci (sc-eQTLs), single-cell APA quantitative trait loci (sc-aQTLs), and single-cell RNA editing quantitative trait loci (sc-edQTLs), using single-cell RNA-seq (scRNA-seq) data from more than 7.27 million peripheral blood mononuclear cells collected from 2,022 individuals with genotype information. Our study identified 32,763 sc-xQTLs, linking at least one molecular phenotype with 17,422 nearby genes. By analyzing context-specific sc-xQTLs, we uncovered that stimulus-

dependent genetic effects primarily drive context-specific disease regulatory effects. Critically, colocalization and single-cell transcriptome-wide association study (sc-xTWAS) pinpointed potentially causal variants and genes of immune-related diseases in the relevant cell types that cannot be explained by expression differences alone. Furthermore, we developed a comprehensive web portal, the sc-xQTL atlas, to make the sc-xQTLs identified in this study widely accessible. Combined, these data provide an atlas of genetic effects for characterizing cell type-specific regulation in immune-related diseases.

A uniformly processed compendium of multiple molecular phenotypes across 7 million immune cells

To comprehensively dissect the gene regulatory cascades in human immune cells, we compiled, to the best of our knowledge, the largest immune single-cell dataset, comprising scRNA-seq data of 7.27 million cells from 2,022 donors of seven independent cohorts (*7-10, 21-23*) (Fig. 1A and figs. S1A–S1C). Our integrated dataset is vastly improved upon previous efforts in multiple aspects, including significantly increased number of cells, a more balanced sex ratio (female: male = 1.15:1), the inclusion of various immune stimulation conditions (Influenza A virus, SARS-CoV-2, and anti-CD3/CD28) and disease states (162 SLE and 126 tuberculosis cases), and more broad ancestral diversity (70.0% European, 12.7% Peruvian, 9.4% East Asian, 6.5% African, and 1.4% other ancestries). Using CellHint (*24*), we accounted for batch effects, potential cell mislabels, and library/technical differences across datasets and successfully integrated these diverse single-cell datasets into a harmonized compendium composed of 31 immune cell types (Fig. 1B and figs. S1D–E). The accuracy of our cell type annotations was validated by expression analysis of well-established marker genes (Fig. 1C).

To characterize molecular traits at both transcriptional and post-transcriptional levels in scRNA-seq data, we performed rigorous annotation of poly(A) and A-to-I RNA editing sites using both the publicly available databases (*25, 26*) and *de novo* identification approaches (fig. S2, see Methods for details). Our efforts resulted in high-confidence sets of 33,831 poly(A) sites and 62,362 RNA editing sites, respectively. We then employed scUTRquant (*27*) and SAMtools (*28*) to quantify poly(A) site usage and RNA editing levels based on pseudo-bulk analysis of each cell type per individual, respectively.

Our analysis characterized transcriptional and post-transcriptional molecular traits across 22,036 genes. Within each immune cell type, we identified an average of 12,285 expressed genes, 9,582 APA events, and 19,698 RNA editing sites (Fig. 1D). Compared to the OneK1K study, which focused solely on gene expression, we have a 5.76-fold increase in the total number of molecular traits (Fig. 1E). We tested the robustness of our molecular traits and found highly consistent measurements across different cohorts that indicated high reproducibility of the results (fig. S3A). Remarkably, 77.10% of APA events and 74.67% of RNA editing events showed insignificant association with gene expression (Pearson's correlation coefficient $|r| < 0.3$), suggesting that these post-transcriptional molecular traits are mostly independent of gene expression (Fig. 1D).

Upon immune stimulation, we observed extensive dynamic changes in APA and RNA editing events that were also largely independent of gene expression changes (Fig. 1F and fig. S3B; table S1). Notably, global 3′UTR shortening was observed upon virus infection, consistent with previous reports (*15*). For example, *CXCL9*, a chemokine induced by interferon-γ that recruits protective $CXCR3^+$ T cells (*29*), exhibited 3′UTR shortening under influenza A virus

(IAV) treatment (Fig. 1G). For Additionally, RNA editing levels were massively increased for most editing sites in SLE patients, likely attributed to the elevated *ADAR1* expression induced by the IFN immune response (*30*). For instance, *CTSS*, a cysteine protease implicated in multiple autoimmune diseases (*31*) exhibited elevated RNA editing levels in SLE patients (Fig. 1H).

In summary, we have constructed a comprehensive compendium of uniformly processed molecular phenotypes across millions of single immune cells as a valuable resource for understanding cell-type specific transcriptional and post-transcriptional regulation in the immune system. In company with our harmonized dataset, we developed SERA (Single-cell gene Expression, RNA editing, and APA), a streamlined computational framework for analyzing multiple transcriptional and post-transcriptional phenotypes in scRNA-seq data. SERA is publicly available at https://github.com/lilab-bioinfo/SERA.

sc-xQTLs significantly expanded cell type-specific genetic effects

To understand the impact of cell type-specific genetic effects in immune cells, we first employed a unified imputation procedure (fig. S4) using the Michigan Imputation Server (*32*), to genotype 11,023,719 common variants with minor allele frequency ≥ 5% across 2,022 individuals (see Methods for details). We used QTLtools (*33*) to identify 32,763 sc-xQTLs significantly associated with 17,422 genes in *cis* via at least one molecular phenotype (Fig. 2A; FDR < 0.05; see Methods for details). Functional pathway analysis revealed a strong enrichment of sc-xQTL-associated genes (hereafter denoted as sc-xGenes, including sc-aGenes, sc-edGenes, and sc-eGenes) for immune-related pathways, underscored by the T cell receptor signaling pathway (3.12-fold enrichment, $P = 2.77 \times 10^{-5}$), regulation of response to biotic stimulus (3.25-fold, $P = 8.59 \times 10^{-9}$), and positive regulation of cytokine production (2.13-fold, $P = 7.39 \times 10^{-8}$) (fig. S5). Notably, compared to matched bulk tissue xQTLs and the OneK1K cohort, sc-xQTLs exhibited mostly consistent directions of effects with significantly larger effect sizes (Fig. 2B and fig. S6).

To characterize the genomic architecture of sc-xQTLs, we employed torus (*34*) to calculate their enrichment across various functional annotations. Our analysis revealed that sc-xQTLs were predominantly enriched in genomic regions of functional relevance, as demonstrated by sc-aQTLs enrichment for 3′UTR regions, sc-edQTLs enrichment for ADAR binding regions, and sc-eQTLs enrichment for promoter regions (Fig. 2C). These distribution patterns were further supported by our meta-gene analysis of the relative positions of the lead sc-xQTLs over their associated genes (fig. S7A). Additionally, fine-mapping of 20,335 sc-xQTLs stratified by posterior inclusion probability (PIP > 0.5) showed that most of the fine-mapped sc-eQTLs (58.71%), sc-edQTLs (63.53%), and sc-aQTLs (71.66%) were located within ±100 kb of their respective transcription start sites, RNA editing sites, or poly(A) sites (Fig. 2D).

To further dissect the complex effects of shared genetic variants across multiple molecular phenotypes, we focused on the 36.6% of sc-aGenes and 20.3% of sc-edGenes that overlapped with sc-eGenes (Fig. 2E and fig. S7B), and employed a Bayesian model-based causal inference approach bmediatR (*35*) to measure the extent to which one molecular phenotype might act as an intermediary for another. Specifically, two relationship models were evaluated: (1) The causal mediation model, where the putative mediator (M) (e.g., APA, RNA editing) fully or partially accounts for SNP effects on the outcome (Y) (e.g., gene expression); and (2) The independent model, where SNP effects on both M and Y are independent (Fig. 2F). Using gene expression as the outcome (Y) and APA or RNA editing as a putative mediator (M), we found a large proportion of sc-xQTLs exhibit independent effects, suggesting weak mediation correlation (Fig.

2G). This is consistent with our multi-layer sc-xQTL colocalization result (fig. S7C). For example, rs15797 is both a sc-eQTL (β = -0.24, $P = 1.08 \times 10^{-15}$) and sc-aQTL (β = 0.42, $P = 5.47 \times 10^{-18}$) of *RHOF* in CD4$^+$T cell type. However, the variation of *RHOF* gene expression was independent of its APA usages (posterior probability = 0.96, Figs. 2H–J). After conditioning on the APA usage of *RHOF*, the residual *RHOF* expression remains significantly associated with rs15797 (β = -0.13, $P = 2.01 \times 10^{-8}$), which demonstrates their independent effects (Fig. 2K). Collectively, our analyses reveal mostly independent regulatory mechanisms among APA, RNA editing, and gene expression.

To delineate cell type-specificity of genetic effects, we employed the multivariate adaptive shrinkage method (mash) (*36*) to compare sc-xQTL effect size across all cell types. We identified 16,653 sc-xQTLs with significant cell-type specific effects (local false sign rate < 0.05) (Fig. 2L; table S2; see Methods for details). It is important to note that these cell type-specific sc-eQTLs showed strong enrichment for open chromatin regions in the matched cell types (2.96-fold enrichment compared to 0.52-fold for unmatched cell types), similar to observations in previous studies (*9, 37*) (fig. S8). Among these cell-type-specific sc-xQTLs, many were previously implicated in immune-related diseases. For example, *ITGA4,* which encodes an α4 integrin crucial for immune cell migration, has been validated as a therapeutic target for MS since a humanized monoclonal antibody, natalizumab, against ITGA4 is successful in disease treatment (*38*). *ITGA4* was found to be associated with a lead monocyte-specific sc-eQTL (rs4667282, lsfr < 0.05). The variant rs1449263 and rs1375493 (in high LD with rs4667282, r^2 > 0.91 and 0.95 in European population, respectively), previously linked to monocyte count and inflammatory bowel disease (IBD) (*39, 40*), modulates *ITGA4* expression specifically in monocytes, but not in other immune cell types (Fig. 2M and fig. S9A). This finding leads to a testable hypothesis that rs1375493 may modulate disease susceptibility by regulating α4 integrin expression that further influences monocyte migration, which agrees with the well-established role of monocytes in IBD etiology. Furthermore, we also uncovered previously uncharacterized, cell-specific genetic effects that linked post-transcriptional regulation with immune-related diseases. For instance, we found a CD4$^+$ T cell-specific sc-aQTL (rs927335) associated with the 3′UTR length of *CD44* (Fig. 2N and fig. S9B). This finding reveals a cell-type-specific regulation of *CD44* via alternative usage of polyadenylation and provides potentially new mechanistic insight into disease susceptibility, given the multifaceted role of CD44 in leukocyte recruitment to inflammatory sites and its strong associations with immune-related diseases such as SLE and RA (*41, 42*). Additionally, we identified a B cell-specific sc-edQTL (rs12582584) in *CLEC2D*, a regulator of interferon-gamma production (*43*) (Fig. 2O and fig. S9C).

Taken together, we have constructed a comprehensive atlas of cell type-specific genetic regulatory cascades across immune cells. Our findings, fully replicated by bulk tissue and previous sc-eQTL analyses, greatly expand the understanding of gene regulation at transcriptional and post-transcriptional levels. The sc-xQTLs identified in this study are publicly available on our dedicated and freely accessible web portal (http://bioinfo.szbl.ac.cn/sc-xQTLatlas).

Immune stimulation primarily drives context-specific regulatory effects

Immune cell activation, critical for immune system function, is often responsive to various environmental factors and contexts (*44*). To elucidate the context-specific role of transcriptional and post-transcriptional regulation of immune cells, we first applied a linear mixed model using

variancePartition (*45*) to delineate the relative contributions of immune stimulation, sex, and genetic ancestry to the inter-individual gene expression variation. Our analysis revealed that immune stimuli—including viral infection (IAV and COV), and T cell activation via anti-CD3/CD28—consistently accounted for the most significant proportion of gene expression variance, surpassing the effects of sex or ancestry (Fig. 3A). This finding was further supported by an unsupervised clustering analysis that predominantly grouped the cells by immune stimuli rather than sex or ancestry (Fig. 3B), highlighting the primary role of stimuli in driving context-specific gene expression. Similar observations were found for APA and RNA editing, further supporting the crucial role of immune stimuli in shaping transcriptional and post-transcriptional phenotypes in immune cells (fig. S10).

To characterize the context-specific genetic regulation, we mapped context-dependent sc-xQTLs across immune cells that were mainly driven by immune stimulation (Fig. 3C). We identified a median of 1,268 stimulus-dependent, 927 ancestry-dependent, and 44 sex-biased sc-xGenes (FDR < 0.05). Notably, stimulus-dependent sc-xQTLs exhibited larger effect sizes under stimulation conditions compared to baseline (Fig. 3D and fig. S11). We performed gene set enrichment analysis on these stimulus-dependent sc-xQTL-associated genes and observed significant enrichment for multiple immune-related pathways (Fig. 3E). For example, rs2130531, which was strongly linked to rs11015435 (lead SNP, $r^2 = 1$) and located in the IFN-stimulated response element (ISRE) of interferon-stimulated gene (ISG) *ARL5B*, significantly decreased *ARL5B* expression upon IAV and COV stimulation but not in non-stimulated cells (Fig. 3F), indicating stimulus-specific genetic regulation (Fig. 3G). Moreover, rs6846553 functions as a strong COV-dependent sc-aQTL to mediate 3′UTR elongation for *TNIP3*, an inhibitor of NF-κB activation (*13*) (Fig. 3H). rs5755840 is significantly associated with RNA editing increase of *APOL6* only upon IAV stimulation (Fig. 3I).

To characterize the ancestry-dependent genetic effects on gene regulation, we mapped i) population-biased sc-xQTLs, defined as genetic variants with different regulatory effects across populations (*6*); and ii) population-specific sc-xQTLs, defined as genetic variants with different allele frequencies across populations (*46*) (fig. S12A). We identified a total of 4,490 sc-eQTLs, 484 sc-aQTLs, and 366 sc-edQTLs across immune cells (FDR < 0.05) (Fig. 3J). For example, the sc-eQTL rs10850097 was significantly associated with the differential expression of the anti-viral double-stranded RNA (dsRNA) sensor *OAS3* in monocytes of European descendants ($P = 3.27 \times 10^{-10}$) but not African descendants ($P = 0.24$), a finding also replicated in an independent cohort (*47*) (fig. S12B). Moreover, the G allele of African-specific sc-aQTL rs7649736 was significantly associated with APA shortening of *SACM1L*, which regulates autophagosomal phosphatidylinositol-4-phosphate for xenophagy-directed bacterial clearance (*48*) (Fig. 3K). We experimentally validated this observation using 3′ rapid amplification of complementary DNA ends (RACE) in HEK-293 cells by transfecting two identical constructs with one nucleotide difference (A-to-G) at rs7649736 (table S3). The alternative G allele construct utilized the proximal poly(A) sites, which resulted in significantly shortened 3′UTR and reduced protein abundance (Student′ s t-test, $P = 0.015$) (Fig. 3L). Additionally, the European (EUR)-specific variant rs62011287 was associated with expression of *USP3*, which controls the innate immune signaling and inflammatory response through the NF-κB pathway (fig. S12C). EAS-specific variant rs148991166 was associated with RNA editing level of RAD51-AS1, which is a lncRNA involved in DNA repair inhibition and elevated in MS patients (*49*) (fig. S12D).

To characterize sex-mediated genetic effects on gene regulation, we identified sex-biased sc-xQTLs across various immune cell types (fig. S13) and identified 125 genes, 59 APA events, and

59 RNA editing sites with significant sex-biased genetic regulation (FDR < 0.05). For example, the allergy-associated risk variant rs7226035 ($P = 1.22 \times 10^{-19}$) exhibited a female-specific association with altered retinoic acid receptor alpha (*RARα)* expression in conventional dendritic cells (cDCs) ($P = 3.35 \times 10^{-5}$). RARα serves as a critical mediator in adaptive $CD4^+$ T cell immunity and additionally modulates the function of antigen-presenting cells (APCs), such as DCs, thereby shaping T cell responses (*50*).

We also performed pseudo-time analysis to characterize the dynamics of genetic effects and identified dynamic sc-xQTLs for 1,636 genes, 798 APA events, and 415 RNA editing sites (FDR < 0.05) using linear mixed models (*8*) that considered both linear (genotype × pseudotime) and quadratic interaction (genotype × pseudotime^2) (fig. S14). For example, during B cell maturation from naive to memory states, we uncovered dynamic *CD44* upregulation and the non-linear changes of genetic effects of sc-eQTL rs1885936. We also observed increased 3′UTR lengthening effects associated with sc-aQTL rs872375 in *PTPRCAP*, encoding CD45-associated protein that activates T- and B-cell signaling pathways, with effect size increasing from $\beta = 0.10$ in naive B cells to $\beta = 0.59$ in memory B cells throughout B cell maturation, consistent with its established role in affecting B lymphocyte proliferation and survival (*51*).

Collectively, our findings reveal that immune stimulation is the primary driver of context-specific genetic regulation in immune cells. Immune stimuli, in particular viral infections, amplify the genetic regulatory effects globally that lead to thousands of significant stimulus-specific sc-xQTLs. This observation extends from transcriptional to post-transcriptional levels, suggesting a universal impact of immune stimuli in regulating gene functions that can potentially provide critical insights into gene-environment interactions in immune-related diseases.

sc-xQTL explains a significant proportion of immune-related disease heritability

Prompted by our finding that a large fraction of the cell-type-specific and context-dependent sc-xQTLs were associated with genes implicated in immune-related disease risk, we sought to systematically evaluate the genetic contribution of sc-xQTL to immune-related diseases. We conducted stratified LD score regression (S-LDSC) analysis (*52*), conditioned on 52 baseline annotations, to assess the heritability enrichment of sc-xQTLs across 16 immune-related diseases and seven non-immune diseases as controls (Fig. 4A; table S4). In line with previous reports (*8, 9*), PBMC sc-eQTLs were highly enriched in most immune-related diseases. Interestingly, sc-aQTLs ($P = 0.0022$) and sc-edQTLs ($P = 3.5 \times 10^{-12}$) exhibited even greater heritability enrichment compared to sc-eQTLs in various immune-relevant cell types (Fig. 4A and figs. S15A–D), highlighting the critical role of post-transcriptional regulatory mechanisms in immune disease susceptibility. For example, primary biliary cholangitis (PBC) GWAS variants showed more substantial enrichment for sc-aQTLs in monocytes (fold-enrichment = 5.72 ± 1.32, $P = 1.51 \times 10^{-4}$, BH-adjusted $P = 0.0041$) than sc-eQTLs (fold-enrichment = 2.99 ± 0.58, $P = 4.21 \times 10^{-4}$, BH-adjusted $P = 0.0019$, from S-LDSC) (Fig. 4B). Quantile-quantile plots (QQ-plots) of GWAS *P*-values for sc-xQTLs and genome-wide SNPs further validated this result (fig. S15E). This finding aligned well with previous research that monocyte cell responses are actively involved in amplifying local inflammatory signals in PBC (*53*). Additionally, sc-xQTLs accounted for a significantly larger proportion of heritability compared to bulk QTLs from GTEx whole blood ($P = 0.03$ by Wilcoxon′s two-tailed test) (fig. S15F).

To further pinpoint potentially causal sc-xQTLs in immune diseases, we employed a colocalization approach utilizing SuSiE-COLOC (*54*) and coloc (*55*) by integrating summary

statistics of immune-related disease with sc-xQTLs, bulk-xQTLs, or pseudo-bulk PBMC-xQTLs, which combined all immune cell types, respectively (fig. S16A). Remarkably, sc-xQTLs colocalized with 62.88% of immune-related disease risk loci, considerably higher than bulk-xQTLs (25.90%) and pseudo-bulk PBMC-xQTLs (40.85%) (table S5). Specifically, 39.71% of loci uniquely colocalized with sc-xQTLs, in contrast to only 2.54% exclusively colocalized with bulk xQTLs (Fig. 4C), underscoring that cell-type specific analysis expanded putatively causal events that would be otherwise missed in bulk tissue analysis. We further assessed the contribution of each molecular phenotype to the disease and observed that sc-eQTLs, sc-aQTLs, and sc-edQTLs were colocalized with 55.67%, 25.37%, and 9.21% of the disease risk loci, respectively (fig. S16B), and varied across different cell types (Fig. 4D). This observation was exemplified by a sc-edQTL of *JAK1* (Janus Kinase 1, Fig. 4E), which was exclusively colocalized with COVID-19 in $CD4^+$ T cells (PP4 = 1.00) upon TB stimulation. JAK1 has been known as a key kinase responsible for phosphorylating STAT proteins to activate the IFN signaling pathway. Inhibitors of JAK1 have been approved for treating various autoimmune diseases (*56*). Our findings shed light on the potential role of post-transcriptional regulation of *JAK1* in mediating disease susceptibility of AITD.

To evaluate the disease relevance of context-specific sc-xQTLs, we performed an overlap analysis between sc-xQTL variants (or their LD proxies, $r^2 > 0.8$) and genome-wide significant GWAS loci ($P < 5 \times 10^{-8}$). This integration identified 286 context-specific sc-xQTLs spanning 238 GWAS risk loci (Fig. 4F; table S6). By including the context-specific sc-xQTLs, the detection power of colocalization events was significantly improved for immune-related diseases ($P = 4.55 \times 10^{-5}$), with milder and non-significant improvement for non-immune-related diseases ($P = 0.08$) (Fig. 4G). Notably, we identified a stimulus-dependent sc-aQTL for *SMC4*, which promotes inflammatory innate immune responses by enhancing the transcription of Nuclear Factor-kb essential modulator (NEMO), colocalized with PBC risk loci (*57*) (Figs. 4H–I).

Our analysis revealed that 84, 40, and 17 shared loci colocalized with at least two immune-related diseases, driven by underlying molecular effects of gene expression, APA, and RNA editing, respectively (fig. S16C). Moreover, among the 471 sc-xQTL colocalized disease risk loci, we observed a median of 86.61% of them were shared for the same disease, but the effects of each are independent (Fig. 4J). For example, *IRF5*, a well-known risk gene for autoimmune diseases like SLE (*20*) and PBC (*58*), colocalized with PBC risk via both sc-eQTLs and sc-aQTLs in stimulated monocytes (Fig. 4K). However, their underlying effects were independent (coloc $PP4_{\text{sc-aQTL and sc-eQTL}} = 0.00017$, Fig.4L).

By integrating GWAS data and fine-mapped sc-xQTLs data, we observed a significant improvement in the prioritization of causal genetic variants. Variants found in both GWAS and sc-xQTL credible sets showed significantly higher GWAS PIP than those absent from the credible sets (fig. S17A). This integration of GWAS and sc-xQTL credible sets substantially increased the number of risk loci with ≤ 5 putatively causal variants (fig. S17B). Collectively, our findings robustly demonstrate that genetic regulation of APA and RNA editing significantly contributes to immune-related disease heritability, at least on par with gene expression. Owing to the inclusion of both transcriptional and post-transcriptional genetic regulation, our harmonized colocalization analysis was able to significantly improve the interpretation of immune-related disease signals with cell-type resolution.

Identification and validation of immune susceptibility genes

To systematically identify and prioritize candidate susceptibility genes associated with immune-related diseases, we conducted a summary-based single-cell transcriptome-wide association study (sc-xTWAS) using FUSION (*59*). To optimize sc-xTWAS performance across 31 immune cell types, we employed five-fold cross-validation to select the best-fitting model (significant out-of-sample R^2; nominal $P < 0.05$) for predicting the expression, APA, or RNA editing usage for each gene event using nearby variants. This analysis generated 100,518 high-quality prediction models for 8,568 genes, 5,725 APA events, and 4,430 RNA editing sites (fig. S18A). To evaluate prediction accuracy, we calculated the correlation between predicted and observed heritable molecular phenotypes. Consistent with previous studies (*60, 61*), the average in-sample prediction accuracy was approximately 71.18%, demonstrating that the fitted models effectively captured most of the signals in predicted genetic effects (fig. S18B). By applying these prediction models to GWAS summary statistics from 16 immune-related diseases, we identified 498 expression-linked genes, 286 APA-linked genes, and 235 RNA editing-linked genes significantly associated with disease susceptibility (Fig. 5A; table S7). Comparative analysis of these susceptibility genes across diseases revealed distinct disease clusters, highlighting shared genetic mechanisms and potential common therapeutic targets (fig. S19A). Among the autoimmune diseases, autoimmune thyroid disease (AITD), type I diabetes (T1D), and IBD were mapped with the highest number of significant susceptibility genes and events, likely owing to their high heritability (*2, 3, 62*) (Fig. 5B).

Our analysis captured known disease risk genes, such as T-cell co-stimulation genes (*CD28*, *CD40*, *CTLA4*, and *ICOS*), interleukin receptor genes (*IL2RA*, *IL6R*, *IL6ST*, *IL7R*, *IL12RB1*, *IL12RB2*, *IL18R1*), IRF family (*IRF1*, *IRF3*, *IRF4*, *IRF5*, and *IRF7*), and pleiotropic 17q12-21 locus associated genes (*IKZF3*, *GSDMB*, *ORMDL3*, and *MED24*). Importantly, 56.02% of sc-eTWAS genes, 65.08% of sc-aTWAS genes, and 52.17% of sc-edTWAS genes were independently validated as susceptibility genes in 16 diseases by the Open Targets Platform (*63*) (fig. S19B) with significantly higher disease association scores than randomly selected genes ($P < 2.6 \times 10^{-16}$, Fig. 5D). Many of the sc-xTWAS genes are also shared across multiple immune diseases (figs. S19C–E). Notably, including APA and RNA editing identifies 187 additional susceptibility genes (Fig. 5C), many of which are important risk genes previously unexplained through expression-based approaches, such as *CD44* and *STAT6* (exclusively through sc-aTWAS), and *PTPN22* (exclusively through sc-edTWAS). Additionally, the sc-xTWAS and colocalization analyses jointly discovered 1,012 susceptibility genes (fig. S19F). Among these genes, 382 are mediated through APA and RNA editing mechanisms, mostly independent of expression level changes (85.08% no overlap with eQTL of LD $r^2 < 0.5$).

Pathway enrichment analysis revealed a substantial overrepresentation of sc-xTWAS susceptibility genes in cytokine-mediated signaling pathways, lymphocyte activation, and immune responses, underscoring the relevance of these pathways in immune disease etiology (Fig. 5E). Using AITD as an example, we identified a distinct class of sc-aTWAS and sc-edTWAS susceptibility genes in several disease-relevant cell types (Fig. 5F). Notably, *CTLA4* was found as a susceptibility gene with the highest effect size in AITD (Z-score = -19.86), especially in $CD4^+$ T cells. It was also identified among the top susceptibility genes in T1D and RA, implying its functional importance in autoimmune disease pathogenesis (Figs. 5A, 5G, and 5H). Cytotoxic T-lymphocyte antigen 4 (*CTLA4*), also known as CD152, is a protein receptor expressed by activated T cells that functions as an immune checkpoint and "off" switch to downregulate immune responses (*64*). Mechanistically, the rs3087243-G risk allele in AITD GWAS ($P = 4.6 \times 10^{-84}$) was associated with shortened 3′UTR length of *CTLA4* through APA regulation, an effect most pronounced in stimulated $CD4^+$ T cells (Fig. 5I). We experimentally

validated the causal effect of shortened 3′UTR length of *CTLA4* in significantly reducing protein abundance (Fig. 5J), which may potentially lead to more potent T cell activation that promotes autoimmune disease risk (Fig. 5K). Our sc-edTWAS approach identified *SIK3* as another candidate gene for AITD risk in $CD4^+$ T cells and monocytes (Figs. 5L and 5M, Z-score = -4.69). SIK3 plays an integral role in innate immunity by preventing the differentiation of macrophages into a potent and stable anti-inflammatory phenotype (*65*).

Overall, our sc-xTWAS analysis transforms our understanding of immune disease susceptibility by uncovering a significant layer of disease genes that operate through post-transcriptional mechanisms, providing new insights into the molecular drivers of immune-related diseases.

sc-xQTLs significantly contribute to the protein abundance

Protein quantitative trait locus (pQTL) data provide crucial insights for proteogenomic research and drug target prioritization (*66*). To elucidate the relationship between cell type-specific genetic regulation of transcriptional phenotypes and protein abundance, we analyzed 2,653 circulating *cis*-pQTLs identified from 79,751 individuals of three cohorts (*66-68*) (Fig. 6A). The effect sizes of pQTLs were in high concordance across studies (Pearson′s r = 0.78-0.92) (fig. S20A). The pQTL-associated genes were significantly enriched in immune-related biological processes such as leukocyte migration, humoral immune response, adaptive immune response, and negative regulation of response to external stimulus (Fig. 6B and fig. S20B).

Colocalization analysis revealed substantial overlap between protein abundance and gene regulation mechanisms, with 32% (849/2,653) of pQTLs significantly colocalized with sc-xQTLs (PP4 > 0.5) (table S8). Vice versa, 44.03% (775/1,757), 31.40% (211/670), and 11.27% (24/213) of sc-eQTLs, sc-aQTLs, and sc-edQTLs of the overlapping genes colocalized with pQTLs, respectively (fig. S20C). Importantly, we identified 68 pQTLs that colocalized with sc-aQTLs but showed no evidence of colocalization (PP4 = 0.001) with sc-eQTLs, demonstrating that APA events can contribute protein abundance variations independently of gene expression (fig. S20D). Previous study reported only six colocalized eQTL-pQTL pairs when using whole-blood eQTL data from the eQTLGen Consortium. In contrast, our single-cell approach substantially increased detection power, identifying 15 sc-eQTL-pQTL pairs and revealing seven additional pQTL colocalizations from post-transcriptional mechanisms (Fig. 6C). For example, the *CCL3* pQTL (rs5820141) exclusively colocalized with sc-eQTLs, while the *SLAMF1* pQTL (rs1572382) specifically colocalized with sc-aQTLs (Fig. 6D and fig. S20E). Overall, our analysis provided transcriptional and post-transcriptional interpretation for 2.5-fold more pQTLs compared to previous bulk tissue studies.

Directional effect analysis uncovered significant correlations between pQTLs and molecular phenotypes. Overall, a strong positive correlation was observed between pQTLs and eQTLs (Fig. 6E), as exemplified by the T-allele of rs484915, which was strongly associated with increased MMP1 protein abundance ($\beta = 0.35$, $P = 5.73 \times 10^{-192}$) and gene expression ($\beta = 0.62$, $P = 1.19 \times 10^{-19}$, fig. S21A). Conversely, sc-aQTLs exhibited a weak negative correlation with pQTLs, as exemplified by variants rs17207042 and rs12969657, which were associated with *CD226* 3′UTR shortening and increased protein abundance (fig. S21B).

To prioritize for putatively disease causal genes, we performed an integrative of pQTLs and sc-xQTLs using HyPrColoc (*69*) and identified 32 genes colocalized with immune disease

GWAS, including 28 and 12 through gene expression and APA, respectively, achieving a two-fold increase over previously bulk tissue analyses (Fig. 6F). Notably, 19 of the 32 genes were previously established as potentially druggable targets (*70, 71*) (Fig. 6G). For example, *STAT6*, a member of the STAT family of transcription factors with a central role in exerting IL-4 and IFN-mediated cellular immune responses, has been found as a susceptibility gene in multiple immune-related diseases (*72, 73*). Here, we found that the allergic disease risk variant rs703817 strongly colocalized with sc-aQTL and pQTL of *STAT6*, independent of sc-eQTL (Fig. 6H and fig. S22A). Mechanistically, the T allele of rs4559 (PIP = 1 and in strong LD with rs703817) results in a G-to-A change (reverse strand) in the 3′UTR of *STAT6* that creates a new proximal polyadenylation motif (AATGTA to AATATA), which significantly promotes 3′UTR shortening presumably through recruiting APA regulators (*74*) (Fig. 6I). To experimentally validate the causal effect of rs4559, we employed a minigene reporter assay to make a single nucleotide G-to-A change to mimic the T allele of rs4559 (Fig. 6J). Compared to the wildtype sequences, this single G-to-A change dramatically shifted the usage of polyadenylation sites and resulted in significant 3′UTR shortening, as evidenced by 3′RACE analysis (Fig. 6K). Further, we showed that this sc-aQTL effect on *STAT6* APA contributes to protein abundance dysregulation, potentially increasing the risk of allergic disease.

In addition to previously identified immune disease-related susceptibility genes, our study uncovered novel APA-related immune disease susceptibility genes, such as dihydrolipoamide dehydrogenase (*DLD*), which encodes a mitochondrial enzyme involved in energy metabolism (*75*) and susceptibility to ulcerative colitis (UC) with unclear underlying mechanism (*76*) (Figs. 6L and 6M). Our analysis revealed strong colocalization between the UC GWAS signal and a sc-aQTL (rs4518) of *DLD* specifically regulating its APA in $CD4^+$ T cells and $CD8^+$ T cells (PP4 = 0.77), thereby potentially affecting *DLD* protein expression as a pQTL (PP4 = 0.81), independent of overall gene expression levels (PP4 = 0.08) (Fig. 6N and fig. S22B). The minigene reporter assay demonstrated that the C to T change of rs4518 alters the polyadenylation site usage, leading to an extended 3′UTR of *DLD* (Fig. 6O). Further dual-luciferase reporter assays confirmed that the long 3′UTR isoform of *DLD* significantly enhances protein expression (Fig. 6P). Therefore, our study establishes a direct causal link between rs4518 and ulcerative colitis disease susceptibility through post-transcriptional regulation of *DLD*. In summary, our results demonstrate that genetic variants can profoundly influence protein abundance through post-transcriptional processes, revealing a critical mechanism underlying immune disease susceptibility.

Discussion

We generated a high-quality expansive resource for the scientific community. Exploiting this unprecedented dataset of over 7 million immune cells, three distinct transcriptional and post-transcriptional factors in gene regulation were explored, namely RNA abundance, alternative polyadenylation, and RNA editing. We showed that *cis*-genetic effects through APA (sc-aQTLs) and RNA editing (sc-edQTLs) operate largely independently of RNA abundance (sc-eQTLs) (63% and 86%, respectively). This independence suggests previously underappreciated layers of gene regulation beyond RNA abundance. Indeed, we found that including sc-aQTLs and sc-edQTLs explained an additional 213 immune-related disease loci, which further expanded upon our previous studies in bulk tissues (*13, 14*). While our data only covers a fraction of the post-transcriptional regulation space (RNA splicing, RNA modifications, RNA stability, etc), our estimates may be conservative for the role of post-transcriptional regulation governing immune-

related gene function. In fact, since the identification of aQTLs and edQTLs is far from saturation compared to eQTLs (*14, 74*), it is expected that future studies of greater sample size and *trans*-effects will account for part of the immune-related disease heritability we attribute to independent genetic effects.

Owing to our large-scale dataset composed of multiple molecular traits at transcriptional and post-transcriptional levels, we identified 32,763 sc-xQTLs in 17,422 xGenes, a significant step forward compared to previous work (*8, 9*) in providing a comprehensive survey of the complex genetic architecture of gene regulation in immune cells. Notably, about half of the sc-xQTLs (16,653/32,763) were identified with significant cell-type specific genetic effects, which were not attributed to cell-type specific expression of xGenes (4,733, or 14.45% of total xGenes) or low statistical power to detect allelic effects in more than one cell type. Compared to bulk-tissue QTLs, cell-type specific *cis*-effects are mostly consistent in the direction of effect but of much greater effect sizes, which likely explains the 2.43-time increase in the number of colocalized immune-disease GWAS loci. These findings supported the notion that genetic effects on gene regulation are largely cell-type specific (*77, 78*).

Another intriguing finding is the dominant role of stimulation in driving context-dependent genetic regulation in immune cells (*21, 22*). Our analysis reveals that immune stimuli exert a substantial influence, more than sex or ethnicity, on the *cis*-genetic effects of gene regulation in immune cells. Consistent with previous studies (*21, 79*), fewer sc-eGenes were identified upon immune stimulation. However, immune stimulation significantly increased the overall magnitude of genetic effects compared to the unstimulated state, which further helped to explain 11.21% more GWAS loci in immune-related diseases. Because immune cells are quiescent in the absence of extrinsic stimulation, it is essential to explore genetic effects in biologically relevant contexts to fully uncover the genes and pathways modulated by disease-risk variants.

With the support of protein-level evidence, our integrative analysis revealed 32 sc-xQTL-pQTL-GWAS triplets, most of which have been independently identified as druggable targets for immune-related diseases (*70, 71*). By directly testing the allelic effects on APA and protein abundance, we experimentally validated the causal effects of disease-risk variants on targeted genes. The observation that the allele-specific changes at the protein level were not due to expression level differences, once again, reaffirmed the notion that post-transcriptional regulation is an independent and important layer in gene regulation. Interestingly, we only identified one sc-edQTL-pQTL-GWAS triplet, likely because RNA editing is known to regulate nucleic acid sensing through affecting the structure and conformation of RNA, which is mostly independent of expression and protein level (*17, 18*).

In summary, our comprehensive landscape of immune cell regulation provides a valuable resource for investigating how genetic variation influences immune function in healthy and diseased individuals. The novel molecular mechanisms we identified, particularly those mediated by APA and RNA editing, reveal promising targets for therapeutic intervention in immune-related diseases. Moving forward, this work establishes a framework for understanding the complex interplay between genetic variation, cellular context, and multi-layered gene regulation in the immune system.

References and Notes

1. C. Selmi, Q. Lu, M. C. Humble, Heritability versus the role of the environment in autoimmunity. *J. Autoimmun.* **39**, 249–252 (2012).
2. S. Saevarsdottir, T. A. Olafsdottir, E. V. Ivarsdottir, G. H. Halldorsson, K. Gunnarsdottir, A. Sigurdsson, A. Johannesson, J. K. Sigurdsson, T. Juliusdottir, S. H. Lund, A. O. Arnthorsson, E. L. Styrmisdottir, J. Gudmundsson, G. M. Grondal, K. Steinsson, L. Alfredsson, J. Askling, R. Benediktsson, R. Bjarnason, A. J. Geirsson, B. Gudbjornsson, H. Gudjonsson, H. Hjaltason, A. B. Hreidarsson, L. Klareskog, I. Kockum, H. Kristjansdottir, T. J. Love, B. R. Ludviksson, T. Olsson, P. T. Onundarson, K. B. Orvar, L. Padyukov, B. Sigurgeirsson, V. Tragante, K. Bjarnadottir, T. Rafnar, G. Masson, P. Sulem, D. F. Gudbjartsson, P. Melsted, G. Thorleifsson, G. L. Norddahl, U. Thorsteinsdottir, I. Jonsdottir, K. Stefansson, *FLT3* stop mutation increases FLT3 ligand level and risk of autoimmune thyroid disease. *Nature* **584**, 619–623 (2020).
3. J. Chiou, R. J. Geusz, M. L. Okino, J. Y. Han, M. Miller, R. Melton, E. Beebe, P. Benaglio, S. Huang, K. Korgaonkar, S. Heller, A. Kleger, S. Preissl, D. U. Gorkin, M. Sander, K. J. Gaulton, Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. *Nature* **594**, 398–402 (2021).
4. K. Ishigaki, S. Sakaue, C. Terao, Y. Luo, K. Sonehara, K. Yamaguchi, T. Amariuta, C. L. Too, V. A. Laufer, I. C. Scott, S. Viatte, M. Takahashi, K. Ohmura, A. Murasawa, M. Hashimoto, H. Ito, M. Hammoudeh, S. A. Emadi, B. K. Masri, H. Halabi, H. Badsha, I. W. Uthman, X. Wu, L. Lin, T. Li, D. Plant, A. Barton, G. Orozco, S. M. M. Verstappen, J. Bowes, A. J. MacGregor, S. Honda, M. Koido, K. Tomizuka, Y. Kamatani, H. Tanaka, E. Tanaka, A. Suzuki, Y. Maeda, K. Yamamoto, S. Miyawaki, G. Xie, J. Zhang, C. I. Amos, E. Keystone, G. Wolbink, I. van der Horst-Bruinsma, J. Cui, K. P. Liao, R. J. Carroll, H. S. Lee, S. Y. Bang, K. A. Siminovitch, N. de Vries, L. Alfredsson, S. Rantapaa-Dahlqvist, E. W. Karlson, S. C. Bae, R. P. Kimberly, J. C. Edberg, X. Mariette, T. Huizinga, P. Dieude, M. Schneider, M. Kerick, J. C. Denny, P. BioBank Japan, K. Matsuda, K. Matsuo, T. Mimori, F. Matsuda, K. Fujio, Y. Tanaka, A. Kumanogoh, M. Traylor, C. M. Lewis, S. Eyre, H. Xu, R. Saxena, T. Arayssi, Y. Kochi, K. Ikari, M. Harigai, P. K. Gregersen, K. Yamamoto, S. Louis Bridges, Jr., L. Padyukov, J. Martin, L. Klareskog, Y. Okada, S. Raychaudhuri, Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nat. Genet.* **54**, 1640–1651 (2022).
5. K. K. Farh, A. Marson, J. Zhu, M. Kleinewietfeld, W. J. Housley, S. Beik, N. Shoresh, H. Whitton, R. J. Ryan, A. A. Shishkin, M. Hatan, M. J. Carrasco-Alfonso, D. Mayer, C. J. Luckey, N. A. Patsopoulos, P. L. De Jager, V. K. Kuchroo, C. B. Epstein, M. J. Daly, D. A. Hafler, B. E. Bernstein, Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
6. G. Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
7. B. Soskic, E. Cano-Gamez, D. J. Smyth, K. Ambridge, Z. Ke, J. C. Matte, L. Bossini-Castillo, J. Kaplanis, L. Ramirez-Navarro, A. Lorenc, N. Nakic, J. Esparza-Gordillo, W. Rowan, D. Wille, D. F. Tough, P. G. Bronson, G. Trynka, Immune disease risk variants regulate gene expression dynamics during $CD4^+$ T cell activation. *Nat. Genet.* **54**, 817–826 (2022).
8. S. Yazar, J. Alquicira-Hernandez, K. Wing, A. Senabouth, M. G. Gordon, S. Andersen, Q. Lu, A. Rowson, T. R. P. Taylor, L. Clarke, K. Maccora, C. Chen, A. L. Cook, C. J. Ye, K. A. Fairfax, A. W. Hewitt, J. E. Powell, Single-cell eQTL mapping identifies cell type-

specific genetic control of autoimmune disease. *Science* **376**, eabf3041 (2022).

9. R. K. Perez, M. G. Gordon, M. Subramaniam, M. C. Kim, G. C. Hartoularos, S. Targ, Y. Sun, A. Ogorodnikov, R. Bueno, A. Lu, M. Thompson, N. Rappoport, A. Dahl, C. M. Lanata, M. Matloubian, L. Maliskova, S. S. Kwek, T. Li, M. Slyper, J. Waldman, D. Dionne, O. Rozenblatt-Rosen, L. Fong, M. Dall'Era, B. Balliu, A. Regev, J. Yazdany, L. A. Criswell, N. Zaitlen, C. J. Ye, Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).
10. H. E. Randolph, J. K. Fiege, B. K. Thielen, C. K. Mickelson, M. Shiratori, J. Barroso-Batista, R. A. Langlois, L. B. Barreiro, Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. *Science* **374**, 1127–1133 (2021).
11. K. R. Simpfendorfer, B. E. Armstead, A. Shih, W. Li, M. Curran, N. Manjarrez-Orduño, A. T. Lee, B. Diamond, P. K. Gregersen, Autoimmune disease-associated haplotypes of *BLK* exhibit lowered thresholds for B cell activation and expansion of Ig class-switched B cells. *Arthritis Rheumatol.* **67**, 2866–2876 (2015).
12. C. Wen, M. Margolis, R. Dai, P. Zhang, P. F. Przytycki, D. D. Vo, A. Bhattacharya, N. Matoba, M. Tang, C. Jiao, M. Kim, E. Tsai, C. Hoh, N. Aygun, R. L. Walker, C. Chatzinakos, D. Clarke, H. Pratt, E. C. Psych, M. A. Peters, M. Gerstein, N. P. Daskalakis, Z. Weng, A. E. Jaffe, J. E. Kleinman, T. M. Hyde, D. R. Weinberger, N. J. Bray, N. Sestan, D. H. Geschwind, K. Roeder, A. Gusev, B. Pasaniuc, J. L. Stein, M. I. Love, K. S. Pollard, C. Liu, M. J. Gandal, E. C. Psych, Cross-ancestry atlas of gene, isoform, and splicing regulation in the developing human brain. *Science* **384**, eadh0829 (2024).
13. L. Li, X. Ma, Y. Cui, M. Rotival, W. Chen, X. Zou, R. Ding, Y. Qin, Q. Wang, L. Quintana-Murci, W. Li, Immune-response 3'UTR alternative polyadenylation quantitative trait loci contribute to variation in human complex traits and diseases. *Nat. Commun.* **14**, 8347 (2023).
14. Q. Li, M. J. Gloudemans, J. M. Geisinger, B. Fan, F. Aguet, T. Sun, G. Ramaswami, Y. I. Li, J. B. Ma, J. K. Pritchard, S. B. Montgomery, J. B. Li, RNA editing underlies genetic risk of common inflammatory diseases. *Nature* **608**, 569–577 (2022).
15. X. Jia, S. Yuan, Y. Wang, Y. Fu, Y. Ge, Y. Ge, X. Lan, Y. Feng, F. Qiu, P. Li, S. Chen, A. Xu, The role of alternative polyadenylation in the antiviral innate immune response. *Nat. Commun.* **8**, 14605 (2017).
16. B. Tian, J. L. Manley, Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell. Biol.* **18**, 18–30 (2017).
17. S. B. Hu, J. B. Li, RNA editing and immune control: from mechanism to therapy. *Curr. Opin. Genet. Dev.* **86**, 102195 (2024).
18. B. J. Liddicoat, R. Piskol, A. M. Chalk, G. Ramaswami, M. Higuchi, J. C. Hartner, J. B. Li, P. H. Seeburg, C. R. Walkley, RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* **349**, 1115–1120 (2015).
19. B. E. Mittleman, S. Pott, S. Warland, T. Zeng, Z. Mu, M. Kaur, Y. Gilad, Y. Li, Alternative polyadenylation mediates genetic regulation of gene expression. *Elife* **9**, e57492 (2020).
20. R. R. Graham, C. Kyogoku, S. Sigurdsson, I. A. Vlasova, L. R. Davies, E. C. Baechler, R. M. Plenge, T. Koeuth, W. A. Ortmann, G. Hom, J. W. Bauer, C. Gillett, N. Burtt, D. S. Cunninghame Graham, R. Onofrio, M. Petri, I. Gunnarsson, E. Svenungsson, L. Rönnblom, G. Nordmark, P. K. Gregersen, K. Moser, P. M. Gaffney, L. A. Criswell, T. J. Vyse, A. C. Syvänen, P. R. Bohjanen, M. J. Daly, T. W. Behrens, D. Altshuler, Three functional variants of IFN regulatory factor 5 (*IRF5*) define risk and protective

haplotypes for human lupus. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6758–6763 (2007).

21. Y. Aquino, A. Bisiaux, Z. Li, M. O'Neill, J. Mendoza-Revilla, S. H. Merkling, G. Kerner, M. Hasan, V. Libri, V. Bondet, N. Smith, C. de Cevins, M. Ménager, F. Luca, R. Pique-Regi, G. Barba-Spaeth, S. Pietropaoli, O. Schwartz, G. Leroux-Roels, C. K. Lee, K. Leung, J. T. Wu, M. Peiris, R. Bruzzone, L. Abel, J. L. Casanova, S. A. Valkenburg, D. Duffy, E. Patin, M. Rotival, L. Quintana-Murci, Dissecting human population variation in single-cell responses to SARS-CoV-2. *Nature* **621**, 120–128 (2023).
22. A. Nathan, S. Asgari, K. Ishigaki, C. Valencia, T. Amariuta, Y. Luo, J. I. Beynor, Y. Baglaenko, S. Suliman, A. L. Price, L. Lecca, M. B. Murray, D. B. Moody, S. Raychaudhuri, Single-cell eQTL models reveal dynamic T cell state dependence of disease loci. *Nature* **606**, 120–128 (2022).
23. B. J. Schmiedel, C. Gonzalez-Colin, V. Fajardo, J. Rocha, A. Madrigal, C. Ramírez-Suástegui, S. Bhattacharyya, H. Simon, J. A. Greenbaum, B. Peters, G. Seumois, F. Ay, V. Chandra, P. Vijayanand, Single-cell eQTL analysis of activated T cell subsets reveals activation and cell type-dependent effects of disease-risk variants. *Sci. Immunol.* **7**, eabm2508 (2022).
24. C. Xu, M. Prete, S. Webb, L. Jardine, B. J. Stewart, R. Hoo, P. He, K. B. Meyer, S. A. Teichmann, Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* **186**, 5876–5891.e5820 (2023).
25. C. Lo Giudice, M. A. Tangaro, G. Pesole, E. Picardi, Investigating RNA editing in deep transcriptome datasets with REDItools and REDIportal. *Nat. Protoc.* **15**, 1098–1131 (2020).
26. H. Zhang, J. Hu, M. Recce, B. Tian, PolyA_DB: a database for mammalian mRNA polyadenylation. *Nucleic Acids Res.* **33**, D116–D120 (2005).
27. M. M. Fansler, S. Mitschka, C. Mayr, Quantifying 3'UTR length from scRNA-seq data reveals changes independent of gene expression. *Nat. Commun.* **15**, 4050 (2024).
28. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
29. J. Menke, G. C. Zeller, E. Kikawada, T. K. Means, X. R. Huang, H. Y. Lan, B. Lu, J. Farber, A. D. Luster, V. R. Kelley, CXCL9, but not CXCL10, promotes CXCR3-dependent immune-mediated kidney disease. *J. Am. Soc. Nephrol.* **19**, 1177–1189 (2008).
30. S. H. Roth, M. Danan-Gotthold, M. Ben-Izhak, G. Rechavi, C. J. Cohen, Y. Louzoun, E. Y. Levanon, Increased RNA editing may provide a source for autoantigens in systemic lupus erythematosus. *Cell Rep.* **23**, 50–57 (2018).
31. K. Zhao, Y. Sun, S. Zhong, J. L. Luo, The multifaceted roles of cathepsins in immune and inflammatory responses: implications for cancer therapy, autoimmune diseases, and infectious diseases. *Biomark. Res.* **12**, 165 (2024).
32. S. Das, L. Forer, S. Schönherr, C. Sidore, A. E. Locke, A. Kwong, S. I. Vrieze, E. Y. Chew, S. Levy, M. McGue, D. Schlessinger, D. Stambolian, P. R. Loh, W. G. Iacono, A. Swaroop, L. J. Scott, F. Cucca, F. Kronenberg, M. Boehnke, G. R. Abecasis, C. Fuchsberger, Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
33. O. Delaneau, H. Ongen, A. A. Brown, A. Fort, N. I. Panousis, E. T. Dermitzakis, A complete tool set for molecular QTL discovery and analysis. *Nat. Commun.* **8**, 15452 (2017).
34. X. Wen, Molecular QTL discovery incorporating genomic annotations using Bayesian false discovery rate control. *The Annals of Applied Statistics* **10**, 1619–1638 (2016).

35. W. L. Crouse, G. R. Keele, M. S. Gastonguay, G. A. Churchill, W. Valdar, A Bayesian model selection approach to mediation analysis. *PLoS Genet.* **18**, e1010184 (2022).

36. S. M. Urbut, G. Wang, P. Carbonetto, M. Stephens, Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat. Genet.* **51**, 187–195 (2019).

37. M. Ota, Y. Nagafuchi, H. Hatano, K. Ishigaki, C. Terao, Y. Takeshima, H. Yanaoka, S. Kobayashi, M. Okubo, H. Shirai, Y. Sugimori, J. Maeda, M. Nakano, S. Yamada, R. Yoshida, H. Tsuchiya, Y. Tsuchida, S. Akizuki, H. Yoshifuji, K. Ohmura, T. Mimori, K. Yoshida, D. Kurosaka, M. Okada, K. Setoguchi, H. Kaneko, N. Ban, N. Yabuki, K. Matsuki, H. Mutoh, S. Oyama, M. Okazaki, H. Tsunoda, Y. Iwasaki, S. Sumitomo, H. Shoda, Y. Kochi, Y. Okada, K. Yamamoto, T. Okamura, K. Fujio, Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. *Cell* **184**, 3006–3021.e3017 (2021).

38. C. H. Polman, P. W. O'Connor, E. Havrdova, M. Hutchinson, L. Kappos, D. H. Miller, J. T. Phillips, F. D. Lublin, G. Giovannoni, A. Wajgt, M. Toal, F. Lynn, M. A. Panzara, A. W. Sandrock, A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N. Engl. J. Med.* **354**, 899–910 (2006).

39. M. A. Nalls, D. J. Couper, T. Tanaka, F. J. van Rooij, M. H. Chen, A. V. Smith, D. Toniolo, N. A. Zakai, Q. Yang, A. Greinacher, A. R. Wood, M. Garcia, P. Gasparini, Y. Liu, T. Lumley, A. R. Folsom, A. P. Reiner, C. Gieger, V. Lagou, J. F. Felix, H. Völzke, N. A. Gouskova, A. Biffi, A. Döring, U. Völker, S. Chong, K. L. Wiggins, A. Rendon, A. Dehghan, M. Moore, K. Taylor, J. G. Wilson, G. Lettre, A. Hofman, J. C. Bis, N. Pirastu, C. S. Fox, C. Meisinger, J. Sambrook, S. Arepalli, M. Nauck, H. Prokisch, J. Stephens, N. L. Glazer, L. A. Cupples, Y. Okada, A. Takahashi, Y. Kamatani, K. Matsuda, T. Tsunoda, T. Tanaka, M. Kubo, Y. Nakamura, K. Yamamoto, N. Kamatani, M. Stumvoll, A. Tönjes, I. Prokopenko, T. Illig, K. V. Patel, S. F. Garner, B. Kuhnel, M. Mangino, B. A. Oostra, S. L. Thein, J. Coresh, H. E. Wichmann, S. Menzel, J. Lin, G. Pistis, A. G. Uitterlinden, T. D. Spector, A. Teumer, G. Eiriksdottir, V. Gudnason, S. Bandinelli, T. M. Frayling, A. Chakravarti, C. M. van Duijn, D. Melzer, W. H. Ouwehand, D. Levy, E. Boerwinkle, A. B. Singleton, D. G. Hernandez, D. L. Longo, N. Soranzo, J. C. Witteman, B. M. Psaty, L. Ferrucci, T. B. Harris, C. J. O'Donnell, S. K. Ganesh, Multiple loci are associated with white blood cell phenotypes. *PLoS Genet.* **7**, e1002113 (2011).

40. K. Kundu, M. Tardaguila, A. L. Mann, S. Watt, H. Ponstingl, L. Vasquez, D. Von Schiller, N. W. Morrell, O. Stegle, T. Pastinen, S. J. Sawcer, C. A. Anderson, K. Walter, N. Soranzo, Genetic associations at regulatory phenotypes improve fine-mapping of causal variants for 12 immune-mediated diseases. *Nat. Genet.* **54**, 251–262 (2022).

41. J. C. Crispín, B. T. Keenan, M. D. Finnell, B. L. Bermas, P. Schur, E. Massarotti, E. W. Karlson, L. M. Fitzgerald, S. Ergin, V. C. Kyttaris, G. C. Tsokos, K. H. Costenbader, Expression of CD44 variant isoforms CD44v3 and CD44v6 is increased on T cells from patients with systemic lupus erythematosus and is correlated with disease activity. *Arthritis Rheum.* **62**, 1431–1437 (2010).

42. L. Eshkar Sebban, D. Ronen, D. Levartovsky, O. Elkayam, D. Caspi, S. Aamar, H. Amital, A. Rubinow, I. Golan, D. Naor, Y. Zick, I. Golan, The involvement of CD44 and its novel ligand galectin-8 in apoptotic regulation of autoimmune inflammation. *J. Immunol.* **179**, 1225–1235 (2007).

43. C. Germain, A. Meier, T. Jensen, P. Knapnougel, G. Poupon, A. Lazzari, A. Neisig, K. Håkansson, T. Dong, N. Wagtmann, E. D. Galsgaard, P. Spee, V. M. Braud, Induction of lectin-like transcript 1 (LLT1) protein cell surface expression by pathogens and

interferon-γ contributes to modulate immune responses. *J. Biol. Chem.* **286**, 37964–37975 (2011).

44. S. L. Swain, K. K. McKinstry, T. M. Strutt, Expanding roles for CD4$^+$ T cells in immunity to viruses. *Nat. Rev. Immunol.* **12**, 136–148 (2012).
45. G. E. Hoffman, E. E. Schadt, variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**, 483 (2016).
46. D. J. Taylor, S. B. Chhetri, M. G. Tassia, A. Biddanda, S. M. Yan, G. L. Wojcik, A. Battle, R. C. McCoy, Sources of gene expression variation in a globally diverse human cohort. *Nature* **632**, 122–130 (2024).
47. L. Shang, J. A. Smith, W. Zhao, M. Kho, S. T. Turner, T. H. Mosley, S. L. R. Kardia, X. Zhou, Genetic architecture of gene expression in European and African Americans: an eQTL mapping study in GENOA. *Am. J. Hum. Genet.* **106**, 496–512 (2020).
48. K. Liu, L. Kong, D. B. Graham, K. L. Carey, R. J. Xavier, SAC1 regulates autophagosomal phosphatidylinositol-4-phosphate for xenophagy-directed bacterial clearance. *Cell Rep.* **36**, 109434 (2021).
49. S. Ghafouri-Fard, M. Gholipour, S. Eslami, B. M. Hussen, M. Taheri, M. Samadian, M. D. Omrani, Abnormal expression of MAPK14-related lncRNAs in the peripheral blood of patients with multiple sclerosis. *Noncoding RNA Res.* **8**, 335–339 (2023).
50. J. A. Hall, J. L. Cannons, J. R. Grainger, L. M. Dos Santos, T. W. Hand, S. Naik, E. A. Wohlfert, D. B. Chou, G. Oldenhove, M. Robinson, M. E. Grigg, R. Kastenmayer, P. L. Schwartzberg, Y. Belkaid, Essential role for retinoic acid in the promotion of CD4$^+$ T cell effector responses via retinoic acid receptor alpha. *Immunity* **34**, 435–447 (2011).
51. E. Kleiman, D. Salyakina, M. De Heusch, K. L. Hoek, J. M. Llanes, I. Castro, J. A. Wright, E. S. Clark, D. M. Dykxhoorn, E. Capobianco, A. Takeda, J. C. Renauld, W. N. Khan, Distinct transcriptomic features are associated with transitional and mature B-cell populations in the mouse spleen. *Front. Immunol.* **6**, 30 (2015).
52. S. Gazal, H. K. Finucane, N. A. Furlotte, P. R. Loh, P. F. Palamara, X. Liu, A. Schoech, B. Bulik-Sullivan, B. M. Neale, A. Gusev, A. L. Price, Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).
53. T. K. Mao, Z. X. Lian, C. Selmi, Y. Ichiki, P. Ashwood, A. A. Ansari, R. L. Coppel, S. Shimoda, H. Ishibashi, M. E. Gershwin, Altered monocyte responses to defined TLR ligands in patients with primary biliary cirrhosis. *Hepatology* **42**, 802–808 (2005).
54. C. Wallace, A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet.* **17**, e1009440 (2021).
55. C. Wallace, Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **16**, e1008720 (2020).
56. A. L. Rachubinski, E. Wallace, E. Gurnee, B. A. Enriquez-Estrada, K. R. Worek, K. P. Smith, P. Araya, K. A. Waugh, R. E. Granrath, E. Britton, H. R. Lyford, M. G. Donovan, N. P. Eduthan, A. A. Hill, B. Martin, K. D. Sullivan, L. Patel, D. J. Fidler, M. D. Galbraith, C. A. Dunnick, D. A. Norris, J. M. Espinosa, JAK inhibition decreases the autoimmune burden in Down syndrome. *eLife* **13**, RP99323 (2024).
57. Q. Wang, C. Wang, N. Li, X. Liu, W. Ren, Q. Wang, X. Cao, Condensin Smc4 promotes inflammatory innate immune response by epigenetically enhancing NEMO transcription. *J. Autoimmun.* **92**, 67–76 (2018).
58. G. M. Hirschfield, X. Liu, Y. Han, I. P. Gorlov, Y. Lu, C. Xu, Y. Lu, W. Chen, B. D. Juran, C. Coltescu, A. L. Mason, P. Milkiewicz, R. P. Myers, J. A. Odin, V. A. Luketic, D. Speiciene, C. Vincent, C. Levy, P. K. Gregersen, J. Zhang, E. J. Heathcote, K. N.

Lazaridis, C. I. Amos, K. A. Siminovitch, Variants at *IRF5-TNPO3*, 17q12-21 and *MMEL1* are associated with primary biliary cirrhosis. *Nat. Genet.* **42**, 655–657 (2010).

59. A. Gusev, A. Ko, H. Shi, G. Bhatia, W. Chung, B. W. Penninx, R. Jansen, E. J. de Geus, D. I. Boomsma, F. A. Wright, P. F. Sullivan, E. Nikkola, M. Alvarez, M. Civelek, A. J. Lusis, T. Lehtimäki, E. Raitoharju, M. Kähönen, I. Seppälä, O. T. Raitakari, J. Kuusisto, M. Laakso, A. L. Price, P. Pajukanta, B. Pasaniuc, Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
60. H. Chen, Z. Wang, L. Gong, Q. Wang, W. Chen, J. Wang, X. Ma, R. Ding, X. Li, X. Zou, M. Plass, C. Lian, T. Ni, G. H. Wei, W. Li, L. Deng, L. Li, A distinct class of pan-cancer susceptibility genes revealed by an alternative polyadenylation transcriptome-wide association study. *Nat. Commun.* **15**, 1729 (2024).
61. Y. Cui, F. J. Arnold, F. Peng, D. Wang, J. S. Li, S. Michels, E. J. Wagner, A. R. La Spada, W. Li, Alternative polyadenylation transcriptome-wide association study identifies APA-linked susceptibility genes in brain disorders. *Nat. Commun.* **14**, 583 (2023).
62. J. Z. Liu, S. van Sommeren, H. Huang, S. C. Ng, R. Alberts, A. Takahashi, S. Ripke, J. C. Lee, L. Jostins, T. Shah, S. Abedian, J. H. Cheon, J. Cho, N. E. Dayani, L. Franke, Y. Fuyuno, A. Hart, R. C. Juyal, G. Juyal, W. H. Kim, A. P. Morris, H. Poustchi, W. G. Newman, V. Midha, T. R. Orchard, H. Vahedi, A. Sood, J. Y. Sung, R. Malekzadeh, H. J. Westra, K. Yamazaki, S. K. Yang, J. C. Barrett, B. Z. Alizadeh, M. Parkes, T. Bk, M. J. Daly, M. Kubo, C. A. Anderson, R. K. Weersma, Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
63. D. Ochoa, A. Hercules, M. Carmona, D. Suveges, J. Baker, C. Malangone, I. Lopez, A. Miranda, C. Cruz-Castillo, L. Fumis, M. Bernal-Llinares, K. Tsukanov, H. Cornu, K. Tsirigos, O. Razuvayevskaya, A. Buniello, J. Schwartzentruber, M. Karim, B. Ariano, R. E. Martinez Osorio, J. Ferrer, X. Ge, S. Machlitt-Northen, A. Gonzalez-Uriarte, S. Saha, S. Tirunagari, C. Mehta, J. M. Roldán-Romero, S. Horswell, S. Young, M. Ghoussaini, D. G. Hulcoop, I. Dunham, E. M. McDonagh, The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. *Nucleic Acids Res.* **51**, D1353–D1359 (2023).
64. L. S. Walker, D. M. Sansom, The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. *Nat. Rev. Immunol.* **11**, 852–863 (2011).
65. N. J. Darling, R. Toth, J. S. Arthur, K. Clark, Inhibition of SIK2 and SIK3 during differentiation enhances the anti-inflammatory phenotype of macrophages. *Biochem. J.* **474**, 521–537 (2017).
66. M. Pietzner, E. Wheeler, J. Carrasco-Zanini, A. Cortes, M. Koprulu, M. A. Wörheide, E. Oerton, J. Cook, I. D. Stewart, N. D. Kerrison, J. Luan, J. Raffler, M. Arnold, W. Arlt, S. O'Rahilly, G. Kastenmüller, E. R. Gamazon, A. D. Hingorani, R. A. Scott, N. J. Wareham, C. Langenberg, Mapping the proteo-genomic convergence of human diseases. *Science* **374**, eabj1541 (2021).
67. B. B. Sun, J. Chiou, M. Traylor, C. Benner, Y. H. Hsu, T. G. Richardson, P. Surendran, A. Mahajan, C. Robins, S. G. Vasquez-Grinnell, L. Hou, E. M. Kvikstad, O. S. Burren, J. Davitte, K. L. Ferber, C. E. Gillies, K. Hedman Å, S. Hu, T. Lin, R. Mikkilineni, R. K. Pendergrass, C. Pickering, B. Prins, D. Baird, C. Y. Chen, L. D. Ward, A. M. Deaton, S. Welsh, C. M. Willis, N. Lehner, M. Arnold, M. A. Wörheide, K. Suhre, G. Kastenmüller, A. Sethi, M. Cule, A. Raj, L. Burkitt-Gray, E. Melamud, M. H. Black, E. B. Fauman, J. M. M. Howson, H. M. Kang, M. I. McCarthy, P. Nioi, S. Petrovski, R. A. Scott, E. N. Smith, S. Szalma, D. M. Waterworth, L. J. Mitnaul, J. D. Szustakowski, B. W. Gibson, M.

R. Miller, C. D. Whelan, Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).

68. J. H. Zhao, D. Stacey, N. Eriksson, E. Macdonald-Dunlop, K. Hedman Å, A. Kalnapenkis, S. Enroth, D. Cozzetto, J. Digby-Bell, J. Marten, L. Folkersen, C. Herder, L. Jonsson, S. E. Bergen, C. Gieger, E. J. Needham, P. Surendran, D. S. Paul, O. Polasek, B. Thorand, H. Grallert, M. Roden, U. Võsa, T. Esko, C. Hayward, Å. Johansson, U. Gyllensten, N. Powell, O. Hansson, N. Mattsson-Carlgren, P. K. Joshi, J. Danesh, L. Padyukov, L. Klareskog, M. Landén, J. F. Wilson, A. Siegbahn, L. Wallentin, A. Mälarstig, A. S. Butterworth, J. E. Peters, Genetics of circulating inflammatory proteins identifies drivers of immune-mediated disease risk and therapeutic targets. *Nat. Immunol.* **24**, 1540–1551 (2023).

69. C. N. Foley, J. R. Staley, P. G. Breen, B. B. Sun, P. D. W. Kirk, S. Burgess, J. M. M. Howson, A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat. Commun.* **12**, 764 (2021).

70. C. Finan, A. Gaulton, F. A. Kruger, R. T. Lumbers, T. Shah, J. Engmann, L. Galver, R. Kelley, A. Karlsson, R. Santos, J. P. Overington, A. D. Hingorani, J. P. Casas, The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **9**, eaag1166 (2017).

71. Y. Zhou, Y. Zhang, D. Zhao, X. Yu, X. Shen, Y. Zhou, S. Wang, Y. Qiu, Y. Chen, F. Zhu, TTD: Therapeutic Target Database describing target druggability information. *Nucleic Acids Res.* **52**, D1465–D1477 (2024).

72. S. A. Leon-Cabrera, E. Molina-Guzman, Y. G. Delgado-Ramirez, A. Vázquez-Sandoval, Y. Ledesma-Soto, C. G. Pérez-Plasencia, Y. I. Chirino, N. L. Delgado-Buenrostro, M. Rodríguez-Sosa, F. Vaca-Paniagua, F. Ávila-Moreno, E. B. Gutierrez-Cirlos, L. E. Arias-Romero, L. I. Terrazas, Lack of STAT6 attenuates inflammation and drives protection against early steps of colitis-associated colon cancer. *Cancer Immunol Res.* **5**, 385–396 (2017).

73. F. Demenais, P. Margaritte-Jeannin, K. C. Barnes, W. O. C. Cookson, J. Altmüller, W. Ang, R. G. Barr, T. H. Beaty, A. B. Becker, J. Beilby, H. Bisgaard, U. S. Bjornsdottir, E. Bleecker, K. Bønnelykke, D. I. Boomsma, E. Bouzigon, C. E. Brightling, M. Brossard, G. G. Brusselle, E. Burchard, K. M. Burkart, A. Bush, M. Chan-Yeung, K. F. Chung, A. Couto Alves, J. A. Curtin, A. Custovic, D. Daley, J. C. de Jongste, B. E. Del-Rio-Navarro, K. M. Donohue, L. Duijts, C. Eng, J. G. Eriksson, M. Farrall, Y. Fedorova, B. Feenstra, M. A. Ferreira, M. B. Freidin, Z. Gajdos, J. Gauderman, U. Gehring, F. Geller, J. Genuneit, S. A. Gharib, F. Gilliland, R. Granell, P. E. Graves, D. F. Gudbjartsson, T. Haahtela, S. R. Heckbert, D. Heederik, J. Heinrich, M. Heliövaara, J. Henderson, B. E. Himes, H. Hirose, J. N. Hirschhorn, A. Hofman, P. Holt, J. Hottenga, T. J. Hudson, J. Hui, M. Imboden, V. Ivanov, V. W. V. Jaddoe, A. James, C. Janson, M. R. Jarvelin, D. Jarvis, G. Jones, I. Jonsdottir, P. Jousilahti, M. Kabesch, M. Kähönen, D. B. Kantor, A. S. Karunas, E. Khusnutdinova, G. H. Koppelman, A. L. Kozyrskyj, E. Kreiner, M. Kubo, R. Kumar, A. Kumar, M. Kuokkanen, L. Lahousse, T. Laitinen, C. Laprise, M. Lathrop, S. Lau, Y. A. Lee, T. Lehtimäki, S. Letort, A. M. Levin, G. Li, L. Liang, L. R. Loehr, S. J. London, D. W. Loth, A. Manichaikul, I. Marenholz, F. J. Martinez, M. C. Matheson, R. A. Mathias, K. Matsumoto, H. Mbarek, W. L. McArdle, M. Melbye, E. Melén, D. Meyers, S. Michel, H. Mohamdi, A. W. Musk, R. A. Myers, M. A. E. Nieuwenhuis, E. Noguchi, G. T. O'Connor, L. M. Ogorodova, C. D. Palmer, A. Palotie, J. E. Park, C. E. Pennell, G. Pershagen, A. Polonikov, D. S. Postma, N. Probst-Hensch, V. P. Puzyrev, B. A. Raby, O. T. Raitakari, A. Ramasamy, S. S. Rich, C. F. Robertson, I. Romieu, M. T.

Salam, V. Salomaa, V. Schlünssen, R. Scott, P. A. Selivanova, T. Sigsgaard, A. Simpson, V. Siroux, L. J. Smith, M. Solodilova, M. Standl, K. Stefansson, D. P. Strachan, B. H. Stricker, A. Takahashi, P. J. Thompson, G. Thorleifsson, U. Thorsteinsdottir, C. M. T. Tiesler, D. G. Torgerson, T. Tsunoda, A. G. Uitterlinden, R. J. P. van der Valk, A. Vaysse, S. Vedantam, A. von Berg, E. von Mutius, J. M. Vonk, J. Waage, N. J. Wareham, S. T. Weiss, W. B. White, M. Wickman, E. Widén, G. Willemsen, L. K. Williams, I. M. Wouters, J. J. Yang, J. H. Zhao, M. F. Moffatt, C. Ober, D. L. Nicolae, Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat. Genet.* **50**, 42–53 (2018).

74. L. Li, K. L. Huang, Y. Gao, Y. Cui, G. Wang, N. D. Elrod, Y. Li, Y. E. Chen, P. Ji, F. Peng, W. K. Russell, E. J. Wagner, W. Li, An atlas of alternative polyadenylation quantitative trait loci contributing to complex trait and disease heritability. *Nat. Genet.* **53**, 994–1005 (2021).

75. M. H. Odièvre, D. Chretien, A. Munnich, B. H. Robinson, R. Dumoulin, S. Masmoudi, N. Kadhom, A. Rötig, P. Rustin, J. P. Bonnefont, A novel mutation in the dihydrolipoamide dehydrogenase E3 subunit gene (*DLD*) resulting in an atypical form of alpha-ketoglutarate dehydrogenase deficiency. *Hum. Mutat.* **25**, 323–324 (2005).

76. D. P. McGovern, A. Gardet, L. Törkvist, P. Goyette, J. Essers, K. D. Taylor, B. M. Neale, R. T. Ong, C. Lagacé, C. Li, T. Green, C. R. Stevens, C. Beauchamp, P. R. Fleshner, M. Carlson, M. D'Amato, J. Halfvarson, M. L. Hibberd, M. Lördal, L. Padyukov, A. Andriulli, E. Colombo, A. Latiano, O. Palmieri, E. J. Bernard, C. Deslandres, D. W. Hommes, D. J. de Jong, P. C. Stokkers, R. K. Weersma, Y. Sharma, M. S. Silverberg, J. H. Cho, J. Wu, K. Roeder, S. R. Brant, L. P. Schumm, R. H. Duerr, M. C. Dubinsky, N. L. Glazer, T. Haritunians, A. Ippoliti, G. Y. Melmed, D. S. Siscovick, E. A. Vasiliauskas, S. R. Targan, V. Annese, C. Wijmenga, S. Pettersson, J. I. Rotter, R. J. Xavier, M. J. Daly, J. D. Rioux, M. Seielstad, Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat. Genet.* **42**, 332–337 (2010).

77. H. M. Natri, C. B. Del Azodi, L. Peter, C. J. Taylor, S. Chugh, R. Kendle, M. I. Chung, D. K. Flaherty, B. K. Matlock, C. L. Calvi, T. S. Blackwell, L. B. Ware, M. Bacchetta, R. Walia, C. M. Shaver, J. A. Kropski, D. J. McCarthy, N. E. Banovich, Cell-type-specific and disease-associated expression quantitative trait loci in the human lung. *Nat. Genet.* **56**, 595–604 (2024).

78. B. J. Schmiedel, D. Singh, A. Madrigal, A. G. Valdovino-Gonzalez, B. M. White, J. Zapardiel-Gonzalo, B. Ha, G. Altay, J. A. Greenbaum, G. McVicker, G. Seumois, A. Rao, M. Kronenberg, B. Peters, P. Vijayanand, Impact of genetic polymorphisms on human Immune cell gene expression. *Cell* **175**, 1701–1715.e1716 (2018).

79. H. Quach, M. Rotival, J. Pothlichet, Y. E. Loh, M. Dannemann, N. Zidane, G. Laval, E. Patin, C. Harmant, M. Lopez, M. Deschamps, N. Naffakh, D. Duffy, A. Coen, G. Leroux-Roels, F. Clement, A. Boland, J. F. Deleuze, J. Kelso, M. L. Albert, L. Quintana-Murci, Genetic adaptation and Neandertal admixture shaped the immune system of human populations. *Cell* **167**, 643–656.e617 (2016).

80. G. X. Zheng, J. M. Terry, P. Belgrader, P. Ryvkin, Z. W. Bent, R. Wilson, S. B. Ziraldo, T. D. Wheeler, G. P. McDermott, J. Zhu, M. T. Gregory, J. Shuga, L. Montesclaros, J. G. Underwood, D. A. Masquelier, S. Y. Nishimura, M. Schnall-Levin, P. W. Wyatt, C. M. Hindson, R. Bharadwaj, A. Wong, K. D. Ness, L. W. Beppu, H. J. Deeg, C. McFarland, K. R. Loeb, W. J. Valente, N. G. Ericson, E. A. Stevens, J. P. Radich, T. S. Mikkelsen, B. J. Hindson, J. H. Bielas, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).

81. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
82. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e289 (2019).
83. Y. Hao, T. Stuart, M. H. Kowalski, S. Choudhary, P. Hoffman, A. Hartman, A. Srivastava, G. Molla, S. Madad, C. Fernandez-Granda, R. Satija, Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
84. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
85. X. Zhang, Y. Lan, J. Xu, F. Quan, E. Zhao, C. Deng, T. Luo, L. Xu, G. Liao, M. Yan, Y. Ping, F. Li, A. Shi, J. Bai, T. Zhao, X. Li, Y. Xiao, CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* **47**, D721-D728 (2019).
86. A. S. E. Cuomo, G. Alvari, C. B. Azodi, D. J. McCarthy, M. J. Bonder, Optimizing expression quantitative trait locus mapping workflows for single-cell studies. *Genome Biol.* **22**, 188 (2021).
87. S. Lee, Y. C. Chen, A. E. Gillen, J. M. Taliaferro, B. Deplancke, H. Li, E. C. Lai, Diverse cell-specific patterns of alternative polyadenylation in Drosophila. *Nat. Commun.* **13**, 5372 (2022).
88. G. W. Li, F. Nan, G. H. Yuan, C. X. Liu, X. Liu, L. L. Chen, B. Tian, L. Yang, SCAPTURE: a deep learning-embedded pipeline that captures polyadenylation information from 3′ tag-based RNA-seq of single cells. *Genome Biol.* **22**, 221 (2021).
89. T. Smith, A. Heger, I. Sudbery, UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
90. Z. Xia, L. A. Donehower, T. A. Cooper, J. R. Neilson, D. A. Wheeler, E. J. Wagner, W. Li, Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat. Commun.* **5**, 5274 (2014).
91. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
92. H. J. Zhou, L. Li, Y. Li, W. Li, J. J. Li, PCA outperforms popular hidden variable inference methods for molecular QTL mapping. *Genome Biol.* **23**, 210 (2022).
93. A. Taylor-Weiner, F. Aguet, N. J. Haradhvala, S. Gosai, S. Anand, J. Kim, K. Ardlie, E. M. Van Allen, G. Getz, Scaling computational genomics to millions of individuals with GPUs. *Genome Biol.* **20**, 228 (2019).
94. N. Siva, 1000 Genomes project. *Nat. Biotechnol.* **26**, 256 (2008).
95. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
96. M. Chen, G. Lyu, M. Han, H. Nie, T. Shen, W. Chen, Y. Niu, Y. Song, X. Li, H. Li, X. Chen, Z. Wang, Z. Xia, W. Li, X. L. Tian, C. Ding, J. Gu, Y. Zheng, X. Liu, J. Hu, G. Wei, W. Tao, T. Ni, 3'UTR lengthening as a novel mechanism in regulating cellular senescence. *Genome Res.* **28**, 285–294 (2018).

Acknowledgments: We thank members of the Li laboratory for helpful discussions. We also thank Shenzhen Bay Laboratory High-Performance Computing and Informatics Core Facility for high-computing support.

Funding:

L.L. was supported by National Natural Science Foundation of China [no. 32370721, 32100533] and Open grant funds from Shenzhen Bay Laboratory [no. SZBL2021080601001].

Author contributions:

Conceptualization: Q.L., L.L.

Methodology: T.Z., H.C., R.D., L.L.

Investigation: T.Z., H.C., S.C., R.D., L.L.

Visualization: T.Z., H.C., S.C., R.D., D.L., Y.W

Funding acquisition: L.L.

Project administration: T.Z., Q.L., L.L.

Supervision: S.H., J.Y., G.W., S.M., Q.L., L.L.

Writing – original draft: T.Z., H.C., X.Z., K.X., L.L.

Writing – review & editing: T.Z., H.C., Q.L., L.L.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability:

The whole transcriptome and genome sequencing data are available through multiple repositories. The popCell project data is deposited in the Institut Pasteur repository (OWEY): scRNA-seq data (https://doi.org/10.48802/owey.e4qn-9190) and genome-wide genotyping data (https://doi.org/10.48802/owey.pyk2-5w22). OneK1K dataset is available at GEO (scRNA-seq: GSE196735; genotype: GSE196829). The scSLE dataset is accessible through GEO (scRNA-seq: GSE174188) and dbGaP (genotype: phs002812.v1.p1). For the scIAV dataset, scRNA-seq data are available at GEO under the accession number GSE162632, the genotype file is available at the Sequence Read Archive (SRA) under accession number PRJNA736483. For scDICE, the scRNA-seq and genotype data are available at dbGap under the accession number phs001703.v4.p1. For scTcell, the scRNA-seq data and genotype data are available at the European Genome-phenome Archive (EGA) with accession numbers EGAD00001008197 and EGAD00010002291, respectively. For scTB, the scRNA-seq and genotype data are available at dbGap under the accession numbers phs002467 and phs002025, respectively. GWAS summary statistics used in this study were obtained from the UK Biobank GWAS (https://www.nealelab.is/uk-biobank), GWAS Catalog (https://www.ebi.ac.uk/gwas/), and COVID-19 GWAS summary statistics are available at https://www.covid19hg.org/results/r7. The details about the GWAS summary statistics are listed in table S4. The human genome sequence reference version is GRCh38.p13, IAV genome is from H1N1/PR/8/1934, and COV genome is from BetaCoV/France/GE1973/2020 (ancestral strain). Gene annotation is available at Genocode v38 (https://www.gencodegenes.org/human/release_38.html), APA annotation is available at

polyA_DB (v2; https://exon.apps.wistar.org/PolyA_DB/v2/), and RNA editing annotation is available at REDIportal (V3.0; http://srv00.recas.ba.infn.it/atlas/). Variants annotation is available in dbSNP(v151) (https://ftp.ncbi.nlm.nih.gov/snp/organisms/human_9606_b151_GRCh38p7/VCF/). GTEx eQTL data are available via the GTEx portal (http://gtexportal.org), GTEx 3′aQTL data are available at 3′aQTL-atlas (https://wlcb.oit.uci.edu/3aQTLatlas), and GTEx edQTL data are available at the GTEx portal (https://gtexportal.org/home/downloads/adult-gtex/qtl). The data described in this study are freely available for querying and visualizing at https://bioinfo.szbl.ac.cn/sc-xQTLatlas. The custom source codes for gene regulation analysis from single-cell RNA-seq data are available under the MIT license at the GitHub repository: https://github.com/lilab-bioinfo/SERA.

List of supplementary materials

Materials and Methods

Figs. S1 to S22

Tables S1 to S8

References (*80–96*)

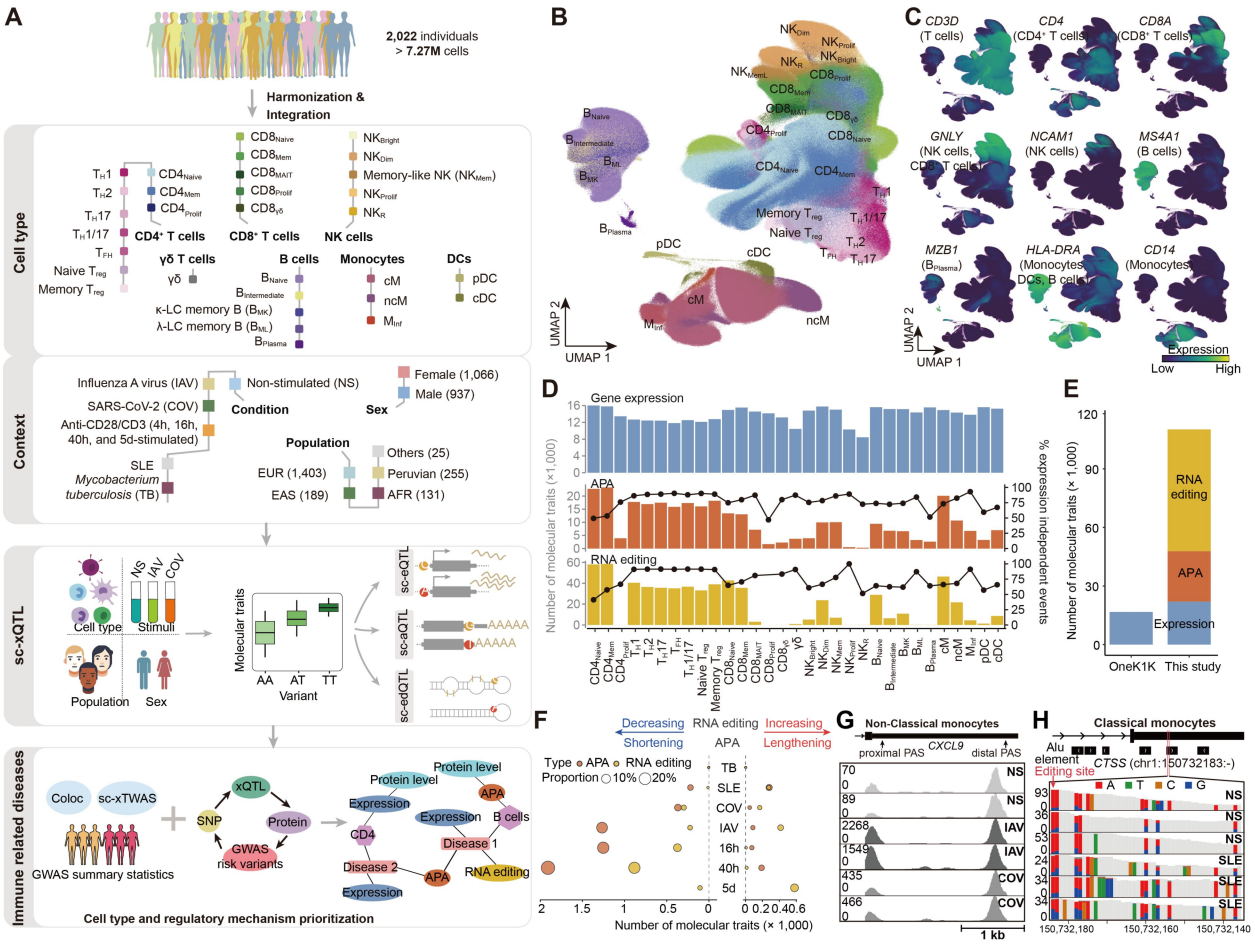

Fig. 1. Single-cell transcriptomic atlas of >7.27 million immune cells. (**A**) Overview of the integrative study. (Top) Compilation and harmonization of 31 immune cell types from 2,022 individuals across seven cohorts, with cross-context integration. (Middle) Characterization of cell type-specific QTLs of expression, APA, and RNA editing. (Bottom) Prioritization of immune-related disease-relevant cell types and genes. (**B**) Uniform manifold approximation and projection (UMAP) visualization of 31 harmonized cell types clustering. (**C**) UMAP plots delineating expression (color scale) of key marker genes in major cell types. (**D**) Molecular phenotypes quantified across immune cell types. Expression-independent fractions of APA and RNA editing events are shown as black points. (**E**) The scale and phenotypic diversity captured by sc-xQTL atlas compared to the OneK1K study. (**F**) Number of context-dependent post-transcriptional events identified, including APA (shortening/lengthening) and RNA editing (up/down-regulation) events in response to: (1) viral (IAV, COV); (2) T cell activation (anti-CD28/CD3 at 16h, 40h, and 5d); and (3) clinical (SLE, TB patients). (**G**) Differential *CXCL9* poly(A) site usages under (Top) baseline condition, (Middle) IAV, and (Bottom) COV stimulation. Arrows indicate proximal and distal poly(A) sites. (**H**) Differential editing event (chr1:150732183:-) of *CTSS* between (Top) healthy controls and (Bottom) SLE patients. Arrow highlights editing site.

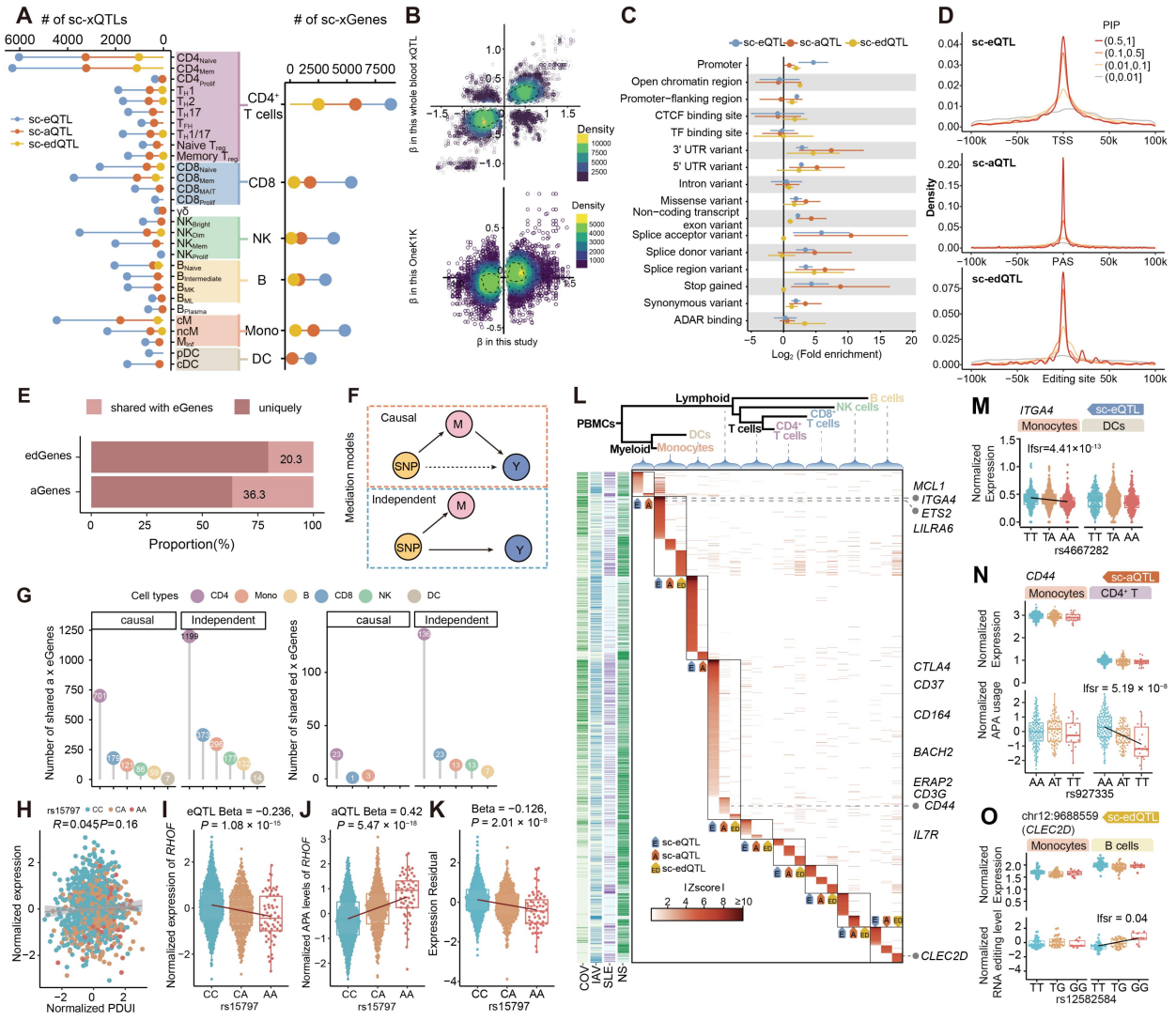

Fig. 2. Identification and characterization of cell-type specific sc-xQTLs. (**A**) Genetically regulated genes across immune cells. (Left) sc-eQTL (blue), sc-aQTL (red), and sc-edQTL (yellow) in 31 cell types. (Right) xGenes in major cell types. (**B**) Density plots comparing: (Top) xQTLs effect sizes (single-cell versus bulk). (Bottom) sc-eQTL effect sizes (OneK1K versus our study). (**C**) Functional enrichment of sc-xQTLs. (**D**) Enrichment of the fine-mapped sc-eQTLs around TSS, sc-aQTLs around PAS, and sc-edQTLs around editing sites. (**E**) Overlapping genes between significant aQTLs-eQTLs and edQTLs-eQTLs. (**F**) Mediation models evaluated in this study: causal forward and independent models. (**G**) Mediation events in each model across major cell types. (**H**) Example of an independent model (posterior probability = 0.965) for sc-eQTL and sc-aQTL of *RHOF*. The plot shows the association between normalized *RHOF* expression, and its APA usage levels (n = 962). The color indicates the genotype of the polymorphism rs15797. The error bars represent standard error. The boxplots show the allelic effect on (**I**) expression, (**J**) APA usages, and (**K**) residual expression after regressing out the APA effects of *RHOF*. (**L**) Heatmap of cell type-specific sc-xQTLs (Z-scores by mashr). (**M**) *ITGA4* monocyte-specific allelic effects on expression. (**N**) Allelic effects on *CD44* APA usage specific to CD4+ T cells. (O) Allelic effect on *CLEC2D* editing site (chr12:9688559) specific to B cells.

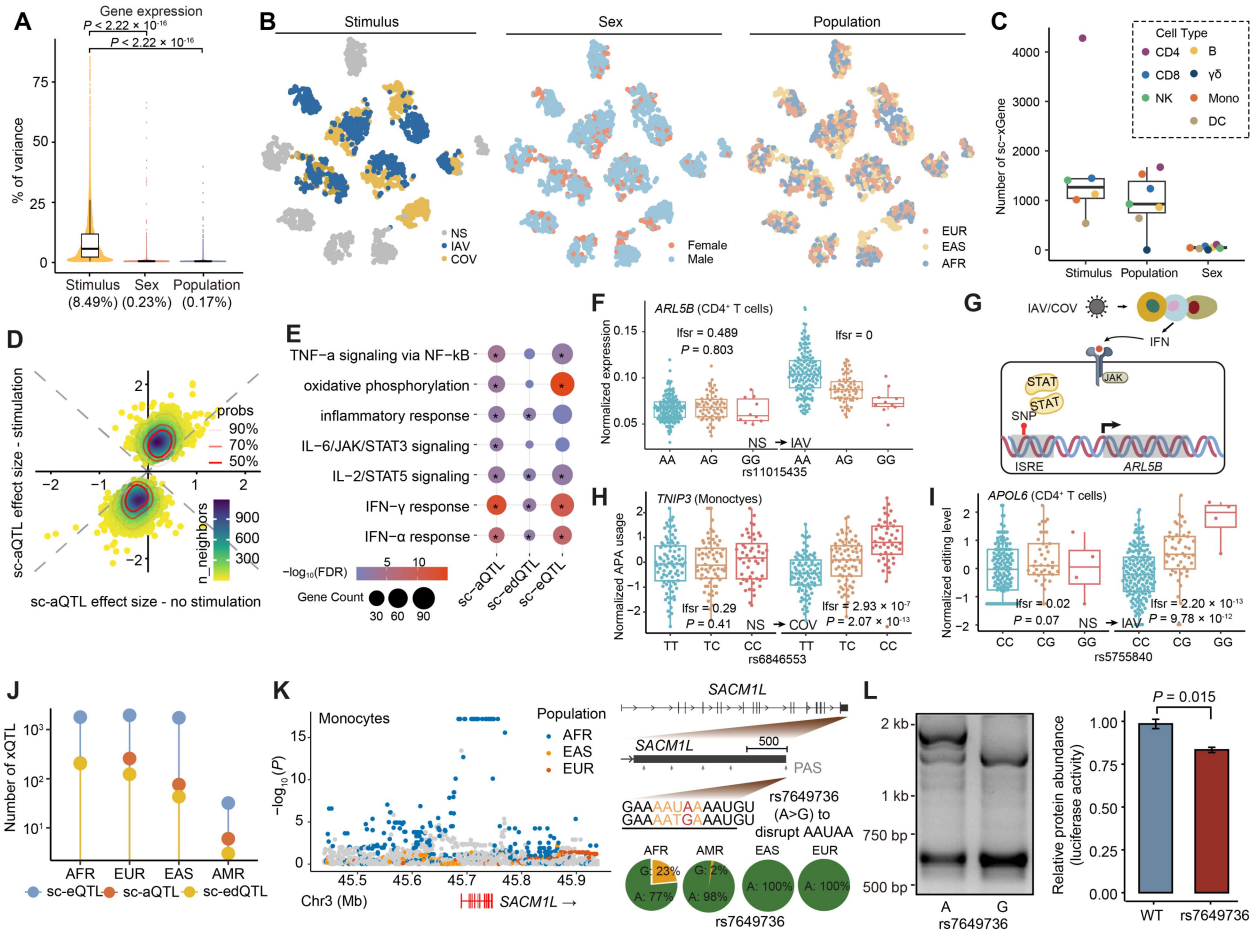

Fig. 3. Context-dependent genetic regulation of immune cell function. (**A**) Relative contribution of factors (stimulus, sex, and population) to gene expression variation in popCell data. *P*-values calculated using two-sided Wilcoxon rank-sum test. (**B**) t-SNE projection of immune cells stratified by stimulus condition, sex, and population. (**C**) Number of stimulus-dependent, population-specific, and sex-biased sc-xQTLs identified in major immune cell types. (**D**) Comparison of sc-aQTL effect sizes between non-stimulated and stimulated conditions. (**E**) Gene Ontology (GO) enrichment for stimulus-dependent sc-xGenes. Tests with FDR < 0.05 are indicated by ‘ * ’. (**F**) Stimulus-dependent regulation of *ARL5B* expression by rs11015435 genotypes. (**G**) Schematic diagram illustrating that rs11015435 modulates *ARL5B* expression via IFN-induced JAK/STAT-ISRE signaling. IFN, Interferon; ISRE, IFN-stimulated response element. (**H**) Stimulus-dependent regulation of *TNIP3* APA usage by rs6846553 genotypes. (**I**) Stimulus-dependent regulation of *APOL6* RNA editing levels (chr22:35661178:+) by rs5755840 genotypes. (**J**) Population-specific/biased sc-xQTLs across four major populations. (**K**) Example of population-specific sc-aQTL. (Left) AFR-specific aQTLs regulating *SACM1L* APA usage in monocytes. (Right) rs7649736 disrupts PAS motif specific to African owing to its high allele-G frequency. (**L**) Experimental validation of rs7649736 effects on *SACM1L* APA. (Left) 3′RACE confirmed G allele of rs7649736 promotes proximal PAS usage of *SAMC1L*. (Right) Dual-luciferase assays comparing *SACM1L* 3′UTR (short versus long) in HEK293 cells (n = 3; P = 0.015 by two-sided t-test).

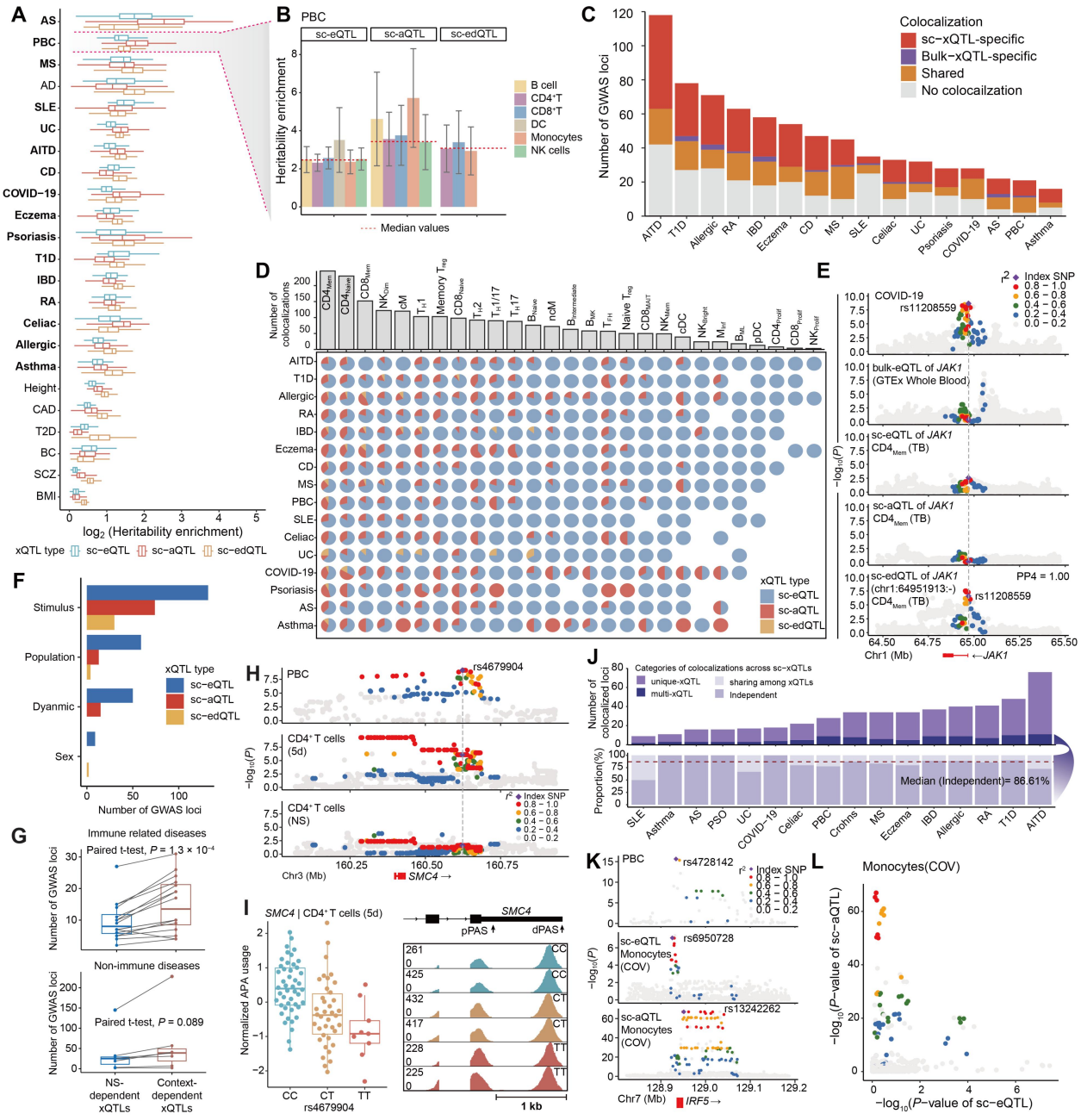

Fig. 4. Cell type-specific genetic regulation contributes to immune disease heritability. (**A**) Heritability enrichment for sc-xQTLs among GWAS associations for seven non-immune-related diseases and 16 immune-related diseases (bold) across cell types and conditions. AD, Alzheimer's disease; BC, breast cancer; CAD, coronary artery disease; SCZ, schizophrenia; BMI, body mass index; T2D, type 2 diabetes; AS, ankylosing spondylitis; PBC, primary biliary cholangitis; COVID-19, coronavirus disease 2019; UC, ulcerative colitis; CD, Crohn's disease; SLE, systemic lupus erythematosus; AITD, autoimmune thyroid diseases; MS, multiple sclerosis; T1D, type 1 diabetes; IBD, inflammatory bowel disease; RA, rheumatoid arthritis. (**B**) Cell type-specific PBC heritability estimates (±SE; red dash = median). (**C**) Colocalizations of sc-xQTLs and bulk xQTLs across immune diseases. **D**) Colocalization events between sc-xQTLs from 33 cell types and 16 immune-related diseases. (Top) The total number of colocalization events per

immune cell type, while (Bottom) the proportion of different sc-xQTLs colocalization across cell types and traits. (**E**) *JAK1* locus associations for: (from top to bottom) COVID-19 GWAS, GTEx whole-blood cis-eQTL, $CD4_{Mem}$ cell (TB) sc-eQTL, $CD4^+$ T_{mem} cell (TB) sc-aQTLs, and $CD4_{Mem}$ cell (TB) sc-edQTLs. TB: tuberculosis. (**F**) Number of context-specific xGenes associated with the immune disease GWAS loci. (**G**) Comparative analysis for: GWAS loci associated with context-dependent xQTLs versus non-stimulus-dependent xQTLs for (Top) immune-related diseases and (Bottom) non-immune-related diseases. (**H**) Colocalization between PBC GWAS associations and stimulus-dependent sc-aQTL for *SMC4* in activated $CD4^+$ T cells (5d). (**I**) Left: Activated $CD4^+$ T cell-specific allelic regulation of *SMC4* APA usages by rs4679904. Right: *SMC4* APA usage differ by rs4679904 genotype. pPAS: proximal PAS; dPAS: distal PAS. (**J**) (Top) Number of unique (single molecular phenotype) versus shared (across different molecular phenotypes) colocalized loci; (Bottom) Proportion of independent (median = 86.61%, red dash) versus shared effects in the multi-xQTLs colocalized loci. (**K**) Colocalization between PBC GWAS signals with sc-eQTL and sc-aQTL in stimulated monocytes (COV), $PP4_{sc\text{-}eQTL\ and\ sc\text{-}aQTL} = 0.00017$. (**L**) Scatter plot at the *IRF5* locus with sc-eQTL and sc-aQTL *P*-values in monocytes (COV).

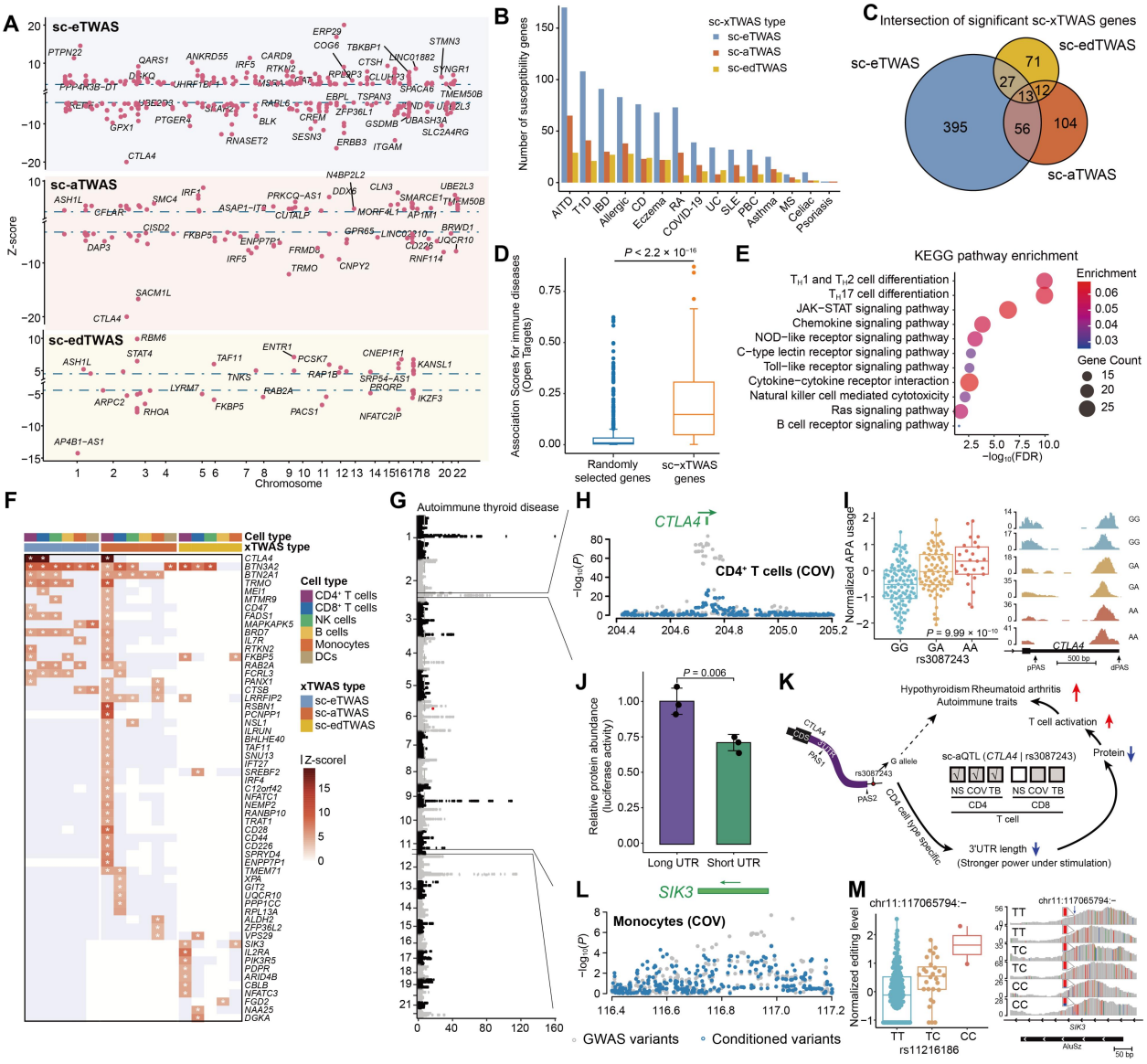

Fig. 5. Single-cell transcriptome-wide association studies reveal cell type-specific disease mechanisms. (**A**) Miami plot displaying Z-score values from (Top) sc-eTWAS, (Middle) sc-aTWAS, and (Bottom) sc-edTWAS across 16 immune-related diseases, with top genes highlighted per chromosome. (**B**) Genes (FDR < 0.05) identified by sc-xTWAS. (**C**) Intersection of genes identified through different sc-xTWAS approaches. (**D**) Comparison of Open Target Platform association scores between TWAS-identified and randomly selected genes. *P*-values calculated using two-sided Wilcoxon rank-sum test. (**E**) KEGG enrichment of sc-xTWAS identified genes. (**F**) Heatmap showing cell type-specific patterns of sc-xTWAS genes in AITD. (**G**) Manhattan plot of AITD associations and highlighting of risk loci. (**H**) *CTLA4* locus showing AITD GWAS signals before (gray) and after (blue) conditioned on predicted APA usage. (**I**) Left: $CD4^+$ T cell (COV)-specific allelic regulation of *CTLA4* APA usages by rs3087243 genotype. Right: *CTLA4* APA usage varies by rs3087243 genotype. (**J**) Dual-luciferase assays comparing *CTLA4* 3′UTR (short versus long) in HEK293 cells ($n = 3$; $P = 0.006$ by two-sided t-test). (**K**) Model of alternative alleles of rs3087243 affecting *CTLA4* APA usage promoting AITD risk. (**L**) *SIK3* locus showing AITD GWAS signals before (gray) and

after (blue) conditioned on predicted APA usage. (**M**) Left: Monnocytes (COV)-specific *SIK3* editing site (chr11:117065794:-) regulated by rs11216186. Right: RNA editing level on *SIK3* differ by rs11216186 genotype.

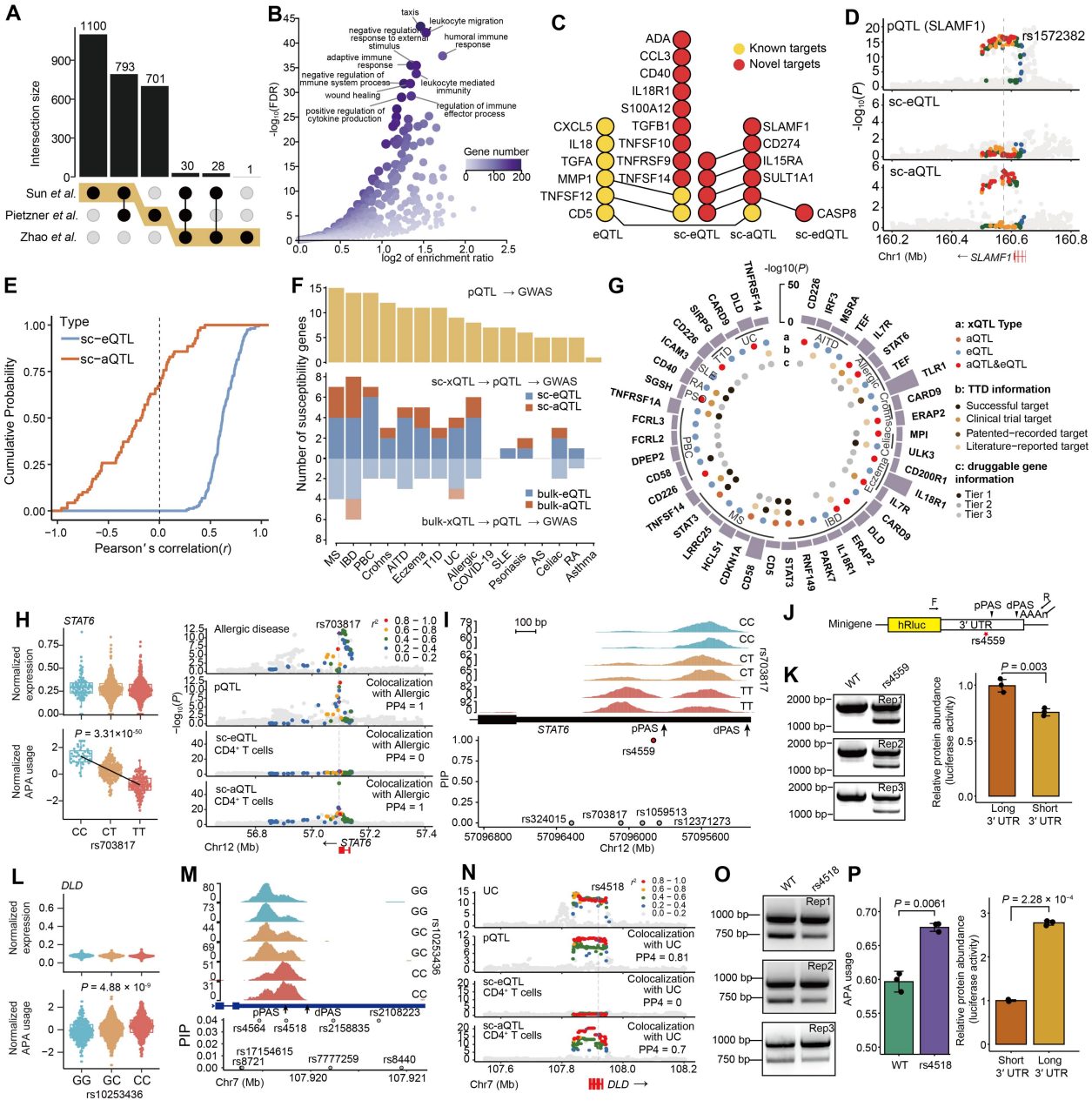

Fig. 6. Integration of protein-level evidence reveals immune disease therapeutic targets. (**A**) Consolidated pQTLs from three datasets (non-redundant in yellow). (**B**) GO enrichment of 2,653 *cis*-pQTL genes. (**C**) Novel (red) and known (yellow) genes with pQTLs colocalized with bulk-eQTL and our sc-xQTLs. Edges connect identical genes. Genes with pQTLs and colocalized bulk-eQTLs were extracted from study of Zhao et al. (**D**) Colocalization between sc-e/aQTL and pQTL for *SLAMF1*. (**E**) Empirical cumulative distribution curves of Pearson correlations between sc-xQTL and pQTL effect sizes. (**F**) Number of colocalizations: (Top) pQTL-GWAS; (Middle) sc-e/aQTL-pQTL-GWAS; (Bottom) GTEx-e/aQTL-pQTL-GWAS. (**G**) Druggable targets among susceptibility genes. Circle a: sc-xQTL type; Circle b: Therapeutic Target Database (TTD) information, Circle c: druggability. Tier 1: targets with approved drugs or clinical-phase candidates; Tier 2: with drug-like compounds (≥ 50% structural similarity to

approved drug targets); Tier 3: with lower(< 50%) similarity to approved drugs. The -log10 (*P*) represents GWAS significance. (**H**) Potential druggable target *STAT6*. Left: Allelic effect of rs703817 associated with differential *STAT6* APA usage but not its expression. Right: Allergic GWAS signal, pQTLs, and sc-aQTLs at the *STAT6* locus. (**I**) Differential *STAT6* 3′UTR usage regulated by SNPs in 95% credible set. pPAS: proximal PAS; dPAS: distal PAS; PIP: posterior inclusion probability. (**J**) Minigene reporter assay to validate the causal effect of rs4559 on *STAT6*. F, forward primer; R: reverse primer. (**K**) Left: 3′RACE confirmed T allele of rs4559 promotes *STAT6* 3′UTR shortening. Right: Luciferase activity assay showing the reduced protein level associated with the shortened 3′UTR of *STAT6* (n = 3; P = 0.003 by two-sided t-test). **(L)** APA-linked risk gene *DLD* in ulcerative colitis. Allelic effect of rs10253436 associated with differential *DLD* APA usage but not its gene expression. (**M**) Differential *DLD* 3′UTR usage regulated by SNPs in 95% credible set. (**N**) Ulcerative colitis GWAS signal, pQTLs, and sc-aQTLs at the *DLD* locus. (**O**) Minigene reporter assay to validate the causal effect of rs4518 on *DLD* 3′UTR region. (**P**) 3′RACE analysis and luciferase activity validating *DLD* APA changes associated with reference and alternative alleles of rs4518 (n = 3; P -values calculated by two-sided t-test).