communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-08115-7

Single-cell multiomic analysis unveils the immune landscape dynamics of graves' ophthalmopathy

Check for updates

Shurui Ke[1,2], Zhong Liu[1,2], Ming Zhou[1], Bing Xiao[1], Qiuling Hu[1], Jinshan Lin[1], Huijing Ye[1] ✉, Lixia Luo[1] ✉ & Xianchai Lin[1] ✉

Graves' ophthalmopathy (GO) is an autoimmune disease that occurs concurrently with Graves' disease, potentially leading to facial disfigurement and irreversible vision loss. However, the molecular characteristics of cellular changes in the peripheral immune environment of GO patients remain unclear. This study presents a multi-omics single-cell analysis of peripheral blood from GO patients and healthy controls, revealing key molecular alterations in immune cell subpopulations. We identify increased chromatin accessibility and regulatory activity of pro-inflammatory factors, such as *FOS* and *NF-κB* family. Immune repertoire analysis shows enhanced diversity in GO patients. Notably, GO-specific clonal expansions are enriched in CD8 effector T (CD8 Te) cells, which exhibit signs of enhanced T cell chemotaxis and exhaustion. Multi-omics integration analysis reveals that changes in *SLC35G1* and *IDNK* expression in CD8 Te are linked to disease phenotypes, with trends aligning with experimental results. We find that this cell cluster could infiltrate into orbital tissues and upregulate fibrosis-related pathways associated with fibroblasts in GO. These findings indicate its potential role in driving disease progression. Our study reveals the epigenetic and transcriptomic landscape in the peripheral blood of GO patients, enhancing the understanding of the pathogenic molecular mechanisms and offering potential directions for prevention and treatment.

Graves' ophthalmopathy (GO) represents a severe and often irreversible autoimmune condition that frequently accompanies Graves' disease (GD)[1]. This disorder is marked by distinctive symptoms, including eyelid retraction, exophthalmos, and optic nerve compression[2], which may progress to vision impairment. The underlying causes of GO are not entirely clear, though it is generally accepted that they are linked to autoantibodies targeting the thyroid-stimulating hormone receptor (TSHR). Additionally, factors such as age, gender, and smoking habits are considered as contributors to the risk of developing GO among individuals with GD[3–5].

Currently, the primary management strategy for GO involves the administration of intravenous glucocorticoids. However, the effectiveness of this treatment is variable, and it carries a risk of side effects and disease recurrence, often necessitating surgical interventions to alleviate symptoms[1,6]. Despite ongoing research and clinical trials investigating alternative therapies, such as inhibitors of lymphocyte proliferation or IGF-IR antagonists, many patients continue to suffer from persistent symptoms[1,7]. This underscores the need for a more profound comprehension of GO's pathogenesis to facilitate the discovery of previously unrecognized therapeutic targets and strategies.

The progression of tissue pathology in GO is influenced by immune cells, with alterations streaming from the aberrant responses of effector T cells, monocytes, plasma cells, and pro-inflammatory cytokines within a genetic predisposed environment[8–14]. Previous studies have indicated that these immune cells infiltrate the orbital tissues, releasing cytokines like interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α). These cytokines activate orbital fibroblasts, which then secrete hyaluronic acid and undergo differentiation into adipocytes and myofibroblasts, leading to the typical pathological manifestations of GO, such as orbital tissue expansion and fibrosis[6,15,16]. Despite these insights, our understanding of the specific pathogenic cells and genetic factors involved in GO remains limited.

Advancements in single-cell sequencing technologies have opened extra avenues for investigating the intricate genetic landscape of disease-affected cells. While previous studies have analyzed the genetic transcriptome of peripheral blood mononuclear cells (PBMCs)[17] and the

[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangzhou, China. [2]These authors contributed equally: Shurui Ke, Zhong Liu.
✉e-mail: yehuijing@qq.com; luolixia@mail.sysu.edu.cn; linxch7@mail.sysu.edu.cn

pathological tissue of the orbit in GO[18], the full extent of heterogeneity in the single-cell epigenomes of PBMCs has yet to be fully explored. By integrating single-cell transposase-accessible chromatin sequencing (scATAC-seq) with single-cell RNA sequencing (scRNA-seq) and single-cell immune repertoire analysis (scTCR-seq and scBCR-seq), we can uncover potential disease-related regulatory programs.

In this study, we have simultaneously generated data on the single-cell transcriptome, single-cell epigenome, and single-cell immune repertoire data of GO patients. This approach has enabled us to identify disease-specific gene regulatory networks and TCR/BCR amplification, pinpoint potential transcription factor (TF) activators of chromatin accessibility and gene expression in GO, and correlate single-cell omics data with known disease-specific genetic variations. Notably, our findings highlight the involvement of specific cell types, such as CD8+ T cells, in the pathophysiology of GO, offering valuable insights into the peripheral immunopathology associated with this disease.

Results

High-resolution single-cell multi-omics peripheral immune cell-type mapping Of Go patients

We performed droplet-based scRNA-seq, single-cell scTCR-seq and scBCR-seq, as well as scATAC-seq (using 10X Genomics technology) to delineate the immune landscape of PBMCs from 21 GO patients and 10 healthy controls (HC) (Fig. 1a) (Supplementary Data 1). After stringent quality control filtering, we retained 202,983 cells from the scATAC-seq dataset and 285,993 cells from the scRNA-seq dataset for downstream analysis (Fig. 1b, S1a, S1b, S1c). We used Harmony[19] to remove batch effects across all data (Fig. S1d). Firstly, we manually annotated seven cell clusters based on the chromatin accessibility of key lineage markers in PBMCs: T cells, B cells, NK cells, monocytes, dendritic cells (DCs), progenitor cells, and cycling cells (Fig. 1b). We assessed chromatin accessibility for key lineage-specific immune cell genes (Fig. 1d). T cells displayed increased accessibility near CD3D, NK cells at NCR1, B cells at MS4A1, monocytes at S100A9, DC cells at CD1C and CLEC4C, progenitors at GATA2, and cycling cells at MKI67. Similarly, for scRNA-seq, we classified cells based on the expression of lineage marker genes: T cells, B cells, NK cells, DCs, progenitor cells, and cycling cells (Fig. 1b) (Supplementary Data 2). Heatmaps show the enrichment of marker gene expression in each cell cluster (Fig. 1e). After removing batch effects, the proportions of annotated major cell clusters did not show significant differences, and Milo analysis showed consistent results (Fig. 1c, S1e). Furthermore, we used chromVAR[20] to calculate TF motif deviations, examining TF motif activity by estimating the enrichment of TF binding motifs in open chromatin regions across immune cells. The results showed *TCF7L2* activity in the T cells, shared *EOMES* factor activity in myeloid cells, *POUF2F2* factor activity in B cells, *SPIC* enrichment in monocytes and DC cells, and *GATA2* factor activity in progenitors (Fig. 1f). In summary, we delineated the high-resolution transcriptomic and epigenomic peripheral immune landscape of GO at single-cell resolution. Additionally, we integrated other scRNA-seq datasets[17,18], including PBMC data and orbital tissue from HC, GD, and GO donors, to provide a more detailed investigation of immune cell heterogeneity (Fig. S1f, S1g).

Multi-omics characterization of T cell heterogeneity in GO

Current research suggests that CD4 + T cells are the major lymphocytes involved in the progression of inflammation and infiltration of orbital tissues in GO[21–23]. The balance between Th1 and Th2 cells has been implicated in the progression of GO[9,10]. The pathogenicity of Th17 cells in GO has also been demonstrated[11,12]. CD4 CTLs are involved in the occurrence and development of GO[8]. However, there has been relatively little in-depth analysis of the overall T cell subtypes in GO patients at the epigenomic level. To further unravel the epigenetic changes in T cell subsets in GO disease, we identified T cell subgroups in the scATAC-seq dataset using gene activity scoring (GAS). Specifically, we divided T cell subtypes into 13 subgroups based on the activity scores of cell-type marker genes (Fig. 2a, S2a), including CD4 naive T cells (CD4 Tn; CD4 + SELL +), CD8 naive T cells (CD8 Tn; CD8 + SELL +), CD8 central memory T cells (CD8 Tcm; CD8 + GPR183 + SELL +), CD8 effector memory T cells (CD8 Tem; CD8 + GPR183 + GNLY +), CD8 effector T cells (CD8 Te; CD8 + GNLY +), CD4 cytotoxic T cells (CD4 CTL; CD4 + GNLY +), CD4 T helper 17 cells (Th17; CD4 + IL7A +), CD4 T helper 1 cells (Th1; CD4 + IFNG +), CD4 T helper 2 cells (Th2; CD4 + CCR4 +), CD4 T follicular helper cells (Tfh; CD4 + ICOS +), CD4 regulatory T cells (CD4 Treg; CD4 + FOXP3 +), mucosal-associated invariant T cells (MAIT; CD8 + KLRB1 +), and γδT cells (gdT; CD8 + KLRK1 +). The same cell subgroups were also identified based on marker gene expression in the scRNA-seq atlas (Fig. S3b, S3c) (Supplementary Data 2).

Our scRNA-seq data combined with Milo[24] analysis revealed an increased proportion of Th1 cells in GO patients, while the proportions of Th17, CD8 Te, and CD4 CTL decreased (Fig. 2b, S3a). These cells have been reported to expand in autoimmune disease patients and are associated with disease progression[25–27]. To further identify specific functional differences in immune cells during disease, we determined the differentially accessible regions (DARs) and differentially expressed genes (DEGs) between GO and HC in the scRNA-seq and scATAC-seq datasets (Fig. 2d, e, S3d, Supplementary Data 3). Among the major effector T cell populations, we found that CD4 CTLs had the higher number of cell-specific upregulated DARs and DEGs, respectively, suggesting increased heterogeneity of these cells in the disease. This is consistent with previous reports indicating that CD4 CTLs can generate strong cytotoxic responses against GO autoantigens. Furthermore, we performed gene ontology analysis of the upregulated DARs and DEGs in each cell population (Fig. 2f, g). The results showed that transcription-related pathways and cellular metabolism pathways were enriched in the majority of T cells' DARs. Additionally, ribosome-related pathways were enriched in the majority of T cells' DEGs, suggesting that T cells in GO patients are in a highly active state, possibly due to stimulation by abnormal autoantigens[8,11,28,29].

TFs play a crucial role in immune cell fate and have been implicated in the pathogenesis of autoimmune diseases, such as the *STAT* family and *BCL6* in systemic lupus erythematosus[30,31]. We performed motif enrichment analysis to further clarify the upregulated TFs in GO effector T cells (Fig. S3d). We found that the *WT1* and the *ZNF* family were upregulated in several major effector cell populations, both of which have been reported to be associated with immune imbalance in autoimmune diseases[32,33]. Furthermore, the loss of immunosuppressive function in Tregs is recognized to be associated with the pathogenesis of autoimmune diseases, so we identified a cluster of TFs with decreased activity in Tregs. We found that *LMO2* showed the most manifest downregulation, which has been reported to be associated with T-lineage development[34]. Thus, we speculated that the decreased activity of the transcription factor *LMO2* may affect the proliferation and differentiation of Tregs, thereby contributing to the development of GO.

In autoimmune states, T cells can produce a specific TCR repertoire clone expansion in response to disease-specific antigens[35,36]. We examined the T cell receptor immune repertoire of GO patients through scTCR-seq. Notably, we found that the proportion of unique TCR clonotypes in GO patients was significantly increased, and similarly, the distribution of TCR repertoire diversity was higher in GO patients (Fig. 2h, S4a, S4b). Moreover, the GO-specific clones expansion was predominantly enriched in CD8 Te cells, which suggests a potential importance of this cell population in disease progression (Figs. 2i, j).

Under normal immune conditions, CD8 Te cells can be activated by specific antigens, leading to a rapid and robust cytotoxic response that maintains health and immune homeostasis[37]. However, in the context of autoimmune disease, CD8 Te cells may undergo clonal expansion, migrate to local lesions, and exhibit abnormal activation or functional exhaustion[38]. In our study, we observed a significant reduction in the proportion of CD8 Te cells in peripheral blood under GO conditions (Figs. 2b, c). Previous studies have reported similar findings in autoimmune diseases such as rheumatoid arthritis, where the proportion of

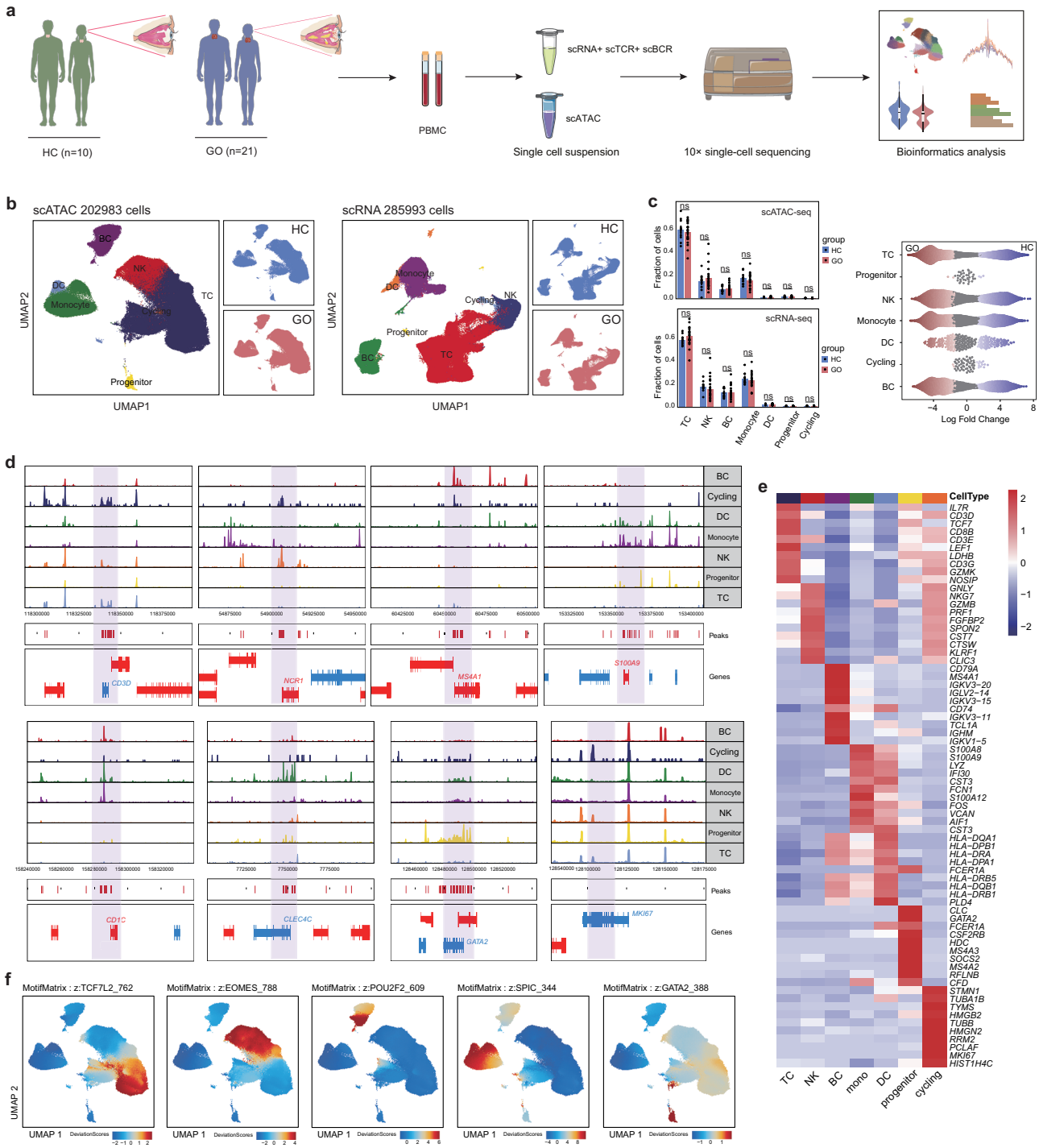
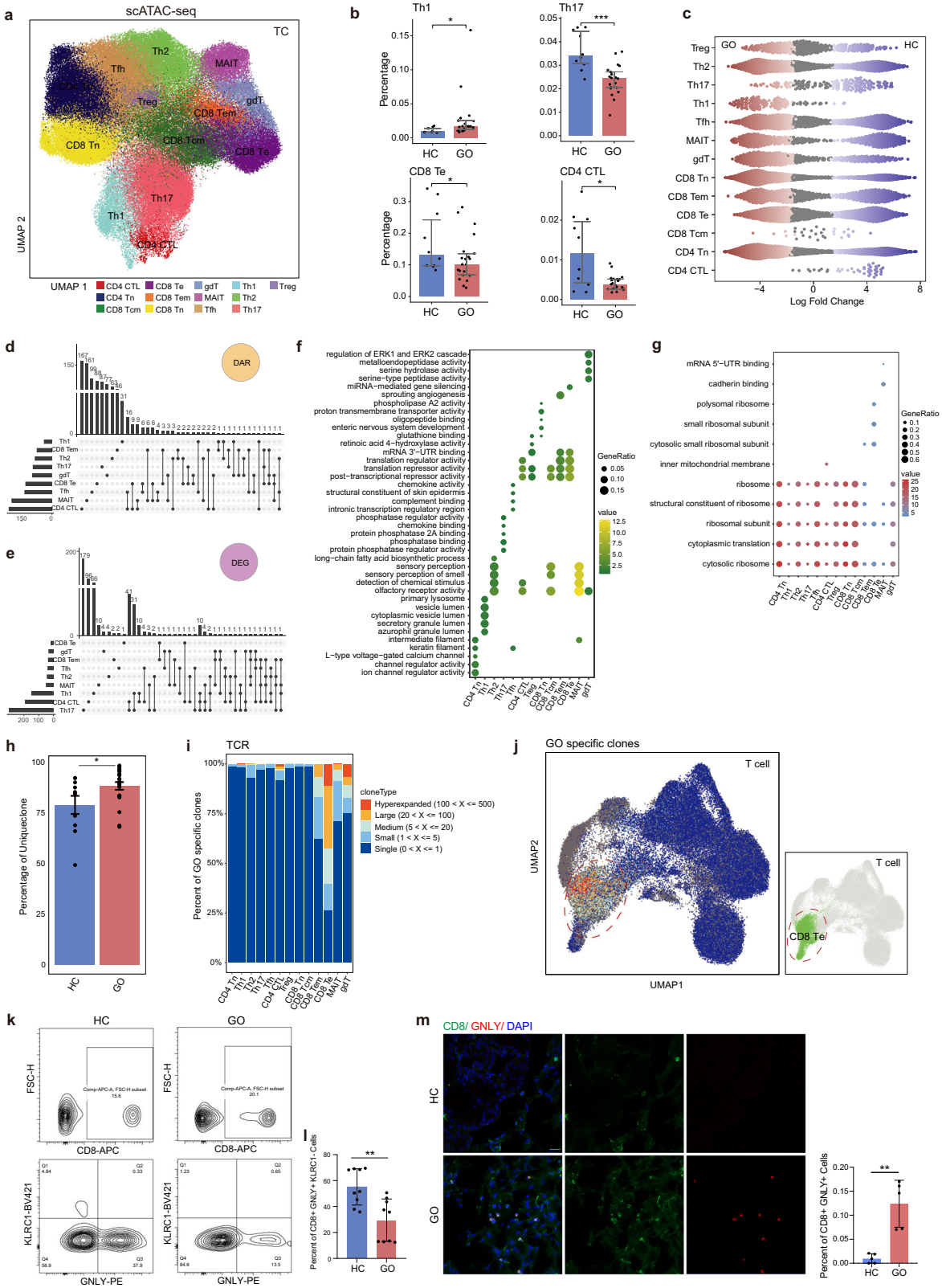


Fig. 1 | High-throughput single-cell multi-omics profiling of GO patients and healthy human peripheral blood immune cells. a Schematic representation of single-cell multi-omics sequencing (scRNA-seq, scATAC-seq, scTCR-seq, scBCR-seq) experimental design. peripheral blood immune cells (PBMCs) were isolated from healthy controls (HC, $n = 10$), and Graves' ophthalmopathy (GO, $n = 21$), followed by processing using the 10X Genomics platform. This image was adapted from Servier Medical Art (licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/). **b** UMAP plots depicting major immune cell types in peripheral blood based on scATAC-seq and scRNA-seq datasets between GO and HC. **c** Left: Proportion of each immune cell type in scATAC-seq and scRNA-seq analyses. Right: Proportion of each immune cell type in scRNA-seq after processing with the Milo algorithm. *P < 0.05 by Mann-Whitney U test. **d** Aggregate accessibility profiles for scATAC-seq main cell type at canonical cell-type marker genes. **e** Row-normalized single-cell gene expression heatmap of main immune cell-type marker genes for scRNA-seq. **f** UMAP projection of scATAC-seq profiles colored by chromVAR TF motif bias-corrected deviations for the indicated factors.

CD8 Te cells decreases in peripheral blood but increases in inflamed joint tissues. Thus, we speculated that CD8 Te cells may migrate from the peripheral blood to the orbital lesions in GO. To further validate this hypothesis, we performed an integrated analysis of T-cell subsets from orbital tissue and simultaneously analyzed T-cell subsets from another PBMC dataset to ensure reliability (Fig. S3e, S3f). The results showed a consistent decrease in the proportion of CD8 Te in PBMCs, while a significant increase was observed in local lesions (Fig. S3g, S3h). Moreover, chemokine activity gene set scores for CD8 Te cells were significantly elevated in GO, along with increased cytotoxicity and



exhaustion scores (Fig. S3i). To accurately identify CD8 Te cells and exclude the interference of gdT cells (Fig. S3c), we utilized flow cytometry to detect the proportion of CD8 + GNLY + KLRC1- cells in PBMCs. The results showed a significant increase in this population in the GO group (Figs. 2k, l) (Supplementary Data 4). Additionally, marked infiltration of CD8 + GNLY+ cells was observed in the orbital tissues of GO patients (Fig. 2m) (Supplementary Data 4). These findings suggested that CD8 Te cells may migrate from the peripheral blood to local orbital lesions, exhibiting increased cytotoxicity and exhaustion, which may be closely related to the development and progression of GO.

Fig. 2 | scATAC-seq, scRNA-seq and scTCR-seq analysis of the changes in T cell subsets in GO. a UMAP projections of T cell subsets of scATAC-seq dataset. **b** Differences in the proportions of Th1, Th17, CD8 Te and CD4 CTL among HC ($n = 10$) and GO ($n = 21$). Data are represented as the median [IQR]. $^*P < 0.05$ by Mann-Whitney U test. **c** Proportion of each T subsets cell type in scRNA-seq after processing with the Milo algorithm. **d** UpSet plot showing the integrated comparative analysis of upregulated DARs in effector T cells of the GO group compared to HC. **e** UpSet plot showing the integrated comparative analysis of upregulated DEGs in effector T cells of the GO group compared to HC. **f** Enriched gene oncology terms in upregulated DARs of T cell subsets in the GO group. **g** Enriched gene oncology terms in upregulated DEGs of T cell subsets in the GO group. **h** Bar chart illustrating the proportion of unique TCR clones in HC and GO groups. Data are represented as median [IQR]. $^*P < 0.05$ by Mann-Whitney U test. **i** Bar chart showing expression abundance of GO-specific clones across T cell subsets in the GO group. **j** UMAP plot mapping GO-specific clones in T cells. Contour plots (**k**) and bar chart (**l**) indicating CD8 + GNLY + KLRC1- cell proportions from HC and GO donors. **m** Immunofluorescence images showing the infiltration of CD8 + GNLY+ cells in the orbits of HC and GO donors (left). Bar chart illustrating the difference in the proportion of CD8 + GNLY+ cells infiltration in HC and GO groups (Right). Data are represented as the median [IQR]. $^*P < 0.05$ by Mann-Whitney U test. scale bar: 20 μm.

Multi-omics characterization of B cell heterogeneity in GO

We investigated the changes in B cells in GO patients, as B cells play a crucial role in the progression of GO[39]. B cells were manually annotated into four subtypes: naive B cells (NB, *TCL1A* + *CCR7* +), memory B cells (MB, *CD27* + *GPR183*+), activated B cells (ActB, *EGR* + *CD69* +), double-negative B cells (DN2B; *ITGAX* +), and plasma cells (PB, *CD38* + *JCHAIN* +) (Fig. 3a, S5a, S5b) (Supplementary Data 2). Notably, the proportion of PB was significantly decreased in GO patients, while the proportion of ActB was increased, with no differences observed in the proportions of other subgroups (Fig. 3b, S5c). Interestingly, unlike the scATAC-seq data, naive B cells were dispersed into nine distinct clusters in the scRNA-seq data. We identified the top marker gene sets for these nine clusters and performed gene ontology analysis, revealing that each NB subgroup has different functional biases (Fig. 3a, S5d) (Supplementary Data 2).

To further identify the molecular mechanisms associated with B cell heterogeneity in GO, we identified the top 2 upregulated TFs in B cells, *SNAI1* and *NR5A2*, which have been reported to be involved in immune regulation and autoimmune diseases[40,41]. We performed footprint analysis of these two TFs to further clarify the situation of B cell subgroups. The results showed that PB had the highest accessibility (Fig. 3d, e). Next, we conducted gene ontology analysis of the GO-upregulated DARs and DEGs in each B cell subgroup (Fig. 3f, g). The results of the DARs indicated that B cell subgroups were enriched in pathways related to signal transduction and microenvironmental responses. The DEGs analysis revealed that the B cell subgroups were enriched in multiple ribosome-related pathways, consistent with findings in T cell subgroups. Additionally, PB showed upregulation of several endoplasmic reticulum-related pathways, suggesting a high level of protein synthesis, such as antibody production. These findings suggested that B cell subtypes in GO patients exhibited more active functions compared to HC, implying enhanced humoral immunity in patients with GO.

Specific expansion BCR repertoire plays an important role in the pathogenesis of autoimmune diseases[42,43]. Based on this, we performed scBCR-seq. The results showed that the proportion of unique clones in GO patients and the distribution of BCR repertoire diversity were significantly increased (Fig. 3h, S5e, S5g). Further analysis of GO-specific clones expansion revealed that the peripheral blood PB exhibited the most significant expansion (Fig. 3i), further indicating the activation of humoral immunity in these individuals. Considering that NB cells are divided into 9 clusters (Fig. 3a), we hypothesized that they may not only be related to functional differences (Fig. S5d) but also represent the expansion of different BCR clones during the pathogenesis of GO. Therefore, we mapped the CT genes of the BCR data to the UMAP plot to explore the differences in clone expansion among the NB subgroups. We observed visually evident single-clone expansion in the naive B cell subgroups NB3, NB5, and NB7, with the most pronounced expansion seen in NB3 among GO patients (Fig. 3j, S5f). Besides, the proportion of NB3 cells was significantly increased in GO (Fig. 3k). Given the enrichment of B cell activation-related pathways in the gene ontology analysis of the markers for NB3 cells (Fig. S5d), we speculated that this group of cells is abnormally activated in the context of GO.

Multi-omics characterization of NK cell heterogeneity in GO

Despite studies indicating the involvement of NK cells in the immune dysregulation of GO, the pathogenic mechanisms of NK cell subtypes in GO still require further exploration[13,14]. As previously demonstrated, based on gene activity scores (GAS) and marker genes, peripheral NK cells were identified as three subtypes: active NK cells (NK1; *CD160*^{high}), resting NK cells (NK2, *CD160*^{low}*SELL* + *XCL1* +), and NKT cells (NKT, *CD3D* +) (Figs. 4a, b, S6b) (Supplementary Data 2). In cell proportion comparison, only NK1 showed a significant increase in GO patients, while NKT exhibited a decrease (Figs. 4c, d, S6a). To further explore the transcriptional and epigenetic remodeling features of NK1 cells, we identified the nearest genes to DARs and overlapped them with DEGs (Supplementary Data 3).We found that the *AP-1* TF family member *FOS*, which showed shared open states in NK1 cells, exhibited increased accessibility in GO patients, consistent with previous reports linking *AP-1* TF to the development of autoimmune thyroid diseases[44] (Fig. 4e). Next, we conducted gene ontology analysis of DARs and DEGs in NK cell subgroups (Fig. S6e, S6f). Interestingly, the DEGs of the NK subgroups were also enriched in pathways related to the ribosome. To further evaluate the TF driving NK1 differences, we identified TFs enriched in the peaks of upregulated differential accessibility in the 3 NK cell subgroups (Fig. S7a, S7b). Comparing the TF footprint analysis of NK1 cells between GO and HC, we found that the *NF-κB* transcription factor family (*REL*, *RELA*, and *RELB*) showed higher accessibility, which has been implicated in the inflammatory development of GD[45–47] (Fig. 4F). These findings suggested that NK1 cells in GO patients might be in an activated state of immune response and exhibited pro-inflammatory features.

Multi-omics characterization of myeloid cell heterogeneity in GO

Previous studies have recognized the pathogenic contribution of myeloid cells in the development of GD into GO[17,48,49]. To increase cellular resolution and dissect myeloid cells, we first clustered Monocytes and DCs separately, identifying two cell subtypes for each: Monocytes were divided into classical monocytes (CM, *CD14* + *S100A12* +) and non-classical monocytes (NCM, *FCGR3A* + *SIGLEC10*+); DC were divided into conventional DC (cDC, *cDC1C*+), and plasmacytoid DC (pDC, *CLEC4C*+) (Fig. 4g, h, i, j, S6c, S6d) (Supplementary Data 2). Although we did not observe significant changes in the percentages of the four cell populations (Fig. S6a), we noticed notable transcriptional and epigenetic remodeling in all cell subtypes, including DEGs and DARs (Fig. S6e, S6f, Supplementary Data 3). To further elucidate the pathogenic pathways, we conducted gene ontology analysis of the GO-upregulated DEGs and DARs for all cell subtypes (Figs. 4k, l). It is noteworthy that the DARs of CM were enriched in multiple pathways related to intermediate filaments and the cytoskeleton, which play critical roles in the migration, morphological changes, and signal transduction of myeloid cells. The results of differentially expressed genes (DEGs) indicated that these subtypes were predominantly enriched in mitochondrial-related pathways, suggesting an active cellular metabolism. Overall, myeloid cell subtypes maintained chromatin reprogramming and transcriptional changes, which may promote inflammatory responses in GO.

Multi-omics integration maps combined with GWAS identified disease-risk CD8 Te subpopulation

We inferred that integrating data from scATAC-seq and scRNA-seq datasets may further identify the chromatin accessibility information corresponding to gene expression of cell subtypes, facilitating the interpretation of key regulatory processes related to GO pathogenesis. Based on the

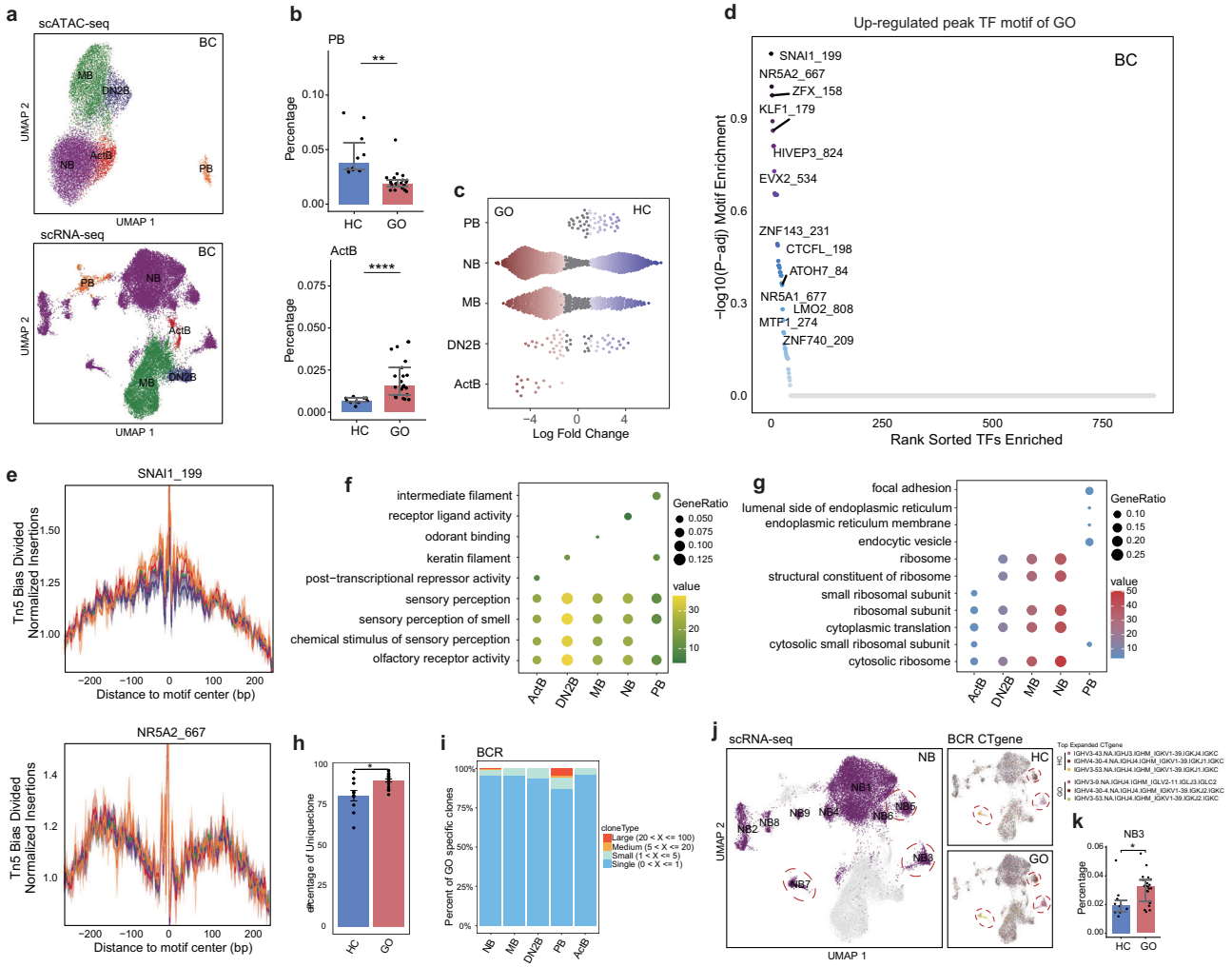


Fig. 3 | scATAC-seq, scRNA-seq and scBCR-seq analysis of the changes in B cell subsets in GO. a UMAP projections of B cell subsets of scATAC-seq and scRNA-seq dataset. **b** Differences in the proportions of PB and ActB cells among HC ($n = 10$) and GO ($n = 21$). Data are represented as the median [IQR]. $^{**}P < 0.001$, $^{****}P < 0.00001$ by Mann-Whitney U test. **c** Proportion of each B subsets cell type in scRNA-seq after processing with the Milo algorithm. **d** FDR-corrected P values from two-sided, two-sample t-tests of motif enrichment in differential peaks showing up-accessible peaks motif enrichment of total B cells in GO compared to HC. **e** Comparison of aggregate TF footprints for SNAI1_199 and NR5A2_667 among all B cell subsets in GO. **f** Representative gene oncology terms enriched in the nearest genes of upregulated DARs of B cell subsets in the GO group. **g** Representative gene oncology terms enriched in the upregulated DEGs of B cell subsets in the GO group. **h** Bar chart illustrating the proportion of unique BCR clones in HC and GO groups. Data are represented as median [IQR]. $^{*}P < 0.05$ by Mann-Whitney U test. **i** Bar chart showing the expression abundance of GO-specific clones across B cell subsets in the GO group. **j** UMAP plot showing the mapping of BCR CTgene in B cells. The purple color represents nine different naive B cell subsets, with NB3, NB5, and NB7 primarily composed of unique BCR expansion each. The CTgene of the top expanded clones in NB3, NB5, and NB7 are sequentially labeled in the HC and GO groups. **k** Differences in the proportions of NB3 cells among HC ($n = 10$) and GO ($n = 21$). Data are represented as the median [IQR]. $^{*}P < 0.05$ by Mann-Whitney U test.

aforementioned sub-clustering and manual cell annotation, we integrated transcriptomic data from the scRNA-seq dataset with chromatin accessibility data from the scATAC-seq dataset (Fig. S9a). As expected, GAS and gene expression matched, allowing us to distinguish 25 immune cell types (Fig. 5a).

Current research suggests that, except for very few cases caused by single gene mutations, most autoimmune diseases have a complex genetic background leading to disease susceptibility[50,51]. However, the resolution of GWAS signals intersecting with bulk data is hindered by cell type heterogeneity[52,53]. Therefore, to gain insights into how GD-related genetic variants affect the progression of GO at the cellular level, we assessed the enrichment of immune cell marker genes in GD-related genetic variants. We found that GD risk variants were significantly positively associated only with CD8 Te cells (Fig. 5b). Furthermore, we investigated which pathways were associated with GD variants enrichment in differentially expressed genes of CD8 Te cells (Fig. 5c). In the GO analysis of Cellular Component, Biological Process, and Molecular Function, we identified terms such as "ubiquitin-like protein ligase binding", "regulation of fibroblast migration", "natural killer cell mediated immunity", and "natural killer cell mediated cytotoxicity". Particularly, the term of "regulation of fibroblast migration" seems to imply the potential interaction between CD8 Te and the main pathogenic cells of GO, fibroblasts[6,15,16].

GO has been reported to occur more frequently in females; however, it tends to be more severe in males[1]. Consequently, we focused our investigation on the molecular characteristics of T cells and B cells within the GO group. Despite the absence of significant differences in cell proportions between sexes (Fig. S8a, S8b), we observed that PB and most T cell subpopulations showed elevated cytokine response scores in males (Fig. S8c, S8d). This indicates a heightened immune stress state in these cells among male patients. Notably, in CD8 Te cells, both chemokine activity scores and exhaustion scores were significantly higher in male patients, alongside increased cytotoxicity scores (Fig. S8e). These findings further support the

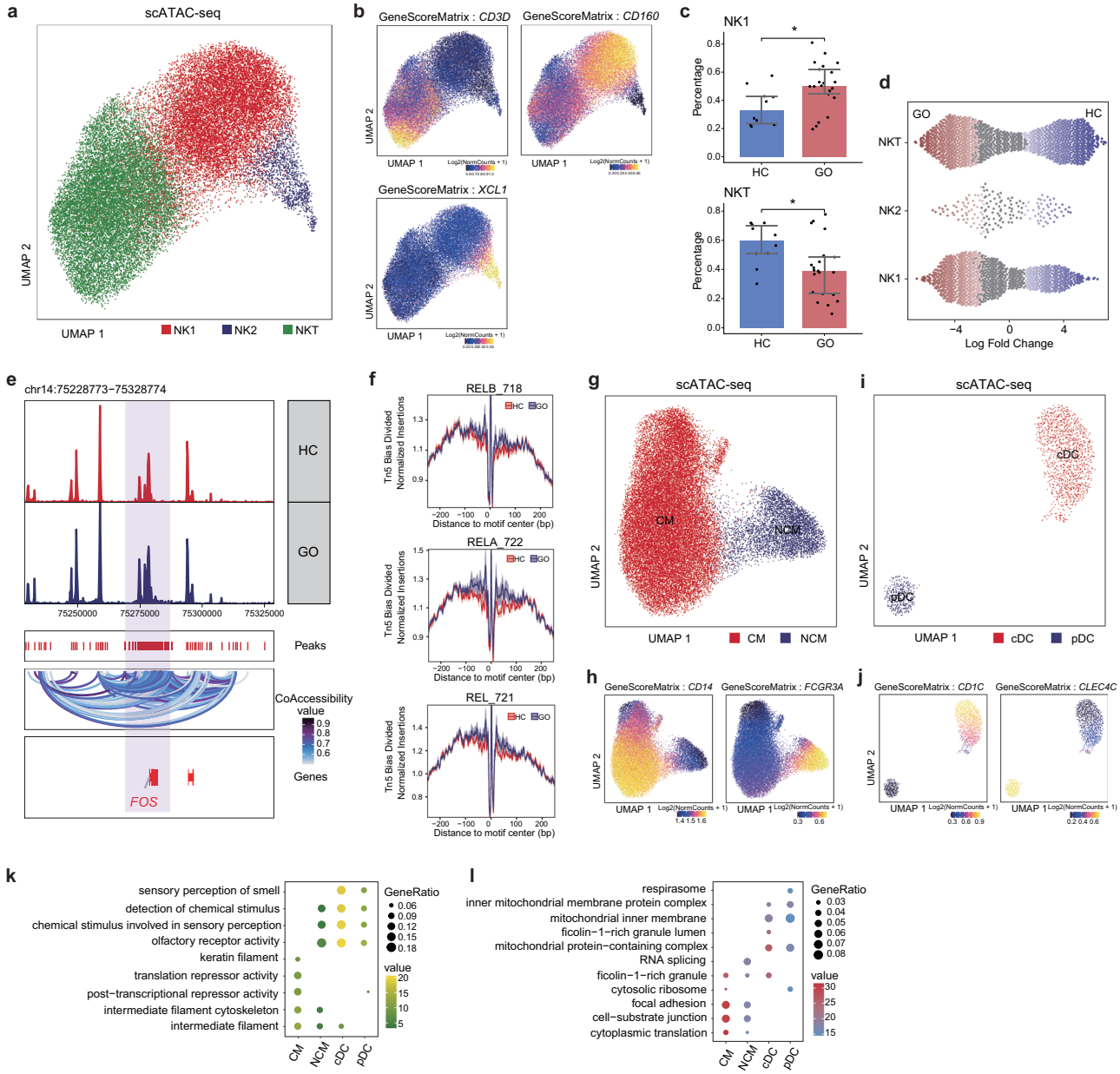


Fig. 4 | scATAC-seq and scRNA-seq analysis of the changes in NK cell and myeloid cell subsets in GO. a UMAP projections of NK cell subsets of scATAC-seq dataset. **b** UMAP projection colored by gene activity scores for the annotated lineage-defining genes in NK cell subsets of scATAC-seq dataset. The minimum and maximum gene activity scores are shown in each panel. **c** Differences in the proportions of NK1 and NKT cells among HC ($n = 10$) and GO ($n = 21$). Data are represented as the median [IQR]. $^*P < 0.05$ by Mann-Whitney U test. **d** Proportion of each NK subsets cell type in scRNA-seq after processing with the Milo algorithm. **e** Browser tracks with peak-to-gene links of *FOS* in NK1 cell. **f** Comparison of aggregate TF footprints for RELB_718, RELA_722, and REL_721 in NK1 cells from GO and HC. UMAP projections of Monocyte subsets (**g**) and DC subsets (**i**) of scATAC-seq dataset. UMAP projection colored by gene activity scores for the annotated lineage-defining genes in Monocyte subsets **h** and DC subsets (**j**) of scATAC-seq dataset. The minimum and maximum gene activity scores are shown in each panel. **k** Representative gene oncology terms enriched in the nearest genes of upregulated DARs of myeloid cell subsets in the GO group. **l** Representative gene oncology terms enriched in the upregulated DEGs of myeloid cell subsets in the GO group.

association between abnormal molecular changes in CD8 Te cells and the phenotypic progression of GO.

The expression trends of disease phenotype-related genes *SLC35G1* and *IDNK* in CD8 Te cells are consistent with the results of SMR analysis

GWAS identified that nearly 90% of disease-causing gene variants are located in non-coding regions, and the genes or functional DNA elements affected by these variants are often unknown[54]. Therefore, we combined gene transcript data from GWAS with expression quantitative trait loci (eQTL) data in blood to identify genes whose expression levels are associated with disease due to pleiotropy[55]. Two genes (*SLC35G1* rs61865961 and *IDNK* rs75397435) passed the Summary data-based Mendelian randomization (SMR) and HEterogeneity In Dependent Instruments (HEIDI) tests, with PSMR = 5.97 × 10^-6, PHEIDI = 0.25 for *SLC35G1* (Fig. 5d), and PSMR = 2.23 × 10^-3, PHEIDI = 0.99 for *IDNK* (Fig. 5g). No studies have yet identified the association of these two genes with GO disease. Furthermore, through SMR analysis, we identified the direction of effect between these two specific genes and our disease phenotype. The effect plots showed the correlation between the GWAS effect size of our discovered genes (*SLC35G1* and *IDNK*) and the meta-analysis cis-eQTL collection. Importantly, from these plots, the expression level of *SLC35G1* showed positive correlation with the disease (Fig. 5e), while *IDNK* showed negative correlation with the disease (Fig. 5h). We observed the gene

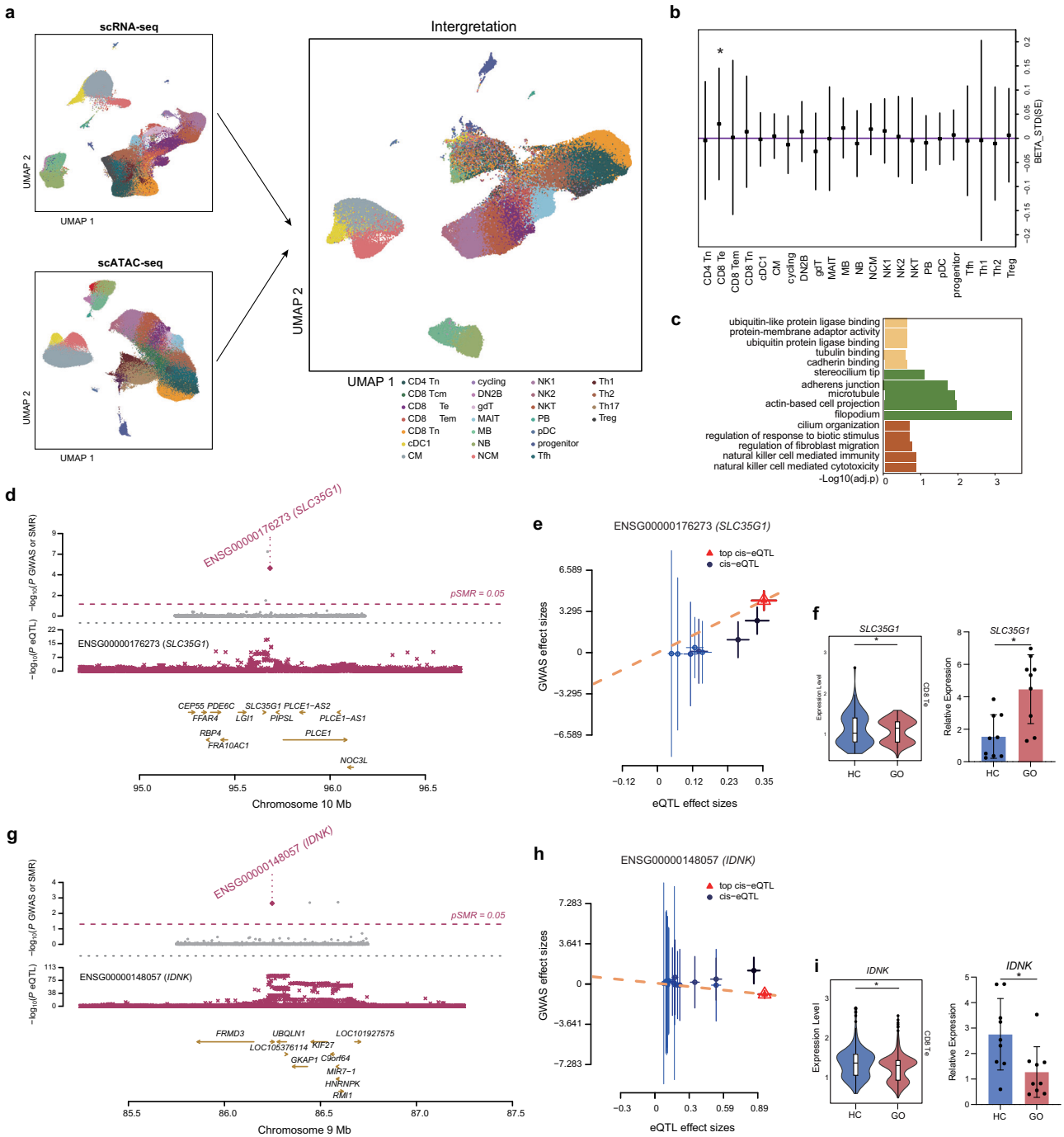


Fig. 5 | Integrative multiomic analysis and cell type association with genetic risk of Graves' disease. a Schematic for multiomic integration strategy for processing the scRNA-seq and scATAC-seq dataset. Following integration and label transfer. Dots represent individual cells, and colors indicate immune cell types. **b** Plots of the BETA_STD for the association between the GD-associated variants and the marker genes of each cell types. This approach describes the enrichment of GD risk variants, taken from the GD GWAS, in genes with cell-type-specific patterns in order to identify disease-risk specific cell types in the PBMC. Only significant association *P*-values were shown (**P* < 0.05). **c** Gene Ontology terms associated respectively with the CD8 Te marker genes responsible for the GD variant enrichment. **d** Locus plot for *SLC35G1*, showing its location on chromosome 10 and negative log *p*-values from SMR analysis. SMR p-value for *SLC35G1* is 5.97×10^-6. Y-axis represents negative log *p*-values; x-axis represents BP location. **e** SMR Effect Plot for *SLC35G1*. X-axis represents cis-eQTL effect sizes while the y-axis represents GWAS effect sizes. **f** The violin plot shows the gene expression levels of *SLC35G1* in CD8 Te from HC and GO groups of scRNA-seq (Left). The bar chart displays the gene expression levels of *SLC35G1* in CD8 Te from HC and GO donors by qRT-PCR (Right). Data are represented as the median [IQR]. ***P* < 0.001 by Mann-Whitney U test. **g** Locus plot for *IDNK*, showing its location on chromosome 9 and negative log *p*-values from SMR analysis. SMR p-value for *IDNK* is 2.23 × 10^-3. Y-axis represents negative log *p*-values; x-axis represents BP location. **h** SMR Effect Plot for *IDNK*. X-axis represents cis-eQTL effect sizes, while y-axis represents GWAS effect sizes. **i** The violin plot shows the gene expression levels of *IDNK* in CD8 Te from HC and GO groups of scRNA-seq (Left). The bar chart displays the gene expression levels of *IDNK* in CD8 Te from HC and GO donors by qRT-PCR (Right). Data are represented as the median [IQR]. ***P* < 0.001 by Mann-Whitney U test.

expression levels and chromatin open peaks of these two genes at the single-cell resolution level in both disease and healthy backgrounds (Fig. S9b, S9c). In detail, we quantitatively compared the gene expression levels in the previously identified disease-risk specific CD8 Te cells and verified with qRT-PCR, showing that the expression level of *SLC35G1* was higher compared to HC (Fig. 5f), while the expression level of *IDNK* was lower (Fig. 5i) (Supplementary Data 4), consistent with the SMR analysis results.

Importantly, the levels of the two genes in CD8 Te cells were higher than in most T cell subpopulations (Fig. S9d, S9e), and this was the only T cell subpopulation in the GO group where the expression level of *SLC35G1* signifcantly exceeded that of HC (Fig. 5f). Furthermore, analyses of additional supplementary datasets revealed that the gene expression levels of CD8 Te cells in peripheral blood from GD patients exhibited changes similar to those observed in GO patients, while expression levels in orbital tissues were consistent with those in the healthy group (Fig. S9f, S9g). Similarly, we found that the expression level of *SLC35G1* in CD8 Te cells was significantly higher in male GO patients compared to female GO patients, while the expression of *IDNK* was lower (Fig. S9h). These findings provided further evidence for the identity of CD8 Te cells as a disease-risk subpopulation. In summary, by combining GWAS and eQTL datasets, we identified a subset of CD8 Te cells associated with disease-related genetic variants, implying anotherpathogenic mechanisms of GO.

Changes of cell communication in the peripheral blood of GO and potential pathogenic pathways related to CD8 Te in localized lesions

Intercellular communication among immune cells plays a crucial role in the development of inflammation and autoimmune diseases through mechanisms such as immune responses and antigen presentation[56]. We further identified intercellular communication changes associated with GO through communication analysis. First, we observed an increase in both the total amount and strength of cell interactions in GO patients compared to HC, with most cell subpopulations exhibiting elevated interaction intensity (Fig. S10a, S10b). We identified four GO-specific pro-inflammatory pathways—EPHB, GRN, PARs, and CD48, which were predominantly enriched in cDC, pDC, NCM, NKT, and NK1 cells (Fig. S10c, S10d, S10e). The analysis of co-receptors in CD8 Te cells revealed an upregulation of the killer cell lectin-like receptor family and chemokine family in GO, which is consistent with the observed changes in gene set scoring (Fig. S10f). To further explore the potential role of CD8 Te in local lesions, we annotated all cell subpopulations in the tissue data, including the classical fibrotic subpopulations associated with GO pathogenesis. The results indicated that CD8 Te cells upregulated several pathways related to cellular fibrosis, such as AREG-EGFR, ITGB2-ICAM1, and TGFB1-TGFBR[57–59] (Fig. S10g).

Dynamic differentiation features of CD8 Te disease risk phenotype

T cell differentiation plays a crucial role in the onset and progression of autoimmune diseases, where abnormal T cell activation and reactivation may lead to the immune system attacking self-antigens, thus contributing to and promoting the development of autoimmune diseases[60–64]. Based on the changes in the expression levels of disease-risk-associated genes observed in CD8 Te cells in the context of GO, we aim to further elucidate the dynamic changes of the CD8 Te disease-specific phenotype in the potential differentiation pathways.

We defined two pseudotime trajectories based on integrated data from scRNAseq and scATACseq to examine the differentiation characteristics of activation (CD8 Tn-CD8 Te) (Fig. 6a) and memory reactivation (CD8 Tcm-CD8 Tem-CD8 Te) (Fig. 6b). The accessibility of *IDNK*-related TF motifs gradually increased during the activation trajectory (Figs. 5i, 6c–e), but decreased during the reactivation trajectory (Figs. 6d–f). Meanwhile, compared to HC, the accessibility of *IDNK*-related TF motifs in the GO disease background showed a greater decrease in the terminal stage (Fig. 6d). Quantitative comparison of gene expression levels showed significant decreases of *IDNK* expression in CD8 Tn and CD8 Tcm at the initial stage, as well as a significant decrease in CD8 Te at the terminal stage in GO (Figs. 5i, 6e, f). No differences were observed in the accessibility of TF motifs of *SLC35G1* and the expression levels at different stages in different disease backgrounds (Fig. S11a, S11b, S11c), implying that the changes in *IDNK* expression levels were more important in the dynamic differentiation characteristics of CD8 Te disease risk phenotypes.

Discussion

GO, which often occurs alongside the autoimmune disorder GD, represents a localized intensification of a multifaceted systemic condition. The specific immune cell types and pathways involved in its pathogenesis remain unclear, hindering the development of targeted therapies for this disorder[1,6,14]. To shed light on the autoimmune landscape of this complex disease at a single-cell level, we employed scATAC-seq and scRNA-seq to precisely map chromatin accessibility and gene transcription. We also integrated single-cell immune repertoire sequencing to identify GO-specific TCR and BCR clonal expansions. Furthermore, we linked GD risk variants with specific PBMC types in both GO patients and controls to pinpoint disease risk-associated subgroups. Our peripheral blood single-cell atlas and multi-omics integrated analysis approach unveiled several key findings: (1) the single-cell epigenomic and transcriptomic profiles of peripheral blood in GO patients; (2) the extensive activation of the autoimmune spectrum in the peripheral blood of GO patients; (3) disease risk-associated cell subgroup, CD8 Te cells, infiltrated local lesions and exerted potential pathogenic effects on fibrosis; (4) GO-related genes expression changes of *SLC35G1* and *IDNK* in CD8 Te cells. Prior single-cell studies of GO have been restricted to scRNA-seq, without the inclusion of scATAC-seq and single-cell TCR/BCR sequencing[8,17,18].

CD4 T cells are regarded as one of the primary pathogenic cell types in the pathogenesis of GO[21–23]. Consistent with previous studies, our dataset reveals an upregulation of the immune stress state in the CD4 T cell population, indicating that they are in an activated state. Interestingly, the TCR repertoire diversity in GO patients was significantly higher than that in healthy controls, which contrasted with previous research conclusions regarding most autoimmune diseases, such as systemic lupus erythematosus and rheumatoid arthritis[65,66]. However, some studies have reported that patients with multiple sclerosis (MS) exhibit significantly greater TCR diversity than HC, possibly because this disease is not caused by a single specific antigen[67]. Therefore, we speculated that the increased diversity observed in the GO group may be due to the involvement of multiple autoantigens in the pathogenesis of GO, including the previously reported thyroid-stimulating hormone receptor (TSHR) and insulin-like growth factor 1 receptor (IGF-1R), which may represent a unique characteristic of the immune repertoire in GO. Of course, the increase in diversity in the GO group suggested a decrease in clonality, which may imply that certain TCR clones in the healthy controls have experienced exhaustion or suppression in the GO group[68]. However, this cannot be analyzed and validated using single-cell immune repertoire sequencing and requires further experimental exploration.

B cells, which produce TSHR autoantibodies and interact with CD4 + T cells, contribute to GO progression[39,69]. However, the pathogenic role has not been well recognized in GO-related studies. Our data shows an obvious decrease in PB, but their functions, including antigen presentation and immune response, are upregulated in the disease context. Additionally, BCRs of PB show clonal expansions in the disease background, suggesting that enhanced humoral immunity may potentially impact GO development. NK cells have been described as important components of the orbital immune microenvironment[13,14], and our dataset identifies a subgroup of *CD160*^{high} NK cells at a more refined resolution. In the GO background, these cells are expanded in number, and their immune activation function is upregulated, potentially contributing to the inflammatory and cytotoxic millieu of the disease. Our analysis associated the *NF-κB* transcription factor family with these cells, which has been implicated in the inflammatory environment of GO progression[45–47]. Additionally, *CD169*+ classical monocytes have been reported to expand in GO patients and participate in

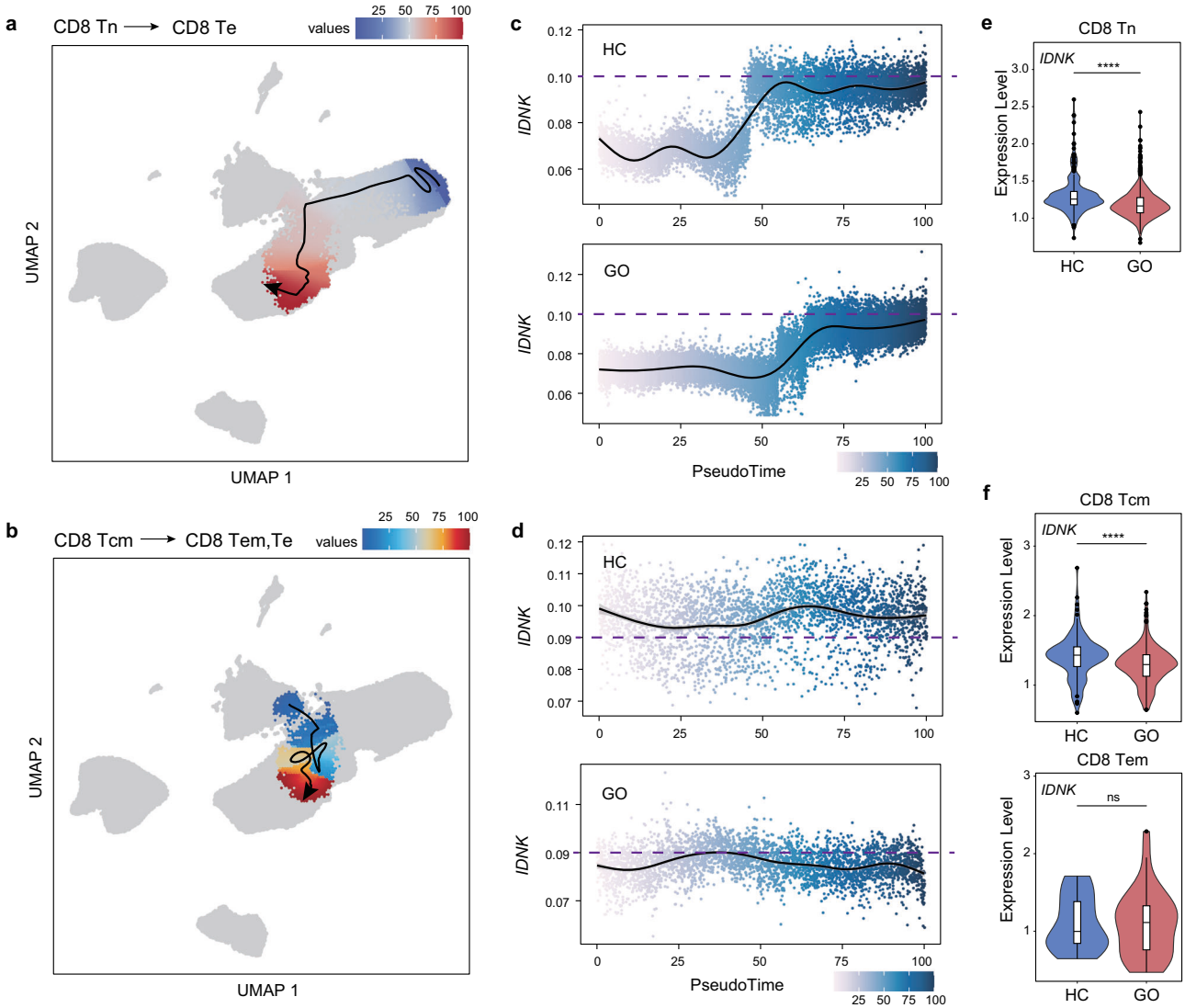


Fig. 6 | Dynamic differentiation features of CD8 Te in initial activation and reactivation. **a** The smoothed arrow represents a visualization of the differentiation trajectory from CD8 Tn to CD8 Te in the UMAP embedding. The color indicates the cellular fate across the inferred trajectory. **b** The smoothed arrow represents a visualization of the differentiation trajectory from CD8 Tcm to CD8 Te in the UMAP embedding. The color indicates the cellular fate across the inferred trajectory. **c** Dot plot of the integrated gene expression of *IDNK* across pseudotime trajectory from CD8 Tn to CD8 Te in HC and GO groups. **d** Dot plot of the integrated gene expression of *IDNK* across pseudotime trajectory from CD8 Tcm to CD8 Te in HC and GO groups. **e** Violin plots of the gene expression levels of *IDNK* in CD8 Tn from HC and GO groups. Data are represented as the median [IQR]. ****$P < 0.00001$ by Mann-Whitney U test. **f** Violin plots of the gene expression levels of *IDNK* in CD8 Tcm (up) and CD8 Tem (down) from HC and GO groups. Data are represented as the median [IQR]. ****$P < 0.00001$ by Mann-Whitney U test.

orbital inflammation, adipogenesis, and fibrosis[17]. Our comprehensive dataset identifies extensive transcriptional changes and epigenomic remodeling in myeloid cells. It is noteworthy that almost all upregulated DEGs in immune cell subpopulations were enriched in ribosomal-related pathways, and the upregulated DARs were enriched in translation-related pathways. This suggested an overall enhancement of immune cell metabolism, leading to a state of excessive activation of the immune system. Overall, our integrated multi-omics dataset has uncovered a wealth of cis-regulatory elements and gene resources, enriching our understanding of the cellular and molecular mechanisms in the immune landscape of GO.

Unlike the extensive research on CD4 T cells in the pathogenesis of autoimmune diseases including GO, the role of CD8 + T cell subsets has received less attention[38]. We found that CD8 Te cells exhibited the highest enrichment of GO-specific clonal expansions and were the only cell subset enriched with GD risk variant genes. Importantly, our data indicate that CD8 Te cells tend to migrate and infiltrate local lesions, potentially participating in pathways that promote fibroblast fibrosis in GO. Current understanding of GO pathogenesis suggests that T cells interact with CD40 on orbital fibroblasts, triggering a cascade that enhances T cell infiltration mediated by sphingosine-1-phosphate (S1P)[70]. This interaction with CD40 stimulates orbital fibroblasts to release pro-inflammatory cytokines and prostaglandin E2. Therefore, we hypothesize that CD8 Te cells may play a crucial role in the development of orbital lesions through their interactions with fibroblasts.

Our analysis also highlighted two disease-related genes identified via SMR analysis that exhibit expression patterns in CD8 Te cells aligned with disease progression. Specifically, *SLC35G1* shows a positive correlation with the disease phenotype, with its expression markedly elevated in CD8 Te cells in GO. Conversely, *IDNK* is negatively correlated with the disease phenotype, and its expression is reduced in CD8 Te cells in GO. Additional data show that CD8 Te cells in GD exhibited expression changes similar to those in GO, but only in PBMCs. This suggests that the orbital lesion phenotype

may be related to the local infiltration of CD8 Te cells and abnormal gene expression changes. These findings could serve as biomarkers for screening the risk of GO in GD patients and may represent potential preventive targets. However, this required further experimental validation and exploration.Our study provided a more comprehensive description of the immune landscape in GO and identified a previously unreported GO disease risk-associated cell subset, CD8 Te, which may serve as another target for prevention and treatment. However, there were limitations in our research. The correlation between immune repertoire characteristics and the disease, as well as the sources and regulatory mechanisms of abnormal cellular molecular changes, require further investigation. Future work could involve using animal models with TCR/BCR gene interference to explore the relationship between specific clonal expansions and disease phenotypes, conducting lineage tracing analyses to identify the origins of abnormal cells, and employing gene knockout models to elucidate the gene expression networks of *SLC35G1* and *IDNK* to clarify specific pathogenic mechanisms. Although the precise molecular mechanisms underlying GO pathogenesis are not yet fully understood, our findings provide fresh insight into the regulatory dynamics of the GO immune environment. This is particularly evident in the comprehensive single-cell multi-omics data related to epigenetics and transcriptional patterns that we have presented. Our dataset constitutes a valuable resource for advancing our understanding of human autoimmunity regulation.

Methods

Human subjects

This study was approved by the Ethics Committee of Zhongshan Ophthalmic Center (Guangzhou, China, Approval No. 2019KYPJ114), and all participants provided written informed consent. All participants were recruited from Zhongshan Ophthalmic Center. The study comprised 31 PBMC samples. Among these, 10 samples were from healthy volunteers, and 21 from GO patients. Samples with comorbidities such as cancer, immunodeficiency, diabetes, and other autoimmune diseases were excluded. All ethical regulations relevant to human research participants were followed.

PBMC isolation

Human peripheral venous blood samples were collected and processed using Ficoll-Hypaque density solution and heparin. Following centrifugation for 30 minutes, the PBMC layer solution was isolated. PBMCs were then diluted with PBS (Sigma, #P2272) and stained with Trypan Blue (Biosharp, #BL627A) to evaluate both cell viability and quantity. Cell viability was maintained above 90% while kept on ice.

scATAC-seq library processing

We followed the manufacturer's guidelines for nuclear extraction, suspending nuclei in an appropriate volume of chilled Diluted Nuclei Buffer (10x Genomics; #PN-2000153). Subsequently, we promptly utilized these nuclei to generate 10× scATAC-seq libraries. These libraries underwent unique barcoding and RT-qPCR quantification. Each sample library was loaded onto an Illumina Novaseq 6000 and sequenced in paired-end mode. Our sequencing goal for the libraries was to achieve 90% saturation or an average of 30,000 unique reads per cell. We strictly adhered to the protocols for sample handling, library preparation, and instrument setup on the 10× Chromium platform. The raw sequencing data were converted to fastq format using Cellranger atac mkfastq (v.1.0.0). Subsequently, the scATAC-seq data reads were aligned to the GRCh38 (hg38) reference genome and quantified using the Cellranger count function.

scATAC-seq quality control

Arrow files were created using ArchR (v0.9.5)[71] by analyzing accessible read fragments from each sample, with default augmentations applied unless specified otherwise. To ensure robust signal quality, we filtered out cells with fewer than 2500 unique fragments and TSS enrichment below 9. The cells in the scATAC-seq dataset had an average of 8,854 unique nuclear fragments and an average TSS enrichment of 17.2 (Fig. S1a, S1b). Doublets were identified and eliminated using ArchR[71]. Furthermore, cells mapping to regions on the ENCODE project's blacklist were excluded.

scATAC-seq dimensionality reduction and clustering

We implemented a layered strategy for dimensionality reduction, employing latent semantic indexing (LSI) and singular value decomposition (SVD). Subsequently, we applied Harmony[19] batch correction individually to each sample. Next, single-cell accessibility profiles underwent clustering using Seurat shared nearest neighbor (SNN) graph clustering algorithm with the 'FindClusters' function, applying a resolution of 1.0 for the T cell subpopulations and a resolution of 0.6 for all other cell subpopulations. Visualization of the data was accomplished using uniform manifold approximation and projection (UMAP)[72] in a two-dimensional space.

Gene activity scores analysis

Gene activity scores were calculated based on accessibility within the gene body, at the promoter, and at distal regulatory elements using ArchR[71] with default parameters. Additionally, we applied the imputed weight method MAGIC[73] to the resulting gene activity scores to reduce noise caused by the sparsity of scATAC-seq data.

Pseudobulk replicate generation and peak calling

To perform differential comparisons between clusters, we generated non-overlapping pseudobulk replicates from cell groups using the 'addGroupCoverages' function with different arguments. These replicates were then utilized to construct the peak matrix (via 'addReproduciblePeakSet'). Furthermore, we employed MACS2[74] for peak calling. And the pseudobulk peak set was subsequently used for downstream analysis.

Motif enrichment and deviation analysis

We conducted motif enrichment and motif deviation analyses on the pseudobulk peak set. For peak annotation, we utilized motifs from the Catalog of Inferred Sequence Binding Preferences (CIS-BP)[75], and JASPAR2020 motif[76]. Furthermore, chromVAR deviation scores for these motifs were computed by ArchR[71].

scATAC–seq differential analysis

The pseudobulked peak set was used for differential analysis between different cell types and different disease states using the 'getMarkerFeatures' function. We randomly sampled an equal number of cells from each group for analysis. Peak intensity was defined as the log2 of normalized read counts, and we used the Wilcoxon test and Benjamini-Hochberg correction to calculate p values and FDR between sample pairs. Differentially accessible distal peaks were defined as those with FDR ≤ 0.1 and log2-fold change ≥ 0.5.

TF Foot-print analysis

To conduct motif footprint analysis, we examined Tn5 insertions within genome-wide motifs and normalized them by subtracting the Tn5 bias from the footprinting signal. For motif position determination, we employed CIS-BP motifs. Normalization of footprints entailed using mean values ± 200–250 from the motif center. Subsequently, we plotted the mean and standard deviation for each footprint pseudo-replicate. Comparison of TF footprints between groups was performed using either the Wilcoxon rank sum test or one-way ANOVA followed by Tukey's multiple comparison test.

Peak to gene linkage analysis

The "addPeak2GeneLinks" function was utilized to compute peak accessibility and gene expression using the parameters of "corCutOff = 0.2" and "reducedDims" as the dimensionality reduction results after batch correction. The resulting "Granges" object was used for visualization.

scRNA-seq library and scRepertoire library processing

The Single-cell RNA libraries were prepared using the Chromium Single Cell 5' Kit v2 (10x Genomics; #PN-1000263) and Chromium Single Cell

V(D)J Reagent kits (10x Genomics, #PN-1000252(TCR), #PN-1000253(BCR)), following the manufacturer's instructions. And scRNA-seq libraries were sequenced on Illumina NovaSeq6000 with PE150 reads, and the quality was assessed using FastQC. Raw data was processed and aligned the GRCh38 reference by the cellranger software with default parameters (v3.1.0) for each sample.

scRNA-seq quality control

We utilized Seurat V4[77] to filter out low-quality cells with mitochondrial gene expression exceeding 20% or cells with detected genes fewer than 200 or exceeding 4,000. The cells in the scRNA-seq dataset had an average of 3,608 UMIs (Fig. S1c). Additionally, we filtered out cell clusters identified as expressing genes related to HBB, HBA1, PPBP, and PF4, representing red blood cells and platelets.

scRNA-seq dimensionality reduction and clustering

Following the standard seurat[77] pipeline, we applied the 'FindVariableFeatures' function to normalize and scale the top 2,000 variable genes, and then conducted principal component analysis (PCA = 30). Subsequently, batch effects were corrected using Harmony[19]. Cell clustering was performed using KNN (The resolution for the T cell subpopulations was set to 0.8, while a resolution of 0.6 was used for all other cell subpopulations.), followed by UMAP[72] analysis. All samples underwent processing with the 'DoubletFinder' function to remove potential doublets, and cell clusters were annotated based on the expression of feature genes.

scRNA–seq differential analysis

For differential gene expression analysis, we utilized Seurat's[77] "FindMarkers" function with the MAST test81. We randomly sampled an equal number of cells from each group for analysis. Genes with an adjusted $|\log FC| > 0.25$ and adjusted P-value < 0.05 were selected as upregulated and downregulated DEGs.

Using Milo to compare cell type distribution analysis

We performed differential cell type distribution analysis on single-cell RNA sequencing (scRNA-seq) data (Milo package (https://github.com/MarioniLab/miloR [24]) by following these steps: Construct a k-nearest-neighbor (k-NN) graph from the scRNA-seq data to group similar cells into "neighborhoods"; Count the number of cells within each neighborhood; Define experimental groups to compare differences in cell-type distribution; Perform differential abundance testing to identify neighborhoods with significant changes in cell representation between groups; Visualize the results with an alpha value of 4, which sets the significance threshold for highlighting differentially abundant neighborhoods.

Integration of scRNA-seq and scATAC-seq data

We followed the procedure outlined on the ArchR[71] and Seurat[77] websites to integrate single-cell RNA sequencing (scRNA-seq) and single-cell chromatin accessibility sequencing (scATAC-seq) datasets. Firstly, the complete dataset was partitioned into smaller cell groups and individual alignments were performed using ArchR[71] to enable parallelization of this process. Subsequently, the FindTransferAnchors() function from the Seurat[77] package was utilized to align the log-normalized and standardized scATAC-seq gene score matrix with the scRNA-seq gene expression matrix using default parameters. Ultimately, each cell in the scATAC-seq space was assigned a gene expression feature to obtain cell matching across the entire dataset.

Analysis workflow of public data

We downloaded the published datasets from Wang et al[17].(GSA: HRA005673) and Li et al[18].(GSA: HRA000870) and performed data quality control and processing using the same workflow as outlined earlier. Immune cell subpopulations from orbital tissue were extracted and integrated with our immune cell data for further analysis, maintaining consistent analysis parameters throughout.

TCR and BCR V(D)J immune complex sequencing and analysis

Full-length TCR/BCR V(D)J segments were obtained by PCR amplification from 5' libraries using a Chromium Single-cell V(D)J Enrichment kit. Demultiplexing, gene quantification, and BCR/TCR clonotype assignment were performed using the Cell Ranger (v.6.0.0) vdj pipeline with GRCh38 as the reference genome. The V(D)J immune repertoire was analyzed using scRepertoire[78] (https://github.com/ncborcherding/scRepertoire). Extract CT genes from the GO and HC groups, and identify GO-specific clones by calculating the difference between the two groups.

Flow cytometry and Fluorescence-activated Cell Sorting

After treatment with Lymphocyte Separation Solution (Solarbio, #P8900, #P8620), PBMCs (peripheral blood mononuclear cells) were collected from the interface between the PBS and Ficoll layers into new tubes. Live/dead cell counting was performed using AO/PI dye (Countstar, #RE010212) with an Automated Fluorescence Cell Analyzer (Countstar Rigel S2, China), resulting in a live cell proportion of approximately 95%. Subsequently, cells were incubated with the appropriate surface antibodies in Cell Staining Buffer (BioLegend, #420201). For intracellular antibody staining, cells were treated with the eBioscience™ Intracellular Fixation & Permeabilization Buffer Set (Thermo Fisher, 88-8824-00) according to the manufacturer's instructions, followed by incubation with the corresponding intracellular antibodies. Fc-block (BioLegend, #422302) was used during the incubation process. Antibodies used in the flow cytometry experiments included CD8 (BioLegend, #980904, 2.5 μl per million cells in 100 μl staining volume), KLRC1 (BioLegend, #375139, 5 μL per million cells in 100 μL staining volume), and GNLY (BioLegend, #348003, 5 μl per million cells in 100 μl staining volume). Isotype control antibodies (BioLegend, #400411, #400157, #981804) were also used. Flow cytometry data were collected using the BD FACSCanto II flow cytometer (BD Biosciences), with single-color compensation for each fluorescence channel. Data analysis was performed using FlowJo (v.10), where obvious dead cells and debris were excluded through FSC/SSC gating. Single cells were selected using FSC-H and FSC-W gating, excluding aggregates and debris. Analysis was conducted within the lymphocyte gate using FlowJo (v.10). Fluorescence activated Cell Sorting was performed to collect target cells using the BD FACSAria™ Fusion Cell Sorter (BD Biosciences).

qRT-PCR

High-quality RNA was extracted using the RNeasy Mini Kit (Qiagen, #74104) according to the manufacturer's instructions. Equal amounts of RNA were extracted as required and added to the corresponding primers and HiScript II One-Step qRT-PCR SYBR Green Kit reaction solution (Vazyme, #Q221) for qRT-PCR reactions, and all expression values were normalized to Actin levels. The Primer sequence: *SLC35G1* (F: TGTCCTGGACTT GGCTTGTTT, R: GCATTTGGAACACACATCGAAA); *IDNK* (F: GG AAAAGGCATACCGCTCAAT, R: AACAGGCTAGAACCACACGC).

Immunofluorescence

In orbital decompression or cosmetic surgery, human orbital tissue was obtained and fixed overnight in 4% paraformaldehyde (Biosharp, #BL539A) at 4°C. The tissue was then dehydrated in 30% sucrose and embedded in Tissue-Tek O.C.T. Compound (Sakura Finetek USA), followed by freezing at -80°C. Using a Leica CM1860 cryostat (Leica Biosciences), frozen tissues were sectioned into 10 μm-thick slices. Antigen retrieval was performed using sodium citrate buffer (Beyotime, #P0081). After blocking and permeabilization, primary and secondary antibody incubations were carried out. The antibodies used for immunostaining included Anti-CD8 antibody (ThermoFisher, #MA1-81692, 1:200), Anti-GNLY (CST, #24253, 1:800) antibody, and secondary antibody (ThermoFisher, #A48269 5 μg/mL& #A31572 4 μg/mL). Immunofluorescence images were captured using a ZEISS LSM 980 microscope, and analysis was conducted using ImageJ software.

Cell type association with genetic risk of GD

We performed an analysis using Multi-marker Analysis of GenoMic Annotation (MAGMA) v1.08[79] to investigate the association between cell

type-specific expressed genes and GD. The aim was to identify cell types associated with the risk of GD. MAGMA[79] is a gene-set enrichment analysis method that tests the joint association of all risk SNPs with the phenotype, while considering the linkage disequilibrium (LD) structure between SNPs. In our study, SNPs and their *P*-values were obtained from the summary statistics of the BioBank Japan (BBJ) whole-genome association study (GWAS) (bbj-a-123)[80]. The publicly available East Asian subset of the 1000 Genomes Phase 3 was used as a reference panel to estimate LD between SNPs. The MAGMA[79] analysis consisted of three steps. First was the annotation step, where SNPs were mapped to genes using the NCBI GRCh38 build. Second, the gene-based association analysis step calculated gene-level *P*-values based on SNP GWAS *P*-values. The third step was the gene-set analysis, implemented as a linear regression model on a gene-set data matrix. Here, the gene sets consisted of cell type-specific expressed genes [filtered for false discovery rate (FDR)-corrected *P*-values < 0.05, percentage of cells in the cluster where the expression was detected > 0.5, and log fold change (logFC) > 0.25]. MAGMA[79] gene-set analysis provided association results for each gene set. The association of genes with cell types was quantified by the variable regression coefficient BETA, while BETA_STD represented the semi-standardized regression coefficient corresponding to the predicted change in Z-value given a change of one standard deviation in the predictor gene set. A *p*-value < 0.05 from the gene-set analysis was considered statistically significant, with high positive BETA_STD indicating the most positively correlated gene sets.

SMR analysis and HEIDI methods

We implemented the SMR & HEIDI methods[55] to test for pleiotropic association between the expression level of a gene and a complex trait of interest using summary-level data from GWAS and eQTL studies[55]. Firstly, we integrated summary statistics of blood cis-eQTL from the eQTLGen Consortium study[81] with GWAS results from BBJ. As the genetic variants provided serve as the primary link between phenotype and gene expression data, the program conducts a quality-control check on allele frequencies to ensure consistency between SNP information used in both eQTL and GWAS studies. Next, to perform Mendelian Randomization analysis, a significant SNP-eQTL relationship is required. Therefore, only cis-eQTLs (defined within a standard 1 Mb radius from the gene probe) with a top PeQTL ≤ 5 × 10−8 are included for the SMR analysis. Additionally, SNPs with eQTL minor, effect, and/or GWAS allele frequencies < 0.01 were removed. Subsequently, only SNPs with eQTL *p*-values surviving a Bonferroni-corrected threshold, as determined by the number of SMR calculations run per command, were fully analyzed. Finally, to address linkage disequilibrium issues, SNPs with an r2 > 0.90 or r2 < 0.05 with the top SNP for that cis-eQTL were excluded, with one result of every pair of SNPs satisfying these LD requirements also being excluded. Our analysis identified two genes through SMR analysis and HEIDI test.

Cell-cell interaction analysis

Communication analysis between immune cells and fibroblast subtypes was performed using the CellChat R package (v.1.6.1)[82].

Statistics and reproducibility

The details regarding the experimental design and statistical methods used in the various data analyses of this study are provided in the respective sections of the results and methods. For the sequencing part, 31 independent biological replicates were used (10 from healthy controls and 21 from the GO group). Flow cytometry was performed with 9 independent biological replicates per group. Immunofluorescence was conducted with 5 independent biological replicates per group. For qRT-PCR, each group included 9 independent biological replicates, with each target in each sample measured in triplicate technical repeats. Detailed information regarding statistical tests can be found in the figure legends and methods, including biological replicates and assay methods. All *p*-values less than or equal to 0.05 were considered statistically significant. In the figures, *p*-values are denoted as $^{*}p \le 0.05$, $^{**}p \le 0.01$, $^{***}p \le 0.001$, $^{****}p \le 0.0001$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The scRNA-seq, scATAC-seq, scTCR-seq, scBCR-seq data analyzed in the article are available from the corresponding author upon request under the Accession No. HRA007561 in https://ngdc.cncb.ac.cn/gsa-human/. Source data underlying graphs can be obtained in Supplementary data(1-4). All other data are available from the corresponding author upon reasonable request.

Code availability

The code of the analyses in the manuscript has been uploaded to GitHub and can be accessed at https://github.com/liuzh295/Graves-ophthalmopathy.git.

Received: 26 May 2024; Accepted: 22 April 2025;
Published online: 12 May 2025

References

1. Bahn, R. S. Graves' ophthalmopathy. *N. Engl. J. Med.* **362**, 726–738 (2010).
2. Li, S. & Liu, Y.-J. Application of quantitative orbital analysis to assess the activity of Graves' ophthalmopathy. *Am. J. Nucl. Med. Mol. Imaging* **13**, 259–268 (2023).
3. Wiersinga, W. et al. Predictive score for the development or progression of Graves' orbitopathy in patients with newly diagnosed Graves' hyperthyroidism. *Eur. J. Endocrinol.* **178**, 635–643 (2018).
4. Moi, L., Hamedani, M. & Ribi, C. Long-term outcomes in corticosteroid-refractory Graves' orbitopathy treated with tocilizumab. *Clin. Endocrinol.* **97**, 363–370 (2022).
5. He, M. et al. The potential markers involved in newly diagnosed graves' disease and the development of active graves' orbitopathy. *Cytokine* **127**, 154998 (2020).
6. Davies, T. F. et al. Graves' disease. *Nat. Rev. Dis. Prim.* **6**, 52 (2020).
7. Hwang, C. J., Nichols, E. E., Chon, B. H. & Perry, J. D. Bilateral dysthyroid compressive optic neuropathy responsive to teprotumumab. *Eur. J. Ophthalmol.* **32**, NP46–NP49 (2022).
8. Wang, Y. et al. A novel CD4+ CTL subtype characterized by chemotaxis and inflammation is involved in the pathogenesis of Graves' orbitopathy. *Cell. Mol. Immunol.* **18**, 735–745 (2021).
9. Han, R. & Smith, T. J. T helper type 1 and type 2 cytokines exert divergent influence on the induction of prostaglandin E2 and hyaluronan synthesis by interleukin-1beta in orbital fibroblasts: implications for the pathogenesis of thyroid-associated ophthalmopathy. *Endocrinology* **147**, 13–19 (2006).
10. Antonelli, A. et al. Extra-ocular muscle cells from patients with Graves' ophthalmopathy secrete α (CXCL10) and β (CCL2) chemokines under the influence of cytokines that are modulated by PPARγ. *Autoimmun. Rev.* **13**, 1160–1166 (2014).
11. Fang, S. et al. Regulation of orbital fibrosis and adipogenesis by pathogenic Th17 cells in graves Orbitopathy. *J. Clin. Endocrinol. Metab.* **102**, 4273–4283 (2017).
12. Jiang, M. et al. Looking beyond Th17 cells: a role for Th17.1 cells in thyroid-associated ophthalmopathy?. *Endocrinology* **164**, bqad004 (2023).
13. Fang, S. et al. Insights into local orbital immunity: evidence for the involvement of the Th17 cell pathway in thyroid-associated Ophthalmopathy. *J. Clin. Endocrinol. Metab.* **104**, 1697–1711 (2019).
14. Gao, Y. & Li, W. Mechanisms of immune-related differentially expressed genes in thyroid-associated ophthalmopathy based on the GEO database. *Ann. Transl. Med.* **10**, 926 (2022).
15. Eckstein, A. K. et al. Thyrotropin receptor autoantibodies are independent risk factors for Graves' ophthalmopathy and help to

predict severity and outcome of the disease. *J. Clin. Endocrinol. Metab.* **91**, 3464–3470 (2006).
16. Kuriyan, A. E., Woeller, C. F., O'Loughlin, C. W., Phipps, R. P. & Feldon, S. E. Orbital fibroblasts from thyroid eye disease patients differ in proliferative and adipogenic responses depending on disease subtype. *Invest. Ophthalmol. Vis. Sci.* **54**, 7370–7377 (2013).
17. Wang, D. et al. CD169+ classical monocyte as an important participant in Graves' ophthalmopathy through CXCL12-CXCR4 axis. *iScience* **27**, 109213 (2024).
18. Li, Z. et al. Single-cell RNA sequencing depicts the local cell landscape in thyroid-associated ophthalmopathy. *Cell Rep. Med.* **3**, 100699 (2022).
19. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
20. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
21. Smith, T. J. TSH-receptor-expressing fibrocytes and thyroid-associated ophthalmopathy. *Nat. Rev. Endocrinol.* **11**, 171–181 (2015).
22. Khong, J. J., McNab, A. A., Ebeling, P. R., Craig, J. E. & Selva, D. Pathogenesis of thyroid eye disease: review and update on molecular mechanisms. *Br. J. Ophthalmol.* **100**, 142–150 (2016).
23. Rotondo Dottore, G. et al. Role of the mononuclear cell infiltrate in Graves' orbitopathy (GO): results of a large cohort study. *J. Endocrinol. Invest.* **45**, 563–572 (2022).
24. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
25. Gensous, N. et al. T Follicular Helper Cells in Autoimmune Disorders. *Front. Immunol.* **9**, 1637 (2018).
26. Seki, N. et al. Th17/IL-17A axis is critical for pulmonary arterial hypertension (PAH) in systemic sclerosis (SSc): SSc patients with high levels of serum IL-17A exhibit reduced lung functions and increased prevalence of PAH. *Cytokine* **176**, 156534 (2024).
27. Lu, C. et al. Single-cell transcriptome analysis and protein profiling reveal broad immune system activation in IgG4-related disease. *JCI Insight* **8**, e167602 (2023).
28. Zhu, X. et al. Noc4L-mediated ribosome biogenesis controls activation of regulatory and conventional T cells. *Cell Rep.* **27**, 1205–1220.e4 (2019).
29. Emmott, E., Jovanovic, M. & Slavov, N. Ribosome stoichiometry: from form to function. *Trends Biochem. Sci.* **44**, 95–109 (2019).
30. Sohrabi, S. et al. STATs signaling pathways in dendritic cells: As potential therapeutic targets?. *Int. Rev. Immunol.* **43**, 138–159 (2024).
31. Scherlinger, M. et al. CaMK4 controls follicular helper T cell expansion and function during normal and autoimmune T-dependent B cell responses. *Nat. Commun.* **15**, 840 (2024).
32. Bollain-Y-Goytia, J. J. et al. Soluble fas and the -670 polymorphism of fas in lupus nephritis. *Int. J. Nephrol.* **2014**, 780406 (2014).
33. Liu, S. et al. ZNF384: A potential therapeutic target for Psoriasis and Alzheimer's disease through inflammation and metabolism. *Front. Immunol.* **13**, 892368 (2022).
34. Wang, W. et al. A comprehensive analysis of LMO2 pathogenic regulatory profile during T-lineage development and leukemic transformation. *Oncogene* **41**, 4079–4090 (2022).
35. Yu, H. et al. Mechanism study of ubiquitination in T cell development and autoimmune disease. *Front. Immunol.* **15**, 1359933 (2024).
36. Song, R., Jia, X., Zhao, J., Du, P. & Zhang, J. A. T cell receptor revision and immune repertoire changes in autoimmune diseases. *Int. Rev. Immunol.* **41**, 517–533 (2022).
37. Gravano, D. M. & Hoyer, K. K. Promotion and prevention of autoimmune disease by CD8+ T cells. *J. Autoimmun.* **45**, 68–79 (2013).
38. Petrelli, A. & van Wijk, F. CD8(+) T cells in human autoimmune arthritis: the unusual suspects. *Nat. Rev. Rheumatol.* **12**, 421–428 (2016).
39. Hansen, M., Cheever, A., Weber, K. S. & O'Neill, K. L. Characterizing the Interplay of Lymphocytes in Graves' Disease. *Int. J. Mol. Sci.* **24**, 6835 (2023).
40. Fang, J. & Ding, Z. SNAI1 is a prognostic biomarker and correlated with immune infiltrates in gastrointestinal cancers. *Aging* **12**, 17167–17208 (2020).
41. Wu, M.-N. et al. Genetic analysis of potential biomarkers and therapeutic targets in ferroptosis from psoriasis. *Front. Immunol.* **13**, 1104462 (2022).
42. Chang, L. et al. B cell receptor repertoire analysis in primary Sjogren's syndrome salivary glands identifies repertoire features associated with clinical activity. *Arthritis Res. Ther.* **26**, 62 (2024).
43. Sun, W. et al. B cell activation and autoantibody production in autoimmune diseases. *Best. Pract. Res. Clin. Rheumatol.* **38**, 101936 (2024).
44. Szabo-Fresnais, N., Blondeau, J. P. & Pomerance, M. Activation of the cAMP pathway synergistically increases IL-1-induced IL-6 gene expression in FRTL-5 thyroid cells: involvement of AP-1 transcription factors. *Mol. Cell. Endocrinol.* **284**, 28–37 (2008).
45. Cui, X. et al. Circulating Exosomes From Patients With Graves' Disease Induce an Inflammatory Immune Response. *Endocrinology* **162**, bqaa236 (2021).
46. Liu, Y. et al. TSAb Modulates Microglia M1 polarization via activation of the TSHR-mediated NF-κB signaling pathway. *Neuroendocrinology* **113**, 470–478 (2023).
47. Lu, Y. et al. Promotion of IL-17/NF-κB signaling in autoimmune thyroid diseases. *Exp. Ther. Med.* **25**, 51 (2023).
48. Eckstein, A. K. et al. Thyroid associated ophthalmopathy: evidence for CD4(+) gammadelta T cells; de novo differentiation of RFD7(+) macrophages, but not of RFD1(+) dendritic cells; and loss of gammadelta and alphabeta T cell receptor expression. *Br. J. Ophthalmol.* **88**, 803–808 (2004).
49. Fernando, R., Caldera, O. & Smith, T. J. Therapeutic IGF-I receptor inhibition alters fibrocyte immune phenotype in thyroid-associated ophthalmopathy. *Proc. Natl Acad. Sci. Usa.* **118**, e2114244118 (2021).
50. Fugger, L., Jensen, L. T. & Rossjohn, J. Challenges, progress, and prospects of developing therapies to treat autoimmune diseases. *Cell* **181**, 63–80 (2020).
51. Chi, X. et al. Innate and adaptive immune abnormalities underlying autoimmune diseases: the genetic connections. *Sci. China Life Sci.* **66**, 1482–1517 (2023).
52. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
53. Maurano, M. T. et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
54. Farh, K. K.-H. et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
55. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
56. Auguste, M., Balbi, T., Ciacci, C. & Canesi, L. Conservation of cell communication systems in invertebrate host-defence mechanisms: possible role in immunity and disease. *Biology* **9**, 234 (2020).
57. My, Z. et al. A critical role of AREG for bleomycin-induced skin fibrosis. *Cell Biosci.* **11**, 40 (2021).
58. Chen, L. et al. Single-cell RNA sequencing reveals reduced intercellular adhesion molecule crosstalk between activated hepatic stellate cells and neutrophils alleviating liver fibrosis in hepatitis B virus transgenic mice post menstrual blood-derived mesenchymal stem cell transplantation. *MedComm* **5**, e654 (2024).

59. Gu, J. et al. Topical application of nitrate ameliorates skin fibrosis by regulating ST2+CD4+ T cells in Systemic Sclerosis Mouse Model. *J. Invest. Dermatol*. S0022-202X(24)01870–0 https://doi.org/10.1016/j.jid.2024.06.1273 (2024).
60. Chang, L. et al. Single-cell clonal tracing of glandular and circulating T cells identifies a population of CD9+ CD8+ T cells in primary Sjogren's syndrome. *J. Leukoc. Biol.* **115**, 804–818 (2024).
61. Talaie, T. et al. Origin, phenotype and autoimmune potential of T cells in human immune system mice receiving neonatal human thymus tissue. *Front. Immunol.* **14**, 1159341 (2023).
62. Lin, F., Sun, X., Lei, J. & Xu, A.-E. Altered circulating memory T cells in vitiligo cases followed NB-UVB therapy. *Photodermatol. Photoimmunol. Photomed.* **38**, 76–82 (2022).
63. Li, Y., Wang, H., Lu, H. & Hua, S. Regulation of memory T cells by Interleukin-23. *Int. Arch. Allergy Immunol.* **169**, 157–162 (2016).
64. Ehlers, M. R. & Rigby, M. R. Targeting memory T cells in type 1 diabetes. *Curr. Diab. Rep.* **15**, 84 (2015).
65. Thapa, D. R. et al. Longitudinal analysis of peripheral blood T cell receptor diversity in patients with systemic lupus erythematosus by next-generation sequencing. *Arthritis Res. Ther.* **17**, 132 (2015).
66. Liu, X. et al. T cell receptor β repertoires as novel diagnostic markers for systemic lupus erythematosus and rheumatoid arthritis. *Ann. Rheum. Dis.* **78**, 1070–1078 (2019).
67. Hayashi, F. et al. A new clustering method identifies multiple sclerosis-specific T-cell receptors. *Ann. Clin. Transl. Neurol.* **8**, 163–176 (2021).
68. Schönland, S. O. et al. Premature telomeric loss in rheumatoid arthritis is genetically determined and involves both myeloid and lymphoid cell lineages. *Proc. Natl. Acad. Sci. Usa.* **100**, 13471–13476 (2003).
69. Lane, L. C., Cheetham, T. D., Razvi, S., Allinson, K. & Pearce, S. H. S. Expansion of the immature B lymphocyte compartment in Graves' disease. *Eur. J. Endocrinol.* **189**, 208–216 (2023).
70. Davies, T. F. et al. Graves' disease. *Nat. Rev. Dis. Prim.* **6**, 1–23 (2020).
71. Granja, J. M. et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
72. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol*. https://doi.org/10.1038/nbt.4314 (2018).
73. van Dijk, D. et al. Recovering gene interactions from single-cell data using data diffusion. *Cell* **174**, 716–729.e27 (2018).
74. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
75. Lambert, S. A. et al. Similarity regression predicts evolution of transcription factor sequence specificity. *Nat. Genet.* **51**, 981–989 (2019).
76. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. **48**, D87–D92 (2020).
77. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
78. Borcherding, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Research* **9**, 47 (2020).
79. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
80. Sakaue, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
81. Võsa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
82. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

Acknowledgements

This study was supported by the National Natural Science Foundation of China (81721003) and Science and Technology Program of Guangzhou, China (2024A03J0258). The funding body had no role in study design, analysis of data, writing of the manuscript and the decision to submit the manuscript for publication.

Author contributions

S.R.K. and Z.L. designed the study, analyzed the single-cell data and wrote the manuscript, which revised by H.J.Y., L.X.L., M.Z. and X.B. prepared the figures and performed the statistical analyses. Q.L.L. and J.S.L. prepared the experiment. X.C.L., H.J.Y. and X.C.L. guided and coordinated the progress of the entire project and revised the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08115-7.

Correspondence and requests for materials should be addressed to Huijing Ye, Lixia Luo or Xianchai Lin.

Peer review information *Communications Biology* thanks Feng Zeng and the other, anonymous, reviewers for their contribution to the peer review of this work. Primary Handling Editors: Dr Chien-Yu Chen and Dr Ophelia Bu. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025