



Title	Single-cell analysis reveals immune cell abnormalities underlying the clinical heterogeneity of patients with systemic sclerosis
Author(s)	Shimagami, Hiroshi; Nishimura, Kei; Matsushita, Hiroaki et al.
Citation	Nature Communications. 2025, 16(1), p. 4949
Version Type	VoR
URL	https://hdl.handle.net/11094/102920
rights	This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Single-cell analysis reveals immune cell abnormalities underlying the clinical heterogeneity of patients with systemic sclerosis

Received: 12 July 2024

Accepted: 9 May 2025

Published online: 17 June 2025

 Check for updates

Hiroshi Shimagami^{1,2,3}, Kei Nishimura^{3,4,5}, Hiroaki Matsushita^{3,4,5},
Shoichi Metsugi^{3,4,5}, Yasuhiro Kato^{1,2,3}, Takahiro Kawasaki^{1,2,3},
Kohei Tsujimoto^{1,2,3}, Ryuya Edahiro^{1,6,7}, Eri Itotagawa^{1,2,3}, Maiko Naito^{1,2,3},
Shoji Kawada^{1,2,3}, Daisuke Nakatsubo^{1,2,3}, Kazuki Matsukawa^{1,2,3},
Tomoko Namba-Hamano⁸, Kazunori Inoue⁸, Atsushi Takahashi⁸,
Masayuki Mizui⁸, Seiya Kato⁹, Hayato Hikita⁹, Shigeaki Nakazawa¹⁰,
Yoichi Kakuta¹⁰, Hachiro Konaka¹¹, Kensuke Mitsumoto¹¹, Nachi Ishikawa¹²,
Jun Fujimoto¹², Shinji Higa¹², Ryusuke Omiya^{4,5}, Yoshitaka Isaka⁸,
Tetsuo Takehara⁹, Norio Nonomura¹⁰, Yukinori Okada^{6,7,13,14,15,16},
Kunihiro Hattori^{4,5}, Masashi Narazaki^{1,2,3}, Atsushi Kumanogoh^{1,2,13,16,17,18,19} ✉ &
Masayuki Nishide^{1,2,3,19} ✉

The autoimmune disease systemic sclerosis (SSc) presents with multiple organ manifestations that often complicate management strategies. To explore variations in immune cell subsets and their link to clinical heterogeneity, here we perform single-cell profiling of peripheral blood mononuclear cells (PBMC) from 21 SSc patients who never received immunosuppressive therapy. We identify a subset of *EGRI*⁺ CD14⁺ monocytes in patients with scleroderma renal crisis (SRC). This subset activates NF- κ B signaling and differentiates into tissue-damaging macrophages, which accumulate at sites of tissue injury. Furthermore, we identify a CD8⁺ T cell subset with type II interferon signature in the peripheral blood and the lung tissue of patients with progressive interstitial lung disease (ILD), suggesting that chemokine-driven migration of these cells contributes to ILD progression. Thus, our single-cell analysis reveals distinct immune cell abnormalities associated with clinical organ manifestations, providing insights into tailored treatment strategies.

The clinical heterogeneity in systemic autoimmune diseases often complicates the management of individual patients¹. Systemic sclerosis (SSc) is primarily characterized by Raynaud's phenomenon and skin sclerosis, with an estimated global prevalence of approximately one million individuals. Patients with SSc present with a particularly diverse range of organ manifestations². These complications directly impact

the daily activities of SSc patients and are associated with a poor prognosis³.

The specific organs affected vary between patients; 50–65% develop interstitial lung disease (ILD), approximately 50% develop digital ulcers, and 1–14% develop scleroderma renal crisis (SRC), which is the most severe acute organ complication leading to end-stage renal

A full list of affiliations appears at the end of the paper. ✉ e-mail: kumanogo@imed3.med.osaka-u.ac.jp; nishide@imed3.med.osaka-u.ac.jp

disease and even death^{2,4}. While vascular damage and tissue fibrosis due to immune dysregulation play a central role in the pathogenesis of SSc², the immunological abnormalities underlying the clinical heterogeneity of the disease and the diversity of organ involvement have not been sufficiently investigated. Therefore, it is of great interest to explore the variation of immune abnormalities underlying the diversity of organ involvement in SSc.

Single-cell RNA sequencing (scRNA-seq) is a technique that comprehensively captures the diversity of individual cells. Since 2018, scRNA-seq studies of SSc patient samples have provided important insights into the pathology of the disease. Skin vascular endothelial cells from SSc patients show significant increases in gene expression related to extracellular matrix formation and angiogenesis inhibition⁵. Looking further into the diseased skin, functional alterations of *LRG5*⁺ fibroblasts⁶, enrichment of *SFRP2*^{high} fibroblasts⁷, the presence of *FCN1*⁺ dendritic cells⁸ and *CXCL13*⁺ T cells⁹ have also been demonstrated. Thus, while studies of the pathological mechanisms at the lesion sites in SSc are rapidly advancing¹⁰, it is still largely unknown how cellular diversity relates to symptom diversity from a single-cell transcriptomic perspective. Therefore, understanding the immune cell abnormalities underlying heterogeneous organ involvement may help to develop optimal strategies for disease monitoring and treatment in individual patients.

In this study, we perform scRNA-seq on peripheral blood mononuclear cells (PBMC) from immunosuppressive therapy-naïve SSc patients and show that *EGRI*⁺ CD14⁺ monocytes and CD8⁺ effector memory T cells with type II interferon signature are key subsets for SRC and ILD, respectively. In addition, renal tissue analysis suggests that *EGRI*⁺ CD14⁺ monocyte-derived tissue-damaging macrophages contribute to severe renal injury in SRC. The diversity of peripheral blood single-cell profiles identifies

pathological subsets of SRC and ILD, which hold potential as biomarkers and therapeutic targets.

Results

Single-cell profiling of PBMCs reveals distinct immune landscapes in SSc patients with different organ complications
PBMCs were obtained from 21 patients with SSc and six age- and sex-matched healthy donors. All recruited patients with SSc fulfilled the 2013 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria¹¹. The clinical characteristics of patients with SSc are summarized in Table 1. Detailed information on individual patients with SSc is provided in Supplementary Tables 1–3. Isolated PBMCs were analyzed on a 10× chromium® platform, and the transcriptome and expression of 43 surface proteins were simultaneously obtained using Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq); the experimental overview is shown in Fig. 1a. A total of 238,924 cells were processed, and each cell was annotated with supervised analysis using existing datasets¹². Uniform Manifold Approximation and Projection (UMAP) plots of PBMCs from SSc patients and healthy donors are shown in Fig. 1b. UMAP plots of PBMCs from each SSc patient or healthy donor are provided in Supplementary Fig. 1a. Highly expressed genes in each cell population are detailed in Supplementary Fig. 1b. The ratio of the number of cells in each subset to the total number of PBMCs was calculated. There were no significant differences in the proportion of each cell population relative to the total number of PBMCs between SSc patients and healthy donors (Fig. 1c, d). Similarly, no significant differences were found in cellular proportions when SSc patients were categorized into diffuse and limited cutaneous groups (Supplementary Fig. 2).

To analyze in-depth profiles of each cell population using single-cell transcriptomes, we further subdivided each cell population using previously reported marker genes^{13–16}. UMAP plots of monocyte subsets and profiles of highly expressed genes are provided in Supplementary Fig. 3a, b, respectively. UMAP plots of monocyte subsets from individual SSc patients or healthy donors are provided in Supplementary Fig. 3c. Similarly, single-cell level gene expression profiles were determined for CD4⁺ T cell subsets (Supplementary Fig. 4a–c), CD8⁺ T cell subsets (Supplementary Fig. 5a–c), B cell subsets (Supplementary Fig. 6a–c), natural killer (NK) cell subsets (Supplementary Fig. 7a–c), and plasmacytoid dendritic cell (pDC) subsets (Supplementary Fig. 8a–c).

Patients with SSc exhibit a variety of organ involvements. To unbiasedly capture the global changes in cell composition across recruited patients, we conducted principal component analysis (PCA) on the relative proportions of each cell subset (defined in the subset analysis shown in Supplementary Figs. 3–8) to the total PBMCs. Each vector representing a specific cell subpopulation is plotted on a two-dimensional graph according to its principal component 1 (PC1) and principal component 2 (PC2) (Fig. 1e). Individual SSc patients and healthy donors were then plotted based on their respective PC1 and PC2 values (Fig. 1f, left). The mean of the vectors representing individual SSc patients and those representing healthy donors pointed in distinct directions. Patients with SSc were further categorized by concomitant organ complications: SRC, ILD without SRC, and neither SRC nor ILD (Fig. 1f, right). The direction of the mean vector of the SRC group was aligned with the vectors representing monocyte and dendritic cell subpopulations, whereas the mean vector of the ILD without SRC group was closely associated with the vectors for T cell subpopulations and plasmablasts. The mean vector of cases with neither SRC nor ILD showed an orientation similar to that of healthy donors.

To further elucidate the compositional changes in PBMCs from SSc patients with SRC or ILD, we performed differential abundance analysis using milo¹⁷, which is a cluster-free and age-adjusted approach designed to detect changes in cell composition between conditions.

Table 1 | Clinical characteristics of systemic sclerosis patients and healthy donors for scRNA-seq analysis

		Systemic sclerosis		Healthy donors	
Overall profiles					
Number		21		6	
Age (Years)		64.0	(56.5–75.0)	65.5	(61.3–73.5)
Male/Female		5/16		2/4	
Duration (years)		7.0	(4.5–18.0)		
Skin and organ involvement					
Skin	Skin sclerosis	21	100%		
	Diffuse/Limited cutaneous	3/18	14/86%		
	Raynaud's phenomenon	21	100%		
	Digital ulcer	5	23.8%		
Interstitial lung disease		12	57.1%		
Scleroderma renal crisis		4	19.0%		
Pulmonary artery hypertension		3	14.3%		
Chronic intestinal pseudo-obstruction		2	9.5%		
Autoantibody					
Anti-RNA polymerase III		9	42.9%		
Anti-topoisomerase I		3	14.3%		
Anti-centromere		7	33.3%		
Anti-nuclear (nucleolar pattern)		2	9.5%		

Values for Age and Duration are presented as median (interquartile range).

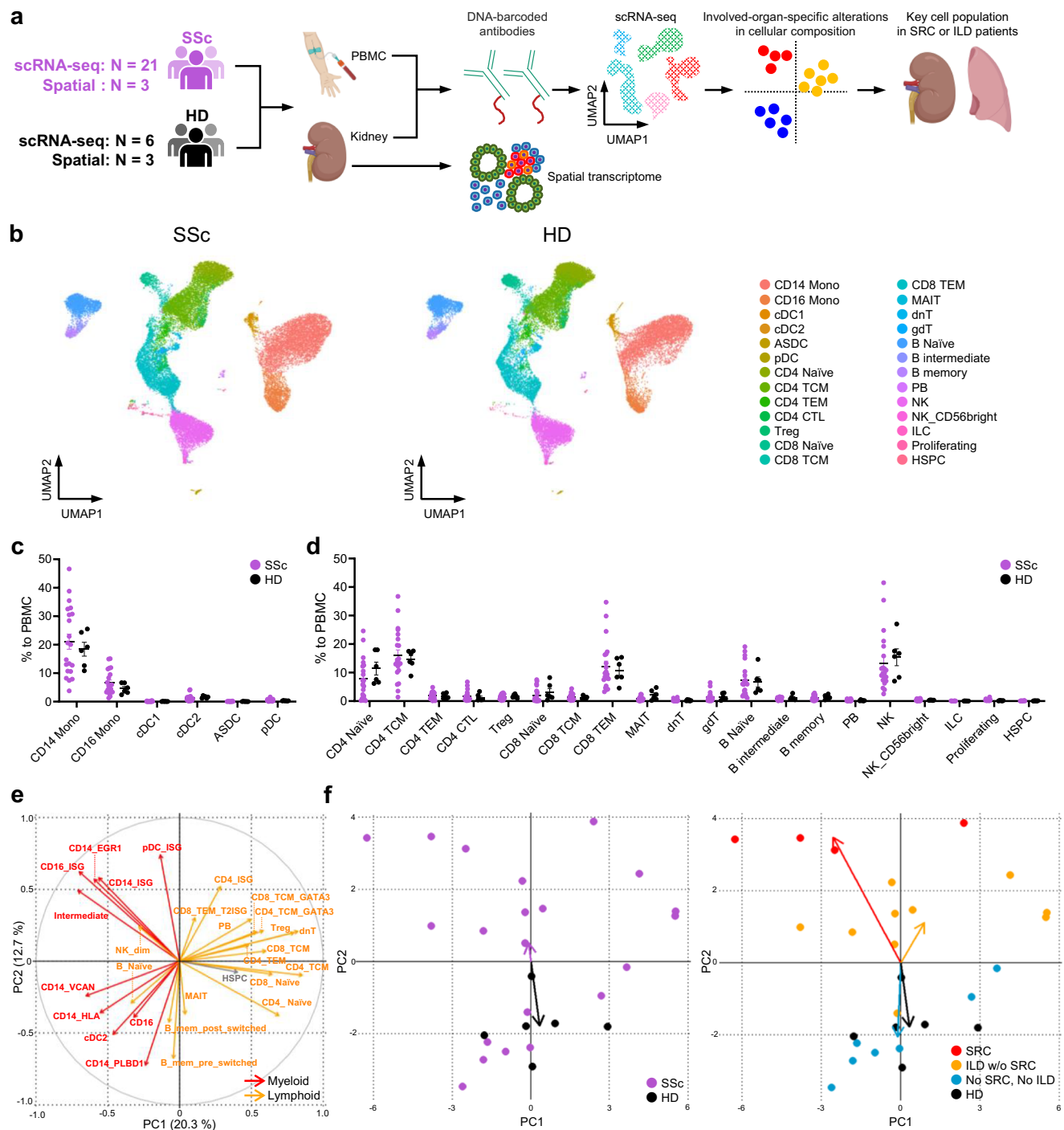


Fig. 1 | Comprehensive single-cell analysis reveals distinct immune profiles in patients with systemic sclerosis and healthy donors. **a** Overview of the experimental workflow. Created in BioRender. Shimagami, H. (2025) <https://BioRender.com/jhwa38f>. SSc systemic sclerosis, HD healthy donor, PBMC peripheral blood mononuclear cells, scRNA-seq single-cell RNA sequencing, SRC scleroderma renal crisis, ILD interstitial lung disease. **b** UMAP plots displaying CITE-seq data of PBMCs from SSc patients (SSc; $n = 21$) and HDs (HD; $n = 6$), annotated with reference mapping. Each plot shows 30,000 randomly selected cells. Mono monocytes, cDC conventional dendritic cells, ASDC AXL^+ dendritic cells, pDC plasmacytoid dendritic cells, TCM central memory T cells, TEM effector memory T cells, CTL cytotoxic T lymphocytes, Treg regulatory T cells, MAIT mucosal associated invariant T cells, dnT double negative T cells, gdT gamma-delta T cells, PB plasmablasts, NK natural killer cells, ILC innate lymphoid cells, Proliferating proliferating cells, HSPC hematopoietic stem and progenitor cells. Percentages of myeloid (c) and lymphoid (d) cell populations relative to the total PBMCs in SSc patients ($n = 21$, purple) and

HDs ($n = 6$, black). Values represent means with standard error of the mean (SEM). Statistical analysis was conducted using the two-sided Kolmogorov–Smirnov test with Bonferroni correction for multiple comparisons. **e** Principal component analysis on the proportion of each cell subset to the total PBMCs. Each subset was defined based on CITE-seq data (Supplementary Figs. 3–8). Vectors represent Principal Component (PC) 1 and PC2 for each cell subset. Red, myeloid cell subset; orange, lymphoid cell subset; gray, HSPC. Vectors with lengths greater than 0.3 are displayed. **f** Distribution of values of PC1 and PC2 for each study participant. In the left panel, dots represent individuals, with SSc patients shown in purple and HDs in black. In the right panel, each SSc patient is indicated by a colored dot: red for SRC patients, orange for ILD patients without SRC (ILD w/o SRC), and blue for patients without both SRC and ILD (No SRC, No ILD); HDs are represented by black dots. Arrows show the mean vectors for each group, colored to match the dots of corresponding group. Source data are provided as a Source Data file.

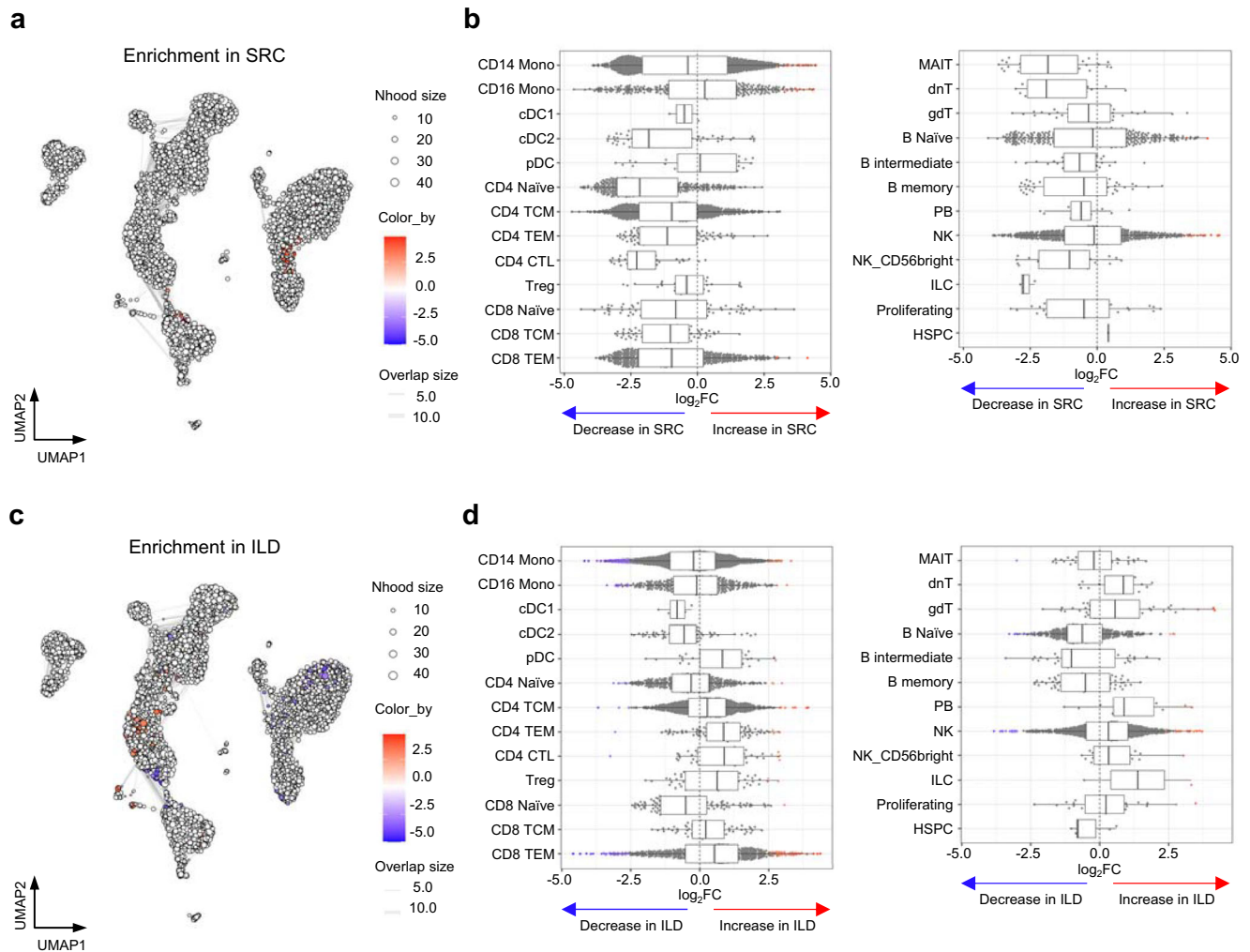


Fig. 2 | Differential abundance analysis on PBMCs derived from SSc patients with scleroderma renal crisis and interstitial lung disease. **a** Neighborhood graph of PBMCs from SSc patients, generated using Milo differential abundance testing. Nodes represent neighborhoods of the PBMCs. The color scale indicates the log₂-fold difference between SRC patients and the other SSc patients. Neighborhoods showing an increase in SRC are colored in red, while those with a decrease are in blue. **b** Beeswarm and box plots showing the distribution of log₂-fold differences between SRC patients ($n = 4$) and the other SSc patients ($n = 17$) in neighborhoods in different cell subsets colored as in (a). Box plots show median

and interquartile range (IQR); the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value that is no further than $1.5 \times \text{IQR}$ from the hinge. The lower whisker extends from the hinge to the smallest value that is at most $1.5 \times \text{IQR}$ from the hinge. Neighborhood graph (c) and Beeswarm plots (d) of PBMCs from ILD without SRC ($n = 10$) and the other SSc patients ($n = 11$), generated in a similar fashion as in (a) and (b), respectively. Neighborhoods showing an increase in ILD without SRC are shown in red, while those with a decrease are in blue.

Differential abundance analysis revealed that CD14⁺ monocytes, CD16⁺ monocytes, and NK cells were particularly enriched in patients with SRC compared to those without SRC (Fig. 2a, b). In contrast, memory T cell subsets such as CD8⁺ effector memory T cells were notably enriched in ILD patients (Fig. 2c, d). These data suggest that a skew in the gene expression profiles within the peripheral blood of SSc patients is associated with their organ complications. Specifically, SRC was linked to myeloid cell subsets, whereas ILD was linked to lymphoid cell subsets. Differential abundance analysis further highlighted the distinct immune abnormalities underlying each complication; enrichment of monocytes in SRC and memory T cells in ILD.

Enrichment of EGRI-expressing CD14⁺ monocytes in the peripheral blood of SRC patients

PCA and differential abundance analysis identified the monocyte subsets as interesting targets in terms of elucidating the pathogenesis of SRC. UMAP plots of monocytes from SSc patients with SRC or without SRC are shown in Fig. 3a. Five cellular clusters of CD14⁺

monocytes were identified by their distinct gene expression profiles: CD14⁺ monocytes with high expression of *EGRI* (CD14_EGRI), interferon signature genes (ISG) (CD14_ISG), *PLBD1* (CD14_PLBD1), *VCAN* (CD14_VCAN), or HLA (CD14_HLA). Other clusters include intermediate monocytes (Intermediate), CD16⁺ monocytes characterized by high expression of ISGs (CD16_ISG), other CD16⁺ monocytes (CD16), conventional type 1 dendritic cells (cDC1), and conventional type 2 dendritic cells (cDC2). Quantitative analysis of the relative distribution of each cellular subpopulation showed that the CD14_EGRI, CD14_ISG, Intermediate, and CD16_ISG subsets were abundant in patients with SRC (Fig. 3b). Differential abundance analysis further identified significant enrichment of CD14_EGRI (median log₂-fold change: +1.9), CD14_ISG (median log₂-fold change: +1.3), and CD16_ISG (median log₂-fold change: +1.6) subsets in patients with SRC, with the interquartile range of neighbors shifting entirely toward SRC, relative to the baseline where the fold change equals 1 (Fig. 3c, d). Similar results were observed when conducting differential abundance analysis comparing SRC patients with healthy donors (Supplementary Fig. 9a, b).

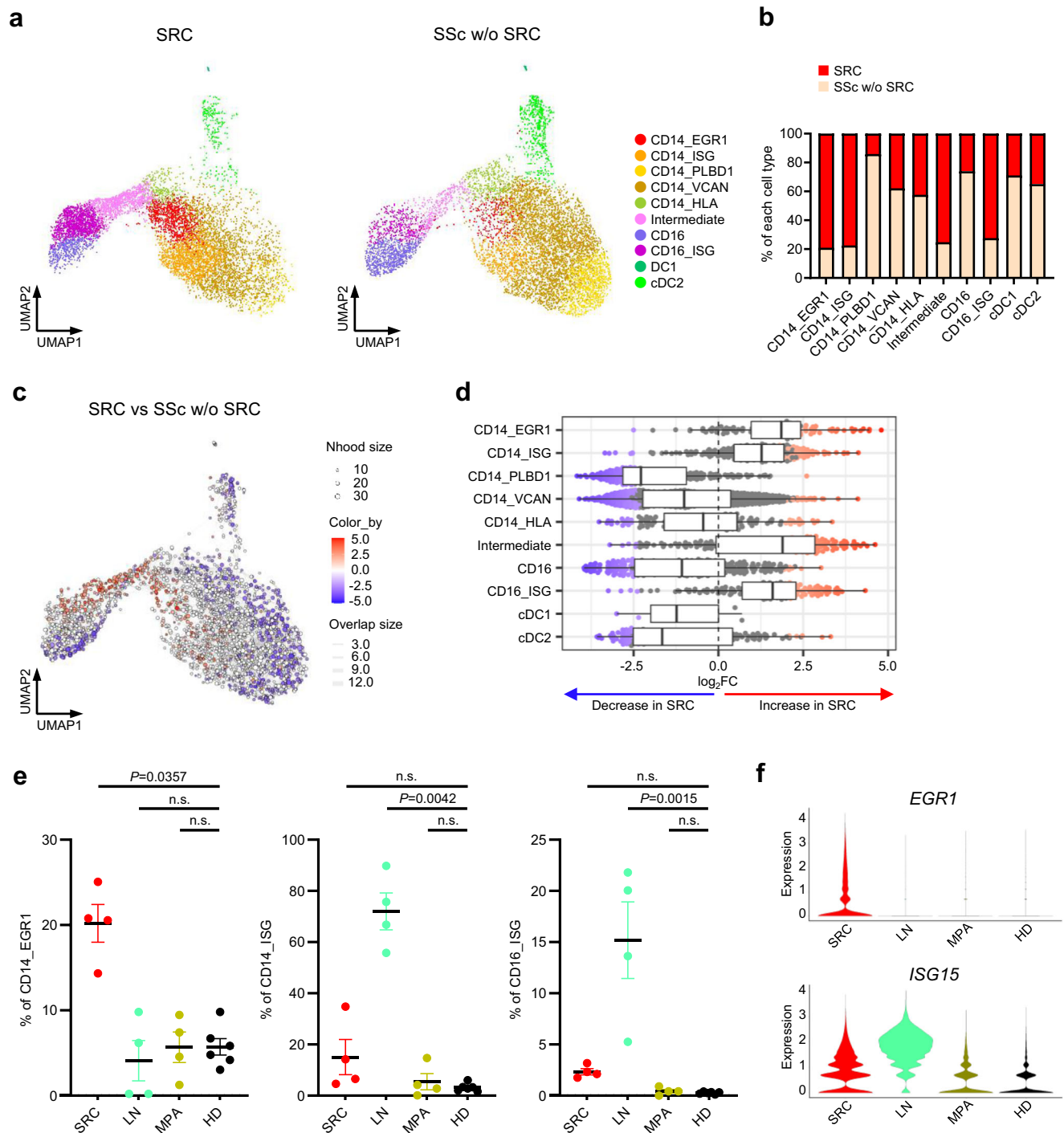


Fig. 3 | CD14⁺ monocytes with *EGR1* expression are specifically enriched in PBMCs from patients with scleroderma renal crisis. **a** UMAP plots of monocytes from SSc patients with SRC (SRC; $n = 4$) or without SRC (SSc w/o SRC; $n = 17$). Each plot shows 8000 randomly selected cells. **b** Stacked bar graph showing the relative contribution of the disease states to the total count of each cell type. The ratio of each cell population to total monocytes was calculated. The ratios were subsequently averaged in each disease state and normalized to a sum of 100% as the relative contribution. **c** Neighborhood graph of monocytes from SSc patients, generated in a similar fashion as in Fig. 2a. The color scale indicates the log₂-fold difference between SRC and SSc w/o SRC. Neighborhoods showing an increase in SRC are colored in red, while those with a decrease are in blue. **d** Beeswarm and box

plots showing the distribution of log₂-fold differences between SRC ($n = 4$) and SSc w/o SRC ($n = 17$) in neighborhoods in different cell subsets, colored as in (c). Box plots are created as in Fig. 2b. **e** Percentages of CD14_EGR1, CD14_ISG, and CD16_ISG subsets in the total monocyte population derived from patients with SRC ($n = 4$), patients with lupus nephritis (LN, $n = 4$), patients with microscopic polyangiitis (MPA, $n = 4$), and HDs ($n = 6$). Values represent means with SEM. P -values were calculated using two-sided Dunn's multiple comparison test between HD group and each of the other groups after Kruskal–Wallis test. **f** Violin plots showing *EGR1* and *ISG15* expression in monocytes derived from patients with SRC, patients with LN, patients with MPA, and HDs. Source data are provided as a Source Data file.

To determine whether the increases in the CD14_EGR1, CD14_ISG, and CD16_ISG subsets are characteristic of SRC patients, we conducted two additional analyses. First, we focused on other vascular phenotypes in SSc, such as digital ulcers (DU) and pulmonary artery hypertension (PAH). Differential abundance analysis revealed a weak but significant enrichment of the CD14_EGR1 (median log₂-fold change: +0.78), CD14_ISG (median log₂-fold change: +0.95), and CD16_ISG (median log₂-fold change: +1.00) subsets in patients with DU compared to those without DU (Supplementary Fig. 10a, b). Although the enrichment levels were not as prominent as those seen in SRC (Fig. 3c, d), this finding aligns with a previous study suggesting that the development of DU may share a similar pathogenic mechanism with SRC¹⁸. No significant enrichment of monocyte subsets was observed in patients with pulmonary artery hypertension compared to those without PAH (Supplementary Fig. 10c, d). Second, we recruited patients with other autoimmune diseases associated with kidney damage, including lupus nephritis (LN) and microscopic polyangiitis (MPA). Using CITE-seq, we compared the proportions of each immune cell subset among four groups: patients with SRC, patients with LN, patients with MPA, and healthy donors. The clinical profiles of patients with LN and MPA are presented in Supplementary Tables 4 and 5, respectively. UMAP plots derived from the subset analysis of monocytes are shown in Supplementary Fig. 11a. Highly expressed genes in each cell subpopulation are detailed in Supplementary Fig. 11b. UMAP plots of monocytes from individual patients and healthy donors are provided in Supplementary Fig. 11c. The proportion of the CD14_EGR1 subset was specifically increased in patients with SRC compared to healthy donors, whereas the proportions of the CD14_ISG and CD16_ISG subsets were higher in patients with LN (Fig. 3e). The expression of *EGR1* in total monocytes was increased only in patients with SRC, while the expression of *ISG15*, a representative ISG, was highly elevated in patients with LN (Fig. 3f). These data suggest that an increased proportion of the CD14_EGR1 subset in the peripheral blood is a key feature of SRC.

Pathological role of CD14_EGR1 monocytes in tissue damage mediated by NF-κB pathway activation

To further explore the characteristics of this specific CD14_EGR1 subset, we identified differentially expressed genes (DEG) in this subset compared to other monocyte subpopulations (Supplementary Table 6). Pathway analysis of the DEGs (log₂-fold change > 0.5) showed predominant activation of the “TNF-alpha signaling via NF-κB” pathway in the CD14_EGR1 subset (Fig. 4a). The expression levels of representative NF-κB-targeted genes such as *EGR1*, *IL1B*, *CCR1*, and *SGK1* are shown in Fig. 4b. The results indicate that the expression of each gene was increased in the SRC group, with a similar distribution to *EGR1*. The expression profiles of cell surface antigens based on CITE-seq indicated that the CD14_EGR1 subset highly expressed CD11b, CD38, and CD4 (Supplementary Fig. 12), suggesting that the subset has enhanced capacity for tissue migration and differentiation^{19–21}.

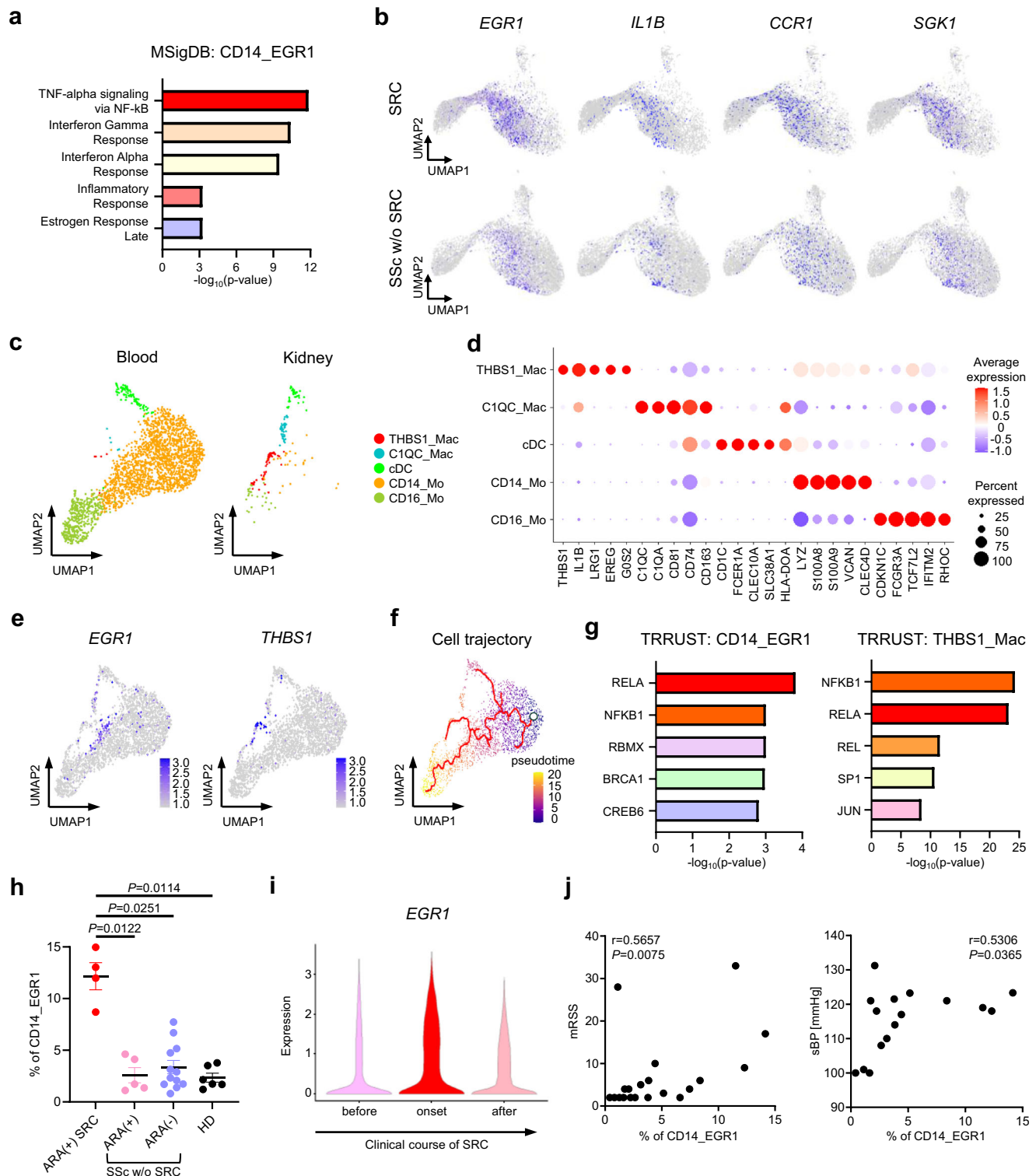
On the basis of these gene expression profiles and surface antigen data, we hypothesized that the CD14_EGR1 subset possesses enhanced migratory capacity in tissues and contributes to renal damage. A sample of kidney tissue was obtained from one patient (patient ID: SSc-19) at the onset of the first episode of SRC. Pathological findings showed intimal thickening and luminal narrowing within arteries, fibrinoid necrosis of arterioles, red blood cell fragmentations, and focal tubular necrosis, which support the diagnosis of SRC (Supplementary Fig. 13a–d). Renal cells and total peripheral white blood cells from the same patient were independently analyzed on the BD Rhapsody® platform. The UMAP plots of the peripheral white blood cells and renal cells are shown in Supplementary Fig. 14a. Profiles of highly expressed genes (Supplementary Fig. 14b) and surface antigens (Supplementary Fig. 14c) for each cell population are also provided. We then subset

monocytes, conventional dendritic cells (cDC), and macrophages. The RNA expression profile of known gene markers²² was used to divide these cells into five groups: CD14⁺ monocytes (CD14_Mo), CD16⁺ monocytes (CD16_Mo), cDCs, macrophages characterized by high expression of *THBS1* (THBS1_Mac), and macrophages characterized by high expression of *CIQC* (CIQC_Mac). The CIQC_Mac subset was considered as kidney resident macrophages characterized by high expression of genes such as *CIQC*, *CD81*, and *CD74*²³. UMAP plots derived from this subset analysis are presented in Fig. 4c, divided by blood and renal cells. Profiles of highly expressed genes in myeloid subsets are provided in Fig. 4d. In the THBS1_Mac subset, genes linked to vascular injury and fibrosis, such as *THBS1*²⁴, *IL1B*²⁵, and *LRG1*²⁶, were highly expressed. DEGs were identified in the THBS1_Mac subset compared to other myeloid cell subpopulations (Supplementary Table 7), and subsequent pathway analysis revealed an enrichment of the “Interleukin-1 regulation of extracellular matrix” pathway in this subset (Supplementary Fig. 15). The expression levels of *EGR1* and *THBS1* are shown as feature plots in Fig. 4e. To estimate the cell trajectory from peripheral blood to kidney tissue, we conducted trajectory analysis using Monocle 3²⁷. The root node was set to immature monocytes characterized by high expression of *S100A12* and low expression of HLA-DR genes²⁸. The expression levels of *S100A12* and *HLA-DRB1* are shown as feature plots in Supplementary Fig. 16. CD14⁺ monocytes, which have high *EGR1* expression, showed a differentiation trajectory leading to the THBS1_Mac subset (Fig. 4f). We next performed transcriptional regulatory relationships unraveled by sentence-based text-mining (TRRUST) analysis²⁹, a transcription factor analysis using DEGs in the CD14_EGR1 and THBS1_Mac subsets. The activation of *RELA* and *NFKB1*, members of the NF-κB family³⁰, was a shared transcriptional characteristic between these subsets (Fig. 4g). These data suggest that the CD14_EGR1 subset may differentiate into the THBS1_Mac subset in the kidney of SRC patients. Gene expression profiles of the THBS1_Mac subset implicate a role in vascular damage and fibrosis in SRC tissue.

Clinical relevance between EGR1 expression in monocytes and bedside observations

Clinically, all patients with SRC who are included in the scRNA-seq analysis in this study tested positive for anti-RNA polymerase III antibody (ARA), the autoantibody strongly associated with SRC³¹. Therefore, we compared the frequency of the CD14_EGR1 subset according to the presence or absence of ARA. The proportion of the CD14_EGR1 subset was significantly higher in SRC patients than in those without, regardless of ARA positivity (Fig. 4h). These data suggest that the enrichment of the CD14_EGR1 subset is due to the presence of SRC rather than to autoantibody profiles. To further validate this concept, we analyzed publicly available scRNA-seq data of PBMCs from SSc patients with SRC, those without SRC (SSc w/o SRC), and healthy donors⁶. UMAP plots of monocytes are shown in Supplementary Fig. 17a. In SRC patients, the expression levels of representative DEGs of the CD14_EGR1 subset in our dataset were generally elevated in CD14⁺ monocytes from the public dataset (Supplementary Fig. 17b). In addition, the module score calculated using the DEGs of the CD14_EGR1 subset was high in CD14⁺ monocytes from SRC patients (Supplementary Fig. 17c), independent of anti-topoisomerase I antibody (ATA)- or ARA-positivity (Supplementary Fig. 17d). These data support our concept that *EGR1* upregulation in CD14⁺ monocytes is a characteristic of SRC regardless of their autoantibody status.

In addition, we assessed changes in *EGR1* expression during the clinical course of SRC. We collected PBMCs from a patient (patient ID: SSc-1) and conducted CITE-seq analysis at three different time points: three months before the onset of SRC, at the onset of SRC, and after the improvement of SRC. The UMAP plots of PBMCs in this



longitudinal analysis are shown in Supplementary Fig. 18. The expression of *EGR1* in total monocytes was highly upregulated at the onset of SRC and decreased following treatment (Fig. 4i). Finally, we evaluated the association of the CD14⁺ *EGR1* subset with clinical features of SSc. The proportion of the CD14⁺ *EGR1* subset in SSc patients showed a positive correlation with the modified Rodnan skin score (mRSS) and systolic blood pressure (sBP), both of which are clinical indicators of SRC (Fig. 4j). These data suggest that *EGR1* expression in CD14⁺ monocytes could serve as a predictive marker for the onset and progression of SRC.

Spatial transcriptomic analysis on kidney tissue from SRC patients

To further investigate the gene expression profiles of immune cells within the disease site, we applied spatial transcriptomic analysis to formalin-fixed, paraffin-embedded (FFPE) kidney tissues from SRC patients and healthy donors. Using the CosMX platform, single-cell level expression profiles of 1000 genes were obtained from kidney tissue samples from three SRC patients at disease onset and three healthy donors. The clinical characteristics of each participant are shown in Supplementary Table 8. UMAP plots of all kidney cells are

Fig. 4 | Involvement of the CD14_EGRI subset in tissue damage and its potential as a biomarker for the progression of scleroderma renal crisis. **a** Gene set enrichment analysis (GSEA) of differentially expressed genes (DEGs) in the CD14_EGRI subset using the Molecular Signatures Database (MSigDB). *P*-values were calculated using Fisher's exact test, with Benjamini–Hochberg method for multiple comparison. **b** The feature plots of *EGRI*, *IL1B*, *CCR1*, and *SGKI* in monocytes from SSc patients with SRC (SRC) or without SRC (SSc w/o SRC). **c** UMAP plots showing the monocytes, macrophages and cDCs from peripheral blood (left) and kidney (right) of a SSc patient (Patient ID: SSc-19) at the onset of SRC. **d** Balloon plot showing highly expressed genes in subpopulations shown in (c). **e** Feature plots showing *EGRI* and *THBS1* expression at the onset of SRC. **f** Cell trajectory analysis on the peripheral blood and kidney tissue cells at the onset of SRC. The red line on the UMAP plot represents the cell trajectory. The root node is marked with a circle. Each cell is colored according to each pseudotime. **g** GSEA of DEGs in CD14_EGRI (left) and THBS1_Mac (right) subsets using transcriptional regulatory

relationships unraveled by sentence-based text-mining (TRRUST) Transcriptional Factor 2019. *P*-values were calculated using Fisher's exact test, with Benjamini–Hochberg method for multiple comparison. **h** Percentage of the CD14_EGRI subset to the total monocytes in SSc patients classified by serum autoantibody status. Patients were stratified into four subgroups: ARA (+) SRC ($n = 4$), ARA (+) SSc w/o SRC ($n = 5$), ARA (–) SSc w/o SRC ($n = 12$). ARA, anti-RNA polymerase III antibody. Values represent means with SEM. *P*-values were calculated using two-sided Dunn's multiple comparison test between ARA (+) SRC group and each of the other groups after Kruskal–Wallis test. **i** The violin plot of *EGRI* expression in monocytes before SRC development (before), at SRC onset (onset), and after treatment (after). **j** Correlation between the percentage of the CD14_EGRI subset to total peripheral monocytes and the modified Rodnan skin score (mRSS) or systolic blood pressure (sBP). Correlations and *P*-values were calculated using Spearman's correlation coefficient (r) with a two-sided test. Source data are provided as a Source Data file.

shown in Supplementary Fig. 19a. Profiles of highly expressed genes in each cell population are also provided (Supplementary Fig. 19b). UMAP plots of kidney cells from individual SRC patients or healthy donors are shown in Supplementary Fig. 19c. Representative CosMX images of kidney tissue from each SRC patient and healthy donor are shown in Fig. 5a. In the kidney tissue of SRC patients, the proportion of immune cells was elevated compared to those in healthy donors (Fig. 5b). Given the susceptibility of the proximal tubule to ischemic injury³², we next performed a subset analysis of proximal tubular cells (PTB). UMAP plots of PTBs are shown in Supplementary Fig. 20a. Profiles of highly expressed genes in each cell population are also provided (Supplementary Fig. 20b). Notably, in SRC patients, the *VCAM1*-expressing injured PTB subset was a predominant population.

As we speculate that the CD14_EGRI subset differentiates into the THBS1_Mac subset in the kidneys of SRC patients (Fig. 4f), we next focused on myeloid cell subsets in SRC kidney tissue. UMAP plots of myeloid cells are shown in Fig. 5c. Profiles of highly expressed genes in each myeloid cell subset are also provided (Fig. 5d). Four cellular clusters of myeloid cells were identified by their gene expression profiles: macrophage 1 (Mac1), macrophage 2 (Mac2), macrophage 3 (Mac3), and cDC. Module score analysis based on DEGs in the THBS1_Mac subset (Supplementary Table 7) showed the similarity of the Mac1 subset to the THBS1_Mac subset (Fig. 5e). DEGs were identified in the Mac1 subset compared to other myeloid cell subpopulations (Supplementary Table 9), and subsequent pathway analysis revealed increased signaling related to extracellular matrix formation (Fig. 5f). TRRUST analysis highlighted the activation of NFKB1 and RELA in the Mac1 subset (Fig. 5g), indicating a similar activation state to the CD14_EGRI and THBS1_Mac subsets. We further assessed the spatial proximity of each cell type within SRC-affected kidneys. Proximity groups were defined by clustering kidney cells based on the type and number of neighboring cells. In SRC kidney tissue, Mac1 and PTB were frequently found within the same proximity group (Fig. 5h), whereas this proximity was not observed in HD kidney tissue (Supplementary Fig. 21). These data suggest that the THBS1-expressing macrophages, located at diseased loci in close proximity to injured PTB, may contribute to the fibrosis observed in damaged SRC tissue.

Enrichment and functional implications of interferon gamma-responsive CD8⁺ T cell subsets in SSc-ILD

The results of the PCA and differential abundance analysis of PBMCs prompted us to conduct further analysis of CD4⁺ T cells and CD8⁺ T cells in patients with SSc-ILD. UMAP plots of CD4⁺ T cells from SSc patients are shown in Supplementary Fig. 22a. Differential abundance analysis indicated the enrichment of CD4⁺ T cells characterized by high expression of ISGs in patients with SSc-ILD compared to those without ILD, but no significant enrichment was found in comparison with healthy donors (Supplementary Fig. 22b–e).

UMAP plots of CD8⁺ T cells from SSc patients with ILD and without ILD are shown in Fig. 6a. Six cellular clusters were identified: naïve CD8⁺ T cells (CD8_T_Naive), central memory CD8⁺ T cells characterized by high expression of *GATA3* (CD8_TCM_GATA3), other central memory CD8⁺ T cells (CD8_TCM), effector memory CD8⁺ T cells characterized by high expression of type 2 ISGs (CD8_TEM_T2ISG), other effector memory CD8⁺ T cells (CD8_TEM), and cytotoxicity active CD8⁺ T cells (CD8_CTL). Quantitative analysis of the relative distribution of each cellular subpopulation showed that CD8⁺ memory T cell subsets, CD8_TCM, CD8_TCM_GATA3, CD8_TEM, and CD8_TEM_T2ISG, were abundant in patients with SSc-ILD (Fig. 6b). Differential abundance analysis revealed that the CD8_TEM_T2ISG subset was most enriched in patients with SSc-ILD (median log₂-fold change: +1.2) (Fig. 6c, d). The increase in CD8_TEM_T2ISG was consistently observed in patients with SSc-ILD when compared to the healthy donors (Supplementary Fig. 23a, b). To further investigate the characteristics of the CD8_TEM_T2ISG subset, DEGs of the CD8_TEM_T2ISG compared to other CD8⁺ T cell subpopulations were identified (Supplementary Table 10). Pathway analysis of the DEGs highlighted significant enrichment of the “Interferon Gamma Response” pathway in the CD8_TEM_T2ISG subset (Fig. 6e). Interferon gamma (IFN- γ) is widely known to enhance tissue migration in CD8⁺ T cells³³. We therefore examined the gene expression profiles of chemokine receptors in each CD8⁺ T cell subset. *CXCR3* and *CCR5*, which are important for T cell migration in pathological conditions³⁴, were highly expressed in the CD8_TEM_T2ISG subset (Fig. 6f).

In our study, the ratio of the CD8_TEM_T2ISG subset to the total CD8⁺ T cell population was significantly higher in patients positive for ATA or ARA compared to those negative for these two antibodies (Fig. 6g). Clinically, patients with SSc-ILD who are positive for ATA or ARA are known to have a greater risk of ILD progression compared to those positive for anti-centromere antibody³⁵. Therefore, we next investigated whether IFN- γ -associated immunological changes in CD8⁺ T cells contribute to lung damage in progressive SSc-ILD. Publicly available scRNA-seq datasets of the lung tissue derived from patients with advanced SSc-ILD and healthy donors³⁶ were analyzed, followed by a detailed subset analysis of CD8⁺ T cells. UMAP plots of CD8⁺ T cells in lung tissues from SSc-ILD patients and healthy donors are shown in Supplementary Fig. 24a. Using previously reported marker genes³⁴, CD8⁺ T cells in the lung tissue were divided into three groups: CD8⁺ T cells characterized by high expression of type 2 ISG (CD8_T_T2ISG), GZMH (CD8_T_GZMH), or GZMK (CD8_T_GZMK). Highly expressed genes in each cell population are shown in Supplementary Fig. 24b. Differential abundance analysis revealed that the CD8_T_T2ISG subset was enriched in the lung tissue of SSc-ILD patients (median log₂-fold change: +0.7), whereas the CD8_T_GZMH (median log₂-fold change: –1.6) and CD8_T_GZMK (median log₂-fold change: –0.7) subsets showed a decrease compared to healthy donors (Fig. 6h, i). The ratio of the CD8_T_T2ISG subset to the total lung CD8⁺ T cell population was

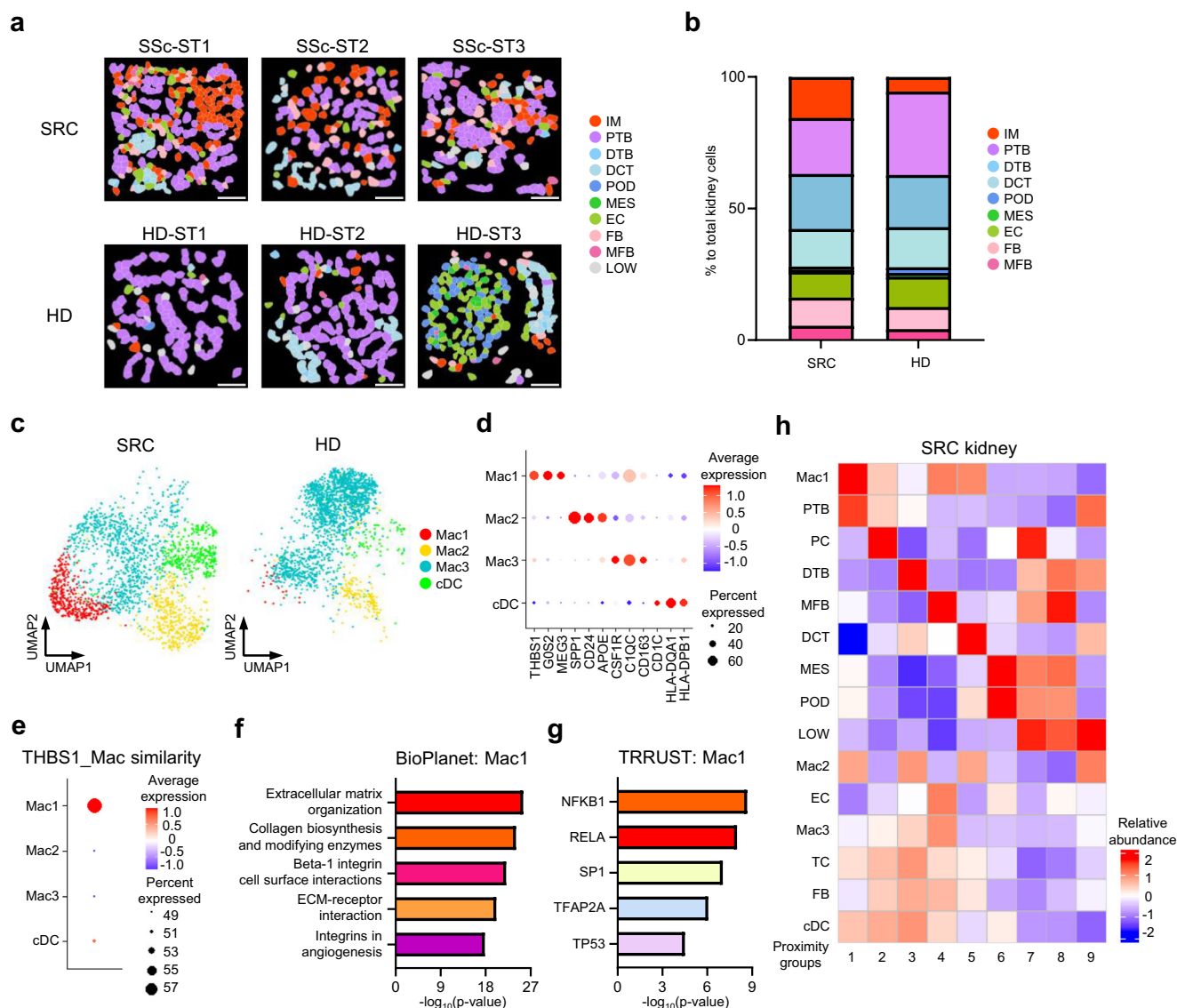


Fig. 5 | Accumulation of fibrosis-associated macrophages at the site of tissue injury in SRC. **a** Close-up CosMX images of kidney tissue from three patients with SRC (top row) and three HDs (bottom row). IM immune cells, PTB proximal tubular cells, DCT collecting duct cells, DTB distal tubular cells, POD podocytes, MES mesangial cells, EC endothelial cells, FB fibroblasts, MFB myofibroblasts, LOW low-quality cells. Scale bar = 50 μ m. **b** Stacked bar graphs showing the composition of cell types in kidney tissue from the SRC patients (left) and HDs (right). The colors correspond to those used in (a). **c** UMAP plots of myeloid cells from the SRC patients and HDs. **d** Balloon plot showing highly expressed genes in each subpopulation shown in (c). **e** Balloon plot showing the module score for

each myeloid cell subset calculated using the highly expressed genes in the THBS1_Mac subset. **f** GSEA of DEGs in the Mac1 subset using the BioPlanet 2019. **g** GSEA of DEGs in the Mac1 subset using TRRUST Transcriptional Factor 2019. *P*-values were calculated using Fisher's exact test, with Benjamini–Hochberg method for multiple comparison. **h** Heatmap showing the abundance of each cell subpopulation in proximity groups in the kidney of SRC patients. Each proximity group was determined by clustering cells based on the similarity of the type and number of neighboring cells, and numbered from 1 to 9. Source data are provided as a Source Data file.

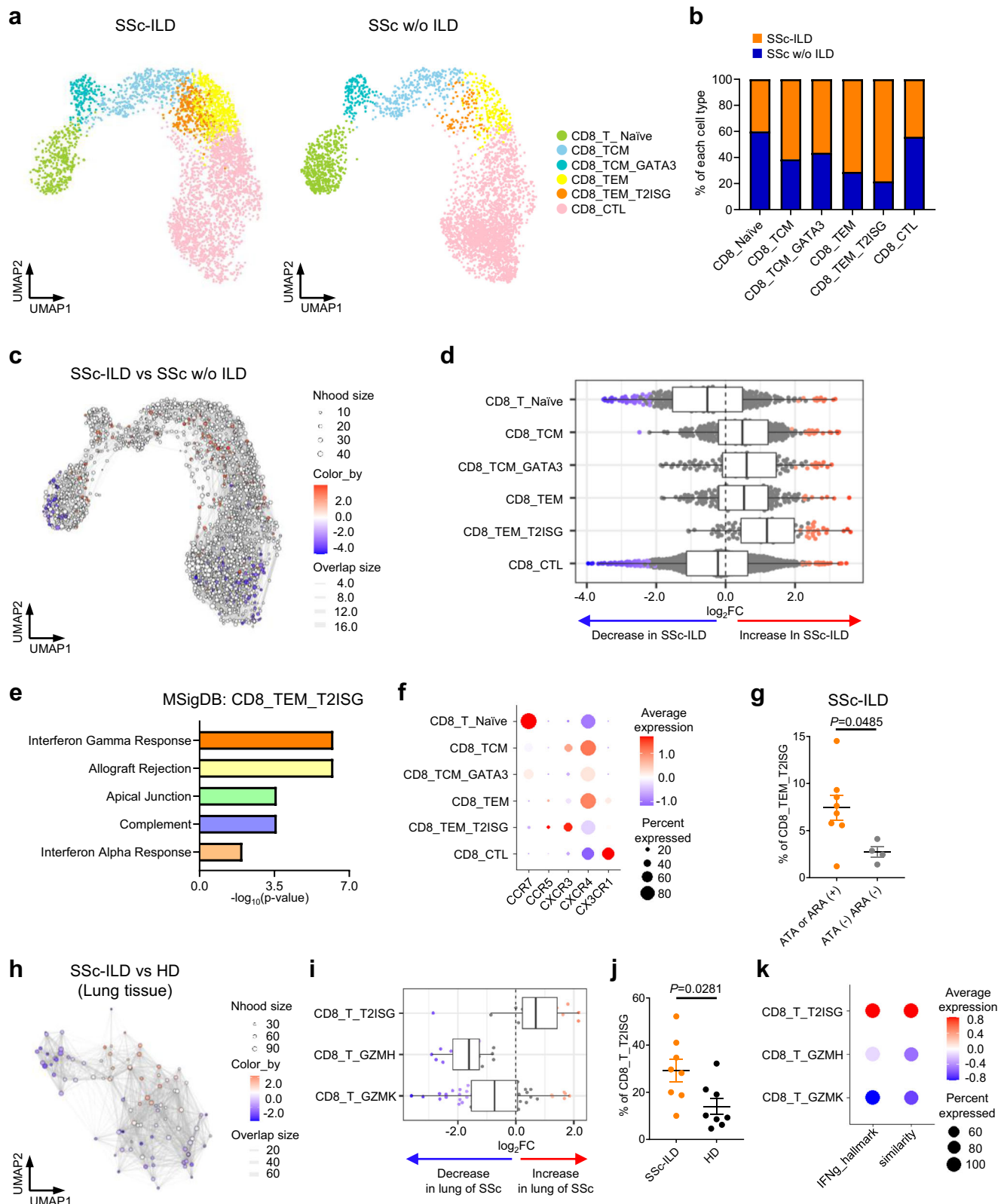
significantly higher in patients with SSc-ILD (Fig. 6j). Among CD8⁺ T cell populations in the lung, the module scores calculated using the set of IFN- γ signature genes from the Molecular Signatures Database (MSigDB)³⁷ were highest in the CD8_T_T2ISG subset (Fig. 6k). The module scores representing the similarity of each cell subset to CD8_T_T2ISG were also highest in the CD8_T_T2ISG subset (Fig. 6k). This similarity in cellular characteristics between peripheral blood and lung may explain that CD8⁺ T cell populations with IFN- γ signature genes, with their high migratory capacity, play a role in the pathophysiology of progressive ILD.

In summary, this study identified distinct immune abnormalities in PBMCs from patients with SSc, underlying SRC or ILD complications. There is characteristic enrichment of the

CD14_EGR1 subset in SRC and the CD8_TEM_T2ISG subset in SSc-ILD. Our findings further suggest that the CD14_EGR1 subset could potentially differentiate into the THBS1_Mac subset and contribute to the fibrosis observed in damaged SRC tissue. Clinically, *EGR1* expression in monocytes may serve as a novel biomarker for disease progression of SRC. In SSc-ILD patients, the CD8_TEM_T2ISG subset demonstrates characteristics of high migratory capacity and can be implicated in progressive lung damage (Fig. 7).

Discussion

The distribution of organ involvement in patients with autoimmune diseases is heterogeneous. In this study, we identified



distinct immune abnormalities underlying the clinical heterogeneity of SSc, based on single-cell transcriptome and surface protein profiles of PBMCs. Patients who had never received immunomodulatory drugs were recruited for scRNA-seq. The enrichments of myeloid subsets in patients with SRC and of lymphoid subsets in patients with SSc-ILD highlight the distinct characteristics associated with these organ complications.

In-depth subset analysis revealed the enrichment of a specific cellular population, CD14⁺ EGFR1⁺, in the peripheral blood of patients with SRC. *EGFR1* is classified as an immediate early gene and encodes a transcription factor crucial for the differentiation of monocytes into macrophages³⁸. TGF- β stimulation upregulates *EGFR1* expression in fibroblasts and increases collagen production³⁹. Therefore, previous studies of *EGFR1* function in SSc have focused on its role in tissue

Fig. 6 | Increase of the CD8⁺ T cell subpopulation with type II interferon signature genes is implicated in progressive SSc-ILD. **a** UMAP plots of peripheral CD8⁺ T cells from the SSc patients with ILD (SSc-ILD; $n = 12$) or without ILD (SSc w/o ILD; $n = 9$). Each plot shows 4000 randomly selected cells. **b** Stacked bar graph showing the relative contribution of the disease states to the total count of each cell subset, generated in a similar fashion as in Fig. 3b. **c** Neighborhood graph of peripheral CD8⁺ T cells from the SSc patients, generated in a similar fashion as in Fig. 2a. The color scale indicates the $\log_2$ -fold difference between SSc-ILD and SSc w/o ILD. Neighborhoods showing an increase in SSc-ILD are colored in red, while those with a decrease are in blue. **d** Beeswarm and box plots showing the distribution of $\log_2$ -fold differences between SSc-ILD ($n = 12$) and SSc w/o ILD ($n = 9$) in neighborhoods in different cell subsets, colored as in (c). Box plots are created as in Fig. 2b. **e** GSEA of DEGs in the CD8_{TEM}T2ISG subset using the MSigDB. P -values were calculated using Fisher's exact test, with Benjamini–Hochberg method for

multiple comparison. **f** Chemokine receptor expression across peripheral CD8⁺ T cell subsets. **g** Percentage of the CD8_{TEM}T2ISG subset in peripheral CD8⁺ T cells of SSc-ILD, stratified by ATA or ARA positivity (ATA or ARA (+), $n = 8$; ATA (–) ARA (–), $n = 4$). Values represent means with SEM. P -values were calculated using two-sided Mann–Whitney U test. Neighborhood graph (**h**) and Beeswarm plots (**i**) of lung CD8⁺ T cells from advanced SSc-ILD ($n = 8$) and HDs ($n = 8$), generated in a similar fashion as in (c) and (d), respectively. Neighborhoods showing an increase in SSc-ILD are colored in red, while those with a decrease are in blue. **j** Percentage of the CD8_{TEM}T2ISG subset in lung CD8⁺ T cells of SSc-ILD ($n = 8$) and HDs ($n = 8$). Values represent means with SEM. P -values were calculated using two-sided Mann–Whitney U test. **k** Module scores for each lung CD8⁺ T cell subset, calculated using the “HALLMARK_INTERFERON_GAMMA_RESPONSE” gene set from MSigDB (IFN γ hallmark) or the upregulated DEGs of the CD8_{TEM}T2ISG subset (similarity). Source data are provided as a Source Data file.

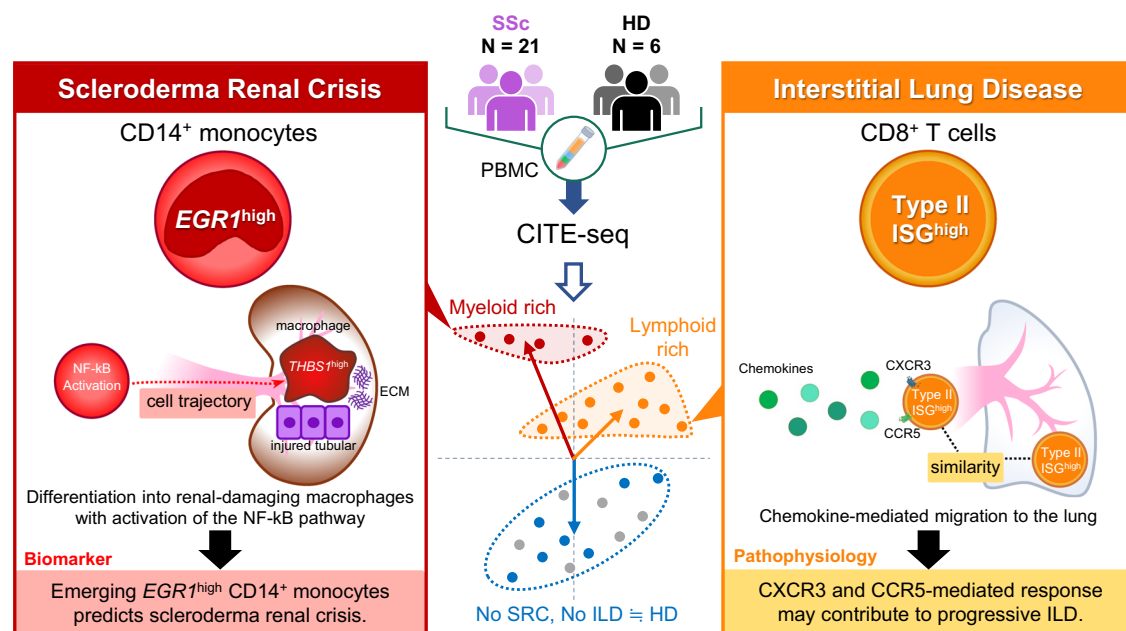


Fig. 7 | Graphical scheme of this study. CITE-seq analysis of PBMCs from patients with SSc ($n = 21$) or HDs ($n = 6$) identified distinct immune abnormalities in SSc patients with SRC or ILD. Patients with SRC demonstrated specific enrichment of CD14⁺ monocytes with increased *EGR1* expression and activation of NF-κB-related pathways. Trajectory analysis indicated their differentiation into macrophages with elevated expression of *THBS1* in the kidney. Clinically, changes in monocyte *EGR1* expression show potential as a biomarker for monitoring the disease progression of SRC. In patients with SSc-ILD, CD8⁺ TEMs with increased type II ISG expression were

enriched in PBMCs. A similar cell population was also enriched in the lung tissue of patients with advanced SSc-ILD, suggesting migration of CD8⁺ T cells from peripheral blood to the lung, mediated by chemokine receptors such as CXCR3 and CCR5. SSc systemic sclerosis, HD healthy donor, PBMC peripheral blood mononuclear cells, CITE-seq Cellular Indexing of Transcriptomes and Epitopes by Sequencing, SRC scleroderma renal crisis, ECM extracellular matrix, ILD interstitial lung disease, ISG interferon signature genes, TEMs effector memory T cells. Created in BioRender. Shimagami, H. (2025) <https://BioRender.com/m7mo04j>.

fibroblasts⁴⁰. In this study, *EGR1* expression in peripheral monocytes was significantly upregulated in SRC patients, and the pathway analysis indicated that NF-κB and type 1/2 interferon-related signaling were enriched in the CD14_{EGR1} subset. The synergistic effect of these signaling pathways results in enhanced activation of monocytes and macrophages, leading to increased production of pro-inflammatory cytokines⁴¹. Thus, the CD14_{EGR1} subset is considered an activated monocyte subpopulation, suggesting another contributory role beyond tissue fibrosis in the pathogenesis of SRC.

Two questions arise here: in the pathogenesis of SRC, how is the CD14_{EGR1} subset induced peripherally, and how does this population contribute to organ damage? The pathogenesis of SRC is hypothesized to involve initial damage to the renal vascular endothelium, followed by activation of the renin–angiotensin–aldosterone system. A harmful cycle of renal artery constriction and increased renin production leads to severe hypertension and significant organ damage³¹. In monocytes

and macrophages, angiotensin II induces the expression of *EGR1*⁴² and activates NF-κB⁴³. In the context of ischemia–reperfusion injury, pattern recognition receptors–mediated signaling triggered by damage-associated molecular patterns (DAMP), such as HMGB1, activates the NF-κB pathway⁴⁴ and induces *EGR1* expression⁴⁵. These observations may answer the first question; factors associated with SRC, including angiotensin II, DAMPs, and ischemia–reperfusion injury, can be involved in the induction of the CD14_{EGR1} subset in peripheral blood.

With regard to the second question, CD11b¹⁹, CD4²⁰, and CD38²¹, which are highly expressed surface antigens on the CD14_{EGR1} subset, are known to facilitate monocyte adhesion, migration, and differentiation. Considering gene expression and surface antigen profiles, the CD14_{EGR1} subset is suggested to have an enhanced capacity for tissue migration and differentiation. In addition, the trajectory analysis showed that the CD14_{EGR1} subset may differentiate into the

THBS1_Mac subset in the kidney. Of note, these subsets share the activation state of several transcription factors which are major components of NF- κ B. THBS1 is a component of the extracellular matrix and plays a role in promoting cell adhesion, regulating angiogenesis, and modulating inflammatory responses²⁴. In an animal model of renal ischemia-reperfusion injury, macrophages that infiltrate the kidney express elevated levels of THBS1, contributing to tubular damage⁴⁶. Other genes highly expressed in the THBS1_Mac subset, such as *LRG1*²⁶ and *IL1B*²⁵, are also implicated in renal damage. Given these findings, the THBS1_Mac subset may contribute to tubular damage and be involved in a part of the pathogenesis of acute kidney injury (AKI) in SRC. More interestingly from a clinical perspective, the expression of *EGRI* in monocytes was highly upregulated at the onset of SRC, and it decreased following treatment. Currently, there are no suitable markers for monitoring disease progression in SRC. The changes in monocyte *EGRI* expression observed over the clinical course of SRC implicate its potential as a biomarker for predicting progression of the disease.

Spatial transcriptomic analysis of SRC kidney tissue revealed an accumulation of the Mac1 macrophage subset, characterized by activated fibrosis-related pathways, alongside PTB. The module score analysis suggested that the Mac1 subset corresponds to the THBS1_Mac subset in the scRNA-seq experiment. These macrophages share the activation states of NF- κ B-related transcription factors, supporting our hypothesis that the CD14_EGRI, THBS1_Mac, and Mac1 subsets are on a common differentiation pathway. Moreover, the injured PTB subset was abundant in the kidney tissue of SRC patients, with highly expressed genes such as *VCAM1*^{47,48}, *MMP7*⁴⁹, and *PIGR*⁵⁰. These genes are recognized as markers of injured tubular cells and serve as hub genes that link AKI to renal fibrosis by promoting immune cell accumulation. Thus, the formation of niches by pathogenic macrophages and injured PTB may contribute to severe renal dysfunction in SRC by promoting excessive fibrosis.

In patients with SSc-ILD, the CD8_TEM_T2ISG subset was enriched in peripheral blood. A subset with a similar gene expression profile, CD8_T_T2ISG, was also enriched in lung tissue. The CD8_TEM_T2ISG subset highly expressed IFN- γ response genes and showed increased expression of chemokine receptors such as *CXCR3* and *CCR5*. *CXCL4*, a ligand for *CXCR3*, is implicated in T cell migration⁵¹. Furthermore, serum levels of *CXCL4* are elevated in patients with SSc and correlate with lung fibrosis⁵². *CCR5* also facilitates the migration of CD8⁺ T cells through its interaction with ligands such as *CCL3* and *CCL4*³⁴. These data suggest that the CD8_TEM_T2ISG subset have high tissue migratory capacity, which is at least partly influenced by the *CXCL4*–*CXCR3* or *CCL3*/*CCL4*–*CCR5* axis during these cells' migration into lung tissue. This hypothesis is further supported by the fact that the gene expression profiles of enriched CD8⁺ T cells in patients with SSc-ILD are similar between the peripheral blood and lung tissue. In addition, CD8⁺ T cells expressing high levels of *CXCR3* contribute to lung injury in SSc-ILD⁵³, implicating another role of the CD8_TEM_T2ISG subset in lung damage. The reduction in GZMH- and GZMB-expressing CD8⁺ T cells in the peripheral blood and lung tissue of patients with SSc-ILD suggests that lung involvement extends beyond inflammation, which may be associated with the minimal efficacy of glucocorticoid therapy in SSc-ILD⁴. Consistent with these findings, a high prevalence of the CD8_TEM_T2ISG subset was observed in patients with SSc-ILD who were positive for ATA or ARA, both of which are autoantibodies linked to a poorer pulmonary prognosis³⁵.

In conclusion, scRNA-seq identified key cellular subpopulations associated with specific organ manifestations. Specifically, the CD14_EGRI subset in SRC and the CD8_TEM_T2ISG subset in SSc-ILD may play crucial roles in the respective organ damage. Time-series analysis and tissue perturbation studies of these cellular subsets in larger cohorts will further clarify the pathogenesis of SSc and demonstrate the potential of these subsets as therapeutic targets.

Methods

Study participants

Samples for scRNA-seq were collected from study participants who provided informed consent, in accordance with the Declaration of Helsinki and with approval from the ethics review board of the Graduate School of Medicine, Osaka University, Japan (No. 855). Samples for spatial transcriptomic analysis were collected in accordance with protocols approved by the ethics committee of Osaka University Hospital, Japan (No.11122 and No.18370). Consent was obtained to publish clinical information potentially identifying individuals, such as age, sex, the name of the medical center, and the diagnosis. The sex of each study participant was determined based on self-report. Study participants received no compensation. Patients were diagnosed with SSc according to the 2013 ACR/EULAR classification criteria¹¹. The diagnosis was confirmed by at least two rheumatologists.

Patient profiles for scRNA-seq

Twenty-one patients diagnosed with SSc (sixteen females and five males) and six healthy donors (four females and two males) were recruited for scRNA-seq. The patients had never received immunosuppressive therapy. All patients were either admitted to or visited Osaka University Hospital, where they underwent a comprehensive assessment to rule out infectious diseases, neoplastic lesions, and any overlap of other systemic autoimmune diseases, before the 2013 ACR/EULAR classification criteria were applied. The presence of ILD was diagnosed based on clinical symptoms, computed tomography (CT) scan, and pulmonary function tests. CT images were reviewed by at least two rheumatologists and one radiologist. SRC was diagnosed according to the UK Scleroderma Study Group guideline⁵⁴. Samples were included in the SRC group if they were collected from SSc patients within three months of SRC onset or at any time after the onset. Blood pressure and mRSS were measured on the day of blood collection. Four patients with LN and four patients with MPA were recruited for some analyses. Systemic lupus erythematosus was diagnosed according to the 2019 EULAR/ACR classification criteria⁵⁵, and LN was confirmed by renal biopsy⁵⁶. MPA was diagnosed according to the 2022 ACR/EULAR classification criteria⁵⁷, and renal involvement was defined by a positive Birmingham vasculitis activity score version 3 'renal involvement' component⁵⁸. Datasets from three MPA patients were previously deposited in the Genomic Expression Archive (GEA) with accession code E-GEAD-635. The patients with LN or MPA had never received immunosuppressive therapy. None of the patients were complicated by infectious diseases, malignancies, or other systemic autoimmune diseases.

Patient profiles for spatial transcriptomic analysis

Three SRC patients (two females and one male) and three healthy donors (two females and one male) were recruited for spatial transcriptomic analysis of kidney tissue. Kidney tissue samples from SRC patients were biopsied within one month of the first episode of SRC. Kidney tissue samples from healthy donors were obtained from donor kidneys at the time of kidney transplantation. The tissues were fixed in 10% neutral buffered formalin and then embedded in paraffin for further analysis. Two of the three patients had received or were receiving immunosuppressive medications, such as prednisolone, cyclophosphamide, or mycophenolate mofetil. Patients were diagnosed with SSc according to the 2013 ACR/EULAR classification criteria¹¹. SRC was diagnosed according to the UK Scleroderma Study Group guideline⁵⁴. The patients underwent a comprehensive assessment to rule out infectious diseases, neoplastic lesions, and any overlap of other systemic autoimmune diseases.

PBMCs preparation

Twenty milliliters of whole blood was collected into Na-heparin blood collection tubes (Terumo, Cat. No. VP-H070K). PBMCs were isolated

using Leucosep (Greiner, Cat. No. 22788-013), then washed and resuspended in Cellbanker 1plus (ZENOAQ, Cat. No. CB023) to a concentration of 1.0×10^7 cells/mL before storage at -150°C .

Single-cell library construction for CITE-seq

Single-cell library construction for CITE-seq was performed using the DNA-barcoded antibodies described in Supplementary Table 11, which were obtained from BioLegend. The antibodies were diluted according to the manufacturer's protocol. Thawed PBMCs were treated with those antibodies, and single-cell suspensions were processed using the 10 $\times$ Genomics Chromium Controller (10 $\times$ Genomics). The libraries were constructed according to the protocol outlined in the user guide of Chromium Single Cell 5' Reagent Kits v2 (Dual Index, Cat. No. PN-1000263) (10 $\times$ Genomics). Briefly, up to 10,000 labeled live cells per sample were individually loaded into the 10 $\times$ Genomics platform without sample mixing to generate a barcoded cDNA library for individual cells. Data quality control was conducted using a Bioanalyzer system (Agilent). Individual libraries were pooled and sequenced on the HiSeq 2500 or Novaseq 6000 platform (Illumina) to analyze gene and surface protein expression. Sequence information of CITE-seq is summarized in Supplementary Table 12.

Reference-based cell annotation and analysis of CITE-seq data

Raw FASTQ files were aligned to the GRCh38 reference genome using CellRanger (version 6.0.6). Filtered HDF5 feature-barcode matrix files were generated using the CellRanger count to create a Seurat object. Data quality control, scaling, transformation, clustering, dimensionality reduction, differential expression analysis, and visualization were performed using the Seurat R package (V4.3.0). Cells with nFeature_RNA values less than 200 or greater than 5,000, or mitochondrial read percentages exceeding 20%, were removed. Data normalization and scaling were performed using the SCTransform function. Except for the time course analysis, only the initial samples collected during the clinical course were integrated. For the comparison between SRC, LN, and healthy donors, samples from patients diagnosed with SRC or LN were integrated with those from healthy donors. For the time course analysis of the SRC patient, samples collected at three different time points from the same individual were included. Each cell subpopulation was identified through two rounds of clustering. Initially, reference-based integration was applied to the query dataset using the public CITE-seq dataset of 211,000 human PBMCs as the reference¹². Shared anchors between the query and reference datasets were identified using the FindTransferAnchors function with precomputed supervised PCA transformation for SCT-normalized data to correct for batch effects. The MapQuery function was used to transfer cell-type labels and protein data from the reference to the query datasets. Platelets and erythrocytes were excluded from the analysis. To identify subpopulations within each cell type, a second round of clustering was conducted on specific populations: monocytes (CD14 Mono, CD16 Mono, cDC1, and cDC2), CD8⁺ T cells (CD8 Naive, CD8 TCM, and CD8 TEM), CD4⁺ T cells (CD4 Naive, CD4 TCM, CD4 TEM, Treg, and CD4 CTL), B cells (B naive, B intermediate, and B memory), NK cells (NK and NK_CD56bright), and pDCs (pDC). The RunUMAP function was used to perform UMAP dimensional reduction with 30 precomputed spca dimensions. A nearest-neighbor graph using the 30 dimensions of the supervised PCA reduction was computed using the FindNeighbors function, followed by the FindClusters function for cell clustering. The resultant UMAP was visualized using the DimPlot function. Each cluster was manually annotated using gene expression and surface protein data. Small cell clusters with apparently high expression of any other cell type-specific protein marker were excluded as doublet clusters. For example, we used CD14 as a monocyte marker, CD3, CD4, and CD8 as T cell markers, CD19 and CD20 as B cell markers, and CD56 as a NK cell marker. The counts of antibodies in each cell were normalized using centered log ratio transformation. A cell cluster was considered a

doublet cluster if the average expression of any other cell type-specific protein marker was approximately more than twice as high as in other clusters. Gene expression information was also used to detect doublet clusters. For example, PF4 and PPBP were used to detect doublets with platelets, while HBB and HBA1 were used to detect doublets with red blood cells.

PCA using the cell composition of PBMCs

PCA was performed using the PCA function in the FactoMineR package, with data scaled to unit variance. First, the proportion of each cell subpopulation defined in the subset analysis was calculated, along with other minor cell populations, all relative to the total PBMCs. The proportion of each cell subset was included as a variable in the PCA, except for populations with fewer than 100 total cells. Vectors representing each subpopulation were shown on a plane defined by the first two principal components (PC1 and PC2) using the fviz_pca_var function. Each study participant was represented as a single plot on the same plane according to their respective PC1 and PC2 values.

Differential abundance analysis

Differential abundance analysis was performed on PBMCs, monocytes, CD4⁺ T cells, and CD8⁺ T cells collected in this study, as well as lung CD8⁺ T cells derived from public data³⁶, using the miloR (version 3.15) package. These analyses were performed to detect groups of cells that are differentially abundant in different conditions by modeling the number of cells within the neighborhoods of a k-nearest neighbor (KNN) graph¹⁷. To adjust for cell number, up to 1000 cells were randomly selected from each sample for differential abundance analysis. The buildGraph function was first used to construct a KNN graph on the basis of precomputed supervised PCA with $k=5$, using 30 principal components. Using the makeNhoods function, cells were then grouped into neighborhoods according to their connectivity over the KNN graph. To test for differential abundance, Milo fitted a negative binomial generalized linear model to the counts for each neighborhood using TMM normalization. Age was used as a covariate in the testNhoods function. The log₂-fold change in cell numbers between two conditions in each neighborhood was calculated.

Differential gene expression analysis and gene set enrichment analysis

In specific cell subsets, DEGs were identified using the FindMarkers function. For the characterization of DEGs, gene set enrichment analysis was conducted on genes exhibiting a log₂-fold change greater than 0.5, using the Enrichr, a web-based tool for analyzing gene sets⁵⁹. The MSigDB Hallmark 2020³⁷ or BioPlanet 2019⁶⁰ were employed as the dataset, and adjusted P-values for each pathway were calculated by the Benjamini–Hochberg method. TRRUST Transcription Factors 2019²⁹ was used to determine the adjusted P-values for each transcription factor-related pathway. Pathways not related to humans were excluded from the analysis.

Preparation of kidney cells and whole-blood leukocytes for Abseq

Renal tissue was obtained by needle biopsy from one patient (patient ID: SSc-19) at the onset of the first episode of SRC. The tissue was immediately washed with phosphate-buffered saline and processed into single-cell suspension using the Tumor Dissociation Kit (human, Miltenyi Biotec). Enzyme R was excluded to preserve the cell-surface epitope. The sample was cut into small pieces with a maximal dimension of approximately 2 mm, combined with the enzyme mixture, and shaken at 37°C in a water bath at 160 rpm for 45 min. The sample was treated with DNase for 5 min at room temperature. No steps were taken to lyse red blood cells or remove dead cells.

Two milliliters of Whole blood was collected from the same patient in an EDTA-2Na blood collection tube (Terumo, Cat. No. VP-Na052K), and the leukocytes were isolated using Polymorphprep (Serumwerk, Cat. No. 1895). The sample was washed and resuspended with Cellbanker 1plus (ZENOAQ, Cat. No. CB023) to a concentration of 2.0×10^6 cell/mL prior to scRNA-seq experiments.

Single-cell library construction for Abseq

Renal cells and total peripheral leukocytes were processed separately through the BD Rhapsody Express System (BD Biosciences, Catalog No. 665915). Thirty-five DNA-barcoded antibodies (detailed in Supplementary Table 13) were obtained from BD Biosciences. These antibodies were diluted according to the manufacturer's protocol and added to each cell suspension. Abseq libraries were constructed using WTA Reagent Kits (BD Biosciences, Cat. No. 665915). Approximately 10,000 peripheral leukocytes were loaded into the BD Rhapsody platform. All obtained renal cells were processed in an identical manner. Data quality control was performed using the Bioanalyzer (Agilent). The libraries were pooled for sequencing on the HiSeq 2500 or Novaseq 6000 platform (Illumina) to analyze gene and surface protein expression. Sequence information of Abseq is summarized in Supplementary Table 14.

Manual cell annotation and analysis of Abseq data

Raw FASTQ files were aligned to the GRCh38 reference genome using the BD Rhapsody Sequence Analysis Pipeline (version 1.12). Distribution-based error correction-adjusted molecule counts were presented as gene expression data tables to establish a Seurat object. The Seurat R package (V4.3.0) was used for data quality control, scaling, transformation, clustering, dimensionality reduction, differential expression analysis, and visualization. A total of 16,081 cells were selected for further analysis based on unique molecular identifiers for each cell and percentages of mitochondrial reads. Peripheral leukocytes with nFeature_RNA values less than 200 or greater than 5000, or mitochondrial read percentages exceeding 20%, were excluded from the analysis. Similarly, kidney cells were removed if they had nFeature_RNA values less than 200 or greater than 5000, or mitochondrial read percentages exceeding 40%. The Abseq data from peripheral blood and kidney samples were normalized and scaled using the SCTransform function. Data integration and batch effect correction across samples were performed by computing a set of anchors with the FindIntegrationAnchors function using reciprocal PCA as the dimension reduction method, followed by applying the IntegrateData function to the anchor set. The RunUMAP function was used for UMAP dimensional reduction with 30 precomputed PCA dimensions. A nearest-neighbor graph was then constructed using the 30 PCA dimensions with the FindNeighbors function, followed by clustering using the FindClusters function. The UMAP was visualized using the DimPlot function. Each cluster was manually annotated using gene expression and surface protein data. Platelets and erythrocytes were removed from the analysis. Doublets were removed using the method described in the section "Reference-based cell annotation and analysis of CITE-seq data".

Pseudotime analysis

The pseudotime of each cell and the cell trajectory were calculated using the Monocle 3 package²⁷. The previously annotated Seurat object was imported into Monocle 3. Using the learn_graph function, we fitted a principal graph and plot the graph through the UMAP coordinates. The order_cells function was employed to calculate the pseudotime for each cell, with the root node set to the immature monocyte population characterized by high expression of S100A12 and low expression of HLA-DRB1²⁸ (Supplementary Fig. 16). The cell trajectory was visualized using the plot_cells function.

Public scRNA-seq data analysis

The scRNA-seq data of PBMCs derived from SSc patients and healthy donors were obtained from the Gene Expression Omnibus (GEO) under accession number GSE195452. To match our recruitment criteria for scRNA-seq, patients receiving immunosuppressants, including mycophenolate mofetil, methotrexate, azathioprine, or leflunomide, were excluded. Patients with unclear medication information were also excluded. Reference-based integration was applied to the query dataset using the public scRNA-seq data of PBMCs¹². FindTransferAnchors function was employed to find anchors between the reference and the query dataset. The MapQuery function was used to transfer cell-type labels from the reference to the query datasets. Subset analysis was conducted on monocytes (CD14 Mono, CD16 Mono, cDC1, and cDC2).

The scRNA-seq data of lung cells derived from SSc patients with advanced ILD and healthy donors³⁶ were obtained from the GEO under accession numbers GSE128169 and GSE128033. We used only lung tissue samples, excluding those labeled with hashtags designated for sample multiplexing. The Seurat object was subsequently created using the CreateSeuratObject function. Lung cells with nFeature_RNA values less than 200 were excluded from the analysis. Data were normalized and scaled using the SCTransform function, followed by data integration using reciprocal PCA. The RunUMAP function was used for UMAP dimensional reduction with 30 precomputed PCA dimensions. A nearest-neighbor graph using the 30 dimensions of the PCA reduction was computed using the FindNeighbors function, followed by clustering using the FindClusters function. The UMAP was visualized using the DimPlot function. Each cluster was manually annotated using gene expression data.

Module scoring

Module scores for each cell were calculated using the AddModuleScore function. For the analysis of public PBMC scRNA-seq data, module scores were calculated using highly expressed genes (fold change >0.5) of the CD14_EGRI subset to evaluate the similarity of CD14⁺ monocytes to the CD14_EGRI subset in each disease group. In the analysis of spatial transcriptomic data, module scores were calculated using highly expressed genes (fold change >0.5) of the THBS1_Mac subset to evaluate the similarity of kidney myeloid cell subsets to the THBS1_Mac subset. In the analysis of public lung scRNA-seq data, to assess the expression of type II ISGs in lung CD8⁺ T cells, module scores were calculated using the "Interferon Gamma Response" gene set from MSigDB Hallmark 2020. To evaluate the similarity of each lung CD8⁺ T cell subset to the peripheral CD8⁺ T cell subpopulation, CD8_TEM_T2ISG, module scores were calculated using highly expressed genes (fold change >0.5) of CD8_TEM_T2ISG. Gene scores for each subset were visualized using the Dotplot function according to cell-based scores.

CosMX SMI instrument run

Five-micron FFPE kidney sections were mounted on Premium Superfrost Plus Microscope Slides (VWR, cat. No. 48311-703) and analyzed using the CosMX spatial molecular imager (SMI) system⁶¹ according to the manufacturer's protocol, CosMX SMI Manual Slide Preparation for RNA Assays (MAN-10184-04, NanoString). CosMx Human Universal Cell Characterization RNA Panel (1,000-plex, NanoString) was used for in situ hybridization. CosMX Human Universal Cell Segmentation Kit (RNA), IO PanCK/CD45 Kit (RNA), and Cytokeratin 8/18 Morphology Marker (RNA) were used for cell segmentation. RNA target readout on the CosMx SMI instrument was performed as described in CosMX SMI Instrument User Manual (Man-10161-07, NanoString).

CosMX data analysis

CosMX data were processed on the AtoMX Spatial Informatics Platform. Cell boundaries were determined using morphological markers

described in the CosMX SMI instrument run subsection. The cell segmentation was performed in AtoMX. The transcripts were assigned to individual cells to obtain count matrix^{61,62}. Data quality control was performed using the following parameters: for negative probe quality control, outlier P -value cutoff = 0.01. For cell quality control, minimal counts per cell = 100, proportion of negative counts = 0.1, count distribution = 1, area outlier = 0.01. For FOV quality control, method was mean, FOV count cutoff = 100. For target quality control, negative control probe quantile cutoff = 0.5, detection = 0.01. After quality control, count matrix flat csv files and cell metadata flat csv files were exported.

Next, the Seurat object was created using the CreateSeuratObject function. Cells with the qcCellsFlagged metadata TRUE were excluded. To compensate for the influence of variations in cross-sectional area on the gene expression profile of each cell, the RNA count for each cell was divided by the ratio of the area of the cell to the median area of all cells in the sample. The Seurat objects were merged using the merge function and normalized using the NormalizeData function. The data were scaled using the ScaleData function, with all 1000 genes included as features. PCA was conducted using the RunPCA function, with npcs set to 30. The data from multiple samples were integrated using IntegrateLayers function, with the method set to CCAIntegration. A nearest-neighbor graph was constructed using the 30 PCA dimensions with the FindNeighbors function, followed by clustering using the FindClusters function. The RunUMAP function was used for UMAP dimensional reduction. Each cluster was manually annotated using the highly expressed genes (Supplementary Fig. 19b).

The analysis of cellular proximity⁶¹ was conducted as follows. For each cell, the neighboring cells are identified using the frNN function, with eps set to 30 μ m. These neighboring cells were aggregated by cell groups. The CreateSeuratObject function was used to create a seurat object that contains information about the type and number of neighboring cells for each cell. The RunPCA, FindNeighbors, and FindClusters functions were used to identify cell groups defined based on the similarity of neighboring cells. The Heatmap function was used to visualize abundance of each cell subset in each cell group.

Statistical Analysis

To compare the cell composition between two groups, two-sided Kolmogorov–Smirnov tests with adjustments for multiple testing using the Bonferroni correction (Fig. 1c, d), or two-sided Mann–Whitney U tests with adjustments for multiple testing using the Bonferroni correction (Fig. 6g, j) were performed. To compare the cell composition among three or more groups, Kruskal–Wallis test was performed (Figs. 3e, 4h and Supplementary Fig. 2). When the result of Kruskal–Wallis test was significant ($P < 0.05$), two-sided Dunn's multiple comparison test was performed (Fig. 3e, and Fig. 4h). In gene set enrichment analysis, P -values for each pathway were calculated using Fisher's exact test, with adjustments for multiple testing using the Benjamini–Hochberg method (Figs. 4a, g, 5f, g, 6e, and Supplementary Fig. 15). Spearman's correlation coefficients (r) and P -values were calculated with a two-sided test, between the percentage of the CD14⁺ EGR1 subset to the total monocytes and mRSS or sBP (Fig. 4j). In Figs. 2a–d, 3c, d, 6c, d, h–i, and Supplementary Figs. 9a, b, 10a–d, 22b–e, 23a, b, milo differential abundance tests were performed. The spatial false discovery rate (FDR) was calculated for each neighbor, with adjustments made to account for multiple testing of neighbors within a comparison. Additional spatial FDR adjustment was not applied for repeated differential abundance analyses across multiple groups. Neighbors with spatial FDR < 0.2 were considered significant.

Reporting summary

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

Data availability

The count matrix of the scRNA-seq data generated in this study has been deposited in the GEA under accession code [E-GEAD-872](#). The fast-q files generated in this study have been deposited in the DNA Data Bank of Japan Sequence Read Archive under accession code [DRA019274](#). The CosMX spatial transcriptome data generated in this study are provided in the Source Data file. The GRCh38 reference genome data used in this study are available in the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/). The reference dataset used for cell type annotation of PBMCs in the scRNA-seq data analysis are available at the following website (https://satijalab.org/seurat/articles/multimodal_reference_mapping.html). The publicly available scRNA-seq data of PBMCs used for validation in this study are available in the GEO under accession code [GSE195452](#). The publicly available scRNA-seq data of lung cells used in this study are available in the GEO under accession codes GSE128169 and GSE128033 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128169>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128033>). Source Data are provided with this paper. All data are included in the Supplementary Information or available from the authors, as are unique reagents used in this Article. The raw numbers for charts and graphs are available in the Source Data file whenever possible. Source data are provided with this paper.

Code availability

Experimental protocols, data analysis pipelines, analysis steps, functions, and parameters used are described in the Methods section. Custom code used in the paper is available at GitHub (https://github.com/ShimagamiH/ShimagamiH_SSc_scRNAseq) and has been archived on Zenodo (<https://doi.org/10.5281/zenodo.15049738>).

References

1. Cho, J. H. & Feldman, M. Heterogeneity of autoimmune diseases: pathophysiologic insights from genetics and implications for new therapies. *Nat. Med.* **21**, 730–738 (2015).
2. Volkman, E. R., Andreasson, K. & Smith, V. Systemic sclerosis. *Lancet* **401**, 304–318 (2023).
3. Rubio-Rivas, M., Royo, C., Simeon, C. P., Corbella, X. & Fonollosa, V. Mortality and survival in systemic sclerosis: systematic review and meta-analysis. *Semin. Arthritis Rheum.* **44**, 208–219 (2014).
4. Pope, J. E. et al. State-of-the-art evidence in the treatment of systemic sclerosis. *Nat. Rev. Rheumatol.* **19**, 212–226 (2023).
5. Apostolidis, S. A. et al. Single cell RNA sequencing identifies HSPG2 and APLNR as markers of endothelial cell injury in systemic sclerosis skin. *Front. Immunol.* **9**, 2191 (2018).
6. Gur, C. et al. LGR5 expressing skin fibroblasts define a major cellular hub perturbed in scleroderma. *Cell* **185**, 1373–1388 e1320 (2022).
7. Tabib, T. et al. Myofibroblast transcriptome indicates SFRP2(hi) fibroblast progenitors in systemic sclerosis skin. *Nat. Commun.* **12**, 4384 (2021).
8. Xue, D. et al. Expansion of fcggamma receptor IIIa-positive macrophages, ficolin 1-positive monocyte-derived dendritic cells, and plasmacytoid dendritic cells associated with severe skin disease in systemic sclerosis. *Arthritis Rheumatol.* **74**, 329–341 (2022).
9. Gaydosik, A. M. et al. Single-cell transcriptome analysis identifies skin-specific T-cell responses in systemic sclerosis. *Ann. Rheum. Dis.* **80**, 1453–1460 (2021).
10. Nishide, M., Shimagami, H. & Kumanogoh, A. Single-cell analysis in rheumatic and allergic diseases: insights for clinical practice. *Nat. Rev. Immunol.* **24**, 781–797 (2024).
11. van den Hoogen, F. et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis Rheum.* **65**, 2737–2747 (2013).

12. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
13. Nehar-Belaid, D. et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020).
14. Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
15. Nishide, M. et al. Single-cell multi-omics analysis identifies two distinct phenotypes of newly-onset microscopic polyangiitis. *Nat. Commun.* **14**, 5789 (2023).
16. Wu, X. et al. Single-cell sequencing of immune cell heterogeneity in IgG4-related disease. *Front. Immunol.* **13**, 904288 (2022).
17. 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).
18. Pestana-Fernandez, M., et al. The incidence rate of pulmonary arterial hypertension and scleroderma renal crisis in systemic sclerosis patients with digital ulcers on endothelin antagonist receptors (ERAs) and phosphodiesterase-5 inhibitors (PDE5i). *Rheumatology* **60**, 872–880 (2021).
19. Lee, S. J. et al. TLR4-mediated expression of Mac-1 in monocytes plays a pivotal role in monocyte adhesion to vascular endothelium. *PLoS ONE* **9**, e104588 (2014).
20. Zhen, A. et al. CD4 ligation on human blood monocytes triggers macrophage differentiation and enhances HIV infection. *J. Virol.* **88**, 9934–9946 (2014).
21. Musso, T. et al. CD38 expression and functional activities are up-regulated by IFN-gamma on human monocytes and monocytic cell lines. *J. Leukoc. Biol.* **69**, 605–612 (2001).
22. Li, C. & Hua, K. Dissecting the single-cell transcriptome network of immune environment underlying cervical premalignant lesion, cervical cancer and metastatic lymph nodes. *Front. Immunol.* **13**, 897366 (2022).
23. Zimmerman, K. A. et al. Single-cell RNA sequencing identifies candidate renal resident macrophage gene expression signatures across species. *J. Am. Soc. Nephrol.* **30**, 767–781 (2019).
24. Kaur, S. et al. Functions of thrombospondin-1 in the tumor micro-environment. *Int. J. Mol. Sci.* **22**, 4570 (2021).
25. Wang, H., Wu, J., Ma, L., Bai, Y. & Liu, J. The role of interleukin -1 family in fibrotic diseases. *Cytokine* **165**, 156161 (2023).
26. Hong, Q. et al. LRG1 promotes diabetic kidney disease progression by enhancing TGF-beta-induced angiogenesis. *J. Am. Soc. Nephrol.* **30**, 546–562 (2019).
27. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
28. Chen, S. T. et al. A shift in lung macrophage composition is associated with COVID-19 severity and recovery. *Sci. Transl. Med.* **14**, eabn5168 (2022).
29. Han, H. et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **46**, D380–D386 (2018).
30. Sun, S. C. The non-canonical NF-kappaB pathway in immunity and inflammation. *Nat. Rev. Immunol.* **17**, 545–558 (2017).
31. Woodworth, T. G., Suliman, Y. A., Li, W., Furst, D. E. & Clements, P. Scleroderma renal crisis and renal involvement in systemic sclerosis. *Nat. Rev. Nephrol.* **12**, 678–691 (2016).
32. Chevalier, R. L. The proximal tubule is the primary target of injury and progression of kidney disease: role of the glomerulotubular junction. *Am. J. Physiol. Ren. Physiol.* **311**, F145–F161 (2016).
33. Colditz, I. G. & Watson, D. L. The effect of cytokines and chemotactic agonists on the migration of T lymphocytes into skin. *Immunology* **76**, 272–278 (1992).
34. Nagarsheth, N., Wicha, M. S. & Zou, W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nat. Rev. Immunol.* **17**, 559–572 (2017).
35. Hoffmann-Vold, A. M. et al. Cardiopulmonary disease development in Anti-RNA polymerase III-positive systemic sclerosis: comparative analyses from an unselected, prospective patient cohort. *J. Rheumatol.* **44**, 459–465 (2017).
36. Valenzi, E. et al. Single-cell analysis reveals fibroblast heterogeneity and myofibroblasts in systemic sclerosis-associated interstitial lung disease. *Ann. Rheum. Dis.* **78**, 1379–1387 (2019).
37. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
38. Nguyen, H. Q., Hoffman-Liebermann, B. & Liebermann, D. A. The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage. *Cell* **72**, 197–209 (1993).
39. Chen, S. J. et al. The early-immediate gene EGR-1 is induced by transforming growth factor-beta and mediates stimulation of collagen gene expression. *J. Biol. Chem.* **281**, 21183–21197 (2006).
40. Bhattacharyya, S. et al. Smad-independent transforming growth factor-beta regulation of early growth response-1 and sustained expression in fibrosis: implications for scleroderma. *Am. J. Pathol.* **173**, 1085–1099 (2008).
41. Qiao, Y. et al. Synergistic activation of inflammatory cytokine genes by interferon-gamma-induced chromatin remodeling and toll-like receptor signaling. *Immunity* **39**, 454–469 (2013).
42. Meng, Y., Chen, C., Tian, C., Du, J. & Li, H. H. Angiotensin II-induced Egr-1 expression is suppressed by peroxisome proliferator-activated receptor-gamma ligand 15d-PGJ(2) in macrophages. *Cell. Physiol. Biochem.* **35**, 689–698 (2015).
43. Kranzhofer, R., Browatzki, M., Schmidt, J. & Kubler, W. Angiotensin II activates the proinflammatory transcription factor nuclear factor-kappaB in human monocytes. *Biochem. Biophys. Res. Commun.* **257**, 826–828 (1999).
44. Alaaeldin, R. et al. Azilsartan modulates HMGB1/NF-kappaB/p38/ERK1/2/JNK and apoptosis pathways during renal ischemia reperfusion injury. *Cells* **12**, 185 (2023).
45. Yan, S. F. et al. Egr-1, a master switch coordinating upregulation of divergent gene families underlying ischemic stress. *Nat. Med.* **6**, 1355–1361 (2000).
46. Wang, W., Ren, X., Chen, X., Hong, Q. & Cai, G. Integrin beta1-rich extracellular vesicles of kidney recruit Fn1+ macrophages to aggravate ischemia-reperfusion-induced inflammation. *JCI Insight* **9**, e169885 (2024).
47. Gerhardt, L. M. S., Liu, J., Koppitch, K., Cippa, P. E. & McMahon, A. P. Single-nuclear transcriptomics reveals diversity of proximal tubule cell states in a dynamic response to acute kidney injury. *Proc. Natl. Acad. Sci. USA* **118**, e2026684118 (2021).
48. Melchinger, I. et al. VCAM-1 mediates proximal tubule-immune cell cross talk in failed tubule recovery during AKI-to-CKD transition. *Am. J. Physiol. Ren. Physiol.* **327**, F610–F622 (2024).
49. Klocke, J. et al. Urinary single-cell sequencing captures kidney injury and repair processes in human acute kidney injury. *Kidney Int.* **102**, 1359–1370 (2022).
50. Krawczyk, K. M. et al. Localization and regulation of polymeric Ig receptor in healthy and diseased human kidney. *Am. J. Pathol.* **189**, 1933–1944 (2019).
51. Mueller, A. et al. CXCL4-induced migration of activated T lymphocytes is mediated by the chemokine receptor CXCR3. *J. Leukoc. Biol.* **83**, 875–882 (2008).
52. van Bon, L. et al. Proteome-wide analysis and CXCL4 as a biomarker in systemic sclerosis. *N. Engl. J. Med.* **370**, 433–443 (2014).
53. Fujii, H., Hasegawa, M., Takehara, K., Mukaida, N. & Sato, S. Abnormal expression of intracellular cytokines and chemokine

- receptors in peripheral blood T lymphocytes from patients with systemic sclerosis. *Clin. Exp. Immunol.* **130**, 548–556 (2002).
54. Lynch, B. M. et al. UK Scleroderma Study Group (UKSSG) guidelines on the diagnosis and management of scleroderma renal crisis. *Clin. Exp. Rheumatol.* **34**, 106–109 (2016).
 55. Aringer, M. et al. 2019 European league against rheumatism/ American college of rheumatology classification criteria for systemic lupus erythematosus. *Arthritis Rheumatol.* **71**, 1400–1412 (2019).
 56. Weening, J. J. et al. The classification of glomerulonephritis in systemic lupus erythematosus revisited. *J. Am. Soc. Nephrol.* **15**, 241–250 (2004).
 57. Suppiah, R. et al. 2022 American college of rheumatology/European alliance of associations for rheumatology classification criteria for microscopic polyangiitis. *Arthritis Rheumatol.* **74**, 400–406 (2022).
 58. Mukhtyar, C. et al. Modification and validation of the Birmingham vasculitis activity score (version 3). *Ann. Rheum. Dis.* **68**, 1827–1832 (2009).
 59. Xie, Z. et al. Gene set knowledge discovery with enrichr. *Curr. Protoc.* **1**, e90 (2021).
 60. Huang, R. et al. The NCATS BioPlanet - An integrated platform for exploring the universe of cellular signaling pathways for toxicology, systems biology, and chemical genomics. *Front. Pharmacol.* **10**, 445 (2019).
 61. He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* **40**, 1794–1806 (2022).
 62. Stankey, C. T. et al. A disease-associated gene desert directs macrophage inflammation through ETS2. *Nature* **630**, 447–456 (2024).

Acknowledgements

We thank Maho Omiya, Riri Furuta, Takako Oya, Natsumi Doi for their excellent technical assistance. We also thank Kenichi Mogi for the contributions to figure illustrations. Figs. 1a and 7 were created using images from Biorender.com. This work was financially supported by research grants from the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP24K11596 to M.Ni., JP 22H00476 to Y.O., and JP18H05282 to A.K.), Japan Science and Technology Agency (JST) FOREST Program (JPMJFR235B to M.Ni.), JST SPRING (JPMJSP2138 to H.S.), The JCR Grant for Promoting Research for FRONTIER (to M.Ni.), Takeda Science Foundation (to M.Ni. and Y.O.), JSPS Core-to-Core Program (JP223fa627002 to A.K.), Japan Agency for Medical Research and Development (AMED) (223fa627002h0001 to A.K., JP24km0405217, JP24ek0109594, JP24ek0410113, JP24kk0305022, JP243fa627002, JP243fa627010, JP243fa627011, JP24zf0127008, JP24tm0524002, JP24wm0625504, JP24gm1810011 to Y.O.), JST Moonshot R&D (JPMJMS2021 and JPMJMS2024 to Y.O.), Ono Pharmaceutical Foundation for Oncology, Immunology, and Neurology (to Y.O.), and Japan Agency for Medical Research and Development–Core Research for Evolutional Science and Technology (AMED–CREST) (22gm1810003h0001 to A.K.). This research was conducted as part of the ALL-Osaka U Research in “The Nippon Foundation - Osaka University Project for Infectious Disease Prevention”.

Author contributions

H.S., M.Ni., K.N., and H.M. designed the project. H.S., M.Ni., Y.Kat., T.K., K.T., E.I., M.Nai., S.Kaw., D.N., K.Ma., H.K., K.Mi., N.I., J.F., S.H., and M.Nar. recruited study participants and assessed clinical data. H.S., K.N., and H.M. conducted experiments. H.S. and K.N. carried out data analysis. K.N., H.M., S.M., R.E., R.O., Y.O., and K.H. contributed to preparation of materials and provided advice on project planning and data interpretation. T.N.H., K.I., A.T., M.M., Y.I., S.N., Y.Kak., and N.N. performed renal biopsies and examined the pathology. S.Kat., H.H., and T.T. supported spatial transcriptomic analysis. A.K. provided funding for the study and supervised the project. H.S. and M.Ni. wrote the manuscript. All authors contributed to the discussion of the results and approved the final version of the manuscript.

Competing interests

A.K. has received grant support from Chugai Pharmaceutical Co, Ltd. K.N., H.M., S.M., R.O., K.H. are employed by Chugai Pharmaceutical Co, Ltd. and K.N., H.M., S.M., R.O., K.H. also hold stocks in the company. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-60034-7>.

Correspondence and requests for materials should be addressed to Atsushi Kumanogoh or Masayuki Nishide.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. 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

¹Department of Respiratory Medicine and Clinical Immunology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ²Department of Immunopathology, World Premier International Research Center Initiative (WPI), Immunology Frontier Research Center (IFReC), Osaka University, Suita, Osaka, Japan. ³Department of Advanced Clinical and Translational Immunology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ⁴Joint Research Chair of Innovative Drug Discovery in Immunology, World Premier International Research Center Initiative (WPI), Immunology Frontier Research Center (IFReC), Osaka University, Suita, Osaka, Japan. ⁵Research Division, Chugai Pharmaceutical Co. Ltd, Yokohama, Kanagawa, Japan. ⁶Department of Statistical Genetics, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ⁷Laboratory for Systems Genetics, RIKEN Center for Integrative

Medical Sciences, Yokohama, Kanagawa, Japan. ⁸Department of Nephrology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ⁹Department of Gastroenterology and Hepatology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ¹⁰Department of Urology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan. ¹¹Nippon Life Hospital, Osaka, Japan. ¹²Osaka International Medical & Science Center, Osaka, Japan. ¹³Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives (OTRI), Osaka University, Suita, Osaka, Japan. ¹⁴Statistical Immunology, World Premier International Research Center Initiative (WPI), Immunology Frontier Research Center (IFReC), Osaka University, Suita, Osaka, Japan. ¹⁵Department of Genome Informatics, Graduate School of Medicine, the University of Tokyo, Bunkyo-ku, Tokyo, Japan. ¹⁶Japan Agency for Medical Research and Development – Core Research for Evolutional Science and Technology (AMED–CREST), Osaka University, Suita, Osaka, Japan. ¹⁷Center for Infectious Diseases for Education and Research (CiDER), Osaka University, Suita, Osaka, Japan. ¹⁸Center for Advanced Modalities and DDS (CAMA-D), Osaka University, Suita, Osaka, Japan. ¹⁹These authors jointly supervised this work: Atsushi Kumanogoh, Masayuki Nishide

✉ e-mail: kumanogoh@imed3.med.osaka-u.ac.jp; nishide@imed3.med.osaka-u.ac.jp